Studies in Neuroscience, Psychology and Behavioral Economics

Christoph Klein
Ulrich Ettinger *Editors*

Eye Movement Research

An Introduction to its Scientific Foundations and Applications

Springer

Studies in Neuroscience, Psychology and Behavioral Economics

Series Editors

This interdisciplinary book series addresses research findings and methods in the intersection of the fields of neuroscience, psychology and behavioral economics. At the core of the series is the biological basis of human decision making, including aspects of an economic, cognitive, affective and motivational nature. Contributions from fundamental and applied sciences are equally represented and cover a broad range of scientific work, reaching from mathematical modeling to scholarly introductions into new research disciplines and related methodologies and technologies (e.g. genetics, psychophysiology, brain imaging, computational neuroscience, animal modeling). The book series presents empirical and methodological scientific innovations bridging the gap from brain to behavior. It includes research works in neuroeconomics and neuro-information systems (e.g. human-machine interaction) as well as social, cognitive and affective neuroscience, which aim at enhancing the understanding of human decision making.

More information about this series at http://www.springer.com/series/11218

Christoph Klein · Ulrich Ettinger
Editors

Eye Movement Research

An Introduction to its Scientific Foundations and Applications

Springer

Editors
Christoph Klein
Department of Child and Adolescent Psychiatry
University of Freiburg
Freiburg, Germany

Department of Child and Adolescent Psychiatry
University of Cologne
Cologne, Germany

Ulrich Ettinger
Department of Psychology
University of Bonn
Bonn, Nordrhein-Westfalen, Germany

ISSN 2196-6605 ISSN 2196-6613 (electronic)
Studies in Neuroscience, Psychology and Behavioral Economics
ISBN 978-3-030-20083-1 ISBN 978-3-030-20085-5 (eBook)
https://doi.org/10.1007/978-3-030-20085-5

© Springer Nature Switzerland AG 2019
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The recording of eye movements is amongst the most versatile methods to study human information processing and behaviour. The range of applications is immense and so is the range of research questions that have been addressed with this method. The eye-tracking community seems to be ever growing and is well represented at conferences or in monographs or handbooks on the science of eye movements.

Despite the growing success of the science of eye movements, it occurred to the editors of this volume that one aspect is conspicuously missing: An introduction to the study of eye movements for beginners in the field, be it Masters or doctoral students, post-docs or professors. Therefore, beginners in eye movement research, but also those who wish to refresh their knowledge or deepen or broaden their understanding of eye movement research, are the target readership of this textbook.

World-leading experts in their respective fields of eye movement research have contributed to this textbook by providing chapters. Additionally, authors, expert reviewers, the editors and a large group of student reviewers have taken care to make the individual chapters didactically accessible to newcomers in this field.

Part I of this textbook is dedicated to the basics of eye movement research and explains the main types of eye movements relevant in this field, including fixation-related eye movements, different types of saccadic eye movements and smooth pursuit. This part also covers the relationship between eye movements and perceptual or attentional processes as well as reading.

Part II explains the methodological versatility of eye movement recordings as a stand-alone or combined technology. After an in-depth introduction to recording techniques, its possibilities and potential pitfalls, two chapters deal with innovative and advanced approaches to model and quantify eye movement data. Further chapters introduce the fields of pupillometric research, the investigation of eye movements in brain imaging, in the field or in research with animals.

Part III of this textbook then provides introductions to the main clinical applications of the eye movement technology in child, adolescent and adult psychiatry, neurology and pharmacology. These chapters illustrate how the investigation of eye movements can identify and help understand deficits and processing biases in psychiatric populations aid the diagnosis of functional impairments in neurological disorders or trace the actions of pharmacological agents in real time.

Finally, Part IV provides examples of industrial applications of the eye movement technology, including its use in marketing research and neuroeconomics, when improving the usability of websites and other products, when investigating vehicle control or as an interface between humans and computers.

Collectively, these chapters provide an insight into the breadth of eye movement research and help the newcomers in these fields in gaining expertise. This volume is the outcome of the work of more than a hundred scientists and students to whom the editors are immensely grateful.

Freiburg, Germany Christoph Klein
Bonn, Germany Ulrich Ettinger

Contributors

Robert G. Alexander State University of New York Downstate Health Sciences University, Brooklyn, USA

Dag Alnaes Department of Psychology, University of Oslo, Oslo, Norway

Isabelle Amado Department of Cognitive Remediation and Rehabilitation, Psychiatry, Centre Hospitalier Sainte Anne, Inserm, Paris, France

Nicola C. Anderson Department of Psychology, University of British Columbia, Vancouver, BC, Canada

Christiane Attig Department of Psychology, Cognitive and Engineering Psychology, Technische Universität Chemnitz, Chemnitz, Germany

Monica Biscaldi Department of Child and Adolescent Psychiatry, University of Freiburg, Freiburg, Germany

Walter F. Bischof Department of Psychology, University of British Columbia, Vancouver, BC, Canada

Giuseppe Boccignone Department of Computer Science, Universitá di Milano, Milano, Italy

Dario Cazzoli Gerontechnology and Rehabilitation Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

Brett A. Clementz University of Georgia, Athens, GA, USA

Jon Dodd Service Design and User Experience, Bunnyfoot Ltd., London, UK

Ulrich Ettinger Department of Psychology, University of Bonn, Bonn, Germany

Stefan Everling Department of Physiology and Pharmacology, Department of Psychology, Brain and Mind Institute, University of Western Ontario, London, Canada;
Robarts Research Institute, London, Canada

Susann Fiedler Max Planck Institute for Research on Collective Goods, Bonn, Germany

Tom Foulsham Department of Psychology, University of Essex, Colchester, UK

Thomas Franke Institute for Multimedia and Interactive Systems, Engineering Psychology and Cognitive Ergonomics, University of Lübeck, Lübeck, Germany

Mark W. Greenlee Institute of Experimental Psychology, University of Regensburg, Regensburg, Germany

Dan Witzner Hansen IT University Copenhagen, Copenhagen, Denmark

S. B. Hutton Sussex University, Brighton, UK

Jukka Hyönä Department of Psychology, University of Turku, Turku, Finland

Aulikki Hyrskykari Tampere University, Tampere, Finland

Chara Ioannou Department of Child and Adolescent Psychiatry, University of Freiburg, Freiburg, Germany

Sharna D. Jamadar Monash Institute for Cognitive and Clinical Neuroscience, School of Psychological Sciences, Monash Biomedical Imaging, Monash University, Melbourne, Australia;
Australian Research Council Centre of Excellence for Integrative Brain Function, Melbourne, Australia

Beth Johnson Monash Institute for Cognitive and Clinical Neuroscience, School of Psychological Sciences, Monash University, Melbourne, Australia

Kevin Johnston Department of Physiology and Pharmacology, Department of Psychology, Brain and Mind Institute, University of Western Ontario, London, Canada

Johanna K. Kaakinen Department of Psychology, Turku Institute for Advanced Studies, University of Turku, Turku, Finland

Tatiana Karpouzian Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Michael Kavšek Department of Psychology, University of Bonn, Bonn, Germany

Hubert Kimmig Schwarzwald-Baar Klinikum, University of Freiburg, Freiburg, Germany

Alan Kingstone Department of Psychology, University of British Columbia, Vancouver, BC, Canada

Christoph Klein Department of Child and Adolescent Psychiatry, University of Freiburg, Freiburg, Germany;
Department of Child and Adolescent Psychiatry, University of Cologne, Cologne, Germany

Marie-Odile Krebs Department of Cognitive Remediation and Rehabilitation, Psychiatry, Centre Hospitalier Sainte Anne, Inserm, Paris, France

Bruno Laeng Department of Psychology, University of Oslo, Oslo, Norway

Rebekka Lencer Department of Psychiatry and Psychotherapy, Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster, Germany

Päivi Majaranta Tampere University, Tampere, Finland

Susana Martinez-Conde State University of New York Downstate Health Sciences University, Brooklyn, USA

Jennifer E. McDowell University of Georgia, Athens, GA, USA

René Müri Department of Neurology, University of Bern, Bern, Switzerland

Thomas Nyffeler Clinic for Neurology and Neurorehabilitation, Luzerner Kantonsspital, Luzern, Switzerland

Jeff B. Pelz Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, USA

Nadine Petrovsky Department of Psychology, University of Bonn, Bonn, Germany

Jordan E. Pierce University of Georgia, Athens, GA, USA

Frederik Platten BMW Group, Munich, Germany

Pierre Pouget CNRS, Paris, France

Kari-Jouko Räihä Tampere University, Tampere, Finland

James Reilly Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

Philipp Reiter eye square GmbH, Berlin, Germany

Agnes Rosner Department of Psychology, Cognitive Decision Psychology, University of Zurich, Zurich, Switzerland

Matthias Rothensee eye square GmbH, Berlin, Germany

Yoanna Schulz-Zhecheva Department of Child and Adolescent Psychiatry, University of Freiburg, Freiburg, Germany

Divya Seernani Department of Child and Adolescent Psychiatry, University of Freiburg, Freiburg, Germany

Nikolaos Smyrnis Psychiatry Department, National and Kapodistrian University of Athens, Medical School, Eginition Hospital, Athens, Greece;
Laboratory of Cognitive Neuroscience and Sensorimotor Control, University Mental Health, Neurosciences and Precision Medicine Research Institute "COSTAS STEFANIS", Athens, Greece

Oleg Špakov Tampere University, Tampere, Finland

Andreas Sprenger Department of Neurology, Institute of Psychology, University of Lübeck, Lübeck, Germany

John A. Sweeney Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati, Cincinnati, USA

Benjamin W. Tatler School of Psychology, University of Aberdeen, Aberdeen, UK

Volker Thoma Psychology, University of East London, London, UK

Peter Trillenberg Department of Neurology, University of Lübeck, Lübeck, Germany

Bernd Weber Institute of Experimental Epileptology and Cognition Research, University Hospital Bonn, Bonn, Germany;
Centre for Economics and Neuroscience, University of Bonn, Bonn, Germany

Advisory Board

Antoniades, Chrystalina A., Nuffield Department of Clinical Neurosciences, Oxford University, Oxford, UK

Bracewell, Martyn, Schools of Medical Sciences and Psychology, Bangor University, Bangor, UK; Walton Centre for Neurology and Neurosurgery, Liverpool, UK

Bremmer, Frank, Department of Neurophysics and Center for Mind, Brain and Behavior—CMBB, University of Marburg, Marburg, Germany

Corr, Philip, School of Arts and Social Sciences, City, University of London, London, UK

Crawford, Trevor, Department of Psychology, Faculty of Science and Technology, Lancaster University, Lancaster, UK

Duchowski, Andrew, School of Computing, Clemson University, Clemson, SC, USA

Einhäuser-Treyer, Wolfgang, Institute of Physics, Chemnitz University of Technology, Chemnitz, Germany

Engbert, Ralf, Department of Psychology, University of Potsdam, Potsdam, Germany

Faiola, Eliana, Department of Psychology, University of Bonn, Bonn, German

Gilchrist, Iain D., School of Psychological Science, University of Bristol, Bristol, UK

Gooding, Diane C., Department of Psychology, University of Wisconsin-Madison, Madison, USA

Guerreiro, João, Business Research Unit (BRU-IUL), Instituto Universitário de Lisboa (ISCTE-IUL), Lisboa, Portugal

Häusler, Alexander Niklas, Centre for Economics and Neuroscience, University of Bonn, Bonn, Germany

Hodgson, Timothy, School of Psychology, University of Lincoln, Lincoln, UK

Holmes, Tim, Tobii Pro Insight UK Ltd.

Holmqvist, Kenneth, Institute of Psychology, University of Regensburg, Regensburg, Germany

Hong, Elliot, School of Medicine, University of Maryland, Baltimore, USA

Kasparbauer, Anna, Department of Psychology, University of Bonn, Bonn, Germany

Kiefer, Peter, Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland

Krajbich, Ian, Department of Psychology, Department of Economics, The Ohio State University, Columbus, USA

Lappe, Markus, Institute for Psychology, University of Münster, Münster, Germany

Laubrock, Jochen, Department of Psychology, University of Potsdam, Potsdam, Germany

Le Meur, Olivier, Institut de recherche en informatique et systèmes aléatoires, Université de Rennes, Rennes, France

Lehmann, Mirko, Department of Psychology, University of Bonn, Bonn, Germany

Lehtonen, Esko P., Traffic Research Unit, University of Helsinki, Helsinki, Finland

Mathôt, Sebastiaan, Cognitive Psychology Department, VU University Amsterdam, Amsterdam, The Netherlands

Meyhöfer, Inga, Department of Psychiatry and Psychotherapy; University of Münster, Münster, Germany

Oranje, Bob, Department of Psychiatry, University Medical Centre Utrecht, Utrecht, The Netherlands

Seernani, Divya, Department of Child and Adolescent Psychiatry, University of Freiburg, Freiburg, Germany

Spillmann, Lothar, Department of Neurology, University of Freiburg, Freiburg, Germany

Steffens, Maria, Department of Psychology, University of Bonn, Bonn, Germany

Talanow, Tobias, Department of Psychology, University of Bonn, Bonn, Germany

Tavakoli, Hamed R., Department of Computer Science, Aalto University, Espoo, Finland

Walker, Robin, Department of Psychology, Royal Holloway University of London, London, UK

Wiese, Eva, Department of Psychology, George Mason University, Fairfax, USA

Contents

Part III Applications in Clinical Settings

Part IV Applications in Industrial Settings

Part I
Basics

Chapter 1
Introduction to the Study of Eye Movements

Pierre Pouget

> *Clearly, the eyes not only allow us to see the world around us but they also present a window to the working of our mind.* (Eye position, what number you have in mind; Loetscher et al. *Current Biology*, 2010).

Contents

Eye movements reflect the visuo-motor output of complex and dynamical interactions between numerous regions of the brain. This distributed brain network, required for eliciting eye movements, provides a powerful research tool to investigate the working functions of one of the most complex organs: the brain. The examination of eye movements has also been shown as being precious during medical exams to reveal cues to detect abnormalities of various brain functions (Leigh and Zee, 2015). Therefore, the study of eye movements has been for decades a source of information to psychologists, neurobiologists, clinicians as well as most of visual scientists (Pouget, 2015; Schütz, Braun, & Gegenfurtner, 2011).

Ecologically the exchange of information between the different areas of the brain gives to humans amazing abilities to perform activities as diverse as moving their eyes while hitting a tennis ball, reading novels or even a scientific textbook on eye movement research. Somehow, and without any conscious effort, our eyes transmit all the information that we need to behave adaptively in the world around us (Buswell, 1922, 1935, 1937; Yarbus, 1967). The development of eye tracking systems in the last few years has made it possible to record where we are looking for during all

---

P. Pouget (✉)
UMR 975, CNRS, Paris, France
e-mail: pierre.pouget@upmc.fr

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_1

these various active tasks, and to infer what information our eyes are effectively transmitting to our brain. The eyes and their positions give some insight into the nature of the systematic selection of the brain, and scientists are extracting these variables to predict choices, decisions with some reliable confidence. The purpose of this introduction is to illustrate briefly some key terms and notions that are going to be described in great details in the different chapters of this textbook. First, we will question why animals and primates in particular often produced numerous eye movements. Thus, we will very briefly introduce some terms associated with the measures and analysis of eye movements. All these questions and terminology will be described in great details in the following chapters.

1.1 Why Do We Need Our Eyes to Move After All?

To answer this simple question, most psychologists or neuroscientists point to the drastic change of density of photoreceptors in the retina (e.g. Adler & Fliegelman, 1934). It is true that early in the processing of visual stream a potent anatomical constrain is present in the retina. Two classes of photosensitive cells are present with massive variations of spatial density. Classically, these two classes of cells have been distinguished by their shapes (cones and rods) and the type of photopigments they contain. While the rods are excessively sensitive to light, the cones have diverse responses mapping to particular spectral content of light. Finally, as previously mentioned, the distribution of these two types of photoreceptors is uneven in the retina: the rods are present throughout the retina with the exception of the small central region of the retina (~0.5 mm in diameter), the fovea. At the level of fovea, the concentration of cones is very high, and their density decreases very strongly outside 10 degrees of visual angle. Therefore, it is only at the level of the fovea that high visual acuity is allowed. So, to have a precise vision of the world and to be able to recognize and/or identify all the elements that surround us, the eye movements are essential. Indeed, it is by moving our eyes (and our head), that the image of the objects which interest us is placed at the level of the fovea and that we can perceive our environment in all its complexity and color. The vision is therefore an active phenomenon whose eye movements are one of the main actors. *Why do we need our eyes to move after all?* We might respond that *we move our eyes to see the world but also because to accurately see the world we must move our eyes.* The biologist Jacob von Uexküll in 1926, was the first to describe this principle of a sensori-motor circular pattern, which he identified not only in humans but in many animal species. This principle was later called the "Gestalt Cycle" (Gestalkreis) by Viktor von Weizsäcker in 1950, a neurologist who recognized in it this loop the indissoluble union of perception and movement in the nervous system. In more recent years, that principle and its operational aspects have been characterized as the action-perception cycle or, where, the emphasis is not so much on motor behavior as on cognition, simply as the perception cycle. In complex organisms, this cycle allows the representations of current action to feedback to sensory structures to modulate further sensory inputs. In the recent field of neuroscience, this conception of feedback loops is at the origin of what at higher neural stages have been referred as efferent copies and/or corollary discharge (Dore-Mazars, Pouget, & Beauvillain, 2004; Orquin & Mueller Loose, 2013; Pouget et al., 2009; Torralba, Oliva, Castelhano, & Henderson, 2006).

1.2 How Do We Move Our Eyes?

Over the past 90 years, eye movements have been used to gain insight into disorders ranging from muscular dystrophy to schizophrenia and autism (see chapters by Smyrnis et al. and Klein et al. in this volume, as well as Klein & Ettinger, 2008). At the level of molecular genetics, distinctive disorders of eye movements have pointed to the nature of involvement in a broad range of genetic disorders and have served as an index to evaluate new treatments. Eye movements are of two main types: those that stabilize gaze and so keep images steady on the retina and those that displace and so redirect the line of sight to a new object of interest. Certain anatomical circuits make distinctive contributions to each functional class of movements. An understanding of the properties of each functional class of eye movement will guide the physical examination, and subsequent neural substrate involved. Although there is a shared substrate for the combined eye movements to control direction of gaze in three-dimensional space during our natural activities, testing each system in isolation helps identify specific processes. Neurophysiological evidence indicates that the position command is generated from the velocity by integration with respect to time. Four types of conjugate eye movements (vestibular, optokinetic, saccade and pursuit) require velocity coded and position coded components. The fifth type of movement which rotate the eyes in different direction (vergence) appears to have a separate coding (Dodge, 1903; Leigh & Zee, 2015).

More precisely, the first movement represents a response to brief rotational and linear movements: namely the vestibulo-ocular reflex (VOR). During head movements, the vestibular system sends information to help to keep accurate vision. The vestibular system helps to optimize vision during head movement in particular during locomotion. The semicircular canals of the vestibular labyrinth signal how fast the head is rotating and the oculomotor system responds to this signal by rotating the eyes in an equal and opposite velocity. This stabilizes the eyes relative to the external world and keeps visual images fixed on the retinal. This reflex is almost always active and without it we would be unable to see much of anything due to constant smear of the retinal image. Because VOR is a particularly crucial correction to keep good vision during our displacement this reflex has evolved to became the fastest reflex of human body (less than 8 ms).

The second movement, the optokinetic reflex (OKR) that is a visual drive that can maintain by compensatory slow-phase eye movement during sustained rotation or background visual motion. Unlike the VOR, this reflex requires a visible retinal image whereas the VOR works in total darkness. Why do we have both OKR and

VOR when they both respond to similar condition of body movements. Both reflexes are important because the OKR supplements the VOR in several ways. The VOR responds to acceleration and deceleration but not to constant velocity. In contrast, OKR responds to constant retinal image velocity caused by constant body rotation or translation. Basically, the VOR controls initial image stabilization and OKR maintains the stabilization. OKR also compensates for a damaged vestibular apparatus. People who have had infections of their inner ear often complain of motion sickness in cars and boats. The best thing for them to do is look out the window at the horizon and allow the OKR a chance to keep their eyes stable with respect to gravity.

The third movement, the voluntary saccades are quick eye movements that are generated under a broad range of conditions. Normal saccades are fast, brief and accurate to avoid any interference with vision (see chapter by Pierce et al., in this volume). Several parameters of this displacement have been investigated in the literature. First, it's amplitude. The amplitude of the saccades is calculated as the difference between the final position and the initial position of the eyes, respectively before and after the saccade. The gain of a saccade can give information on its accuracy. It corresponds to the ratio between the amplitude of the saccade and retinal error (distance between the eyes before the saccade and the target to be reached). A gain of 1 indicates that the saccade has allowed the eyes to reach their target, it is said that the saccade is normometric. A saccade with a gain less than 1 is hypometric whereas a saccade with a gain greater than 1 is hypermetric (the eyes have exceeded the target). Saccades are precise movements, with a gain of around 0.95 (Becker, 1989), that is to say, they have a slight tendency to be hypometric (highlighted in 1904 by Dearborn (1904), then by Becker and Fuchs (1969) and many other authors).

A second set of parameters can be extracted according to the duration and peak speed of the movements. In this context, the duration of a saccade is defined as the difference between the time at the end of the movement and the time at the beginning of the saccade. The peak speed corresponds to the maximum speed of the eyes during the movement. There is a stereotypical relationship between the duration of saccades and their amplitude and between peak speed and their amplitude called main sequence (Bahill, Clark & Stark, 1975). Jumps are the fastest and shortest movements duration that an organism can produce. These high performances and the main sequence reflect the characteristics of the saccadic generator located at the level of the brainstem.

If the duration of the displacement of the eye has been largely examined, the latency corresponding of the time between the instruction to make a saccade (start signal, appearance of a new target, end of the previous saccade, etc.) and the actual start of the movement has quickly emerged has one of the major sources of analysis in the recent decades (Carpenter & Williams, 1995). The latency of the saccades is generally of the order of 150–500 ms and is used as one of the criteria to classify saccades according to two categories. The first category comprises the saccades that have the shorter latencies and are classified as reactive saccades (also called reflex jerks). These eye movements are the least common in everyday life, yet these are the ones most studied in the laboratory because simple protocols and instructions can induce them in monkeys, healthy subjects and patients. They are caused by the sudden appearance of a new object in the peripheral visual field of an individual. They also allow you to follow a fast-moving object, such as a tennis ball or a baseball. In the controlled environment of the laboratory, this type of saccades can be induced by instructing a subject to fixate a point presented on a computer screen. After a random delay, this point disappears and reappears simultaneously in the peripheral field of the subject. In response to the sudden appearance of this new target, the subject produces a saccade towards this target with a latency of the order of 150–200 ms. The so-called express saccades are included in the category of reactive saccades, which are triggered even earlier than conventional reactive saccades. Indeed, these saccades have latencies of the order of 70–90 ms in the monkey (Fischer & Boch, 1983) and 90–120 ms in humans (Fischer & Ramsperger, 1984).

The second category of saccades existing is that of voluntary saccades. These saccades are differentiated from reactive saccades by higher latencies, usually greater than 250 ms (Carpenter & Williams, 1995; Fischer & Boch, 1983; Fischer & Ramsperger, 1984) and by a slightly greater accuracy (Lemij & Collewjin, 1989). Voluntary saccades can be broken down into different subcategories. The saccades of exploration are the most frequent in our daily lives. They are triggered according to the intentions of an individual and allow, as indicated by their name, the exploration of objects of interest already present in our environment.

The fourth type of movement, with the functional constraints of the fovea also came the need to track a moving object smoothly. Because of the delay inherent to the visual system, the brain generates predictive smooth-pursuit movements to maintain smooth eye movements on the moving target. Primates are also capable to follow an object that moves in the visual field. What is named a pursuit occurs when the eyes continue to jointly fix the same point as it moves on the retina (see for details, see chapter by Lencer et al., in this volume). The movement of a target on the retina may be due either to the actual movement of the target or to the movement of the observer (locomotion, rotation of the head, etc.). This type of movement is slow and cannot be initiated voluntarily in the absence of displacement on the retina. The goal is to maintain a point on the fovea as it moves. Eye tracking actually reduces the movement of a target on the retina but without ever completely abolishing it (Kowler, 1990). Since the movement of the eyes is relatively slow, a high spatial resolution is maintained and it is always possible to extract information from the pursued target. Of course, the movement of the eyes can be done in conjunction as in disjunction (Yarbus, 1967).

Finally, the fifth, the vergence is permitting primates with frontal vision to direct their two foveas at one object of interest. This requires the disjunctive or vergence movements of the eyes (moving the eyes in opposite directions). The fusional vergence movements are generated in response to disparity between the locations of images of a single target on the retina of each eye. This fusional vergence is accompanied by the phenomenon of accommodation of the lens that allows the focus on the object and pupillary constriction that increases the depth of field and improves the sharpness of the retinal image.

An extra behavior that might not directly referred to an eye movement per se is fixation. Fixations during which the eye remains relatively still and the visual system extracts detailed information around the fixation point. By definition fixation it is not strictly speaking a type of eye movement since fixation is the activity of the eyes when they remain more or less long positioned on the same point. It is therefore the activity of the eyes when they do not move. At this point, the target of interest of the environment is reflected on the fovea of both eyes and can therefore be analyzed with a maximum of spatial discrimination (Yarbus, 1967). However, even during a fixation, the eyes continue to move according to three types of movements: drift, tremor and micro-saccades (for details see chapter by Alexander and Martinez-Conde, in this volume). Briefly, micro-saccades, are saccades of very small amplitude, similar for both eyes. They can have a minimum dimension of 2–5 min of angle and occur unintentionally. These movements seem to have as main utility to ensure a slight, but constant, variation of light stimulation on the retina. It has indeed been observed a cessation of perception during the continuous projection of the same stimulation on the retina. Since photoreceptors are always used in the same way (with the same intensity), physiological habituation ensues, and the information is no longer transmitted to the visual cortex (Kowler, 1990; Yarbus, 1967).

All these eye movements are amongst the best-studied forms of motor control and are widely used as research or diagnostic tools in psychology, neurosciences and medicine (see chapters by Klein et al., Smyrnis et al., and Müri et al., in this volume). For many years, animal models for studying eye movements in health and disease have been developed in species ranging from fish to Rhesus monkey making it possible to relate eye movements to biophysical properties of ion channels and neurotransmitters. At the level of cognition, eye movements have provided insights into faculties such as memory, decision-making, task switching as well as psychotic thought disorders. Indeed, as functional imaging has demonstrated, activity related to eye movements can be found in almost every part of the brain. This should come as no surprise because we are creatures that depended on clear vision and we need to focus our attention in order to make prompt correct responses to what is happening around us. Eye movements both facilitate and reflect this central role for vision in survival. The singular value of studying eye movements stem from certain advantages that make them easier to interpret than other types of movements. The first is those eyes movements are essentially restricted to rotations of the globes (linear displacements) are negligible. This facilitates precise measurement, which is a prerequisite for quantitative analysis. A second advantage is the relatively simple relationship between the discharge of moto-neurons and rotation of the eyes, which appear mainly attributable to the simplifying effect of the mechanics of the orbital tissues. Third, different classes of eye movements can be distinguished on the basis of how they aid vision, their physiological properties and their anatomical substrates. Fourth, many abnormalities of eye movements are distinctive and often point to a specific pathophysiological disturbance. Finally, eye movements are readily accessible to clinical observation and systematic examination.

As we have seen in this brief introduction, one advantage of studying eye movements is that their simplified properties and their identified neural substance lend them to formulating quantitative hypotheses, mathematical models that make testable predictions. Thus, the classic hypothesis testing method of science can easily be applied to eye movements. This introductory textbook provides an overview of the history and the current status of research on eye movements and introduces the reader to relevant findings, theories and methods of their measurement, their underlying neural and cognitive control, their development, their alterations in disease and their study in applied contexts.

References

Adler, F. H., & Fliegelman, M. (1934). Influence of fixation on the visual acuity. *Archives of Ophthalmology, 12*, 475–483.

Bahill, T. A., Clark, M. R., & Stark, L. (1975). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences, 24*(3–4), p191–p204.

Becker, W. (1989). The neurobiology of saccadic eye movements. Metrics. *Reviews of Oculomotor Research, 1989*(3), 13–67.

Becker, W., & Fuchs, A. F. (1969). Further properties of the human saccadic system: Eye movements and correction saccadic with and without visual fixation points. *Vision Research, 9*(1969), 1247–1258.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago, IL: University of Chicago Press.

Buswell, G. T. (1935). *How people look at pictures* (pp. 137–155). Chicago: University of Chicago Press. Hillsdale, NJ: Erlbaum.

Buswell, G. T. (1937). *How adults read*. Chicago, IL: University of Chicago Press.

Carpenter, R. H., & Williams, M. L. (1995, September 7). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377*(6544), 59–62.

Dearborn, W. F. (1904). Retinal local signs. *Psychological Review, 11*(4–5), 297–307.

Dodge, R. (1903). Five types of eye movement in the horizontal meridian plane of the field of regard. *American Journal of Physiology, 8,* 307–329.

Dore-Mazars, K., Pouget, P., & Beauvillain, C. (2004). Attentional selection during preparation of eye movements. *Psychological Research, 69*(1–2), 67–76.

Fischer, B., & Boch, R. (1983, January 31). Saccadic eye movements after extremely short reaction times in the monkey. *Brain Research, 260*(1), 21–6.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: Extremely short reaction times of goal directed eye movements. *Experimental Brain Research, 57*(1), 191–195.

Klein, C., & Ettinger, U. (2008). A hundred years of eye movement research in psychiatry. *Brain and Cognition, 68*(3), 215–218.

Kowler, E. (1990). The role of visual and cognitive processes in the control of eye movement. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes* (pp. 1–70). Amsterdam: Elsevier.

Leigh, J.R., & Zee, D.S. (2015). *The neurology of eye movements* (5th ed.). Oxford: Oxford University Press.

Lemij, H. G., & Collewjin, H. (1989). Differences in accuracy of human saccades between stationary and jumping targets. *Vision Research, 29,* 1737–1748.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206.

Pouget P. (2015, February). The cortex is in overall control of 'voluntary' eye movement. *Eye (Lond), 29*(2):241–245.

Pouget, P., Stepniewska, I., Crowder, E. A., Leslie, M. W., Emeric, E. E., Nelson, M. J., Schall, J. D. (2009, May 22). Visual and motor connectivity and the distribution of calcium-binding proteins in macaque frontal eye field: Implications for saccade target selection. *Frontiers in Neuroanatomy, 3*, 2.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11,* 9.

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: The role of global features in object search. *Psychological Review, 1134,* 766–786.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Chapter 2
Saccades: Fundamentals and Neural Mechanisms

Jordan E. Pierce, Brett A. Clementz and Jennifer E. McDowell

Contents

Abstract Saccadic eye movements allow humans to explore the visual environment, quickly moving the fovea and attention to points of interest for detailed visual processing. Despite these rapid changes in visual input, the brain maintains a stable visual representation through saccadic suppression of blurred input and predictive remapping of receptive fields. In the laboratory, visually-guided saccades are driven

---

J. E. Pierce (✉) · B. A. Clementz · J. E. McDowell
University of Georgia, Athens, GA, USA
e-mail: jordan.pierce@unige.ch

J. E. McDowell
e-mail: jemcd@uga.edu

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_2

primarily by the onset of new stimuli, while volitional saccades are driven by instructions or internal motivation. The distinction between these two classes of saccades is not definitive, but used to illustrate different cognitive loads that fall upon a spectrum based on the balance of visual input and cognitive control processes required to perform simple behavioral tasks. Visually-guided prosaccades depend upon a simple stimulus-response mapping whereby the eye movement is directed toward the peripheral stimulus quickly and accurately. Decision-making processes nonetheless slow the latencies of these basic visual-motor responses to ensure selection of an appropriate saccade target. To dissect the time course of saccade programming, researchers can include a secondary stimulus step or distractor stimulus in the saccade paradigm that alters response characteristics based on additional cognitive processes that occur. Attentional processes, for example, have a strong influence on the way in which participants respond: A pre-trial location cue attracts attention and reduces saccade latencies at that location and increases latencies at uncued locations. If attention is re-directed to central fixation, however, latencies to the originally cued location will be slower than other parts of the visual field, an effect known as inhibition of return. Furthermore, attentional preparation may contribute to the production of fast latency "express" saccades that occur in gap paradigms. The offset of fixation warns the participant of the upcoming stimulus appearance and disengages fixation neurons, allowing saccade-related activity to increase in expectation of the peripheral stimulus. On the other end of the spectrum, "volitional" forms of saccades are elicited via instructional manipulations that require participants to suppress an immediate saccade to a visual stimulus or otherwise rely upon endogenous goals to produce a saccade at a certain time. Antisaccades, ocular motor delayed response saccades, and predictive saccades all fall into this category. Antisaccade tasks require a saccade to the mirror image location of a visual stimulus, depending upon suppression of the visually-driven saccade tendency and a sensorimotor transformation of spatial coordinates to generate the volitional saccade successfully. In this task, the internal saccade programs representing the antisaccade and visually-guided saccade responses are modelled as competing activations racing toward a threshold for motor generation. An ocular motor delayed response task includes a delay period between the presentation of the peripheral stimulus and the time when participants are instructed to respond. Spatial memory processes thus are needed to maintain the target location after the visual stimulus itself is extinguished. Typically, both tasks result in slower latencies, more errors, and poorer spatial accuracy than their visually-guided counterparts. Finally, a predictive saccade task consists of a visual stimulus alternating between two locations at regular intervals. After a few trials, participants learn the spatiotemporal pattern and begin to anticipate the movement, generating saccades to the target location based on internal timing mechanisms before the visual stimulus appears. The aforementioned saccade tasks are supported by a well characterized and widespread neural saccade circuitry. This circuitry includes occipital cortex, posterior parietal cortex, frontal and supplementary eye fields, thalamus, basal ganglia, cerebellum, and the superior colliculus. Visual input initiates neural activation in occipital cortex which spreads to parietal and frontal regions for attentional processing, visuospatial calculations, and saccade motor preparation. Volitional saccade tasks show greater strength and/or extent of activation in this circuitry, and recruitment of new regions, including prefrontal cortex and anterior cingulate cortex, to support additional cognitive control requirements, such as inhibition, working memory, and motor learning. Along with input from the cerebellum, thalamus, and basal ganglia, cortical signals ultimately are integrated within the retinotopic maps of the superior colliculus where a single target location is selected and a motor program triggered via the brainstem ocular motor nuclei. Numerous neuroimaging studies in healthy humans using fMRI, EEG, MEG, and PET report activations in these brain areas during saccades and support findings from lesion studies and intracranial recordings in non-human primates. This consistency makes the saccade network a valuable model for studying cognitive control and understanding how the visual system integrates sensory input with internal goals to create a stable visual representation that guides efficient ocular motor behavior.

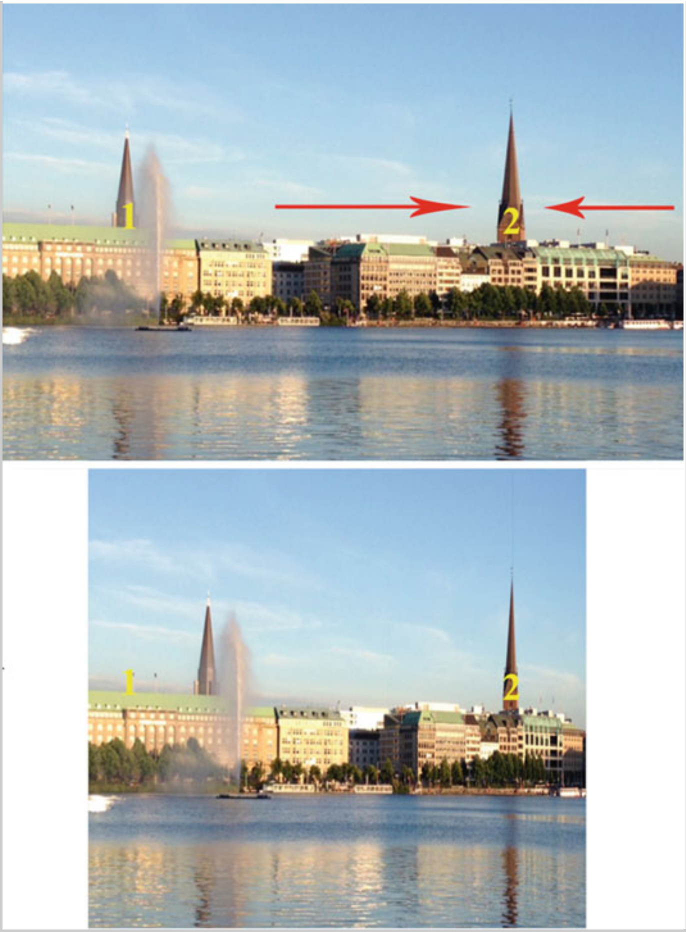
2.1 Introduction and Learning Objectives

Saccades are a class of eye movements that involve a rapid shift of gaze generated to a location of interest and allow individuals to explore the visual environment. Humans make thousands of these ballistic eye movements every day during activities such as driving, reading, and conversing with a friend. Usually one does not consider these movements or feel as if their completion requires effort. Successful execution of this frequent behavior (2–3 times per second), however, requires a complex series of neural processes. The brain constructs a stable perception of the visual world (Box 1) so one may feel that the eye is moving steadily while reading this line of text when, in fact, it is making many quick jumps (*saccades*) and pauses (*fixations*) on individual words along the way (see chapter by Hyönä and Kaakinen, in this volume). Visual acuity (the ability to see fine detail) is greatest within the center of the retina at the fovea and saccades act to accurately bring a salient stimulus onto this area for a period of fixation. During fixation, the eyes are not perfectly still but make many “microsaccades” (small, rapid movements covering <1° of visual angle; see chapter by Alexander and Martinez-Conde, in this volume) to prevent loss of visual input due to neural adaptation (a weakened response to constant stimulation). This allows the visual system to extract detailed information about the object on the fovea, build a mental representation of space, and guide behavior.

Box 1: Visual Stability During Saccades

The visual world is perceived as a stable, cohesive place despite the occurrence of multiple saccades per second that redirect the center of vision from one small patch of space to another. The brain builds a mental representation based on this ever-changing input without disrupting ongoing cognitive functioning. The visual system therefore must contain a means of blocking the distorted visual input during saccades (and other events such as blinks, Stevenson, Volkmann, Kelly, & Riggs, 1986; Volkmann, 1986) from being perceived, and subsequently reconstructing the visual representation once fixation is established. The first of these mechanisms is achieved in part through saccadic suppression, a dampening of the visual signal from just before to shortly after eye movements occur (Wurtz, 2008), noted experimentally by Dodge as early as 1898 in reading lines of text (Dodge, 1900; Erdmann & Dodge, 1898; Volkmann, 1986). More recent work has shown that suppression occurs primarily for the low spatial frequency magnocellular pathway that is sensitive to luminance differences and transmits signals to motion processing regions; high frequency and color information is perceived throughout saccade execution (Burr, Morrone, & Ross, 1994). This suppression appears to be instantiated quite early in visual processing, possibly within the lateral geniculate nucleus of the thalamus (Burr et al., 1994), although other evidence indicates cortical involvement in suppression as well (Ibbotson & Krekelberg, 2011). Since the relative positions of objects typically are maintained across the brief periods during saccades, the brain assumes that objects remain stable within the mental visual representation (Matin & Pearce, 1965), although certain transaccadic changes to targets that violate the visual memory trace are detected (Henderson & Hollingworth, 2003). The visual information received before/after a saccade then may mask the distorted input during the movement, creating the illusion of constant visual space (Matin, 1974; Volkmann, 1986).

In addition to weakening the distorted visual input during saccades, the visual system also must account for the changing positions of objects on the retina (Klier & Angelaki, 2008). A reconfiguration of visual receptive fields (the portion of space to which a neuron responds most sensitively) is necessary for updating retinotopic (based on retinal organization) maps when the eyes move. This reconfiguration happens in anticipation of an intended saccade and is known as predictive remapping or spatial updating. Just before a saccade is generated, a neuron responds not only to its normal receptive field, but also to the region where its receptive field will fall after the saccade is completed (Heiser & Colby, 2006; Wurtz, 2008). This mechanism helps establish continuity in the visual signal by establishing that the change in retinal stimulation occurred because of a saccade and not external motion when a single neuron gets matching input from before and after the saccade. The receptive fields of neurons in the lateral intraparietal (LIP) area, frontal eye fields (FEF), and superior colliculus all are remapped in this manner prior to saccade production (Duhamel, Colby, & Goldberg, 1992; Heiser & Colby, 2006; Joiner, Cavanaugh, & Wurtz, 2013; Merriam, Genovese, & Colby, 2003; Umeno & Goldberg, 1997; Walker, Fitzgibbon, & Goldberg, 1995) based on the corollary discharge of the planned saccade (Wurtz, 2008).



In conjunction with predictive remapping of receptive fields there is an apparent compression of visual space toward the saccade target, which causes a spatial mislocalization of briefly presented stimuli (Honda, 1989; Matin & Pearce, 1965). This compression leads to multiple adjacent stimuli being perceived as a single unit and natural scenes appearing shortened parallel to the saccade vector (Morrone, Ross, & Burr, 1997; Ross, Morrone, & Burr, 1997). Ross, Morrone, and Burr demonstrated that when participants are instructed to saccade from point 1 to point 2 in a flashed scene like that shown in the upper left, it appears compressed toward the saccade target (arrows) so that they perceive something akin to the lower image. Further work by this group showed that time is compressed in a similar fashion such that the perceived interval between two stimuli is shortened when presented around the time a saccade is generated. Curiously, the order of stimuli may even be reported as reversed if the interval between the two is less than 100 ms (Morrone, Ross, & Burr, 2005). These carefully designed experiments reveal quirks in the neural mechanisms designed to maintain visual consistency across saccade execution and demonstrate that what one perceives as a stable visual world is only an internal reconstruction built from both sensory input and cognitive manipulation of this input.

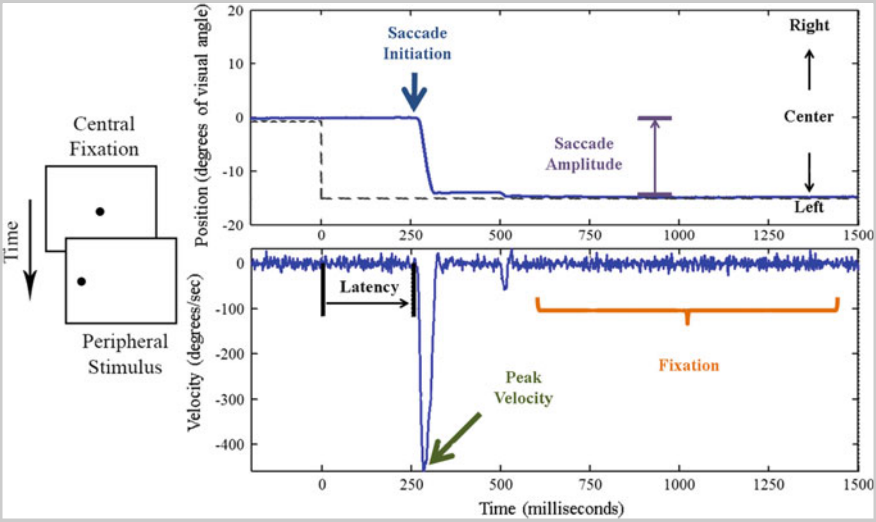
Research on the ocular motor system and the cognitive mechanisms underlying saccade programming has revealed a hierarchical distinction between so-called visually-guided saccades triggered by an external stimulus and voluntary (or endogenous) forms of saccades initiated by an internal goal. Visually-guided saccades are stimulus-driven and occur rapidly and frequently in everyday situations. Volitional saccades are endogenously generated and rely more heavily upon cognitive processes such as attention, inhibition, and working memory, as when searching for a set of keys on a cluttered desktop. The difference between these types of saccades is more a matter of degree than a categorical division, as internal cognitive motivations and decision-making processes influence both visually-guided and volitional saccades. In natural viewing conditions, exploratory saccades rely upon visual cues and volitional motivation to search and respond to the visual world in an efficient manner.

Experimentally, saccades are evoked by simple visual, auditory, or tactile stimuli. A basic saccade task consists of the presentation of a central fixation point followed by the presentation of a peripheral stimulus that serves as a target (see Box 2). Saccade targets may be presented horizontally (right or left of central fixation), vertically (above or below central fixation) or in a circular array. Participants are instructed to look toward the stimulus as quickly and accurately as possible [although variation of the instruction set can alter response patterns (Hutton, 2008; Mosimann, Felblinger, Colloby, & Muri, 2004)]. Methodologies developed to track movement of the eyes allow for precise quantification of numerous features of the saccadic response including: latency (time from onset of the peripheral stimulus to onset of saccade), accuracy/gain (amplitude of eye position relative to stimulus position), duration (amount of time the eye is in motion) and velocity (speed of eye movement; see Box 2 and chapter by Boccignone, in this volume). These characteristics provide detailed information about the nature of the basic saccade response. Visually-guided saccade paradigms and relevant cognitive processes are reviewed in Sect. 2.3 of this chapter.

The cognitive demands of this simple saccade task can be modified easily by changing the instruction set, creating a paradigm that no longer relies primarily on a visually-guided response, but instead requires a greater level of volitional control. For example, in a so-called "antisaccade" task participants are instructed to look away from the target toward the mirror image location of the peripheral stimulus (i.e., opposite direction, same distance from center). Although this is reasonably easy to perform, occasionally reflex-like glances are directed toward the peripheral stimulus. Such antisaccade "errors" are considered a failure of cognitive control linked to the ability to inhibit a prepotent response to the stimulus (the *visual grasp reflex*—Hess, Burgi, & Bucher, 1946), and can be an indicator of frontal lobe impairments and/or some psychiatric disorders (see chapter by Smyrnis et al., in this volume; Everling & Fischer, 1998). An alternate modification of a saccade task changes the paradigm into an ocular motor delayed response/memory-guided saccade task: here the participant is instructed to withhold making a saccade to a briefly appearing peripheral stimulus, remember its location during a delay period (on the order of seconds), and ultimately generate a saccade to the now unmarked location. Voluntary saccades also can be directed to a peripheral location on the basis of a symbolic cue such as an arrow (Henik, Rafal, & Rhodes, 1994; Walker, Walker, Husain, & Kennard, 2000). A further example is found in the predictive tracking paradigm, wherein the visual stimulus periodically alternates between two peripheral locations. Participants learn to anticipate the rhythmic movement of the stimulus and eventually generate saccades with extremely short latencies, potentially enabling foveation of the target location before the stimulus appears.

In general, volitional saccades are characterized as those which utilize cognitive control to: manage the processes involved in analyzing salient sensory information; select the correct motor response given current goals; maintain spatial information in memory; inhibit distracting stimuli; and/or generate an eye movement to the desired location at the appropriate time. Execution of a saccade may depend upon any combination of these processes as well as many others. Untangling the contribution of each of these processes and trying to isolate the mechanisms through which they function in the ocular motor system is an important goal of eye movement research. While the concept of cognitive control itself may be applied to many domains, the saccadic system serves as a basic and straightforward model that has been studied effectively in the laboratory with human participants. The reliability of saccade measures also allows these paradigms to be useful for investigation of individual differences or with clinical populations (Klein & Berg, 2001; Klein & Fischer, 2005b; Li et al., 2012; Meyhöfer, Bertsch, Esser, & Ettinger, 2016). Saccade metrics are largely stable within and between testing sessions (with the possible exception of antisaccade directional error rate which may be reduced after repeated performance; Dyckman & McDowell, 2005; Ettinger et al., 2003; Smyrnis et al., 2002), and thus can be used to assess trait-like differences in cognitive control or sensorimotor reactivity (Clementz et al., 2016). Volitional saccade tasks and the cognitive control functions they represent are discussed in Sect. 2.4, including antisaccade, ocular motor delayed response, and predictive tracking tasks.

Box 2: Saccade Behavioral Response Metrics



The first scientific reports of monitoring saccadic movements relied on personal accounts or simply observing another person's eyes. Apparatus were later developed to accurately record eye movements and subsequently quantify the number, duration, and velocity of saccades; over the last hundred years, numerous technological improvements to these methods have been made. Today, many laboratories utilize digital infrared camera systems to record pupil movement or the corneal reflection automatically and even provide immediate feedback to participants on their performance; other methods include scleral coils and electro-oculograms (see chapter by Hutton, in this volume). Modern digital eye tracking can sample eye position with accuracy on the order of milliseconds (250–1000 Hz sampling rate), which provides abundant information on the nature of each individual saccade. The figure shows an example of a single trial in a visually-guided saccade task where a participant is instructed to follow the dot when it moves (to the left). The upper graph displays the position of the stimulus (dashed line) that moves from center fixation to the periphery and the participant's gaze (blue line) over the course of the trial; the lower graph shows the velocity of eye movements made during this window. Using such information recorded during a task, researchers can "score" the trial as to whether the response was made in the correct direction (right or left based on the appearance of the peripheral stimulus—which is illustrated as up or down in the figure), then calculate the latency of the response (difference between the peripheral target onset and initiation of the saccade, here about 250 ms) and the saccade amplitude (here 15 degrees of visual angle to the left). The time of saccade initiation and its duration are typically defined by the velocity (the first derivative of the position function) of the movement exceeding a given threshold (around 20–40°/s) that distinguishes saccades from other slower eye movements such as smooth pursuit (see chapter by Lencer et al., in this volume). Saccadic gain (eye position/target position * 100) is generally quite accurate (>90%), although with a slight tendency to undershoot the intended target (Becker, 1989; Pélisson, Alahyane, Panouillères, & Tilikete, 2010). Additional characteristics of the saccade such as peak velocity may also be analyzed depending on the researcher's specific questions for the study.

Although experimental manipulations modulate saccadic response characteristics, similar neural mechanisms underlie the production of saccades across the diverse range of paradigms. The neural circuitry involved in saccade generation has been well characterized in humans using modern brain imaging techniques and extensively studied with physiological recordings in non-human primates. Studies of patients with cortical lesions or psychiatric disorders also have added to the understanding of the functions of these regions. This chapter will focus on findings from healthy humans across multiple saccade paradigms, as well as from non-invasive neuroimaging techniques such as functional magnetic resonance imaging (fMRI, see chapter by Jamadar et al., in this volume), electro-encephalography (EEG), and magneto-encephalography (MEG). The final section of this chapter will discuss neuroimaging studies using these techniques that have identified a widespread ocular motor network that encompasses a number of cortical and subcortical regions involved in both visually-guided and volitional saccade tasks. To anticipate the major points of this discussion, these regions include visual cortex, posterior parietal cortex, frontal and supplementary eye fields, prefrontal cortex, anterior cingulate cortex, cerebellum, thalamus, basal ganglia, and superior colliculus, and they perform numerous visual attention, motor planning, and cognitive control operations to ensure effective behavior. This circuitry may be activated to a greater or lesser strength or extent depending on the type of saccade task performed and the cognitive demands required. Although considerable knowledge has been generated by the study of the saccadic system thus far, there remain many open research questions that will further refine theories on the properties and interactions of eye movement neural circuitry. Prior to exploring the present state of cognitive research on saccades, however, the origins of scientific inquiry in this field will be discussed in the following section.

Learning Objectives

- *Identify various types of saccades and how they may be elicited and analyzed in a research laboratory.*
- *Discuss the cognitive processes involved in saccade performance.*

- *Compare the mechanisms underlying different saccade paradigms.*
- *Detail the neural correlates of saccade generation in human subjects.*

2.2 Historical Annotations

Long before empirical techniques existed for studying saccades, some astute individuals observed that the eyes jumped from one relevant visual feature of the world to the next rather than moving smoothly across all aspects of a scene (Dodge, 1916; Hering, 1879; Javal, 1879). It was not until the last half of the 19th century, however, that researchers began to apply the scientific method systematically to recording these eye movements and developed a means of accurately measuring their characteristics (see Wade, Tatler, & Heller, 2003). Research by Lamare (reported by Javal) and by Hering focused on the role of saccades in reading, utilizing methodologies where tubes were affixed to the face in order to detect the movement of the extraocular muscles (Hering, 1879; Javal, 1879; Wade et al., 2003). In the early 20th century the French term "saccade" (literally a jerk or twitch) was adopted by Dodge to describe these rapid eye movements (Dodge, 1916; Wade et al., 2003). Dodge also recognized that people do not necessarily perceive saccadic movements as such, but rather experience a continuous visual image. He suggested that visual input is not processed during the execution of a saccade, but that the retinal image prior to the saccade is maintained throughout the interval of movement (see Box 1; Dodge, 1900, 1903, 1905; Wade et al., 2003).

Since these early studies, interest in saccades and the ocular motor system has increased dramatically. Advances in technology for recording and analyzing the precise characteristics of eye movement data have opened the field for researchers who can investigate saccades in a large number of individuals across a multitude of paradigms. The relatively recent development of neuroimaging techniques has provided insight into the neural correlates of saccade production and control. The timeline (Table 2.1) highlights some of the seminal works in saccade research, although it cannot do justice to the full array of studies that have advanced the field through the years.

2.3 Fundamental Characteristics of Visually-Guided Saccades

Visually-guided saccades are eye movements made to foveate a salient stimulus. This redirection of gaze occurs countless times daily, seemingly without willful exertion. Visual input is disrupted during saccadic movements, so saccades have brief durations and are separated by periods of inter-saccadic fixations. In real world

Table 2.1 Timeline of selected notable advances in saccade research

Year	Researcher(s)	Contribution
1879	Lamare/Hering	The first scientific papers using saccade recordings during reading are reported
1903 1916	Dodge	Rapid eye movements with a consistent relationship between duration and visual angle are described and later, the term ‘sacccade’ is adopted
1967	Saslow	The “gap” paradigm is introduced
1968	Robinson	The oculomotor control system is described in terms of the physiological organization and signal processing necessary to generate eye movements
1973	Holzman and colleagues	Saccadic intrusions during smooth pursuit in schizophrenia are identified, initiating the use of ocular motor measures in psychiatric disorders
1975	Bahill, Stark and Clark	A series of papers describing fundamental characteristics of saccades is published
1978	Hallett	The antisaccade task is introduced
1979	Becker and Jürgens	The double step paradigm is discussed relative to a seminal model of parallel programming of saccades
1979	Melamed and Larsen Orgogozo and Larsen	The first papers describing neural imaging (regional cerebral blood flow) of saccades are published
1980s	Fischer and colleagues	The term “express saccades” is introduced and discussed in relation to the gap paradigm and visual attention
1982	Kurtzberg and Vaughan	An early EEG study of saccades in humans is published
1985	Fox, Fox, Raichle and Burde	An early positron emission tomography (PET) study of saccades in healthy individuals is published

(continued)

Table 2.1 (continued)

Year	Researcher(s)	Contribution
1980s–90s	Pierrot-Deseilligny and colleagues	Human lesion studies are performed that highlight the role of cortical regions in saccade production
1990s	Goldman-Rakic and colleagues	The functions of prefrontal cortex during saccades are demonstrated in primate studies
1994	Land and Lee	An early example of real world eye tracking is utilized to determine the focus of gaze during driving
1995	Carpenter and Williams	The LATER model of variable saccade latency with activity rise to a critical motor threshold is developed
1995–6	Hoffman and Subramaniam Deubel and Schneider	Experiments supporting an obligatory relationship between saccade preparation and attention are published
1996	Darby and colleagues Muri and colleagues	Early reports of fMRI investigations of saccade-related activity are published
1997	Ross, Morone and Burr	Evidence of compression of visual space and perceived time preceding saccade execution is reported
2004	Martinez-Conde and colleagues	The importance of microsaccades in sustaining visual perception during fixation is revealed (see Chap. 3)

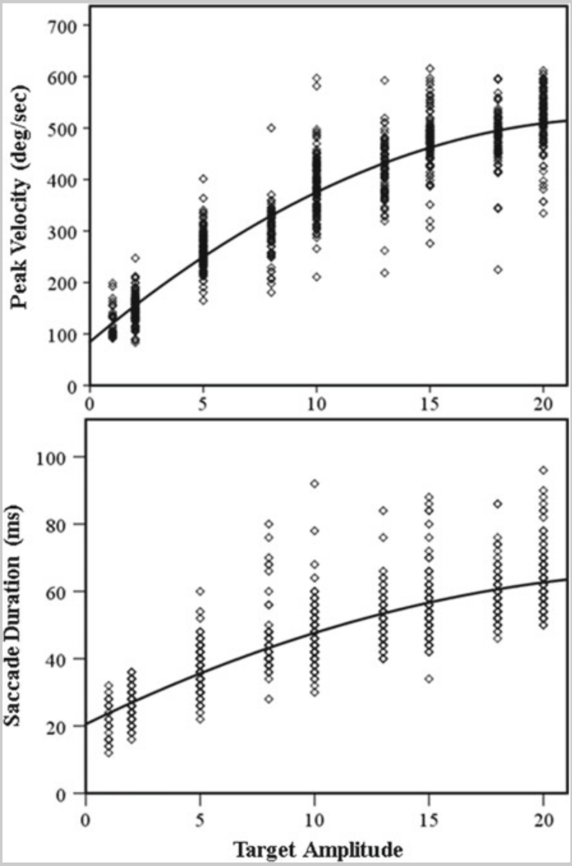
This list is not exhaustive, but emphasizes pioneering research as well as influential researchers situations, saccades often occur in coordination with head and body movements where proprioceptive feedback is necessary to compensate for the continual change in the visual system's frame of reference. This section, however, will focus on saccades generated in controlled laboratory settings while head position remains stable.

2.3.1 Visually-Guided Responses

In a standard visually-guided saccade task, an eye movement is directed *toward* a stimulus of interest. This simple task (also known as a "prosaccade" task) generally results in fewer directional errors, higher spatial accuracy, and shorter latencies than voluntary saccade tasks (Hutton, 2008; Mort et al., 2003). A saccade is initiated around 150 milliseconds (ms) after the visual stimulus appears and is completed within 50 ms or less depending on the distance to the peripheral stimulus. (For a discussion on the relationship between saccade amplitude, duration, and velocity see Box 3.) This latency is moderately variable from trial to trial and across individuals, and is influenced by paradigm design (see Sect. 2.3.4 below). In some individuals, deviation of the typical saccade trajectory may be observed and is known as dynamic overshoot. The eye movement overshoots the stimulus location and immediately rebounds to the correct amplitude in one fluid action. This monocular effect may be caused by rapidly switching neural signals sent to opposing ocular muscles during a saccade and can persist across several trials (Bahill, Clark, & Stark, 1975a).

Typical latencies for visually-guided saccade trials, although fast, actually are much slower than direct signal conduction speeds from the visual input to the motor output require. This suggests that even "reflexive" saccades involve a degree of higher cognitive processing (Hutton, 2008; Reddi & Carpenter, 2000), perhaps to ascertain the identity of the stimulus and access the appropriate response set. In the real world, most visual scenes contain multiple competing stimuli – preventing reflex-like responses toward each one the moment it appears. Selection of relevant stimuli therefore must occur to determine which stimuli are worth foveating, at the expense of other potential targets. This selection process necessarily is time-consuming and adds to the observed latency of saccades.

Box 3: Main Sequence



Fixed relationships exist between saccadic response amplitude and the duration and peak velocity of the response (Bahill, Clark, & Stark, 1975c; Becker & Fuchs, 1969; Dodge, 1903). Saccade execution to a given location is ballistic and stereotyped in that the movement is produced rapidly and consistently in a pre-programmed manner without time for adjustments based on visual changes during the saccade (although occasionally saccades can be interrupted by visual input just before the movement begins, Leigh & Zee, 2015). The figures show responses to visually-guided saccade trials which illustrate that as the amplitude of the saccade target increases, both the duration of the saccade and its peak velocity increase. The correlation between duration and amplitude of saccades was recognized as early as 1903 by Dodge in one of the first systematic descriptions of the various types of eye movements (Dodge, 1903; Wade et al., 2003). This pattern of responses has been described as the "main sequence" and may represent a speed-accuracy trade-off evolved to minimize time spent with poor visual input while maximizing the accuracy of the final eye position (Harris & Wolpert, 2006). During saccadic movement, the retinal image is distorted and little useful visual information can be obtained. The brain must therefore pre-program saccades with appropriate amplitude and corresponding contraction of the eye movement muscles. This pre-programming is largely responsible for the stereotyped velocities and durations of main sequence saccades. Once the feedback from these motor movements (the efference copy or corollary discharge) matches the desired amplitude, the saccade is terminated (Carpenter, 2000). One theory suggests that noise in the motor command to move the eyes accumulates with respect to the length of the movement such that saccades to more distant targets are of faster velocity in order to minimize the potential interference inherent in a longer motor command while maintaining reasonable spatial accuracy (Harris & Wolpert, 2006).

The sudden onset of a visual stimulus can attract attention automatically (*visual grasp reflex*, Hess et al., 1946) and result in the generation of a short latency saccade or a saccade to an irrelevant distractor (Theeuwes, Kramer, Hahn, & Irwin, 1998). Yet in most instances, decision-making processes and other cognitive factors influence saccade performance. Researchers can provide instructions for a given paradigm that confound the degree to which a saccade is truly visually-guided versus more volitionally controlled (Walker et al., 2000), and individual differences in response strategy can affect the behavioral outcome (saccade latency) and the resultant theoretical interpretation (Reddi & Carpenter, 2000). Some individuals are more deliberate in their saccade execution and slower to respond, making performance on the task susceptible to cognitive influences or secondary task demands (Stuyven, Van der Goten, Vandierendonck, Claeys, & Crevits, 2000); others may be highly reactive to all new stimuli and faster to respond, generating saccades that are largely impervious to additional cognitive factors (Schaeffer et al., 2015). In the former case, although the saccade ultimately is directed toward the stimulus, it necessarily must differ cognitively from a response in which the stimulus immediately triggers an eye movement.

One influential model developed by Carpenter and colleagues (Carpenter & Williams, 1995; Reddi & Carpenter, 2000) seeks to explain the variability observed for latencies in a single individual in a simple visually-guided saccade task. The LATER (Linear Approach to Threshold with Ergodic Rate) model posits a baseline level of activity that rises with a variable rate to a critical threshold for saccade generation. The rate of the linear rise changes randomly across trials, due perhaps to physiological noise inherent in neural networks, taking a variable amount of time to reach the decision threshold and leading to a distributed range of saccade latencies (see discussion of antisaccade programming in Sect. 2.4.1 and Fig. 2.7). By biasing the prior expectation of a saccade target appearing at a given location (Carpenter &

Williams, 1995) or the “urgency” of responding quickly (Reddi & Carpenter, 2000), the baseline or threshold levels can be altered such that the mean of the latency distribution shifts slower or faster. This fundamental concept has been extended to several models of saccade generation and tasks that involve competition between multiple saccade programs or locations. Such competition models will be discussed further below and later in relation to volitional saccades.

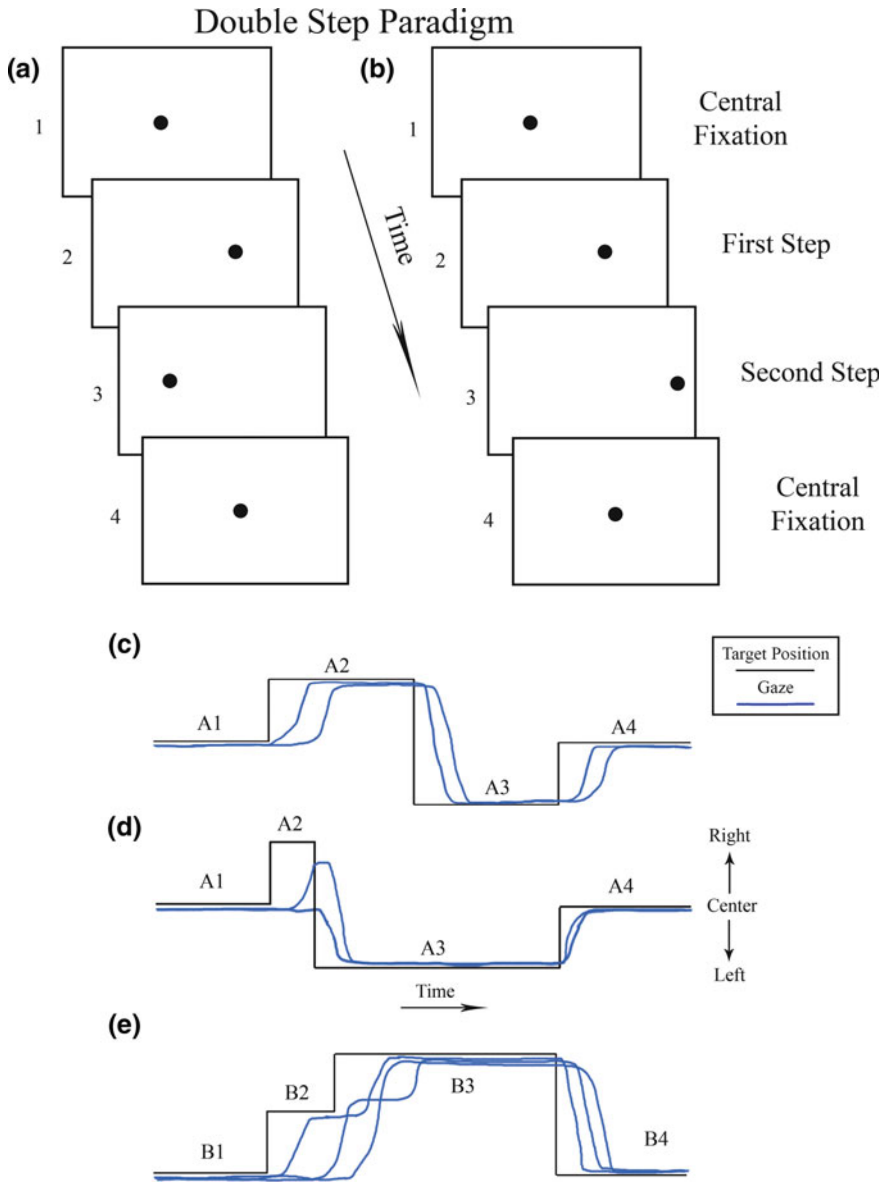
Study Questions:

1. *Is any saccade entirely “visually-guided”? What other processes contribute to the execution of this basic eye movement?*
2. *What is the “main sequence” as it relates to saccades?*

2.3.2 Secondary Stimuli Inform Models of Saccade Generation

In visually-guided saccade paradigms, stimulus features such as eccentricity/amplitude (Kalesnykas & Hallett, 1994; Weber, Aiple, Fischer, & Latanov, 1992) or the number of potential targets influence saccade response characteristics (Findlay, 2009; Hutton, 2008). For example, in a “double-step” task, two peripheral locations are illuminated sequentially, giving the appearance of a single stimulus “stepping” from central fixation to one location and then stepping a second time to another location; participants are instructed to saccade only to the stimulus location that is currently visible (Fig. 2.1). Early studies of this paradigm found that when the interval between the two stimuli was brief (less than 100 ms) individuals often cancelled the initial saccade successfully and looked directly to the second stimulus location (albeit with a slowed latency; Bartlett, 1961; Wheeless, Boynton, & Cohen, 1966). When the interval between stimulus steps was longer, however, individuals completed a saccade to the first stimulus followed by a distinct saccade to the second. Furthermore, if the second stimulus was presented while the participant was performing a saccade to the first stimulus, the latency to the second stimulus was delayed by an amount suggesting that no visual processing of the second stimulus was possible during the first eye movement (a phenomenon known as “saccadic suppression”; e.g., Burr et al., 1994; Matin, 1974; Wurtz, 2008). Saccadic suppression indicates that visual input of additional stimuli is obscured during the motor component of saccades to prevent individuals from experiencing a distorted visual world as the eyes move. The brain instead largely inhibits this blurred visual input until the eye position is stable and clear information about the visual scene is obtained from the new fixation point.

In addition to addressing saccadic suppression, the double-step paradigm is used to investigate mechanisms of saccade adaptation (McLaughlin, 1967; Pélisson et al., 2010). Early work by McLaughlin (1967) demonstrated that when the target stepped to a second location one degree of visual angle closer to central fixation during the


Double Step Paradigm
(a)
(b)
1
2
3
4
Time
Central Fixation
First Step
Second Step
Central Fixation
(c)
A1
A2
A3
A4
Target Position
Gaze
(d)
A1
A2
A3
A4
Right
Center
Left
Time
(e)
B1
B2
B3
B4

◂**Fig. 2.1** Double-step paradigm. Trials begin with a central fixation point, then the stimulus jumps to one location then a second location, and finally back to the center. **a** In some cases, the stimulus moves to opposite sides of central fixation during the two steps (see also **c** and **d**). **b** In other cases, the stimulus moves within the same visual hemifield (see also **e**). **c** A position trace showing the stimulus moving from one side of the screen to the other at a long interval, such that saccades can be made separately to each location. **d** A position trace with the stimulus moving from one side to the other with a short interval, such that slower latency saccades are likely to be directed only to the second location, while some faster latency saccades may be initiated towards the first location before a second saccade is initiated. **e** A position trace with the stimulus moving two steps in the same direction which results in faster latency saccades being near to the first location, slower latency saccades being near to the second location, and a range of intermediate saccade latencies and endpoints in between. Labels overlaid on the position traces refer to the stimuli being presented in (**a**) and (**b**)

initial saccade (so that the participant was unaware of the shift), participants began after several trials to adjust their response to the initial stimulus such that a single 10° target elicited a 9° saccade. Thus, the saccade system monitors the difference in amplitude between an intended target and the actual motor performance and "adapts" the saccade vector based on any resulting overshoot or undershoot error (whether due to a true internal miscalculation or external manipulation). It has been proposed that the adaptation occurs based on the visual error of the target after the saccade, while the adjustment itself affects the motor vector planning (Hopp & Fuchs, 2004; Wallman & Fuchs, 1998). Interestingly, adaptation occurs more quickly (over fewer trials) when the displacement is toward initial fixation (backward adaptation) rather than away from it (forward adaptation), resulting in a smaller saccade amplitude (Miller, Anstis, & Templeton, 1981; Panouilleres et al., 2009; Pélisson et al., 2010). This adaptation may persist for many trials after being conditioned and rely on partially distinct pathways for visually-guided and volitional saccades (Deubel, 1995).

Later work (Becker & Jürgens, 1979) showed that if both stimuli in the double-step paradigm were in the same visual hemifield (half of the visual world that is input to the left vs. right of the fovea of each eye), initial response amplitudes were distributed between the two peripheral locations and varied with latency. Faster responses were closer to the first stimulus and slower responses were closer to the second. If the stimuli were in opposite directions from central fixation, however, saccade endpoints were close to either the first or the second stimulus with few intermediate amplitude responses. Saccades to the second stimulus occasionally were completed with inter-saccadic intervals approaching 0 ms, suggesting to the authors a parallel programming capability of the saccade system. In Becker and Jürgens' interpretation, when the second stimulus was presented early enough to be seen before an eye movement but late enough not to cancel the first saccade, preparation for the second saccade could proceed during the motor portion of the first saccade and thus be initiated almost immediately following completion of the first saccade. The amplitude of the second saccade, remarkably, was highly accurate even at these negligible inter-saccadic intervals, demonstrating the ability of the ocular motor system to incorporate the forthcoming primary saccade into the calculation of the appropriate size of the second saccade from the first stimulus location (Becker & Jürgens, 1979). Thus, they concluded that the saccade system is capable of parallel programming of multiple movements and that the influence of a secondary stimulus depends on the timing relative to motor execution. More recent research has continued to implement this paradigm to investigate motor planning, saccade adaptation, and inhibition during saccade sequences with reference to the nature of the brain's representation of and response to changing visual input (e.g., Camalier et al., 2007; Pélisson et al., 2010; Sharika, Ramakrishnan, & Murthy, 2014; Thakkar, Schall, Logan, & Park, 2015).

In the double-step paradigm a single stimulus jumps from one location to another, while in a "double-target" paradigm a secondary stimulus appears simultaneously with the primary target stimulus (Fig. 2.2a). In this task individuals tend to direct their saccades to the approximate "center of gravity" of the objects, an observation that has been termed the global effect (Coren & Hoenig, 1972; Findlay, 1982; Walker,

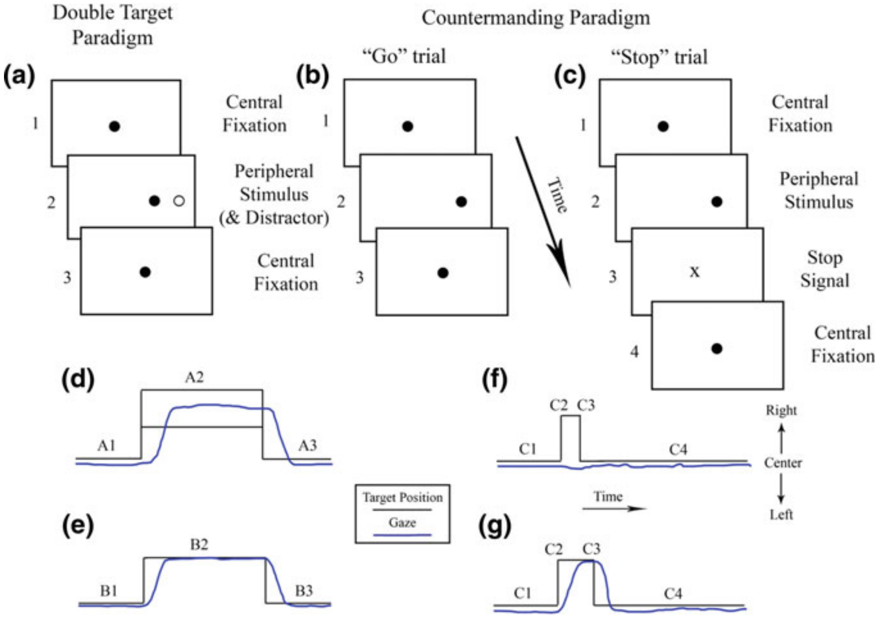


Fig. 2.2 Double-target and countermanding paradigms. **a** In a "double-target" paradigm following central fixation a distractor (open circle) appears simultaneously with the target stimulus at peripheral locations in the same region of the visual field. Typically, participants direct the gaze to a location midway between the two stimuli, as shown in the position trace in (**d**). **b** In a countermanding paradigm, many "go" trials are mixed with a few "stop" trials. On go trials, as in a basic prosaccade task, the stimulus moves from central fixation to the periphery and back to center; the participant must follow it wherever it goes, as illustrated in (**e**). **c** On rare stop trials, a stop signal (here an X) appears shortly after the peripheral stimulus and directs the participant to cancel or "countermand" the saccade being programmed. Depending on the delay between the peripheral stimulus and the stop signal, the participant may (**f**) or may not (**g**) be able to correctly cancel the saccade. Labels overlaid on the position traces refer to the stimuli being presented in (**a**–**c**)

Deubel, Schneider, & Findlay, 1997). The global effect refers to the apparently low resolution of the spatial maps used to calculate the saccade endpoint in the periphery, with multiple stimulus locations being averaged so that the participant acts only upon the global information of the stimuli. In the global effect situation both stimuli are presented in the same hemifield and saccade latency is comparable to that of single target trials. By contrast, in the "remote distractor situation" where the target and distractor are more widely separated, saccade latency is increased by up to 40 ms without influencing landing position (Findlay, 1982). Walker et al. (1997) further modified this directional effect by showing that distractors within a 20° window around the primary peripheral stimulus location affected the amplitude of the saccadic response (global effect), while distractors outside this region increased latencies (remote distractor effect) due to inhibitory influences operating between competing saccade programs. Godijn and Theeuwes (2002) have argued that when distractors are present in a stimulus array, the primary saccade is highly goal-driven in contrast to the externally-driven visual capture of an abruptly onset distractor. Whether a saccade is directed to the target stimulus, the distractor, or somewhere in between depends, therefore, on the relative timing and position of the stimuli. Interestingly, the global effect is most prominent for fast latency saccades when bottom-up processes strongly influence movement planning (Findlay, 1982). When participants are slower to respond, top-down processes can counteract the global effect and direct the saccade more accurately toward the correct target (Heeman, Theeuwes, & Van der Stigchel, 2014). The double-target paradigm thus demonstrates that spatial and temporal characteristics of even "visually-guided" saccade target selection can be biased by both visual factors (presence of multiple stimuli) and volitional processes (current task instructions).

Another means of investigating inhibitory influences in visually-guided saccade tasks is the use of a countermanding or stop-signal paradigm (Boucher, Palmeri, Logan, & Schall, 2007; Hanes & Carpenter, 1999; Logan & Irwin, 2000). This task consists of standard prosaccade trials ("go" trials) intermixed with rare trials during which a signal directing the participant to withhold a saccade is presented after the peripheral stimulus appears ("stop-signal" trials; Fig. 2.2b, c). The delay between the appearance of the stimulus and the stop signal is critical for determining whether or not the saccade can be "countermanded" (stopped) successfully (Hanes & Carpenter, 1999). The relationship between the length of this delay and whether the saccade is stopped indicates how far into saccade preparation an inhibitory signal still can interrupt the motor program and how long the inhibitory signal itself takes to execute. This competition can be modelled as a race between the go and stop signals in the saccade system. Hanes and Carpenter (1999) estimated that with an average latency on go trials of about 230 ms, the stop signal required on average 140 ms to inhibit the saccade, suggesting that the inhibitory stop signal is completed faster than the go signal and can affect the saccade program even when it is more than halfway to completion.

In summary, the use of secondary target stimuli in saccade tasks, whether presented sequentially (double-step) or simultaneously (double-target), demonstrates the flexibility of visually-guided responses and elucidates the time course over which saccade programs are prescribed and enacted. A failure to select the correct response location or suppress competing motor plans creates a distinct response pattern that informs researchers about the limits of the saccade system. Countermanding tasks further describe participants' ability to cancel entirely any saccade motor program and maintain central fixation. All of these tasks show that even "simple" visually-guided saccades involve multiple preparatory processes and are subject to top-down control.

Study Questions:

1. *What is the difference between a double-step and double-target paradigm? How do these tasks reveal the parallel programming capability of the saccade system?*
2. *What is the global effect?*
3. *Can a saccade command be cancelled once it is initiated? What task investigates the timing of saccade cancellation?*

2.3.3 The Role of Visual Attention in Selecting Saccadic Endpoints

Within each of the many paradigms that explore visually-guided saccade generation, one of the central cognitive processes involved with ocular motor behavior is visual attention (see Chap. 7 for a full discussion of attentional processes and the visual system). Much like saccades, attention may be directed volitionally to a new location or captured automatically by a highly salient stimulus (Yantis & Jonides, 1984, 1990). In everyday activities, visual attention and the direction of gaze are closely integrated, with covert attention only being employed if saccades are voluntarily inhibited for a particular reason. Interestingly, it has been shown that one can direct attention covertly to a certain location without moving one's eyes, but one cannot move one's eyes without a concordant shift in attention (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995). Saccade studies of visual attention have made use of dual-task paradigms wherein participants are required to identify a letter at one location and perform a saccade to the same or a different location. When the letter and saccade locations are the same, discrimination of the letter's identity is enhanced compared to when the saccade is directed to another location in the visual field (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Kowler, Anderson, Dosher, & Blaser, 1995). These studies demonstrate that visual attention moves to the target location during preparation for a saccade and thus perception of the letter is facilitated.

Visual attention and saccade preparation also are affected by pre-trial cues that orient the individual toward certain locations as the (possible) target of an upcoming saccade (Posner, 1980; Posner & Cohen, 1984). When a cue appears at a certain location (drawing attention), saccade latencies to that location are faster for valid cues than non-cued or invalid cue trials. The pre-motor theory of attention postulated that this effect is due to an obligatory relationship between attention allocation and saccade preparation (Rizzolatti, Riggio, Dascola, & Umilta, 1987). According to this theory, once visual attention has been directed to a given location any subsequent saccade to another location necessitates cancellation of the motor program that was established for the cued location as well as preparation of a new saccade program and corresponding attention shift to the new saccade destination.

A corollary to this attentional cueing effect is a phenomenon termed "inhibition of return," which was originally observed in manual response tasks. When attention is directed toward a peripheral location, processing and responding to stimuli at that location are facilitated initially; if attention then is redirected to the central location, however, responses to the originally cued location are inhibited relative to other potential target locations (R. M. Klein, 2000). For saccade paradigms also, it has been shown that if a trial requires a "return" to the same location that was cued or saccaded to previously, then latencies are slowed (Massen, 2004; Posner & Cohen, 1984; Rafal, Egly, & Rhodes, 1994). The visual system thus possesses a mechanism that automatically biases processing away from any single location (or a few degrees of visual space around that location; Hooge & Frens, 2000; Maylor & Hockey, 1985) that was prioritized previously. This mechanism may have evolved to encourage visual exploration of the environment, with novel stimuli having precedence over familiar stimuli for generating a motor response (Klein, 2000).

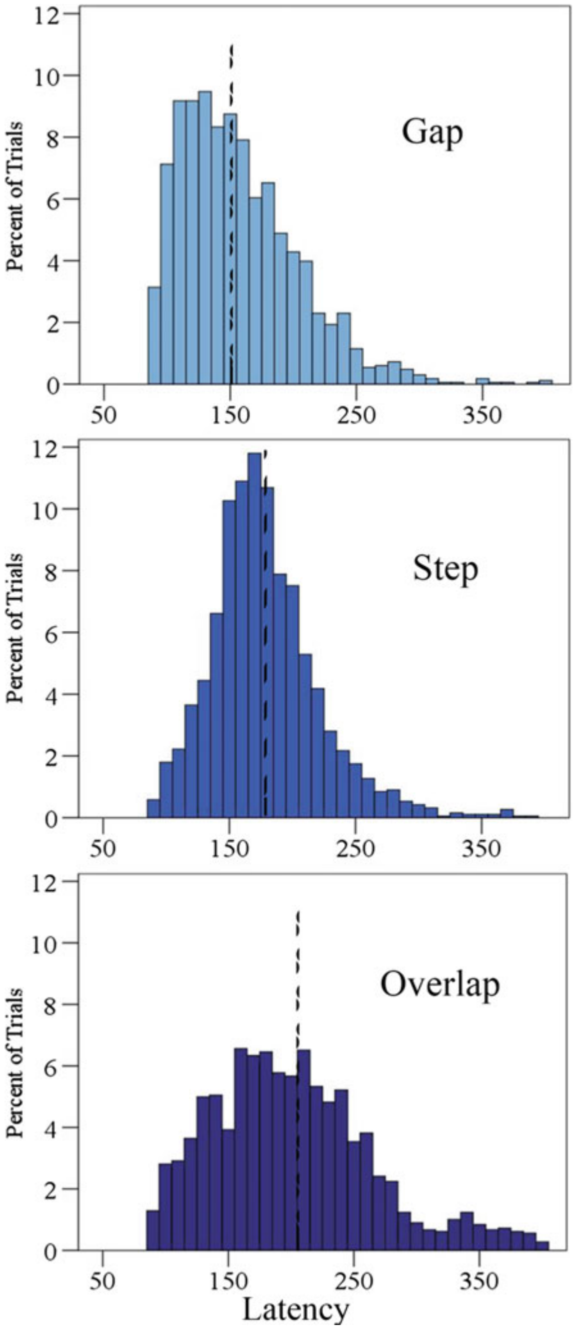
Study Questions:

1. *Can attention be deployed to a location without an eye movement? Can a saccade be generated without a shift of attention?*
2. *Why would visual inhibition of return provide an advantage during exploration of an environment?*

2.3.4 Impact of Stimulus Onset Timing on Saccade Latency

In laboratory tasks, the interval between fixation offset and peripheral stimulus onset can be varied in several ways that impact saccade programming and latency. The slowest responses are observed when using an "overlap" paradigm where the central stimulus remains visible for a period along with the peripheral stimulus (temporal overlap). Latencies decrease if the stimulus simultaneously "steps" from center to the periphery and may be decreased further with the introduction of a blank "gap" period on the order of 200 ms between the central and peripheral stimuli (Fig. 2.3; Fischer & Weber, 1993; Forbes & Klein, 1996; Saslow, 1967). The reduction in latencies associated with presentation of a gap period, the so-called "gap effect," may be due to release of fixation or motor preparation processes that occur between the offset of central fixation and the onset of the peripheral stimulus (Fischer & Weber, 1993; Kingstone & Klein, 1993a; Reuter-Lorenz, Hughes, & Fendrich, 1991). Fixation offset effects are likely specific to the ocular motor system as well as distinct

Fig. 2.3 Impact of stimulus onset timing on latency. Latency distribution from 65 healthy individuals (Clementz et al., unpublished results) showing a shift in mean latency and overall distribution of visually-driven responses for gap, step, and overlap paradigms



from pure visual attention processes that affect early sensory processing (Kingstone & Klein, 1993a; Reuter-Lorenz et al., 1991). The gap effect appears to be more than a forewarning of the peripheral stimulus appearance because presentation of an auditory cue 200 ms before the stimulus does not yield an equivalently large reduction in latency (Pratt, Bekkering, & Leung, 2000; Reuter-Lorenz, Oonk, Barnes, & Hughes, 1995).

The decrease in saccade latency in the gap paradigm may represent an overall shift in the latency distribution or the emergence of a unique class of saccades with a distinct neurophysiological pathway, termed "express saccades". Some studies have reported a bimodal distribution of latencies with an express saccade peak around 100–130 ms and a regular saccade peak around 150–200 ms (although exact latencies depend upon the specifics of the paradigm; Fischer & Breitmeyer, 1987; Fischer & Ramsperger, 1984; Hamm, Dyckman, Ethridge, McDowell, & Clementz, 2010). There even may exist a third peak in latencies that distinguishes fast (140–180 ms) and slow (200–240 ms) regular saccade categories in addition to express saccades. All three peaks may not necessarily be observed in the same subject for a single task (Fischer & Breitmeyer, 1987; Gezeck, Fischer, & Timmer, 1997). Express saccades may represent a visual grasp reflex with minimal attentional engagement (Fischer & Breitmeyer, 1987; Fischer & Weber, 1993), or simply may be an artifact of the overall latency shift (Kingstone & Klein, 1993b). A bimodal distribution with express and regular saccade peaks is not ubiquitously observed (Kingstone & Klein, 1993b) or is seen only in a subset of participants who produce an unusually large proportion of express saccades (Biscaldi, Fischer, & Stuhr, 1996; Wolohan & Knox, 2014). Further research is necessary to make unequivocal conclusions about the origin and nature of express saccades, an issue that is considered again in Sect. 2.5 below on the neural correlates of saccades.

Study Questions:

1. *Through what mechanisms does the gap paradigm allow for faster saccade latencies?*
2. *Do you think express saccades constitute a distinct class of saccades? Why or why not?*

2.4 Volitional Saccade Paradigms Evoking Cognitive Control

In some saccade paradigms the level of cognitive control required to respond is manipulated, generating "volitional" saccades guided by internal goals to a desired or instructed location at a given time. They are more complex than refixations of a new stimulus and require additional cognitive control processes such as inhibition, prediction and working memory. Often there are competing potential responses and participants must facilitate activation of the correct task response while suppressing inappropriate responses. These tasks include novel paradigms designed in the laboratory to isolate particular cognitive control components and test the flexibility of the saccadic response, as opposed to the more natural response required for visually-guided saccade tasks. Many paradigms present a peripheral stimulus comparable to prosaccade trials with only a change in instructions to differentiate performance; others may require generation of an endogenous saccade based on a cue such as the orientation of a central arrow (Abrams & Dobkin, 1994; Rafal et al., 1994) or measure the influence of external factors such as social context of another's gaze location (Friesen & Kingstone, 1998; Hermens & Walker, 2015; Koval, Thomas, & Everling, 2005). Saccades generated during natural viewing conditions also contain a volitional component that guides visually-driven processing (see Box 4).

Experimentally, the increased complexity of the cognitive processes elicited by volitional saccade paradigms results in more variable saccade metrics and more frequent erroneous responses than during comparable prosaccade trials. Certain saccade features are observed during volitional saccade errors that further inform models of ocular motor functioning. For example, initial saccades may miss the intended target via undershoot or overshoot of the peripheral stimulus location. Corrections may occur in the form of secondary saccades or slower, smooth glissades (Bahill, Clark, & Stark, 1975b; Bahill & Stark, 1975; Weber & Daroff, 1972), either of which may have short inter-saccadic intervals or even temporally overlap with the primary saccade. Researchers use diverse stimuli to elicit volitional saccades, but here we will review three standard tasks that take advantage of simple experimental designs to probe the saccade system's ability to manipulate, retain, and anticipate visual information: an antisaccade task, an ocular motor delayed response task, and a predictive tracking task.

2.4.1 Antisaccades: Goal-Directed Responses in Competition with Stimulus-Triggered Impulses

One commonly studied type of volitional saccade is an antisaccade: a task requiring redirection of gaze to the mirror image location of a peripheral stimulus (Hallett, 1978). The antisaccade task was introduced by Hallett to probe participants' ability to follow instructions to arbitrarily change the direction of a saccadic response so the stimulus itself does not fall on the fovea, but rather in the peripheral visual field (Hallett, 1978; Hallett & Adams, 1980). This early work demonstrated that the saccadic system is not restricted to simple stereotyped responses toward a visual stimulus, but instead can be modified by current goals to select a new saccade endpoint. Successful execution of an antisaccade involves suppressing the natural tendency to look at a peripheral stimulus and internally generating a saccade to an unmarked location in visual space. [These processes are not mutually dependent and may be disrupted separately across individuals (Fischer, Gezeck, & Hartnegg, 2000).] The simple instructional manipulation behind an antisaccade task allows for direct comparisons with a prosaccade task, where identical stimuli motivate opposing motor responses (Figs. 2.4, 2.5 and 2.6). Researchers can analyze differences in saccadic behavior that are attributable to internal processes underlying an antisaccade that differ from a prosaccade.

In healthy adults, antisaccades have slower latencies (by 50–100 ms; Fig. 2.6) and reduced peak velocities compared to prosaccades (Evdokimidis et al., 2002; Hutton, 2008; Smit, Van Gisbergen, & Cools, 1987). An initial glance toward a peripheral stimulus on an antisaccade trial constitutes a directional error and is construed as a failure of cognitive control. Directional errors occur on around 10–20% of antisaccade trials and often are followed by a corrective saccade with a fast latency (Fig. 2.5f; Fischer et al., 2000; Hallett, 1978; Hutton, 2008). Initial error corrections tend to be hypometric (i.e., fall short of the target location) and multiple saccades may be performed until the final amplitude is reached; however, error rates, correction rates, and correction latencies are all variable across individuals (Fischer et al., 2000). Individuals with lesions of the frontal lobe, including critical saccade generating and inhibitory regions (described in Sect. 2.5), commit a larger proportion of errors on antisaccade trials as they are unable to suppress the visually-driven saccade (Guit-

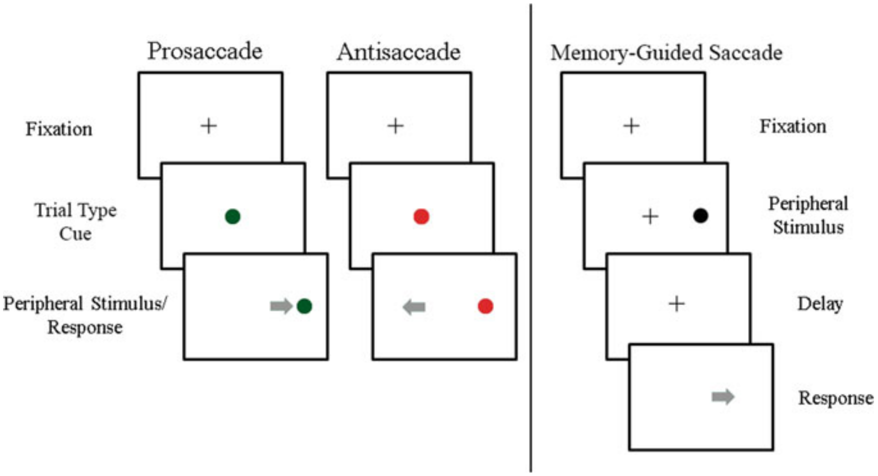


Fig. 2.4 Comparison of visually-guided and volitional saccade paradigms. Prosaccade tasks require a participant to make a saccade toward an abruptly appearing peripheral stimulus onset as quickly and accurately as possible. Antisaccade tasks require a saccade to be directed away from the target, to the mirror image location (opposite direction, equal amplitude). These tasks may be performed in separate blocks of a single trial type or interleaved with trial type cues (e.g., green vs. red circle) that provide the instructions for the upcoming trial. Memory-guided saccades require a participant to withhold an initial response to the stimulus presentation and remember its location throughout a delay period. When a signal is provided (e.g., offset of central fixation or an auditory tone), a saccade must be made to the remembered location. Arrows demonstrate correct eye movements and do not appear during the trials

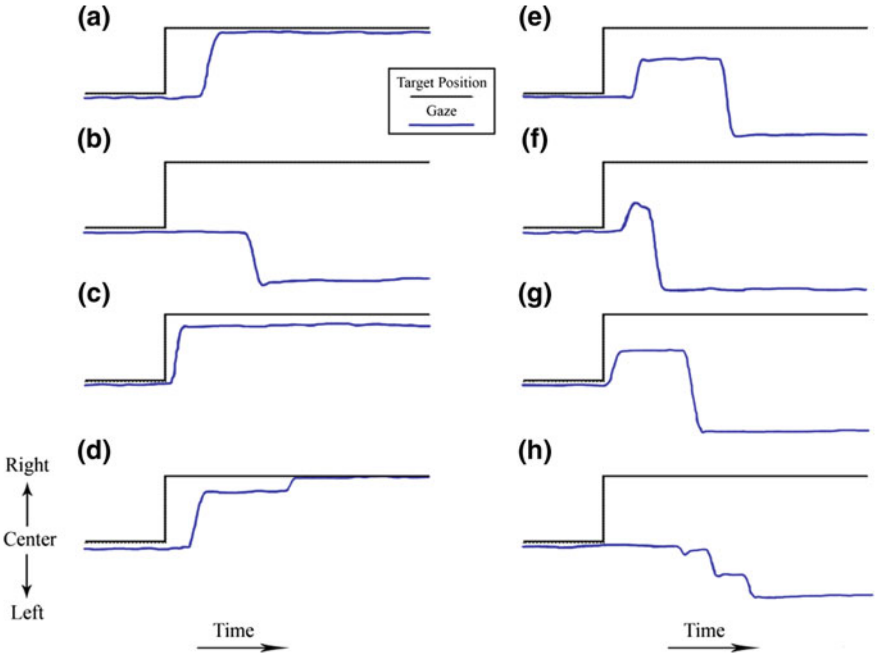


Fig. 2.5 Saccade position traces showing distinct response patterns. In each case, the solid line represents the movement of the stimulus to the periphery and the blue line represents the saccade made either toward or away from this stimulus according to task instructions. **a** A representative visually-guided prosaccade response. **b** An antisaccade, showing a slower latency than the prosaccade. **c** An anticipatory saccade with latency faster than plausible processing times. **d** An initial saccade that undershot its target, followed by a second saccade to achieve the correct final amplitude. **e** An antisaccade directional error (looking at the peripheral stimulus) followed by a corrective saccade of larger magnitude in the opposite direction. **f** An antisaccade error correction with short latency suggesting parallel saccade programming. **g** An erroneous anticipatory saccade followed by a corrective saccade. **h** An antisaccade response consisting of multiple saccades to reach the final amplitude. See Box 2 for further details on the position traces and quantifying of saccadic responses

ton, Buchtel, & Douglas, 1985; Pierrot-Deseilligny et al., 2003; Pierrot-Deseilligny, Rivaud, Gaymard, & Agid, 1991).

Both antisaccade and prosaccade tasks also can be compared to “no-go” trials on which participants are instructed to maintain central fixation when the peripheral stimulus appears, without generating any kind of saccade. No-go trials putatively require the same suppression mechanisms as antisaccades without the need to transform spatial information about the stimulus or execute a motor response (Barton, Raoof, Jameel, & Manoach, 2006; Brown, Goltz, Vilis, Ford, & Everling, 2006). This allows researchers to isolate neural correlates of saccade suppression and observe the frequency of erroneous saccades toward the visual stimulus in conditions without a competing eye movement program. When an antisaccade “go” trial is to be performed, participants must rely upon cognitive control not only for suppression of the unwanted glance toward the stimulus but also for covert orienting

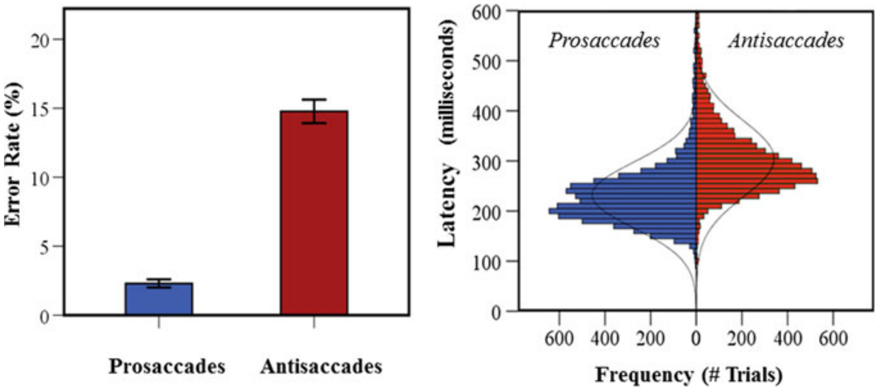


Fig. 2.6 Comparison of error rates and latencies for prosaccade and antisaccade trials. The figure shows data from 54 healthy college-aged individuals in an overlap paradigm with several blocks of mixed pro- and anti-saccades (based on data from Pierce, McCardel, & McDowell, 2015). Both error rates and latencies demonstrate the typical lower values for visually-guided compared to antisaccade trials, reflective of the increased complexity of an antisaccade task

of attention to identify the stimulus location (Connolly, Goodale, Desouza, Menon, & Vilis, 2000). Alternately, during a prosaccade, overt attention and an eye movement both are directed immediately toward the stimulus. The critical difference lies in the stimulus-response mapping (Munoz & Everling, 2004): for prosaccades, a direct correspondence between sensory stimulation and motor response allows for rapid programming of the saccade; for antisaccades, the sensory location must be transformed into the opposite visual hemifield and the motor vector inverted. Finally, during an antisaccade task, control mechanisms must facilitate the generation of a volitional saccade to the newly calculated location.

The relationship between antisaccade and prosaccade task sets has been described by a number of theoretical models on the ocular motor system. A task set refers to the collection of sensory, cognitive, and motor processes required to execute the correct response to a given instruction (Rogers & Monsell, 1995). Traditionally, it was assumed that a serial process of (1) suppressing the tendency to look toward the peripheral stimulus, (2) transforming its spatial location into the opposite visual field, and (3) generating a volitional saccade to the (unmarked) mirror image location was necessary to complete an antisaccade and responsible for the slower latencies. More recent saccade generation models, however, have shifted toward the notion of parallel programming (cf. double-step visually-guided saccades, Becker & Jürgens, 1979) and a race between simultaneous anti- and pro-saccade task activation to reach a critical threshold for triggering a motor response (Fig. 2.7; Cutsuridis, Smyrnis, Evdokimidis, & Perantonis, 2007; Massen, 2004; Noorani & Carpenter, 2013; Trappenberg, Dorris, Munoz, & Klein, 2001). In such models, both anti- and pro-saccade response sets are initiated by the appearance of a peripheral stimulus. Top-down cognitive influences and bottom-up visual input bias the time it takes each program to reach the threshold at which it initiates a saccade and the faster saccade

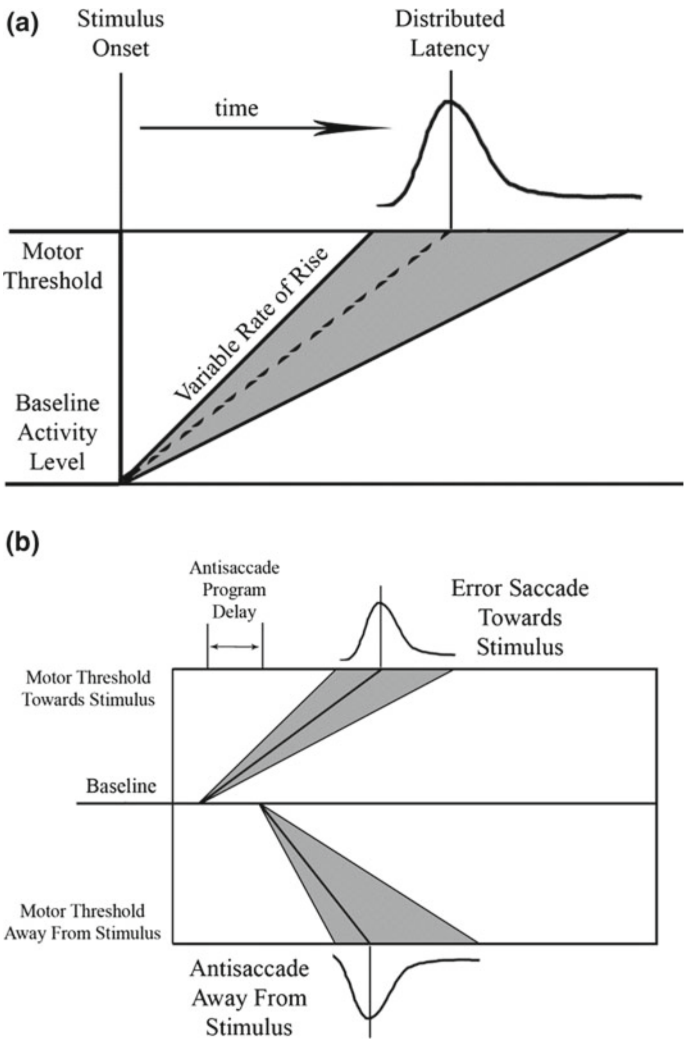


Fig. 2.7 Latency variability and saccade competition. **a** An example of a saccade motor program rising from baseline activity toward a motor threshold with a variable rate and resulting in a range of distributed latencies (as in the LATER model, Carpenter & Williams, 1995). **b** Model of an antisaccade trial showing the competition between a visually-driven saccade toward the stimulus and the correct antisaccade response to the mirror image location. The antisaccade response is initiated with a delay to allow for sensorimotor inversion before the volitional saccade is triggered. On correct trials, the antisaccade response should be facilitated by task goals, have a steeper rate of rise and reach motor threshold faster. Alternatively, the baseline or threshold levels may be adjusted prior to the trial to favor correct performance. Once the antisaccade is completed, a stop program or lateral inhibition cancels the opposing prosaccade program

program is the one performed. Often an antisaccade program is modelled as similar in form to a prosaccade program, but is initiated with a delay representing additional time for the sensorimotor transformation or volitional decision-making processes to occur. Alternately, the base level of activity or rate of rise may differ between the task sets depending on instructional manipulations or prior expectations.

When an erroneous response is generated on an antisaccade trial, it is often not recognized by the participant (Mokler & Fischer, 1999) yet an antisaccade response almost always will still be performed as a corrective saccade. As mentioned above, these corrections can have brief inter-saccadic intervals, even approaching 0 ms, making it improbable that a whole new saccade program was generated following visual recognition of the initial prosaccade error. Instead, the originally programmed antisaccade is enacted based on task instructions regardless of the occurrence of the erroneous response, whereas the reverse is rarely observed (Massen, 2004). When the antisaccade program is performed first, evidently the prosaccade is cancelled automatically; however, researchers have not yet resolved whether this mechanism relies primarily on top-down influences, a specific "stop" program acting on the prosaccade, or lateral inhibition between the competing saccades (Cutsuridis et al., 2007; Everling & Johnston, 2013; Noorani & Carpenter, 2013). It may be that top-down suppression of the inappropriate prosaccade program occurs based on task goals, the generation of a volitional motor response supersedes any other active programs via a potent "stop" signal and/or the inherent opposition of an antisaccade response to a prosaccade response necessitates that the prosaccade program be inhibited. It is likely that multiple factors and pathways contribute to the suppression process, with different mechanisms being favored based on task demands and the participant's current neurophysiological state.

Although a general antisaccade program is activated in response to task instructions to "look to the mirror image location of a stimulus," specific antisaccade metrics are susceptible to changes in numerous experimental features, including stimulus parameters, timing and/or trial history (Table 2.2). Just as prosaccade latencies are

Table 2.2 Examples of effects of task parameters on saccade metrics for either visually-guided or volitional saccades

Parameter manipulation	Saccade behavioral effect
Gap between central fixation and peripheral target	Reduced latency, increased proportion of express saccades
Reduced luminance/contrast	Increased latency
Mixing multiple saccade tasks within a single block	Increased latency and task maintenance errors
Shift in target location during saccade	Adaptation or adjustment of saccade amplitude to the initial stimulus
Instructions	Increased gain accuracy or reduced latency, for example
Distractor stimuli	Global effect (reduced accuracy) or increased latency

decreased in a gap paradigm, so are antisaccade latencies, albeit to a lesser degree (Fischer & Weber, 1992; Reuter-Lorenz et al., 1995). Directional errors on antisaccade trials can yield latencies similar to express prosaccades, which may represent a distinct class of response from regular latency errors/prosaccades (Klein & Fischer, 2005a). Additionally, when anti- and pro-saccade tasks are performed in a single study, response characteristics differ depending on whether the trials are presented in separate blocks or together. If anti- and pro-saccade tasks are performed alone, then error rates are lower and latencies are faster than if both trial types are intermixed within a single block (Ethridge, Brahmbhatt, Gao, McDowell, & Clementz, 2009). Cueing conditions also can be constructed that result in trivial differences in latencies between anti- and pro-saccade trials (Ethridge et al., 2009), indicating that latency differences between anti- and pro-saccade trials do not simply capture cognitive resources associated with suppression.

Within mixed blocks, saccade trials are influenced by the previous trial history and the amount of preparation time between trials (Barton, Greenzang, Hefter, Edelman, & Manoach, 2006; Cherkasova, Manoach, Intriligator, & Barton, 2002; Manoach et al., 2007; Pierce et al., 2015). Trial history effects reflect differences in task switching or task repetition demands, when a participant must prepare a different or identical task set as the previous trial (Meiran, Chorev, & Sapir, 2000). With intermixed pro- and anti-saccade trial presentation, goal neglect may impact "correct" performance (e.g., subjects respond incorrectly on some anti- and pro-saccade trials without making a corrective saccade, as if they believed they made the proper response; Ethridge et al., 2009). Additionally, the need to reconfigure the saccade task set between trials could alter the baseline levels of activity of each program in the race model, favoring the saccade type performed in the previous trial, while cognitive control mechanisms required for an antisaccade may rely upon similar neural resources as task switching processes and have similar latency costs. These control mechanisms implemented during antisaccade performance also may influence participants' responsiveness to visual cues, generating lasting suppression against the prosaccade tendency (Barton, Greenzang et al., 2006; Weiler & Heath, 2012). Indeed, repeated antisaccade testing or extended practice of the task can lead to lower error rates and faster latencies (Dyckman & McDowell, 2005; Ettinger et al., 2003), suggesting that the antisaccade program is facilitated, while the competing prosaccade program may be delayed or weakened.

Study Questions:

1. *What cognitive processes does an antisaccade task require that make it more volitional than a prosaccade task? How does this affect behavior?*
2. *Can multiple saccade commands be programmed simultaneously? In what ways might the correct command compete with erroneous alternatives?*

2.4.2 *Ocular Motor Delayed Response Tasks and Visuospatial Memory*

A second type of volitional saccade task that will be discussed, an ocular motor delayed response (ODR), involves some similar cognitive control processes as an antisaccade task (such as suppressing an immediate response to a visual stimulus), yet its unique complexity arises from increased demands on spatial memory. The task requires participants to maintain the visuospatial location of a peripherally presented stimulus throughout a delay period (that lasts on the order of seconds). After the delay period, participants are signaled to generate a saccade to the remembered stimulus location, which results in a "memory-guided saccade" (Fig. 2.4). Successful completion of this task requires suppression of a glance toward the presentation of the peripheral stimulus throughout the delay period, maintenance of the visuospatial location of the stimulus, and generation of a volitional saccade to the now unmarked location in the visual field (Brown et al., 2004; Sweeney et al., 1996). Originally introduced by Becker and Fuchs (1969) as a comparison of saccades made with visible stimuli to saccades in the dark, the interval between extinguishing the stimulus and initiating the saccade was varied to determine how long "normal" saccade responses could persist after the stimulus light disappeared. They found that after more than a 350 ms delay saccades to remembered stimuli were slower and of longer duration than visually-guided saccades. This finding implies that the visual system can retain the spatial coordinates of a saccade goal for a brief period without continued visual stimulation. Beyond this period, however, saccade endpoint information must be transferred to memory systems and the need to recall visual information from memory alters performance. Errors can be observed in both the generation of premature saccades (immediately responding to the stimulus or anytime throughout the delay) and saccades to an inaccurate location, which occur more frequently than in visually-guided saccade tasks (Anderson et al., 1994). Memory saccades have greater variability in final eye position than visually-guided saccades and tend toward a systematic upward displacement relative to the original stimulus location (Gnadt, Bracewell, & Andersen, 1991).

The delay between the presentation of the peripheral stimulus and the generation of the saccade allows for separate analysis of sensory processing of the stimulus and motor preparation for the response (Brown et al., 2004). To further separate these processes the delayed saccade may also be an antisaccade, so the stimulus and response locations are in opposite hemifields as well as temporally distinct. Neuroimaging studies looking at cortical lateralization during memory saccades indicate that the spatial location of the saccade endpoint is maintained, rather than a visual memory of the peripheral stimulus itself (Medendorp, Goltz, & Vilis, 2005; Van Der Werf, Jensen, Fries, & Medendorp, 2008). The topographic map of visual space encoded in memory evidently loses accuracy throughout increasing delay periods, perhaps as stimulus activation spreads to nearby locations (Godijn & Theeuwes, 2002) or neural noise interferes with location maintenance. Without the continued presence of a visual stimulus, the precise stimulus location is not strengthened and the memory saccade eventually generated has an increased likelihood of amplitude error.

Study Questions:

1. *What cognitive process is especially important for performance in the ODR task?*
2. *How is the ODR task similar to and different from the antisaccade task?*

2.4.3 Internal Monitoring of Spatiotemporal Accuracy in Predictive Saccades

A final type of volitional saccade task requires a different response pattern: tracking of a periodically moving target. This is known as a predictive saccade and the experimental procedure involves periodic stimulus switching between two peripheral locations equidistant from the center of the field of view. Because the movement of the stimulus is constant, participants learn the pace of the movement and automatically start to generate faster saccades which eventually occur in anticipation of the appearance of the stimulus (Stark, Vossius, & Young, 1962). In predictive tracking paradigms, saccadic latencies decrease and the proportion of anticipatory saccades increases over the course of the trial as procedural learning allows participants to prepare the saccade in advance of the stimulus onset (McDowell, Clementz, & Wixted, 1996; Simó, Krisky, & Sweeney, 2005). The amplitude of predictive saccades is often slightly hypometric compared to visually-driven saccades (Bronstein & Kennard, 1987), but remains relatively consistent over a large number of trials (Wong & Shelhamer, 2011). This behavioral profile is suggestive of the saccade being preprogrammed before the actual stimulus appearance. Such a mechanism would allow the motor system to respond to the probable future location of a stimulus, rather than waiting for the stimulus movement to visually trigger a saccade, and would provide a latency advantage when any predictable series of events occurs.

The optimal rate for stimulus presentation uses a square wave function that oscillates at 0.5–1.0 Hz between the two stimulus locations (Ross & Ross, 1987; Stark et al., 1962). Longer intervals between stimulus shifts are typically too great for individuals to successfully predict and anticipate the movement timing, and they instead yield saccades with latencies in the normal visually-guided range (Joiner & Shelhamer, 2006; Shelhamer & Joiner, 2003). Internal monitoring of the frequency of stimulus appearance and sensorimotor feedback on the spatiotemporal accuracy of previous saccades allow participants to adjust their responses over the course of the task (Zorn, Joiner, Lasker, & Shelhamer, 2007). After several trials of predictable stimulus movements an internal clock is engaged that monitors and controls saccade timing, and which persists with high precision for a brief period after the stimuli disappear or switch to a different timing scheme (Joiner & Shelhamer, 2006). This internal clock can be utilized in a number of motor programs with predictable time courses and relies upon a counting mechanism of rhythmic neural oscillations to coordinate initiation of each new saccade program. Noise in the neural system accumulates when the stimulus alternation rate is longer than two seconds, making the endogenous timing mechanism and predictive movements inaccurate and leading individuals to switch back to visually-guided saccades (Joiner & Shelhamer, 2006). The predictive tracking response, therefore, constitutes an optimized saccade program in which anticipation of the visual stimulus allows volitional execution of a predictive saccade when stimulus parameters fall within a certain range where the latency and spatial accuracy of the saccade surpass that of a visually-guided movement.

Study Question:

1. *What type of real world tasks are comparable to the predictive tracking saccade task? Is it beneficial to anticipate future movements based on past patterns?*

Box 4: Saccades in Natural Settings

While much of the main text of this chapter focuses on simple stimuli presented in a laboratory setting, another line of saccade research addresses more ecologically valid conditions such as visual search in natural scenes or gaze control in real world settings (Hayhoe & Ballard, 2005; Kowler, 2011; Luke & Henderson, 2016). Early work by Yarbus (Yarbus, 1967) suggested that eye movements across a natural scene could be used to deduce the instructions a participant received, although this view has been debated (Borji & Itti, 2014; Taya, Windridge, & Osman, 2012). The figure shows two of his traces for free examination (green) and estimating the age of the individuals (red). Clearly, for both tasks the participant's fixation points and saccade scan paths did not cover the entire scene, but jumped to the most salient objects—the faces of the individuals. Recent work has suggested that saliency maps are constructed for all visual scenes one encounters in order to guide attention mechanisms to select appropriate saccade targets (see Chap. 6). These maps can be influenced by bottom-up visual characteristics such as luminance, color, or motion (Koch & Ullman, 1985; Kowler, 2011) as well as by top-down processes such as goal pursuit or emotional content (Mackay, Cerf, & Koch, 2012; Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011; Nummenmaa, Hyona, & Calvo, 2009; Risko, Anderson, Lanthier, & Kingstone, 2012). When a subject has a specific task to accomplish, eye movements and fixations focus on useful objects at the time when they are needed, such as individual ingredients or dishes while cooking (Hayhoe & Ballard, 2005). In contrast, exploratory saccades made to process a scene as a whole may be subject to more basic stimulus properties and the influence of inhibition of return (see Sect. 2.3.3 above), with individuals less likely or slower to saccade to previously fixated targets to encourage viewing of novel locations. Variables of interest in studies of scene viewing include fixation times and locations as well as the frequency and amplitude of saccades (Mills et al., 2011; von Wartburg et al., 2007).

2.5 Cortical and Sub-cortical Neural Circuitry Involved in Saccade Generation

Saccadic circuitry is well documented on many levels from single- or multi-cell recordings in non-human primates (see chapter by Everling and Johnston, in this volume) to lesion and neuroimaging studies in human participants (Table 2.3). The broad array of techniques allows for construction of a detailed map of the functional and structural organization of the ocular motor system (Fig. 2.8). The following discussion focuses on the findings from healthy human participants using non-invasive imaging methods (with reference to neurophysiology studies where appropriate) that address the questions of where, when, and to what degree different regions of the saccade circuitry respond to visually-guided and volitional saccade tasks. These findings illuminate the role of individual brain areas integrated into a network for eye movement generation and cognitive control and provide a foundation for studies of disrupted ocular motor behavior in disease or psychiatric disorders (see chapters by Klein et al. and Smyrnis et al., in this volume). Saccades are supported by widespread cortical and subcortical regions, including occipital cortex, posterior parietal cortex, frontal and supplementary eye fields, superior colliculus, thalamus, striatum/basal ganglia, and cerebellum (see Fig. 2.9 for a visualization of several of these regions

Table 2.3 Summary of neuroimaging techniques used to investigate saccade neural circuitry in humans

Methodology	Neural measure	Spatial resolution	Temporal resolution
Functional magnetic resonance imaging (fMRI)	Concentration of deoxygenated hemoglobin in cerebral blood vessels. Depends on the relationship between cerebral blood flow and local oxygen metabolism	Moderate (millimeters)	Moderate (seconds)
Electro-encephalography (EEG)	Combined electrical signals from neuron action potential firing. Recorded by sensors on the scalp	Moderate to low (centimeters)	Excellent (milliseconds)
Magneto-encephalography (MEG)	Combined magnetic fields perpendicular to the electrical neural signals. Recorded by sensors above the scalp	Moderate to low (centimeters)	Excellent (milliseconds)

within the brain). With the transition from simple to more complex behavior, neural circuitry may show stronger or more extensive activation within this basic circuitry and recruitment of new neural regions to support task performance. Volitional saccades of various types invoke one or more of these patterns of increased activation, with neuroimaging studies often comparing complex saccades to visually-guided saccades or fixation.

Table 2.4 provides an overview of various cognitive processes involved in saccade generation and associated neural regions. Cognition and behavior are possible because of coordinated activity in large scale networks of the brain; therefore, while specific neural populations respond to certain tasks more strongly, no brain region performs only a single function or acts in isolation from the rest of the brain. The neural correlates described in the table and the text below represent regions that have shown consistent activation in neuroimaging studies, but it should be understood that multiple hubs of the saccade network function together to perform numerous cognitive processes simultaneously and a serial, modular conception of saccade circuitry is a poor approximation of the true complexity and interconnectedness of the functional brain.

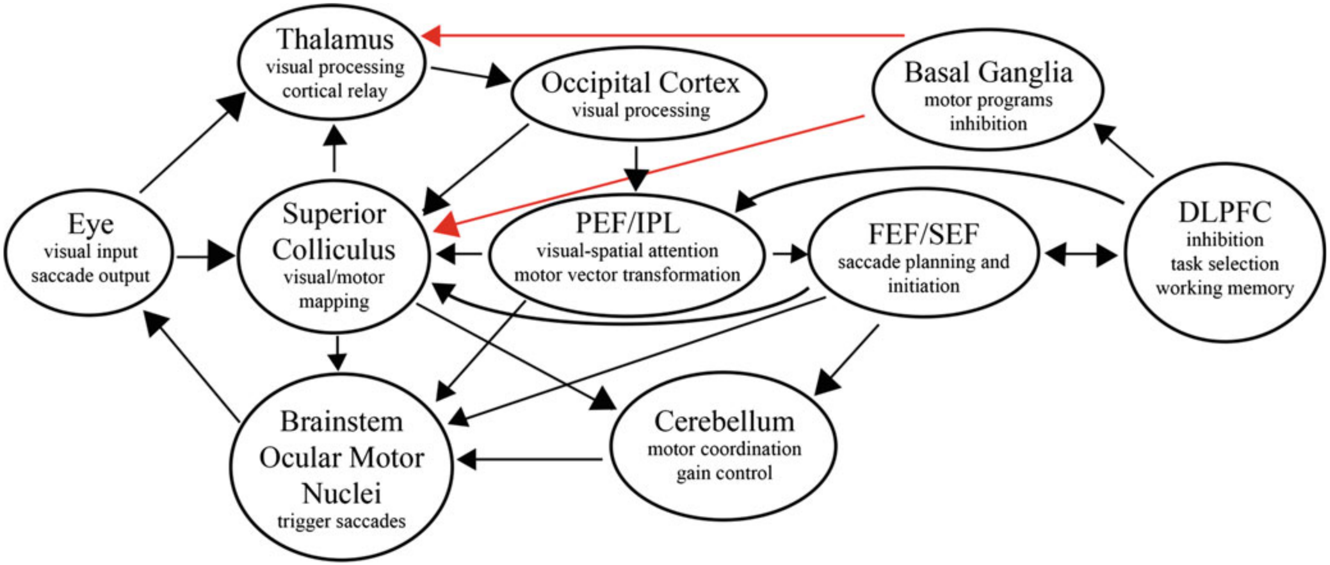


Fig. 2.8 Diagram of saccade circuitry organization and information flow. This schematic shows important regions of saccade circuitry and the connections between them (based on Munoz & Everling, 2004). Black arrows show excitatory pathways and red arrows show inhibitory pathways originating in basal ganglia. Primary functions are listed for each region and see Table 2.4. *PEF* parietal eye field, *IPL* inferior parietal lobule, *FEF* frontal eye fields, *SEF* supplementary eye fields, *DLPFC* dorsolateral prefrontal cortex

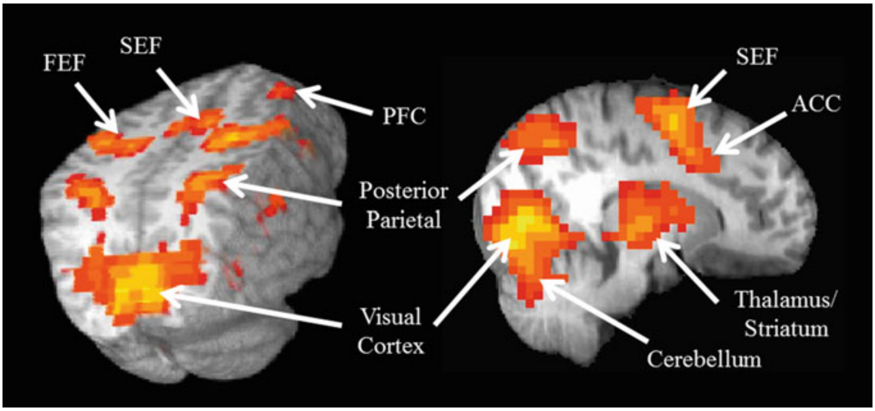


Fig. 2.9 Saccade circuitry averaged across 35 healthy young individuals. This image shows brain regions supporting antisaccade performance from an fMRI study using a block design that alternated between fixation, prosaccades, and antisaccades (Pierce & McDowell, 2016). The colored regions indicate greater activity during antisaccade trials compared to fixation, with brighter hues indexing a stronger degree of activation. Many of the same regions are associated with visually-guided saccade production, although typically with weaker activation. *FEF* frontal eye fields; *SEF* supplementary eye fields; *PFC* prefrontal cortex; *ACC* anterior cingulate cortex

2.5.1 Visual Input to Occipital Cortex

Visual input from the stimulus in a saccade task travels from the retina through the lateral geniculate nucleus of the thalamus to primary visual cortex in the occipital lobe contralateral to the location of the stimulus in the visual field. Basic stimulus features such as color and shape are extracted in secondary visual areas and passed along the ventral "what" stream in the temporal lobe to ascertain stimulus identity and to the dorsal "where" stream in parietal cortex to determine stimulus location in space (Milner & Goodale, 1995; Mishkin, Ungerleider, & Macko, 1983). Regardless of the subsequent task demands, visual processing is essential across all visually-based saccade tasks for initiating the necessary saccade programs (Darby et al., 1996; Luna et al., 1998; McDowell, Dyckman, Austin, & Clementz, 2008; Melamed & Larsen, 1979). Interestingly, occipital cortex may be one of few regions to show greater activation during prosaccades compared to volitional saccades due to increased visual input associated with foveation of the peripheral stimulus or perhaps top-down influences that suppress activity prior to antisaccade or ODR trials (Clementz et al., 2010; Dyckman, Camchong, Clementz, & McDowell, 2007; Krafft et al., 2012).

Table 2.4 List of primary processes involved in saccade tasks and the neural regions associated with these processes

Neural region										
	Occipital cortex	Parietal cortex	FEF	SEF	PFC	ACC	Basal ganglia	Cerebellum	SC	Brainstem
Visual stimulus perception	✓									
Selective attention		✓			✓					
Stimulus identification	✓									
Working memory maintenance		✓			✓					
Motor vector calculation		✓	✓							
Top-down goal facilitation		✓		✓	✓					
Performance monitoring					✓	✓				
Competitive suppression				✓	✓				✓	
Fixation disengagement			✓						✓	
Motor preparation/timing			✓	✓			✓	✓		
Signal integration									✓	✓
Saccade generation			✓						✓	✓

This list is not exhaustive, but highlights some of the important aspects of saccade generation and primary brain regions reported for each process
FEF frontal eye fields, *SEF* supplementary eye fields, *PFC* prefrontal cortex, *ACC* anterior cingulate cortex, *SC* superior colliculus

2.5.2 Attention and Visual-Motor Transformation in Parietal Cortex

When visual information reaches the parietal lobe, visual attention mechanisms are engaged by processes including target selection, calculation of the appropriate saccade movement vector, suppression of a prepotent saccade toward the peripheral stimulus, and working memory maintenance. Multiple sub-regions of posterior parietal cortex [e.g., intraparietal sulcus; inferior parietal lobule (BA40); precuneus (BA7)] serve various saccadic functions and distinct parietal eye fields may exist for visually-guided and volitional saccades. Volitional saccades often elicit greater activation in posterior parietal cortex (Jamadar, Fielding, & Egan, 2013; McDowell et al., 2008), although a few studies (Ford, Goltz, Brown, & Everling, 2005; Mort et al., 2003) reported greater activation in the posterior inferior parietal lobule for stimulus-driven saccades compared to volitional saccades, which may be associated with attentional orienting to targets.

Neurophysiological studies of non-human primates have concluded that the lateral intraparietal (LIP) area constitutes the parietal eye field region most crucial to saccade generation. Stimulation of LIP in animals can directly trigger saccades (Gottlieb & Goldberg, 1999; Johnston & Everling, 2008) and the receptive fields of these neurons are sensitive to impending saccade execution, remapping to the future location prior to the eye movement as part of the mechanisms that support visual continuity (see Box 1; Duhamel et al., 1992; Wurtz, 2008). Neuroimaging studies in humans, however, remain inconclusive as to the precise homologue of this region. The intraparietal sulcus has been hypothesized as the human parietal eye field, yet there are conflicting reports as to where and whether it is activated to a greater extent by basic or complex saccades (Connolly, Goodale, Menon, & Munoz, 2002; Ford et al., 2005; Konen, Kleiser, Wittsack, Bremmer, & Seitz, 2004; Krafft et al., 2012; Medendorp et al., 2005; Mort et al., 2003; Muri, Iba-Zizen, Derosier, Cabanis, & Pierrot-Deseilligny, 1996). One study found that intraparietal sulcus did not exhibit increased activation during preparation for upcoming saccade trials, while other saccade regions did (Connolly, Goodale, Goltz, & Munoz, 2005), which may indicate that this region is more involved in visual attention processes associated with stimulus presentation than motor preparation for saccades. Other studies showed greater activation for endogenous saccades cued via a central arrow (Mort et al., 2003), memory-guided saccades (Brown et al., 2004; Connolly et al., 2002), and antisaccades (Brown et al., 2006; Curtis & D'Esposito, 2003; Ford et al., 2005) as compared to visually-guided saccades. Several EEG/MEG studies indicate that the human homologue of the parietal eye field may be more medial and that it is associated with saccade generation (Clementz, Brahmbhatt, McDowell, Brown, & Sweeney, 2007; Hamm et al., 2010; McDowell et al., 2005).

A similar pattern of activation has been reported for the superior parietal lobule, with greater activation during more volitionally-driven saccades (Kimmig et al., 2001; O'Driscoll et al., 1995). A recent meta-analysis of fMRI studies reporting parietal activation during simple and complex saccade tasks (Krafft et al., 2012), however, found a heterogeneous distribution of saccade-related activity with adjacent parietal regions being activated by volitional, visually-guided, or both types of saccades. This suggests that ambiguities in accounts of parietal cortex activation may be due to the use of broad terms to describe functional activation peaks whereas smaller, distinct locations within these regions do differentiate between saccade functions. For example, one study (Konen et al., 2004) reported that several saccade types activated posterior intraparietal sulcus, while spatially predictable saccades particularly activated an anterior location.

Some insight into the *timing* of parietal activation during saccades was provided by a study (Hamm et al., 2010) of the neural correlates of express saccades using dense array EEG. Three temporally linear neural events predicted express versus regular saccade latency generation. First, in the 1000 ms pre-trial period, particular occipital alpha phases were anti-correlated with express saccade occurrence. The probability that subjects were in these phases also decreased as the proportion of express saccades increased over trials, suggesting that alpha phase may have been a proxy for top-down modulation associated with enhanced saccadic performance. Second, 80 ms into the 200 ms gap period there was an increase in superior parietal cortex activity prior to express saccade generation, which may indicate timely detection of fixation offset to facilitate preparation of saccade-generating circuitry. Third, 170 ms into the gap period there was simultaneous enhanced activation on eventual express saccade trials in BA19 (primary visual cortex) and BA7 (in the putative location of human parietal eye field), which is consistent with priming of saccade-related neural architecture in preparation for rapid response generation.

Another open question regarding parietal cortex function revolves around the lateralization of saccade-related activation. Some studies suggest that visuospatial attention processes are supported by right posterior parietal cortex (Konen et al., 2004), while others argue that parietal cortex encodes the stimulus location in the contralateral hemisphere. To investigate this issue, Van Der Werf and colleagues conducted an MEG study of oscillatory neural activity during memory-guided prosaccade and antisaccade tasks (Van Der Werf et al., 2008). Their analyses focused on changes in the gamma (40–120 Hz) and alpha (7–13 Hz) activity bands, which roughly correspond to local neuronal processing and long range inhibitory activity, respectively. The results revealed an initial increase in gamma power and decrease in alpha power in posterior parietal and occipital cortex in the hemisphere contralateral to the stimulus location, consistent with stimulus encoding. During the delay period, however, gamma band power remained elevated in a discrete parietal region within the contralateral hemisphere for prosaccade trials, whereas on antisaccade trials gamma activity switched to the ipsilateral hemisphere (see also Clementz et al., 2007), indicating that this activity represented the location of the saccade goal in preparation for the motor response (and/or allocation of covert attention to that location). Therefore, the role of parietal cortex may differ throughout the course of a saccade trial and the temporal resolution of MEG/EEG, but not typical fMRI, may be helpful in clarifying the function of this brain region over time for multiple types of saccades.

Study Questions:

1. *What roles does parietal cortex play in saccade generation?*
2. *What are some advantages/disadvantages of using animal physiology studies to understand the human ocular motor system?*

2.5.3 Saccade Planning and Generation in Frontal/Supplementary Eye Fields

Parietal saccade-related regions are reciprocally connected to motor regions in the frontal lobe such as frontal and supplementary eye fields (FEF/SEF), which are important for preparing and initiating saccades. The FEFs are one of the most robustly active regions in the saccade circuitry and have been observed consistently from the earliest neuroimaging studies (Anderson et al., 1994; Connolly et al., 2002; Darby et al., 1996; Fox, Fox, Raichle, & Burde, 1985; Kurtzberg & Vaughan, 1982; Luna et al., 1998; Orgogozo & Larsen, 1979). They are located bilaterally within the precentral gyrus (BA6) in the individual, although group averaging may spread the activation into nearby regions, giving the appearance of involvement of a larger region (Luna et al., 1998; Simó et al., 2005). Anatomical connections exist between the FEFs, the superior colliculus, and brainstem saccade-triggering nuclei, suggesting that build-up of activity in this region contributes directly to saccadic initiation (Hanes & Wurtz, 2001). Furthermore, preparatory activity in the FEFs has been shown to correlate inversely with saccade latencies, indicating that greater activation can facilitate rapid saccade triggering (Connolly et al., 2005; Hanes & Schall, 1996).

Volitional saccades usually elicit greater activation in the FEFs for selecting the appropriate saccade task set and initiating the motor response (Brown et al., 2004, 2006; Curtis & D'Esposito, 2003; DeSouza, Menon, & Everling, 2003; Dyckman et al., 2007; Kimmig et al., 2001; McDowell et al., 2005; Mort et al., 2003; Sweeney et al., 1996). Several studies have reported increased fMRI activation in FEF during the preparatory period for saccades, the magnitude of which is modulated by the subsequent trial type (prosaccade or antisaccade) to be performed (Connolly et al., 2002; Curtis & Connolly, 2008). Although the FEFs often are treated as a single unit, in some cases differences in saccade-related activation are observed along a medial-lateral axis. Stronger activation in FEFs is reported for volitional saccades at medial coordinates while lateral locations are reported for visually-guided saccades or for both saccade types (Jamadar et al., 2013; McDowell et al., 2008), although the spatial resolution of many studies precludes definite identification of such a division in humans at present.

Located medially and anteriorly to the FEFs in the anterior portion of the supplementary motor area on the dorsomedial surface of the paracentral sulcus, the SEF is involved in saccadic initiation, with greater strength or extent of activation observed for volitional compared to visually-guided saccades (Amiez & Petrides, 2009; Grosbras, Lobel, Van de Moortele, LeBihan, & Berthoz, 1999; Jamadar et al., 2013; Leigh & Zee, 2015). Several early studies that used PET imaging to compare memory saccades and/or antisaccades to visually-guided saccades and fixation found stronger activation in the SEF (Anderson et al., 1994; O'Driscoll et al., 1995; Sweeney et al., 1996), a result that has been consistently replicated in later fMRI studies (Brown et al., 2004, 2006; Curtis & D'Esposito, 2003; Ford et al., 2005). In one study addressing the impact of paradigm design on saccade-related fMRI activation, increased SEF activity (along with FEF and precuneus) was found for anti- compared to pro-saccade trials, both when the saccade types alternated with central fixation and when participants had to switch between blocks of prosaccades and antisaccades (Dyckman et al., 2007). Other antisaccade-related regions only showed significant activation during the single saccade type versus fixation blocks, indicating that their involvement may be associated with general differences in task demands between antisaccade and prosaccade trials. The SEF, however, appears to be especially critical for antisaccade generation regardless of the context of the saccade task, being involved in the suppression of a movement toward the peripheral stimulus or enhancing fixation-related activity prior to the stimulus in the FEFs or superior colliculus (Dyckman et al., 2007).

Another instance during which the SEF is activated is the performance of a sequence of saccades (Grosbras et al., 2001; Heide et al., 2001), a task wherein participants must learn and execute a pre-determined series of saccades. In addition to general procedural learning and memory processes involved in saccade sequence tasks, the SEF contributes to the planning of each volitional saccade in the sequence. Further evidence for the SEF's role in voluntary saccade preparation comes from the observation that antisaccade-related activation increases prior to saccade generation (Ford et al., 2005; Jamadar et al., 2013). Interestingly in SEF (as well as several other saccade-related regions), activation decreases following extended practice of an antisaccade task (Lee et al., 2013) as participants improve performance and perhaps reduce demands on neural resources over time.

Study Question:

1. *What functions do the FEF and SEF serve in visually-guided and volitional saccade tasks?*

2.5.4 Top-Down Cognitive Control in Prefrontal Cortex

The prefrontal cortex (PFC) is associated with myriad executive function/cognitive control processes across multiple cognitive domains (D'Esposito, Postle, & Rypma, 2000; Miller & Cohen, 2001). During saccade tasks PFC exhibits increased activation in response to the greater cognitive demands associated with generation of volitional saccades. It is involved in maintaining appropriate stimulus-response mappings, selecting the proper response based on the specific task context, and biasing sensory and motor regions to generate the desired action. Specifically, dorsolateral PFC (DLPFC, BA 9/46) plays a role in the suppression mechanisms necessary for antisaccades and other tasks that require withholding a saccade at the appearance of a new stimulus (DeSouza et al., 2003; Funahashi, Bruce, & Goldman-Rakic, 1991; Pierrot-Deseilligny, Muri, Nyffeler, & Milea, 2005; Sweeney et al., 1996). This top-down control may exert its effects on early visual processing of the peripheral stimulus, dampening the visual response or facilitating central fixation to increase the likelihood of performing a correct antisaccade or ODR response (Clementz et al., 2010; McDowell et al., 2005). For example, a negative correlation was observed between PFC and occipital EEG activity 200 ms after stimulus presentation in an antisaccade task, presumably reflecting this top-down modulation (Clementz et al., 2007). If the initial activation for a visual stimulus is reduced, then the volitional saccade program has a better chance of reaching the motor threshold first and fewer stimulus-driven errors will be generated. Hamm et al. (2012) also demonstrated that pre-trial low alpha phase predicted correct versus error responses during an antisaccade task, with this alpha activation occurring simultaneously in DLPFC and precuneus (Hamm, Dyckman, McDowell, & Clementz, 2012); this result may reveal a mechanism by which top-down control operates. DLPFC shows greater activation preceding correct antisaccade trials versus those on which an error is committed (Ford et al., 2005). Additionally, PFC contributes to the working memory components of ODR trials in conjunction with parietal cortex (Brown et al., 2004; Camchong, Dyckman, Austin, Clementz, & McDowell, 2008; Sweeney et al., 1996).

Activation in other regions of prefrontal cortex, including ventrolateral PFC and inferior frontal cortex, also has been reported during volitional saccades, particularly with respect to response suppression (Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007; Dyckman et al., 2007; Jamadar et al., 2013). Using an event-related fMRI design, Brown et al. (2004) found greater activity on memory saccades as compared to visually-guided saccades during the delay period in right posterior inferior frontal gyrus. This activity is likely not saccade-specific, but generally engaged across paradigms where inappropriate responses must be withheld, such as in the manual go/no-go task (e.g., Aron & Poldrack, 2006).

In many studies, a right hemisphere bias is evident in frontal cortex activation and occasionally in parietal cortex, regardless of the location of the peripheral stimulus (Brown et al., 2004; DeSouza et al., 2003; Ford et al., 2005). This manifests as stronger activation levels in the right hemisphere or an absence of significant left hemisphere activation. Some researchers have suggested that this effect is related to a right lateralization of visuospatial processing in attention and working memory (Konen et al., 2004; Mort et al., 2003), although this question remains unresolved.

2.5.5 Performance Monitoring in Anterior Cingulate Cortex

Another neural region crucial to cognitive control of saccades, particularly during volitional tasks, is the anterior cingulate cortex (ACC). The ACC exhibits increased activation in conjunction with its known role in conflict and error monitoring, as well as preparatory performance optimization (Brown et al., 2006; Ford et al., 2005; Ito, Stuphorn, Brown, & Schall, 2003; Jamadar et al., 2013; Paus, Petrides, Evans, & Meyer, 1993; Polli et al., 2005). The dorsal extent of the ACC plays a key role in cognitive control across paradigms (Shenhav, Botvinick, & Cohen, 2013), coordinating with PFC to determine when greater control is needed and how much (e.g., when cognitive demands are high or conflicting task sets must be maintained). In reference to saccade tasks, higher cognitive control may be needed to perform voluntary saccades that conflict with automatic responses to visual stimuli, to switch between multiple types of saccade performance, or to adjust behavior following an erroneous response. Hamm et al. (2012), using dense array EEG, demonstrated that pre-trial ACC activity was accentuated for a few hundred milliseconds before antisaccade trials to which correct responses were subsequently generated, and was specifically present in the 20 Hz (beta) frequency range, a result consistent with the thesis that stable beta rhythms in motor-related cortices support response suppression. The rostral ACC, on the other hand, was reported in one study to be deactivated during preparation for antisaccade trials that were subsequently performed correctly, while it failed to show this deactivation for trials on which an error occurred (Polli et al., 2005). This study demonstrated that rostral and dorsal ACC exhibited strong responses after commission of an antisaccade error, in line with a role for both sub-regions in performance evaluation (see also Ford et al., 2005). In ODR tasks, increased ACC activation has been reported as well, potentially related to increased demands of spatial memory and motivation (Anderson et al., 1994).

Study Questions:

1. *What processes do PFC and ACC support?*
2. *Why are these regions typically more activated during volitional tasks?*

2.5.6 Motor Control in Cerebellum, Basal Ganglia, and Thalamus

Beyond the neocortex, several other regions are involved in saccade generation. For example, the cerebellum, primarily the oculomotor vermis, is critical for saccade adaptation, motor planning and timing, and amplitude control (Hayakawa, Nakajima, Takagi, Fukuhara, & Abe, 2002; Sweeney et al., 1996; Voogd, Schraa-Tam, van der Geest, & De Zeeuw, 2012). Saccade adaptation refers to a change in amplitude of responses following a consistent, small displacement of the stimulus from its original position (from 10° to 12° from center, for example). As the stimulus moves during execution of the saccade, the originally planned 10° saccade ends up 2° from the stimulus location. This error in amplitude subsequently is incorporated into saccade planning such that a 10° stimulus will elicit a 12° saccade; the cerebellum controls this amplitude adjustment (Hopp & Fuchs, 2004; Pélisson et al., 2010). In a PET study of predictive saccades, the cerebellum (along with the FEF) was shown to have greater blood flow than during smooth pursuit movements (see chapter by Lencer et al., in this volume), suggesting a frontal-cerebellar loop that coordinates motor preparation and timing of saccades (O'Driscoll et al., 2000). While increased activation in the cerebellar vermis is commonly reported during saccade tasks, the cerebellar hemispheres additionally are involved in memory saccade performance (Nitschke et al., 2004).

Subcortically, the thalamus and basal ganglia show increased activation for volitional saccades in relation to attentional and motor learning processes (O'Driscoll et al., 1995; Sweeney et al., 1996). The thalamus, particularly the pulvinar (Robinson & McClurkin, 1989), is reactive to changes in visual attention demands during both visually-guided and volitional saccades. The basal ganglia receive motor signals from frontal cortex and provide inhibitory input (via the substantia nigra pars reticulata) to superior colliculus and thalamus (Munoz & Everling, 2004). For higher order saccades, such as in predictive saccade tracking, the basal ganglia and cerebellum are critically important for motor timing and correctly anticipating the stimulus movement (Joiner & Shelhamer, 2006), while the basal ganglia and striatum (caudate and putamen) are associated with motor learning and reward processing (Hikosaka, Nakamura, & Nakahara, 2006; Shires, Joshi, & Basso, 2010).

2.5.7 Visual and Motor Mapping in Superior Colliculus

One subcortical region, the superior colliculus, plays a major role in selecting visual targets and initiating saccades as it receives direct input from multiple cortical regions as well as from the retina. Superficial layers of the superior colliculus process contralateral visual input in a retinotopic fashion and deeper layers process multiple sensory, motor, and association inputs (Krebs et al., 2010). Visual neurons show effects of predictive remapping with receptive fields shifting prior to saccade production, as in parietal and frontal eye fields (J. Ross et al., 1997; Walker et al., 1995). The intermediate layers contain a motor map corresponding to saccade vectors of increasing amplitude along a rostral-caudal axis (Johnston & Everling, 2008). As the activity from all cortical and subcortical sites is integrated within this map, a target location is selected when activity surpasses the necessary motor threshold (McPeek & Keller, 2002; Trappenberg et al., 2001). Ultimately, physical movement of the eyes by the extraocular muscles is triggered by three cranial nerves via nuclei in the brainstem (Schiller & Tehovnik, 2005). A "pulse" or burst of activity quickly shifts the eye muscles to the new position, after which a “step” or sustained increase in tonic neural activity maintains the muscle tension at that position.

The retinotopic motor map within the superior colliculus may be the site of saccade competition as multiple locations are activated by distractor stimuli or opposing task programs (Cutsuridis et al., 2007; Heeman et al., 2014). Fixation-related and saccade-related neurons mutually inhibit each other, as do saccade neurons for opposing directions. This competition must result in selection of a single target, although the influence of competing locations can slow responses or decrease spatial accuracy. The fixation neurons in superior colliculus also may contribute to latency reductions in the gap paradigm: with the disappearance of a fixation point these neurons are deactivated, allowing activity to build in saccade neurons and trigger movements more quickly when the peripheral stimulus appears (Neggers, Raemaekers, Lampmann, Postma, & Ramsey, 2005). Express saccade generation similarly might occur via a more direct path from visual input to motor output in superior colliculus that could bypass much cortical influence (Schiller, Sandell, & Maunsell, 1987; Schiller & Tehovnik, 2005). Additionally, inhibition of movement neurons for a given location after a saccade could contribute to inhibition of return effects that slow responses to a previously attended location, although both inhibition of return and express saccade production are likely influenced by parietal and FEF inputs (Klein, 2000). Most of these findings are derived from studies using implanted electrodes in non-human primates rather than in human neuroimaging because of technical difficulties in obtaining a reliable signal from this small, deep region with fMRI or EEG. One noteworthy exception using high magnetic field fMRI has shown the role of the superior colliculus in humans to be comparable to that described in animal models (Krebs et al., 2010).

Through coordinated activity within the ocular motor network, saccades are generated efficiently as the brain responds to changing cognitive demands. Activation in occipital and parietal cortex, frontal and supplementary eye fields, prefrontal and anterior cingulate cortex, cerebellum, basal ganglia, thalamus, and superior colliculus allows for processes related to visual input, motor planning and execution, response suppression, and working memory to be executed successfully. The appropriate saccade location can be selected and timing information transmitted to the brainstem and extraocular muscles. Although there remain some discrepancies in specific function and localization of saccade-related regions based on imaging technique or species (e.g., Kagan, Iyer, Lindner, & Andersen, 2010), the overall circuitry is well characterized and neuroimaging studies of healthy humans broadly agree with findings from lesion studies and direct cell recordings in non-human primates.

The scientific study of saccades began more than a hundred years ago and has steadily advanced. Improvements in eye tracking technology and non-invasive neuroimaging approaches have expanded the accessibility of saccade research in humans to numerous laboratories and participant populations, increased the accuracy and resolution of saccade data, and provided valuable insight into the neural mechanisms behind these deceptively simple eye movements. Future research will continue to build on this fundamental knowledge about the cognitive processes required for visually-guided and volitional saccades and further clarify the timing and interactions between neural populations within the saccade circuitry. The tracking of saccades is being implemented in an ever-larger range of paradigms, including real world scenarios and social contexts, that will be integrated with the basic findings described above to expand the current description of the characteristics of the ocular motor system and broaden the practical applications of saccade research.

Study Questions:

1. *How does the superior colliculus integrate information to select a saccade target?*
2. *Do you think cortical or sub-cortical pathways are more important for saccade generation? For volitional tasks?*

2.6 Suggested Readings

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Res, 19*(9), 967–983.

- *An early study of the double step paradigm with a description of how competing saccade programs may be managed by the brain.*

Fischer, B., & Weber, H. (1993). Express saccades and visual attention. *Behavioral and Brain Sciences, 16*, 553–610.

- *A discussion of several studies reporting express saccades and considering theories on the role of disengagement of visual attention during saccade preparation.*

McDowell, J. E., Dyckman, K. A., Austin, B. P., & Clementz, B. A. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades: evidence from studies of humans. *Brain and Cognition, 68*(3), 255–270.

- *A review of recent neuroimaging literature highlighting the cognitive processes and neural circuitry involved in saccade performance.*

Munoz, D. P., & Everling, S. (2004). Look away: the anti-saccade task and the voluntary control of eye movement. *Nature Reviews. Neuroscience, 5*(3), 218–228.

- *A detailed review of the antisaccade task and the organization of the neural circuitry controlling eye movements.*

Reddi, B. A., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience, 3*(8), 827–830.

- *A study investigating the impact of instructions ("respond quickly" vs. "respond accurately") on saccade reaction time. Results are discussed in reference to the LATER model of saccade programming.*

Wade, N. J., Tatler, B. W., & Heller, D. (2003). Dodge-ing the issue: Dodge, Javal, Hering, and the measurement of saccades in eye-movement research. *Perception, 32*(7), 793–804.

- *A review of the origins of saccade research and early methods of quantifying eye movements.*

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Res, 48*(20), 2070–2089.

- *A review of receptive field shifting, retinal displacement, and saccade suppression.*

2.7 Questions Students Should Be Able to Answer

1. *What functions do saccades perform in daily life? Under what circumstances might antisaccades be generated in real life?*
2. *How is vision disrupted during saccades? What mechanisms prevent this instability from being perceived?*
3. *What insight into cognitive processes can studies of saccades provide?*
4. *What types of paradigms can be used to investigate saccade generation?*
5. *How do basic visual cues and complex volitional control interact to generate saccade responses in various contexts?*
6. *How have technological advances allowed for greater understanding of saccade behavior and neural mechanisms?*
7. *What brain areas contribute to stimulus processing, attentional deployment, motor planning, cognitive control, competition, target selection, and signal integration during saccade tasks?*

Bibliography

Abrams, R. A., & Dobkin, R. S. (1994). Inhibition of return: Effects of attentional cuing on eye movement latencies. *Journal of Experimental Psychology: Human Perception and Performance, 20*(3), 467–477.

Amiez, C., & Petrides, M. (2009). Anatomical organization of the eye fields in the human and non-human primate frontal cortex. *Progress in Neurobiology, 89*(2), 220–230. https://doi.org/10.1016/j.pneurobio.2009.07.010.

Anderson, T., Jenkins, I. H., Brooks, D. J., Hawken, M. B., Frackowiak, R. S., & Kennard, C. (1994). Cortical control of saccades and fixation in man. A PET study. *Brain, 117*(Pt 5), 1073–1084.

Aron, A. R., & Poldrack, R. A. (2006). Cortical and subcortical contributions to Stop signal response inhibition: Role of the subthalamic nucleus. *Journal of Neuroscience, 26*(9), 2424–2433. https://doi.org/10.1523/jneurosci.4682-05.2006.

Bahill, A. T., Clark, M. R., & Stark, L. (1975a). Dynamic overshoot in saccadic eye movements is caused by neurological control signal reversals. *Experimental Neurology, 48*(1), 107–122.

Bahill, A. T., Clark, M. R., & Stark, L. (1975b). Glissades—Eye movements generated by mismatched components of the saccadic motoneuronal control signal. *Mathematical Biosciences, 26*(3), 303–318.

Bahill, A. T., Clark, M. R., & Stark, L. (1975c). The main sequence, a tool for studying human eye movements. *Mathematical Biosciences, 24*(3), 191–204.

Bahill, A. T., & Stark, L. (1975). Overlapping saccades and glissades are produced by fatigue in the saccadic eye movement system. *Experimental Neurology, 48*(1), 95–106.

Bartlett, N. R. (1961). Latency of ocular fixation upon the second of two successive stimuli. *Perceptual and Motor Skills, 13*(3), 259–268.

Barton, J. J., Greenzang, C., Hefter, R., Edelman, J., & Manoach, D. S. (2006a). Switching, plasticity, and prediction in a saccadic task-switch paradigm. *Experimental Brain Research, 168*(1–2), 76–87. https://doi.org/10.1007/s00221-005-0091-1.

Barton, J. J., Raoof, M., Jameel, O., & Manoach, D. S. (2006b). Task-switching with antisaccades versus no-go trials: A comparison of inter-trial effects. *Experimental Brain Research, 172*(1), 114–119. https://doi.org/10.1007/s00221-005-0313-6.

Becker, W. (1989). The neurobiology of saccadic eye movements. Metrics. *Reviews of Oculomotor Research, 3,* 13–67.

Becker, W., & Fuchs, A. F. (1969). Further properties of the human saccadic system: Eye movements and correction saccades with and without visual fixation points. *Vision Research, 9*(10), 1247–1258. https://doi.org/10.1016/0042-6989(69)90112-6.

Becker, W., & Jürgens, R. (1979). An analysis of the saccadic system by means of double step stimuli. *Vision Research, 19*(9), 967–983.

Biscaldi, M., Fischer, B., & Stuhr, V. (1996). Human express saccade makers are impaired at suppressing visually evoked saccades. *Journal of Neurophysiology, 76*(1), 199–214.

Borji, A., & Itti, L. (2014). Defending Yarbus: Eye movements reveal observers' task. *Journal of Vision, 14*(3), 29. https://doi.org/10.1167/14.3.29.

Boucher, L., Palmeri, T. J., Logan, G. D., & Schall, J. D. (2007). Inhibitory control in mind and brain: An interactive race model of countermanding saccades. *Psychological Review, 114*(2), 376–397. https://doi.org/10.1037/0033-295x.114.2.376.

Bronstein, A. M., & Kennard, C. (1987). Predictive eye saccades are different from visually triggered saccades. *Vision Research, 27*(4), 517–520.

Brown, M. R. G., DeSouza, J. F., Goltz, H. C., Ford, K., Menon, R. S., Goodale, M. A., Everling, S. (2004). Comparison of memory- and visually guided saccades using event-related fMRI. *Journal of Neurophysiology, 91*(2), 873–889. https://doi.org/10.1152/jn.00382.2003.

Brown, M. R. G., Goltz, H. C., Vilis, T., Ford, K. A., & Everling, S. (2006). Inhibition and generation of saccades: Rapid event-related fMRI of prosaccades, antisaccades, and nogo trials. *Neuroimage, 33*(2), 644–659. https://doi.org/10.1016/j.neuroimage.2006.07.002.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature, 371*(6497), 511–513.

Camalier, C. R., Gotler, A., Murthy, A., Thompson, K. G., Logan, G. D., Palmeri, T. J., Schall, J. D. (2007). Dynamics of saccade target selection: Race model analysis of double step and search step saccade production in human and macaque. *Vision Research, 47*(16), 2187–2211. https://doi.org/10.1016/j.visres.2007.04.021.

Camchong, J., Dyckman, K. A., Austin, B. P., Clementz, B. A., & McDowell, J. E. (2008). Common neural circuitry supporting volitional saccades and its disruption in schizophrenia patients and relatives. *Biological Psychiatry, 64*(12), 1042–1050. https://doi.org/10.1016/j.biopsych.2008.06.015.

Carpenter, R. H. (2000). The neural control of looking. *Current Biology, 10*(8), R291–R293.

Carpenter, R. H., & Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature, 377*(6544), 59–62. https://doi.org/10.1038/377059a0.

Cherkasova, M. V., Manoach, D. S., Intriligator, J. M., & Barton, J. J. (2002). Antisaccades and task-switching: Interactions in controlled processing. *Experimental Brain Research, 144*(4), 528–537. https://doi.org/10.1007/s00221-002-1075-z.

Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across response modalities. *Journal of Cognitive Neuroscience, 19*(1), 69–80. https://doi.org/10.1162/jocn.2007.19.1.69.

Clementz, B. A., Brahmbhatt, S. B., McDowell, J. E., Brown, R., & Sweeney, J. A. (2007). When does the brain inform the eyes whether and where to move? An EEG study in humans. *Cerebral Cortex, 17*(11), 2634–2643. https://doi.org/10.1093/cercor/bhl171.

Clementz, B. A., Gao, Y., McDowell, J. E., Moratti, S., Keedy, S. K., & Sweeney, J. A. (2010). Top-down control of visual sensory processing during an ocular motor response inhibition task. *Psychophysiology, 47*(6), 1011–1018. https://doi.org/10.1111/j.1469-8986.2010.01026.x.

Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D. … Tamminga, C. A. (2016). Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry, 173*(4), 373–384. https://doi.org/10.1176/appi.ajp.2015.14091200

Connolly, J. D., Goodale, M. A., Desouza, J. F., Menon, R. S., & Vilis, T. (2000). A comparison of frontoparietal fMRI activation during anti-saccades and anti-pointing. *Journal of Neurophysiology, 84*(3), 1645–1655.

Connolly, J. D., Goodale, M. A., Goltz, H. C., & Munoz, D. P. (2005). fMRI activation in the human frontal eye field is correlated with saccadic reaction time. *Journal of Neurophysiology, 94*(1), 605–611.

Connolly, J. D., Goodale, M. A., Menon, R. S., & Munoz, D. P. (2002). Human fMRI evidence for the neural correlates of preparatory set. *Nature Neuroscience, 5*(12), 1345–1352. https://doi.org/10.1038/nn969.

Coren, S., & Hoenig, P. (1972). Effect of non-target stimuli upon length of voluntary saccades. *Perceptual and Motor Skills, 34*(2), 499–508.

Curtis, C. E., & Connolly, J. D. (2008). Saccade preparation signals in the human frontal and parietal cortices. *Journal of Neurophysiology, 99*(1), 133–145. https://doi.org/10.1152/jn.00899.2007.

Curtis, C. E., & D'Esposito, M. (2003). Success and failure suppressing reflexive behavior. *Journal of Cognitive Neuroscience, 15*(3), 409–418. https://doi.org/10.1162/089892903321593126.

Cutsuridis, V., Smyrnis, N., Evdokimidis, I., & Perantonis, S. (2007). A neural model of decision-making by the superior colicullus in an antisaccade task. *Neural Networks, 20*(6), 690–704. https://doi.org/10.1016/j.neunet.2007.01.004.

D'Esposito, M., Postle, B. R., & Rypma, B. (2000). Prefrontal cortical contributions to working memory: Evidence from event-related fMRI studies. *Experimental Brain Research, 133*(1), 3–11.

Darby, D. G., Nobre, A. C., Thangaraj, V., Edelman, R., Mesulam, M. M., & Warach, S. (1996). Cortical activation in the human brain during lateral saccades using EPISTAR functional magnetic resonance imaging. *Neuroimage, 3*(1), 53–62. https://doi.org/10.1006/nimg.1996.0006.

DeSouza, J. F., Menon, R. S., & Everling, S. (2003). Preparatory set associated with pro-saccades and anti-saccades in humans investigated with event-related FMRI. *Journal of Neurophysiology, 89*(2), 1016–1023.

Deubel, H. (1995). Separate adaptive mechanisms for the control of reactive and volitional saccadic eye movements. *Vision Research, 35*(23–24), 3529–3540. https://doi.org/10.1016/0042-6989(95)00058-m.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837.

Dodge, R. (1900). Visual perception during eye movement. *Psychological Review, 7*(5), 454.

Dodge, R. (1903). Five types of eye movement in the horizontal meridian plane of the field of regard. *American Journal of Physiology-Legacy Content, 8*.

Dodge, R. (1905). The illusion of clear vision during eye movement. *Psychological Bulletin, 2*(6), 193–199. https://doi.org/10.1037/h0070501.

Dodge, R. (1916). Visual motor functions. *Psychological Bulletin, 13*(11), 421–427. https://doi.org/10.1037/h0074240.

Duhamel, J. R., Colby, C. L., & Goldberg, M. E. (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science, 255*(5040), 90–92.

Dyckman, K. A., Camchong, J., Clementz, B. A., & McDowell, J. E. (2007). An effect of context on saccade-related behavior and brain activity. *Neuroimage, 36*(3), 774–784. https://doi.org/10.1016/j.neuroimage.2007.03.023.

Dyckman, K. A., & McDowell, J. E. (2005). Behavioral plasticity of antisaccade performance following daily practice. *Experimental Brain Research, 162*(1), 63–69. https://doi.org/10.1007/s00221-004-2105-9.

Erdmann, B., & Dodge, R. (1898). *Psychologische Untersuchungen über das Lesen auf experimenteller Grundlage*. Halle, AS: Niemeyer, M.

Ethridge, L. E., Brahmbhatt, S., Gao, Y., McDowell, J. E., & Clementz, B. A. (2009). Consider the context: Blocked versus interleaved presentation of antisaccade trials. *Psychophysiology, 46*(5), 1100–1107. https://doi.org/10.1111/j.1469-8986.2009.00834.x.

Ettinger, U., Kumari, V., Crawford, T. J., Davis, R. E., Sharma, T., & Corr, P. J. (2003). Reliability of smooth pursuit, fixation, and saccadic eye movements. *Psychophysiology, 40*(4), 620–628.

Evdokimidis, I., Smyrnis, N., Constantinidis, T. S., Stefanis, N. C., Avramopoulos, D., Paximadis, C. … Stefanis, C. N. (2002). The antisaccade task in a sample of 2,006 young men. I. Normal population characteristics. *Experimental Brain Research, 147*(1), 45–52. https://doi.org/10.1007/s00221-002-1208-4.

Everling, S., & Fischer, B. (1998). The antisaccade: A review of basic research and clinical studies. *Neuropsychologia, 36*(9), 885–899.

Everling, S., & Johnston, K. (2013). Control of the superior colliculus by the lateral prefrontal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 368*(1628), 20130068. https://doi.org/10.1098/rstb.2013.0068.

Findlay, J. M. (1982). Global visual processing for saccadic eye movements. *Vision Research, 22*(8), 1033–1045.

Findlay, J. M. (2009). Saccadic eye movement programming: Sensory and attentional factors. *Psychological Research, 73*(2), 127–135. https://doi.org/10.1007/s00426-008-0201-3.

Fischer, B., & Breitmeyer, B. (1987). Mechanisms of visual attention revealed by saccadic eye movements. *Neuropsychologia, 25*(1A), 73–83.

Fischer, B., Gezeck, S., & Hartnegg, K. (2000). On the production and correction of involuntary prosaccades in a gap antisaccade task. *Vision Research, 40*(16), 2211–2217.

Fischer, B., & Ramsperger, E. (1984). Human express saccades: Extremely short reaction times of goal directed eye movements. *Experimental Brain Research, 57*(1), 191–195.

Fischer, B., & Weber, H. (1992). Characteristics of "anti" saccades in man. *Experimental Brain Research, 89*(2), 415–424.

Fischer, B., & Weber, H. (1993). Express saccades and visual attention. *Behavioral and Brain Sciences, 16,* 553–610.

Forbes, K., & Klein, R. M. (1996). The magnitude of the fixation offset effect with endogenously and exogenously controlled saccades. *Journal of Cognitive Neuroscience, 8*(4), 344–352. https://doi.org/10.1162/jocn.1996.8.4.344.

Ford, K. A., Goltz, H. C., Brown, M. R., & Everling, S. (2005). Neural processes associated with antisaccade task performance investigated with event-related FMRI. *Journal of Neurophysiology, 94*(1), 429–440. https://doi.org/10.1152/jn.00471.2004.

Fox, P. T., Fox, J. M., Raichle, M. E., & Burde, R. M. (1985). The role of cerebral cortex in the generation of voluntary saccades: A positron emission tomographic study. *Journal of Neurophysiology, 54*(2), 348–369.

Friesen, C. K., & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review, 5*(3), 490–495.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1991). Neuronal activity related to saccadic eye movements in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 65*(6), 1464–1483.

Funahashi, S., Chafee, M. V., & Goldman-Rakic, P. S. (1993). Prefrontal neuronal activity in rhesus monkeys performing a delayed anti-saccade task. *Nature, 365*(6448), 753–756. https://doi.org/10.1038/365753a0.

Gezeck, S., Fischer, B., & Timmer, J. (1997). Saccadic reaction times: A statistical analysis of multimodal distributions. *Vision Research, 37*(15), 2119–2131.

Gnadt, J. W., Bracewell, R. M., & Andersen, R. A. (1991). Sensorimotor transformation during eye movements to remembered visual targets. *Vision Research, 31*(4), 693–715.

Godijn, R., & Theeuwes, J. (2002). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1039.

Gottlieb, J., & Goldberg, M. E. (1999). Activity of neurons in the lateral intraparietal area of the monkey during an antisaccade task. *Nature Neuroscience, 2*(10), 906–912. https://doi.org/10.1038/13209.

Grosbras, M. H., Leonards, U., Lobel, E., Poline, J. B., LeBihan, D., & Berthoz, A. (2001). Human cortical networks for new and familiar sequences of saccades. *Cerebral Cortex, 11*(10), 936–945.

Grosbras, M. H., Lobel, E., Van de Moortele, P. F., LeBihan, D., & Berthoz, A. (1999). An anatomical landmark for the supplementary eye fields in human revealed with functional magnetic resonance imaging. *Cerebral Cortex, 9*(7), 705–711.

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research, 58*(3), 455–472.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18,* 1279–1296.

Hallett, P. E., & Adams, B. D. (1980). The predictability of saccadic latency in a novel voluntary oculomotor task. *Vision Research, 20*(4), 329–339.

Hamm, J. P., Dyckman, K. A., Ethridge, L. E., McDowell, J. E., & Clementz, B. A. (2010). Preparatory activations across a distributed cortical network determine production of express saccades in humans. *Journal of Neuroscience, 30*(21), 7350–7357. https://doi.org/10.1523/jneurosci.0785-10.2010.

Hamm, J. P., Dyckman, K. A., McDowell, J. E., & Clementz, B. A. (2012). Pre-cue fronto-occipital alpha phase and distributed cortical oscillations predict failures of cognitive control. *Journal of Neuroscience, 32*(20), 7034–7041. https://doi.org/10.1523/jneurosci.5198-11.2012.

Hanes, D. P., & Carpenter, R. (1999). Countermanding saccades in humans. *Vision Research, 39*(16), 2777–2791.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science, 274*(5286), 427–430.

Hanes, D. P., & Wurtz, R. H. (2001). Interaction of the frontal eye field and superior colliculus for saccade generation. *Journal of Neurophysiology, 85*(2), 804–815.

Harris, C. M., & Wolpert, D. M. (2006). The main sequence of saccades optimizes speed-accuracy trade-off. *Biological Cybernetics, 95*(1), 21–29. https://doi.org/10.1007/s00422-006-0064-x.

Hayakawa, Y., Nakajima, T., Takagi, M., Fukuhara, N., & Abe, H. (2002). Human cerebellar activation in relation to saccadic eye movements: A functional magnetic resonance imaging study. *Ophthalmologica, 216*(6), 399–405.

Hayhoe, M., & Ballard, D. (2005). Eye movements in natural behavior. *Trends in Cognitive Sciences, 9*(4), 188–194. https://doi.org/10.1016/j.tics.2005.02.009.

Heeman, J., Theeuwes, J., & Van der Stigchel, S. (2014). The time course of top-down control on saccade averaging. *Vision Research, 100,* 29–37. https://doi.org/10.1016/j.visres.2014.03.007.

Heide, W., Binkofski, F., Seitz, R. J., Posse, S., Nitschke, M. F., Freund, H. J., Kompf, D. (2001). Activation of frontoparietal cortices during memorized triple-step sequences of saccadic eye movements: An fMRI study. *European Journal of Neuroscience, 13*(6), 1177–1189.

Heiser, L. M., & Colby, C. L. (2006). Spatial updating in area LIP is independent of saccade direction. *Journal of Neurophysiology, 95*(5), 2751–2767. https://doi.org/10.1152/jn.00054.2005.

Henderson, J. M., & Hollingworth, A. (2003). Eye movements and visual memory: Detecting changes to saccade targets in scenes. *Perception and Psychophysics, 65*(1), 58–71.

Henik, A., Rafal, R., & Rhodes, D. (1994). Endogenously generated and visually guided saccades after lesions of the human frontal eye fields. *Journal of Cognitive Neuroscience, 6*(4), 400–411. https://doi.org/10.1162/jocn.1994.6.4.400.

Hering, E. (1879). Uber Muskelgerausche des Auges. *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften in Wien. Mathematisch-naturwissenschaftliche Klasse. Abt III, 79*, 137–154.

Hermens, F., & Walker, R. (2015). The influence of social and symbolic cues on observers' gaze behaviour. *British Journal of Psychology*. https://doi.org/10.1111/bjop.12159.

Hess, W., Burgi, S., & Bucher, V. (1946). Motor function of tectal and tegmental area. *Monatsschrift fur Psychiatrie und Neurologie, 112,* 1–52.

Hikosaka, O., Nakamura, K., & Nakahara, H. (2006). Basal ganglia orient eyes to reward. *Journal of Neurophysiology, 95*(2), 567–584. https://doi.org/10.1152/jn.00458.2005.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57*(6), 787–795.

Holzman, P. S., Proctor, L. R., & Hughes, D. W. (1973). Eye-tracking patterns in schizophrenia. *Science, 181,* 179–181.

Honda, H. (1989). Perceptual localization of visual stimuli flashed during saccades. *Perception and Psychophysics, 45*(2), 162–174.

Hooge, I. T., & Frens, M. A. (2000). Inhibition of saccade return (ISR): Spatio-temporal properties of saccade programming. *Vision Research, 40*(24), 3415–3426.

Hopp, J. J., & Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Progress in Neurobiology, 72*(1), 27–53. https://doi.org/10.1016/j.pneurobio.2003.12.002.

Hutton, S. B. (2008). Cognitive control of saccadic eye movements [Review]. *Brain and Cognition, 68*(3), 327–340. https://doi.org/10.1016/j.bandc.2008.08.021.

Ibbotson, M., & Krekelberg, B. (2011). Visual perception and saccadic eye movements. *Current Opinion in Neurobiology, 21*(4), 553–558. https://doi.org/10.1016/j.conb.2011.05.012.

Ito, S., Stuphorn, V., Brown, J. W., & Schall, J. D. (2003). Performance monitoring by the anterior cingulate cortex during saccade countermanding. *Science, 302*(5642), 120–122. https://doi.org/10.1126/science.1087847.

Jamadar, S. D., Fielding, J., & Egan, G. F. (2013). Quantitative meta-analysis of fMRI and PET studies reveals consistent activation in fronto-striatal-parietal regions and cerebellum during antisaccades and prosaccades. *Frontiers in Psychology, 4,* 749. https://doi.org/10.3389/fpsyg.2013.00749.

Javal, E. (1879). Essai sur la Physiologie de la Lecture. *Annales D'Oculistique*.

Johnston, K., & Everling, S. (2008). Neurophysiology and neuroanatomy of reflexive and voluntary saccades in non-human primates. *Brain and Cognition, 68*(3), 271–283. https://doi.org/10.1016/j.bandc.2008.08.017.

Joiner, W. M., Cavanaugh, J., & Wurtz, R. H. (2013). Compression and suppression of shifting receptive field activity in frontal eye field neurons. *Journal of Neuroscience, 33*(46), 18259–18269. https://doi.org/10.1523/jneurosci.2964-13.2013.

Joiner, W. M., & Shelhamer, M. (2006). An internal clock generates repetitive predictive saccades. *Experimental Brain Research, 175*(2), 305–320. https://doi.org/10.1007/s00221-006-0554-z.

Kagan, I., Iyer, A., Lindner, A., & Andersen, R. A. (2010). Space representation for eye movements is more contralateral in monkeys than in humans. *Proceedings of the National Academy of Sciences of the United States of America, 107*(17), 7933–7938. https://doi.org/10.1073/pnas.1002825107.

Kalesnykas, R. P., & Hallett, P. E. (1994). Retinal eccentricity and the latency of eye saccades. *Vision Research, 34*(4), 517–531.

Kimmig, H., Greenlee, M. W., Gondan, M., Schira, M., Kassubek, J., & Mergner, T. (2001). Relationship between saccadic eye movements and cortical activity as measured by fMRI: Quantitative and qualitative aspects. *Experimental Brain Research, 141*(2), 184–194. https://doi.org/10.1007/s002210100844.

Kingstone, A., & Klein, R. M. (1993a). Visual offsets facilitate saccadic latency: Does predisengagement of visuospatial attention mediate this gap effect? *Journal of Experimental Psychology: Human Perception and Performance, 19*(6), 1251–1265.

Kingstone, A., & Klein, R. M. (1993b). What are human express saccades? *Perception and Psychophysics, 54*(2), 260–273.

Klein, C., & Berg, P. (2001). Four-week test-retest stability of individual differences in the saccadic CNV, two saccadic task parameters, and selected neuropsychological tests. *Psychophysiology, 38*(4), 704–711.

Klein, C., & Fischer, B. (2005a). Developmental fractionation and differential discrimination of the anti-saccadic direction error. *Experimental Brain Research, 165*(1), 132–138. https://doi.org/10.1007/s00221-005-2324-8.

Klein, C., & Fischer, B. (2005b). Instrumental and test-retest reliability of saccadic measures. *Biological Psychology, 68*(3), 201–213. https://doi.org/10.1016/j.biopsycho.2004.06.005.

Klein, R. M. (2000). Inhibition of return. *Trends in Cognitive Sciences, 4*(4), 138–147.

Klier, E. M., & Angelaki, D. E. (2008). Spatial updating and the maintenance of visual constancy. *Neuroscience, 156*(4), 801–818. https://doi.org/10.1016/j.neuroscience.2008.07.079.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219–227.

Konen, C. S., Kleiser, R., Wittsack, H.-J., Bremmer, F., & Seitz, R. J. (2004). The encoding of saccadic eye movements within human posterior parietal cortex. *Neuroimage, 22*(1), 304–314.

Koval, M. J., Thomas, B. S., & Everling, S. (2005). Task-dependent effects of social attention on saccadic reaction times. *Experimental Brain Research, 167*(3), 475–480. https://doi.org/10.1007/s00221-005-0206-8.

Kowler, E. (2011). Eye movements: The past 25 years. *Vision Research, 51*(13), 1457–1483. https://doi.org/10.1016/j.visres.2010.12.014.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916. https://doi.org/10.1016/0042-6989(94)00279-u.

Krafft, C. E., Schwarz, N. F., Chi, L., Li, Q., Schaeffer, D. J., Rodrigue, A. L., McDowell, J. E. (2012). The location and function of parietal cortex supporting of reflexive and volitional saccades, a meta-analysis of over a decade of functional MRI data. In A. Costa & E. Villalba (Eds.), *Horizons of neuroscience research* (Vol. 9, pp. 131–153). Hauppauge, NY: Nova Science Publishers.

Krebs, R. M., Woldorff, M. G., Tempelmann, C., Bodammer, N., Noesselt, T., Boehler, C. N. … Schoenfeld, M. A. (2010). High-field FMRI reveals brain activation patterns underlying saccade execution in the human superior colliculus. *PLoS ONE, 5*(1), e8691. https://doi.org/10.1371/journal.pone.0008691.

Kurtzberg, D., & Vaughan, H. G., Jr. (1982). Topographic analysis of human cortical potentials preceding self-initiated and visually triggered saccades. *Brain Research, 243*(1), 1–9.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature, 369*(6483), 742–744. https://doi.org/10.1038/369742a0.

Lee, J., Park, C., Dyckman, K. A., Lazar, N. A., Austin, B. P., Li, Q., McDowell, J. E. (2013). Practice-related changes in neural activation patterns investigated via wavelet-based clustering analysis. *Human Brain Mapping, 34*(9), 2276–2291. https://doi.org/10.1002/hbm.22066.

Leigh, R. J., & Zee, D. S. (2015). *The neurology of eye movements*. USA: Oxford University Press.

Li, Q., Amlung, M. T., Valtcheva, M., Camchong, J., Austin, B. P., Dyckman, K. A., McDowell, J. E. (2012). Evidence from cluster analysis for differentiation of antisaccade performance groups based on speed/accuracy trade-offs. *International Journal of Psychophysiology, 85,* 274–277. https://doi.org/10.1016/j.ijpsycho.2012.03.008.

Logan, G. D., & Irwin, D. E. (2000). Don't look! Don't touch! Inhibitory control of eye and hand movements. *Psychonomic Bulletin & Review, 7*(1), 107–112.

Luke, S. G., & Henderson, J. M. (2016). The influence of content meaningfulness on eye movements across tasks: Evidence from scene viewing and reading. *Front Psychol, 7,* 257. https://doi.org/10.3389/fpsyg.2016.00257.

Luna, B., Thulborn, K. R., Strojwas, M. H., McCurtain, B. J., Berman, R. A., Genovese, C. R., & Sweeney, J. A. (1998). Dorsal cortical regions subserving visually guided saccades in humans: An fMRI study. *Cerebral Cortex, 8*(1), 40–47.

Mackay, M., Cerf, M., & Koch, C. (2012). Evidence for two distinct mechanisms directing gaze in natural scenes. *Journal of Vision, 12*(4), 9. https://doi.org/10.1167/12.4.9.

Manoach, D. S., Thakkar, K. N., Cain, M. S., Polli, F. E., Edelman, J. A., Fischl, B., & Barton, J. J. (2007). Neural activity is modulated by trial history: A functional magnetic resonance imaging study of the effects of a previous antisaccade. *Journal of Neuroscience, 27*(7), 1791–1798. https://doi.org/10.1523/jneurosci.3662-06.2007.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5*(3), 229–240.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron, 49*(2), 297–305.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: A neurophysiological analysis. *Trends in Neurosciences, 32*(9), 463–475.

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience, 14*(2), 83–96.

Massen, C. (2004). Parallel programming of exogenous and endogenous components in the antisaccade task. *Quarterly Journal of Experimental Psychology A, 57*(3), 475–498. https://doi.org/10.1080/02724980343000341.

Matin, E. (1974). Saccadic suppression: A review and an analysis. *Psychological Bulletin, 81*(12), 899.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. *Science, 148*(3676), 1485–1488. https://doi.org/10.1126/science.148.3676.1485.

Maylor, E. A., & Hockey, R. (1985). Inhibitory component of externally controlled covert orienting in visual space. *Journal of Experimental Psychology: Human Perception and Performance, 11*(6), 777–787.

McDowell, J. E., Clementz, B. A., & Wixted, J. T. (1996). Timing and amplitude of saccades during predictive saccadic tracking in schizophrenia. *Psychophysiology, 33*(1), 93–101.

McDowell, J. E., Dyckman, K. A., Austin, B. P., & Clementz, B. A. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades: Evidence from studies of humans. [Review]. *Brain and Cognition, 68*(3), 255–270.

McDowell, J. E., Kissler, J. M., Berg, P., Dyckman, K. A., Gao, Y., Rockstroh, B., & Clementz, B. A. (2005). Electroencephalography/magnetoencephalography study of cortical activities preceding prosaccades and antisaccades. *NeuroReport, 16*(7), 663–668.

McLaughlin, S. C. (1967). Parametric adjustment in saccadic eye movements. *Perception and Psychophysics, 2*(8), 359–362.

McPeek, R. M., & Keller, E. L. (2002). Saccade target selection in the superior colliculus during a visual search task. *Journal of Neurophysiology, 88*(4), 2019–2034.

Medendorp, W. P., Goltz, H. C., & Vilis, T. (2005). Remapping the remembered target location for anti-saccades in human posterior parietal cortex. *Journal of Neurophysiology, 94*(1), 734–740. https://doi.org/10.1152/jn.01331.2004.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*(3), 211–253. https://doi.org/10.1006/cogp.2000.0736.

Melamed, E., & Larsen, B. (1979). Cortical activation pattern during saccadic eye movements in humans: Localization by focal cerebral blood flow increases. *Annals of Neurology, 5*(1), 79–88.

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2003). Spatial updating in human parietal cortex. *Neuron, 39*(2), 361–373.

Meyhöfer, I., Bertsch, K., Esser, M., & Ettinger, U. (2016). Variance in saccadic eye movements reflects stable traits. *Psychophysiology, 53*(4), 566–578. https://doi.org/10.1111/psyp.12592.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24,* 167–202. https://doi.org/10.1146/annurev.neuro.24.1.167.

Miller, J. M., Anstis, T., & Templeton, W. B. (1981). Saccadic plasticity: Parametric adaptive control by retinal feedback. *Journal of Experimental Psychology: Human Perception and Performance, 7*(2), 356–366.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*(8), 17–17. https://doi.org/10.1167/11.8.17.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action* (Vol. 27). England.

Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences, 6,* 414–417.

Mokler, A., & Fischer, B. (1999). The recognition and correction of involuntary prosaccades in an antisaccade task. *Experimental Brain Research, 125*(4), 511–516.

Morrone, M. C., Ross, J., & Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. *Nature Neuroscience, 8*(7), 950–954.

Morrone, M. C., Ross, J., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *Journal of Neuroscience, 17*(20), 7941–7953.

Mort, D. J., Perry, R. J., Mannan, S. K., Hodgson, T. L., Anderson, E., Quest, R. … Kennard, C. (2003). Differential cortical activation during voluntary and reflexive saccades in man. *Neuroimage, 18*(2), 231–246. https://doi.org/10.1016/s1053-8119(02)00028-9.

Mosimann, U. P., Felblinger, J., Colloby, S. J., & Muri, R. M. (2004). Verbal instructions and top-down saccade control. *Experimental Brain Research, 159*(2), 263–267. https://doi.org/10.1007/s00221-004-2086-8.

Munoz, D. P., & Everling, S. (2004). Look away: The anti-saccade task and the voluntary control of eye movement. *Nature Reviews Neuroscience, 5*(3), 218–228. https://doi.org/10.1038/nrn1345.

Muri, R. M., Iba-Zizen, M. T., Derosier, C., Cabanis, E. A., & Pierrot-Deseilligny, C. (1996). Location of the human posterior eye field with functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery and Psychiatry, 60*(4), 445–448.

Neggers, S., Raemaekers, M., Lampmann, E., Postma, A., & Ramsey, N. (2005). Cortical and subcortical contributions to saccade latency in the human brain. *European Journal of Neuroscience, 21*(10), 2853–2863.

Nitschke, M. F., Binkofski, F., Buccino, G., Posse, S., Erdmann, C., Kömpf, D., & Heide, W. (2004). Activation of cerebellar hemispheres in spatial memorization of saccadic eye movements: An fMRI study. *Human Brain Mapping, 22*(2), 155–164.

Noorani, I., & Carpenter, R. H. (2013). Antisaccades as decisions: LATER model predicts latency distributions and error responses. *European Journal of Neuroscience, 37*(2), 330–338. https://doi.org/10.1111/ejn.12025.

Nummenmaa, L., Hyona, J., & Calvo, M. G. (2009). Emotional scene content drives the saccade generation system reflexively. *Journal of Experimental Psychology: Human Perception and Performance, 35*(2), 305–323. https://doi.org/10.1037/a0013626.

O'Driscoll, G. A., Alpert, N. M., Matthysse, S. W., Levy, D. L., Rauch, S. L., & Holzman, P. S. (1995). Functional neuroanatomy of antisaccade eye movements investigated with positron emission tomography. *Proc Natl Acad Sci U S A, 92*(3), 925–929.

O'Driscoll, G. A., Wolff, A. L., Benkelfat, C., Florencio, P. S., Lal, S., & Evans, A. C. (2000). Functional neuroanatomy of smooth pursuit and predictive saccades. *NeuroReport, 11*(6), 1335–1340.

Orgogozo, J., & Larsen, B. (1979). Activation of the supplementary motor area during voluntary movement in man suggests it works as a supramotor area. *Science, 206*(4420), 847–850.

Panouilleres, M., Weiss, T., Urquizar, C., Salemme, R., Munoz, D. P., & Pelisson, D. (2009). Behavioral evidence of separate adaptation mechanisms controlling saccade amplitude lengthening and shortening. *Journal of Neurophysiology, 101*(3), 1550–1559. https://doi.org/10.1152/jn.90988.2008.

Paus, T., Petrides, M., Evans, A. C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: A positron emission tomography study. *Journal of Neurophysiology, 70*(2), 453–469.

Pélisson, D., Alahyane, N., Panouillères, M., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience & Biobehavioral Reviews, 34*(8), 1103–1120. https://doi.org/10.1016/j.neubiorev.2009.12.010.

Pierce, J. E., McCardel, J. B., & McDowell, J. E. (2015). Trial-type probability and task-switching effects on behavioral response characteristics in a mixed saccade task. *Experimental Brain Research, 233*(3), 959–969. https://doi.org/10.1007/s00221-014-4170-z.

Pierce, J. E., & McDowell, J. E. (2016). Modulation of cognitive control levels via manipulation of saccade trial-type probability assessed with event-related BOLD fMRI. *Journal of Neurophysiology, 115*(2), 763–772. https://doi.org/10.1152/jn.00776.2015.

Pierrot-Deseilligny, C., Muri, R. M., Nyffeler, T., & Milea, D. (2005). The role of the human dorsolateral prefrontal cortex in ocular motor behavior. *Annals of the New York Academy of Sciences, 1039,* 239–251. https://doi.org/10.1196/annals.1325.023.

Pierrot-Deseilligny, C., Muri, R. M., Ploner, C. J., Gaymard, B., Demeret, S., & Rivaud-Pechoux, S. (2003). Decisional role of the dorsolateral prefrontal cortex in ocular motor behaviour. *Brain, 126*(Pt 6), 1460–1473.

Pierrot-Deseilligny, C., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of reflexive visually-guided saccades. *Brain, 114*(Pt 3), 1473–1485.

Polli, F. E., Barton, J. J., Cain, M. S., Thakkar, K. N., Rauch, S. L., & Manoach, D. S. (2005). Rostral and dorsal anterior cingulate cortex make dissociable contributions during antisaccade error commission. *Proceedings of the National Academy of Sciences of the United States of America, 102*(43), 15700–15705. https://doi.org/10.1073/pnas.0503657102.

Posner, M. I. (1980). Orienting of attention. *The Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. *Attention and performance X: Control of language processes, 32,* 531–556.

Pratt, J., Bekkering, H., & Leung, M. (2000). Estimating the components of the gap effect. *Experimental Brain Research, 130*(2), 258–263.

Rafal, R., Egly, R., & Rhodes, D. (1994). Effects of inhibition of return on voluntary and visually guided saccades. *Canadian Journal of Experimental Psychology, 48*(2), 284–300.

Reddi, B. A., & Carpenter, R. H. (2000). The influence of urgency on decision time. *Nature Neuroscience, 3*(8), 827–830. https://doi.org/10.1038/77739.

Reuter-Lorenz, P. A., Hughes, H. C., & Fendrich, R. (1991). The reduction of saccadic latency by prior offset of the fixation point: An analysis of the gap effect. *Perception and Psychophysics, 49*(2), 167–175.

Reuter-Lorenz, P. A., Oonk, H., Barnes, L., & Hughes, H. (1995). Effects of warning signals and fixation point offsets on the latencies of pro-versus antisaccades: Implications for an interpretation of the gap effect. *Experimental Brain Research, 103*(2), 287–293.

Risko, E. F., Anderson, N. C., Lanthier, S., & Kingstone, A. (2012). Curious eyes: Individual differences in personality predict eye movement behavior in scene-viewing. *Cognition, 122*(1), 86–90. https://doi.org/10.1016/j.cognition.2011.08.014.

Rizzolatti, G., Riggio, L., Dascola, I., & Umilta, C. (1987). Reorienting attention across the horizontal and vertical meridians: Evidence in favor of a premotor theory of attention. *Neuropsychologia, 25*(1A), 31–40.

Robinson, D. A. (1968). The oculomotor control system: A review. *Proceedings of the IEEE, 56*(6), 1032–1049.

Robinson, D. A. (1981). The use of control systems analysis in the neurophysiology of eye movements. *Annual Review of Neuroscience, 4*(1), 463–503.

Robinson, D. L., & McClurkin, J. W. (1989). The visual superior colliculus and pulvinar. *Reviews of Oculomotor Research, 3,* 337–360.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207–231. https://doi.org/10.1037/0096-3445.124.2.207.

Ross, J., Morrone, M. C., & Burr, D. C. (1997). Compression of visual space before saccades. *Nature, 386*(6625), 598–601.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences, 24*(2), 113–121.

Ross, S. M., & Ross, L. E. (1987). Children's and adults' predictive saccades to square-wave targets. *Vision Research, 27*(12), 2177–2180.

Saslow, M. G. (1967). Effects of components of displacement-step stimuli upon latency for saccadic eye movement. *J Opt Soc Am, 57*(8), 1024–1029.

Schaeffer, D. J., Chi, L., Krafft, C. E., Li, Q., Schwarz, N. F., & McDowell, J. E. (2015). Individual differences in working memory moderate the relationship between prosaccade latency and antisaccade error rate. *Psychophysiology, 52*(4), 605–608. https://doi.org/10.1111/psyp.12380.

Schiller, P. H., Sandell, J. H., & Maunsell, J. H. (1987). The effect of frontal eye field and superior colliculus lesions on saccadic latencies in the rhesus monkey. *Journal of Neurophysiology, 57*(4), 1033–1049.

Schiller, P. H., & Tehovnik, E. J. (2005). Neural mechanisms underlying target selection with saccadic eye movements. *Progress in Brain Research, 149,* 157–171. https://doi.org/10.1016/s0079-6123(05)49012-3.

Sharika, K. M., Ramakrishnan, A., & Murthy, A. (2014). Use of exocentric and egocentric representations in the concurrent planning of sequential saccades. *Journal of Neuroscience, 34*(48), 16009–16021. https://doi.org/10.1523/jneurosci.0328-14.2014.

Shelhamer, M., & Joiner, W. M. (2003). Saccades exhibit abrupt transition between reactive and predictive; predictive saccade sequences have long-term correlations. *Journal of Neurophysiology, 90*(4), 2763–2769. https://doi.org/10.1152/jn.00478.2003.

Shenhav, A., Botvinick, M. M., & Cohen, J. D. (2013). The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron, 79*(2), 217–240. https://doi.org/10.1016/j.neuron.2013.07.007.

Shires, J., Joshi, S., & Basso, M. A. (2010). Shedding new light on the role of the basal ganglia-superior colliculus pathway in eye movements. *Current Opinion in Neurobiology, 20*(6), 717–725. https://doi.org/10.1016/j.conb.2010.08.008.

Simó, L. S., Krisky, C. M., & Sweeney, J. A. (2005). Functional neuroanatomy of anticipatory behavior: Dissociation between sensory-driven and memory-driven systems. *Cerebral Cortex, 15*(12), 1982–1991.

Smit, A., Van Gisbergen, J., & Cools, A. (1987). A parametric analysis of human saccades in different experimental paradigms. *Vision Research, 27*(10), 1745–1762.

Smyrnis, N., Evdokimidis, I., Stefanis, N. C., Constantinidis, T. S., Avramopoulos, D., Theleritis, C. … Stefanis, C. N. (2002). The antisaccade task in a sample of 2,006 young males. II. Effects of task parameters. *Experimental Brain Research, 147*(1), 53–63. https://doi.org/10.1007/s00221-002-1207-5.

Stark, L., Vossius, G., & Young, L. R. (1962). Predictive control of eye tracking movements. *Institute of Radio Engineers Trans in Human Factors and Electronics, 3,* 52–57.

Stevenson, S. B., Volkmann, F. C., Kelly, J. P., & Riggs, L. A. (1986). Dependence of visual suppression on the amplitudes of saccades and blinks. *Vision Research, 26*(11), 1815–1824.

Stuyven, E., Van der Goten, K., Vandierendonck, A., Claeys, K., & Crevits, L. (2000). The effect of cognitive load on saccadic eye movements. *Acta Psychologica (Amst), 104*(1), 69–85.

Sweeney, J., Mintun, M., Kwee, S., Wiseman, M., Brown, D., Rosenberg, D., & Carl, J. (1996). Positron emission tomography study of voluntary saccadic eye movements and spatial working memory. *Journal of Neurophysiology, 75*(1), 454–468.

Taya, S., Windridge, D., & Osman, M. (2012). Looking to score: The dissociation of goal influence on eye movement and meta-attentional allocation in a complex dynamic natural scene. *PLoS ONE, 7*(6), e39060. https://doi.org/10.1371/journal.pone.0039060.

Thakkar, K. N., Schall, J. D., Logan, G. D., & Park, S. (2015). Response inhibition and response monitoring in a saccadic double-step task in schizophrenia. *Brain and Cognition, 95,* 90–98. https://doi.org/10.1016/j.bandc.2015.01.007.

Theeuwes, J., Kramer, A. F., Hahn, S., & Irwin, D. E. (1998). Our eyes do not always go where we want them to go: Capture of the eyes by new objects. *Psychological Science, 9*(5), 379–385.

Trappenberg, T., Dorris, M., Munoz, D., & Klein, R. (2001). A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the superior colliculus. *Journal of Cognitive Neuroscience, 13*(2), 256–271.

Umeno, M. M., & Goldberg, M. E. (1997). Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *Journal of Neurophysiology, 78*(3), 1373–1383.

Van Der Werf, J., Jensen, O., Fries, P., & Medendorp, W. P. (2008). Gamma-band activity in human posterior parietal cortex encodes the motor goal during delayed prosaccades and antisaccades. *Journal of Neuroscience, 28*(34), 8397–8405. https://doi.org/10.1523/jneurosci.0630-08.2008.

Volkmann, F. C. (1986). Human visual suppression. *Vision Research, 26*(9), 1401–1416.

von Wartburg, R., Wurtz, P., Pflugshaupt, T., Nyffeler, T., Luthi, M., & Muri, R. M. (2007). Size matters: Saccades during scene perception. *Perception, 36*(3), 355–365.

Voogd, J., Schraa-Tam, C. K., van der Geest, J. N., & De Zeeuw, C. I. (2012). Visuomotor cerebellum in human and nonhuman primates. *The Cerebellum, 11*(2), 392–410.

Wade, N. J., Tatler, B. W., & Heller, D. (2003). Dodge-ing the issue: Dodge, Javal, Hering, and the measurement of saccades in eye-movement research. *Perception, 32*(7), 793–804.

Walker, M., Fitzgibbon, E. J., & Goldberg, M. E. (1995). Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *Journal of Neurophysiology, 73*(5), 1988–2003.

Walker, R., Deubel, H., Schneider, W. X., & Findlay, J. M. (1997). Effect of remote distractors on saccade programming: Evidence for an extended fixation zone. *Journal of Neurophysiology, 78*(2), 1108–1119.

Walker, R., Walker, D. G., Husain, M., & Kennard, C. (2000). Control of voluntary and reflexive saccades. *Experimental Brain Research, 130*(4), 540–544.

Wallman, J., & Fuchs, A. F. (1998). Saccadic gain modification: Visual error drives motor adaptation. *Journal of Neurophysiology, 80*(5), 2405–2416.

Weber, H., Aiple, F., Fischer, B., & Latanov, A. (1992). Dead zone for express saccades. *Experimental Brain Research, 89*(1), 214–222.

Weber, R. B., & Daroff, R. B. (1972). Corrective movements following refixation saccades: Type and control system analysis. *Vision Research, 12*(3), 467–475.

Weiler, J., & Heath, M. (2012). Task-switching in oculomotor control: Unidirectional switch-cost when alternating between pro- and antisaccades. *Neuroscience Letters, 530*(2), 150–154. https://doi.org/10.1016/j.neulet.2012.10.007.

Wheeless, L. L., Jr., Boynton, R. M., & Cohen, G. H. (1966). Eye-movement responses to step and pulse-step stimuli. *Journal of the Optical Society of America, 56*(7), 956–960.

Wolohan, F. D., & Knox, P. C. (2014). Oculomotor inhibitory control in express saccade makers. *Experimental Brain Research*. https://doi.org/10.1007/s00221-014-4076-9.

Wong, A. L., & Shelhamer, M. (2011). Sensorimotor adaptation error signals are derived from realistic predictions of movement outcomes. *Journal of Neurophysiology, 105*(3), 1130–1140.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research, 48*(20), 2070–2089. https://doi.org/10.1016/j.visres.2008.03.021.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*(5), 601–621.

Yantis, S., & Jonides, J. (1990). Abrupt visual onsets and selective attention: Voluntary versus automatic allocation. *Journal of Experimental Psychology: Human Perception and Performance, 16*(1), 121–134.

Yarbus, A. L. (1967). *Eye movements during perception of complex objects.* Berlin: Springer.

Zorn, A., Joiner, W. M., Lasker, A. G., & Shelhamer, M. (2007). Sensory versus motor information in the control of predictive saccade timing. *Experimental Brain Research, 179*(3), 505–515. https://doi.org/10.1007/s00221-006-0806-y.

Chapter 3
Fixational Eye Movements

Robert G. Alexander and Susana Martinez-Conde

Contents

Abstract There is too much going on around us to see everything at once, or to simultaneously process all the information in our field of view. Instead, we normally direct our gaze to parts of the scene that are particularly meaningful or important. Yet, even when we think that we are keeping our eyes still on an object of interest, our eyes remain in continuous motion. This chapter reviews the different kinds of "fixational" eye movements (the eye movements that occur when our gaze is "fixed" on an object or a point in space). We also discuss the effects that fixational eye movements have on vision and perception, their potential adaptive advantages, and their generation mechanisms. After reading this chapter, you will be able to articulate the importance of fixational eye movements in facilitating and influencing visual perception (for instance, in helping us to see the world and assisting us as we perform specific tasks). You will be able to explain the role of microsaccades (the largest of these small fixational eye movements) in counteracting neural adaptation and visual fading (the perceptual vanishing of an unchanging stimulus). You will also be able to outline the

---

R. G. Alexander (✉) · S. Martinez-Conde
State University of New York Downstate Health Sciences University, Brooklyn, USA
e-mail: robert.alexander@downstate.edu

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_3

neural generation of microsaccades. In addition, this chapter will help you understand the relationship between fixational eye movements and shifts of attention. Finally, you will be able to describe, in general terms, the potential relevance of fixational eye movements to the diagnosis of neurological disorders.

3.1 Introduction: Tremor, Drift, and Microsaccades, the Three Kinds of Fixational Eye Movements

Generally, we think of our eyes as either being still (as when we fix our gaze on a specific object or point in space) or moving (as when our gaze darts from one object of interest to the next, or smoothly tracks moving objects such as cars or airplanes). Periods of relative gaze stillness are called "fixations." The term "fixation" is a misnomer, however. Our eyes are never perfectly at rest, even when we think we are holding our gaze steady on a single location—see Fig. 3.1. Small, involuntary, and typically unconscious eye movements known as "fixational eye movements" constantly interrupt visual fixation.

There are three types of fixational eye movements: Drift (sometimes called "slow control"), oculomotor microtremor (which we will refer to here simply as "tremor"), and microsaccades. Microsaccades, the largest of the fixational eye movements, are quick jerks of the eye that typically occur once or twice a second—see Figs. 3.1 and 3.2. Researchers have used many different criteria to detect microsaccades, generally identifying them from their velocities (i.e. as micro-saccadic motions with 2d velocities above a predetermined threshold). Detection criteria differ as to the maximum angular distance that microsaccades may travel across the visual field (i.e. the maximum microsaccadic amplitude). Vision scientists measure angular distances in degrees of "visual angle," using a retinal reference frame to keep such distances invariant across varying experimental scenarios. The current (though arbitrary) consensus is that microsaccades have typical amplitudes smaller than one degree, which is a lesser span than the width of your thumbnail at arm's length (Martinez-Conde et al., 2013). Regular saccades, by comparison, can be tens of degrees in amplitude, and thus they are substantially larger (and also faster) than microsaccades. There is no clear distinction between microsaccades and saccades based solely on amplitude or velocity, however, and converging research indicates that microsaccades and saccades are generated by the same underlying circuitry—see Sect. 3.5 in the current chapter. Microsaccades are therefore essentially the same kind of eye movement as regular saccades (Martinez-Conde, 2006; Mergenthaler & Engbert, 2010; Otero-Millan et al., 2008)—see the text box on measuring fixational eye movements in this chapter for more details.

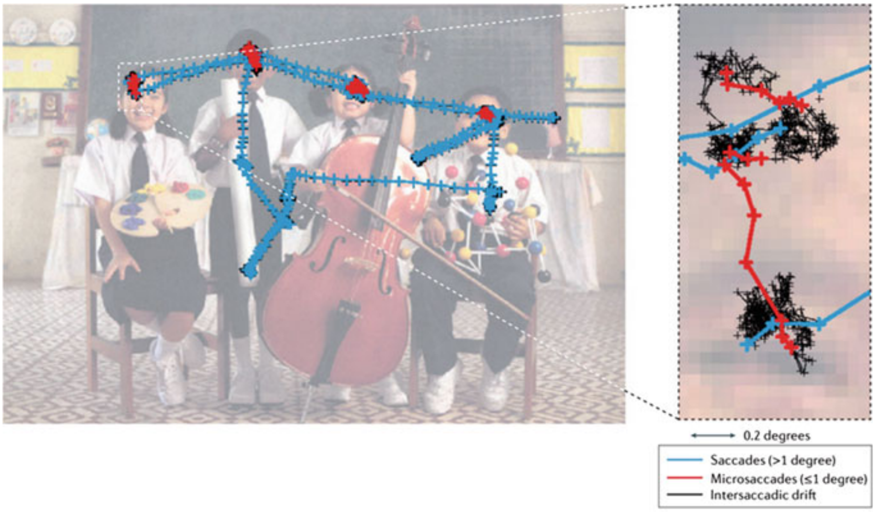
Drift is a slow, constant movement of the eye that takes place between (micro)saccades (Barlow, 1952; see also Rolfs, 2009). Though the amplitude and speed of recorded drift varies across studies (see Martinez-Conde, Macknik, & Hubel, 2004), its amplitude is generally a fraction of that of a microsaccade (usually less than a tenth of a degree) and its speed is typically around 50 arcminutes per second (Rucci & Poletti, 2015). Drifts do not always occur in both eyes at the same time, or in the same direction in each eye (Yarbus, 1967; Martinez-Conde et al., 2004). Drift trajectories can be statistically modelled as a self-avoiding random walk (Engbert, Mergenthaler, Sinn, & Pikovsky, 2011).

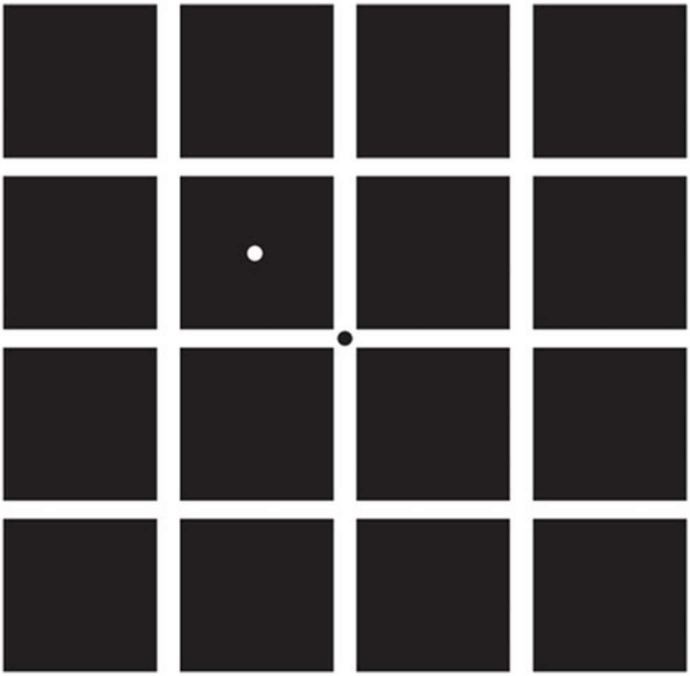
Tremor, the smallest of all fixational eye movements, is a high frequency, very small-amplitude oscillation that occurs simultaneously with drift. Tremor has typical amplitudes of about four-thousandths of a degree—a fraction of the amplitude of drift—and a frequency up to at least 100 Hz (Ditchburn & Ginsborg, 1953; Martinez-Conde et al., 2004; Spauschus, Marsden, Halliday, Rosenberg, & Brown, 1999). Tremor might be a byproduct of the firing of motor neurons (Eizenman, Hallett, & Frecker, 1985, but see Ko, Snodderly, & Poletti, 2016), or it might occur as a result of balancing the eye between pairs of antagonistic muscles (Riggs & Ratliff, 1951). One main reason tremor is not well-understood is that it is particularly difficult to study: tremor not only overlaps in time with drift, but it has such tiny amplitudes that its accurate detection requires special high-resolution equipment (Yarbus, 1967; see also McCamy, Collins, et al., 2013 and Ko, Snodderly, and Poletti, 2016 for recent examples).

Converging evidence points to a variety of functions served by fixational eye movements. Notably, fixational eye movements are important to the prevention and reversal of perceptual fading during fixation (Martinez-Conde, Macknik, Troncoso, & Dyar, 2006; McCamy, Macknik, & Martinez-Conde, 2014; McCamy et al., 2012), as well as to the acquisition and processing of visual information (Ahissar & Arieli, 2012; Aytekin, Victor, & Rucci, 2014; Donner & Hemila, 2007; Ko, Poletti, & Rucci, 2010; Kuang, Poletti, Victor, & Rucci, 2012; McCamy, Otero-Millan, Di Stasi, Macknik, & Martinez-Conde, 2014b; Otero-Millan, Macknik, Langston, & Martinez-Conde, 2013; Otero-Millan et al., 2008; Rucci, Iovin, Poletti, & Santini, 2007; Rucci & Victor, 2015).

Yet, each kind of fixational eye movement—drift, tremor, and microsaccades—creates motion to different degrees. Some of the perceptual benefits from fixational eye movements may arise from the presence of any kind of motion, rather than from the specific motion generated by a particular type of eye movement (i.e. tremor vs. drift vs. microsaccades). This poses a difficulty for determining which eye movements underlie specific perceptual and neural functions or effects.

The following discussion emphasizes microsaccades. Due to their larger amplitudes, microsaccades create more retinal motion than drifts or tremor, and thus they may be more efficient at driving certain perceptual and physiological effects than the other fixational eye movements types. We note that the various types of eye movements may work together most of the time, however, and that optimal visual performance may require the joint action of (micro)saccades, drifts and tremor.


0.2 degrees
Saccades (>1 degree)
Microsaccades (≤1 degree)
Intersaccadic drift



◂**Fig. 3.1** **Top** Eye positions of an observer while exploring this image freely for 45 s. The eye makes large saccades around the image, momentarily resting in some places (particularly regions with interesting or meaningful details, such as faces). However, even during these "fixation" periods, when gaze seems to be at rest, the eye still moves around, producing microsaccades, drift, and tremor movements. The inset shows a fraction of the same data in greater detail. While the eye was "at rest" on this region of the image, microsaccades occurred frequently, along with intersaccadic drift. Tremor was not analyzed in this study. Reproduced from Martinez-Conde, Otero-Millan and Macknik (2013), Otero-Millan, Troncoso, Macknik, Serrano-Pedraza and Martinez-Conde (2008). This kind of research—the analysis of eye movement traces from observers as they view different images—owes its historical foundations to the work of Yarbus, who demonstrated that patterns of fixation locations vary as a function of the images presented and the visual task at hand (Yarbus, 1967). Similar conclusions have been drawn from research on fixational eye movements (Martinez-Conde & Macknik, 2015). **Bottom** You can experience your own fixational eye movements by looking at this pattern. Stare at the central black dot for about a minute, and then immediately shift your gaze to the white dot. As you try to fixate on the white dot, the afterimage of the crisscrossing lines will move constantly, due to your fixational eye movements. Reproduced from Verheijen (1961)

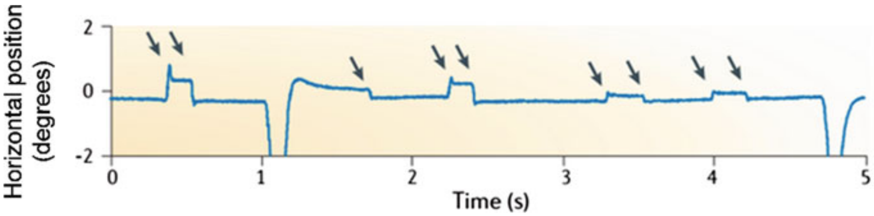


Fig. 3.2 Horizontal eye position trace. Positive values on the y axis indicate movements to the right (relative to the start position) and negative values indicate movements to the left. The arrows pinpoint microsaccades. Reproduced from Martinez-Conde et al. (2013)

3.2 Historical Background

James Jurin, an English physician and scientist, made the first known reference to fixational eye movements in 1738, when he wrote about "trembling of the eye" (Jurin, 1738). Fixational small movements were further described and studied through the 1800s and early 1900s (Adler & Fliegelman, 1934), but it was not until the 1950s and 60s that they were linked to perception, as researchers began to explore vision in the absence of any eye motion (fixational or otherwise; Barlow, 1952; Ditchburn & Ginsborg, 1952; Riggs & Ratliff, 1952; Yarbus, 1957).

Differences in how scientists measure eye movements today, compared to prior studies, may have led to shifts in fixational eye movement descriptions over time. The most prominent of these shifts is that microsaccades in recent studies are reported as having higher amplitudes than in earlier work. Thus, whereas microsaccades used to be considered tiny eye movements (i.e. less than 0.2° in amplitude), studies conducted after ~1980 have tended to report larger microsaccade sizes (most often up to 1 degree of visual angle, that is, five times as large as the original definition).

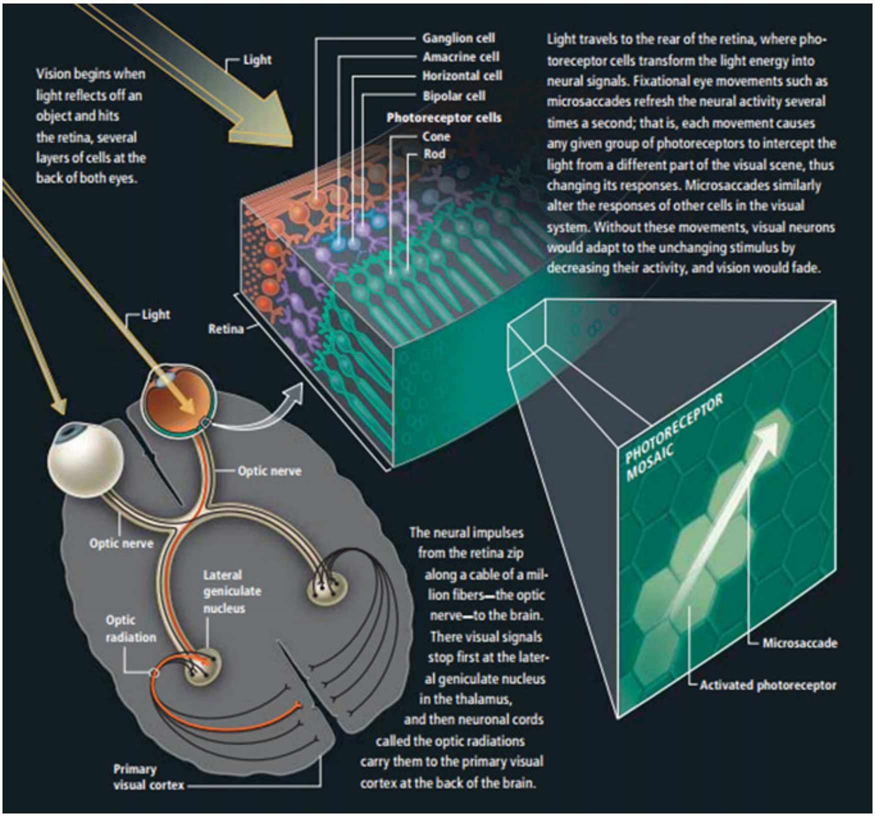
The precise reason for the change is not well understood, but could be partly due to methodological differences between classic and contemporary research (e.g. Collewijn & Kowler, 2008; Nyström, Hansen, Andersson, & Hooge, 2014). Current fixational eye movement research usually employs video-based eye-tracking techniques, whereas older studies used methods that required direct contact with, and physical attachment to, the eye. The latter may have hindered eye movements and decreased their amplitudes (McCamy et al., 2015; Van der Geest & Frens, 2002). Further, contemporary video-based eye-tracking methods track the *pupil* rather than the eyeball: thus, because the pupil can move (decentration of the pupil relative to the iris) after saccadic movements, video trackers may produce an overestimation of the amplitude of microsaccades (Nyström et al., 2014). Many other differences between experiments may contribute to the historical shift in microsaccade amplitudes. These discrepancies could include the use of bite bars in earlier experiments (Moller, Laursen, Tygesen, & Sjolie, 2002), the participation of experienced subjects in the earlier studies versus naïve ones in more recent research (Rolfs, 2009), and other factors such as changes in the stimuli used, task instructions, and room lighting. Because there is no single method that is free of all sources of error, the "true" amplitude of microsaccades remains unclear. Having said this, current consensus has largely consolidated around a definition of microsaccades that includes magnitudes up to 1° (Martinez-Conde, Macknik, Troncoso, & Hubel, 2009; Martinez-Conde et al., 2013).

Drift movements are also described differently in early and more recent research: Early work measured drifts at much slower speeds, below 10 arcmin/s, but these speeds were underestimated for several reasons. Drift measurements were originally measured only in terms of averages over a large time period, and often on a single axis (Cherici, Kuang, Poletti, & Rucci, 2012; Rucci & Poletti, 2015). Drift was therefore treated as linear, when it actually frequently changes direction. Ignoring these other velocity components necessarily results in an underestimation of drift velocity. As with microsaccades, the use of bite bars may have also resulted in differences in measurements (Aytekin et al., 2014). Task differences between early and more recent studies, and the common use of very experienced participants in early studies, may also have driven some of the discrepancies (Rucci & Poletti, 2015).

Differences in experimental setups between early fixational eye movement work and more recent studies may have also influenced debates about the very nature (or even the existence) of microsaccades. For instance, because early studies required subjects to keep their heads perfectly immobilized (even using bite bars to prevent as much head motion as possible), it was hypothesized that microsaccades were an artificial response to unnatural experimental conditions. Indeed, researchers speculated that the laboratory set-up might create so little motion of the retinal image (relative to that produced by normal motions of the head and neck) that microsaccade production increased as a result. Thus, some thought of microsaccades as a laboratory artifact without any real-life use (e.g. Kowler & Steinman, 1980).

Starting in the late 1990s and continuing through the 2000s, human eye-tracking studies revealed that microsaccades not only occurred when participants' heads were restrained, but also in head-unrestrained conditions (Martinez-Conde, 2006;

Martinez-Conde et al., 2006). In addition, neurophysiological research showed that microsaccades generated neural responses throughout the primate visual pathway, suggesting that microsaccade production could impact perception—see Fig. 3.3. Subsequent human psychophysics experiments demonstrated a variety of effects of microsaccades on perception (discussed later in this chapter). These combined studies showed that microsaccades were not merely laboratory artifacts or "nervous tics", but veridical eye movements with real-world effects (see Martinez-Conde et al., 2009 for a review; Martinez-Conde et al., 2013; Rolfs, 2009). Recent work has started to take microsaccade research out of the lab, and extend it to naturalistic, ecologically valid scenarios (Benedetto, Pedrotti, & Bridgeman, 2011; Di Stasi et al., 2015; see Alexander, Macknik, & Martinez-Conde, 2018 for a review).



Vision begins when light reflects off an object and hits the retina, several layers of cells at the back of both eyes.

Light travels to the rear of the retina, where photoreceptor cells transform the light energy into neural signals. Fixational eye movements such as microsaccades refresh the neural activity several times a second; that is, each movement causes any given group of photoreceptors to intercept the light from a different part of the visual scene, thus changing its responses. Microsaccades similarly alter the responses of other cells in the visual system. Without these movements, visual neurons would adapt to the unchanging stimulus by decreasing their activity, and vision would fade.

The neural impulses from the retina zip along a cable of a million fibers—the optic nerve—to the brain. There visual signals stop first at the lateral geniculate nucleus in the thalamus, and then neuronal cords called the optic radiations carry them to the primary visual cortex at the back of the brain.

Fig. 3.3 Microsaccades trigger neural responses along the visual pathway. Reproduced from Martinez-Conde and Macknik (2007)

3.3 Prevention and Reversal of Neural Adaptation and Perceptual Fading

Our sensory systems have evolved to detect change in the world around us. When there is no variation in our sensory input, our neurons discount it. Thus, we do not constantly notice our clothes, or steady noises in our auditory environment. Constant, steady information carries little meaning for us: Once we have processed such inputs, we do not concern ourselves again with them again—unless they change. The physiological mechanism by which our sensory systems stop responding to unchanging stimulation is called *neural adaptation*. Our visual system, like our other sensory systems, prefers change and entropy, while ignoring stable (and thus redundant) information—thereby increasing its preparedness to process new stimuli. Most retinal ganglion cells, for example, respond primarily to illumination transients and many will *only* respond to transients (Hubel, 1988).

Visual perception depends so strongly on change that if you never moved your eyes, and nothing moved around you, you would not see anything! Various researchers pioneered the use of "retinal stabilization" techniques to achieve constant visual stimulation—see Fig. 3.4. In the 1950s, scientists mounted tiny projectors on contact lenses to achieve retinal stabilization. The image projected onto the retina moved with the eye (i.e. it never changed relative to the eye). In such conditions, perceptual fading quickly ensued, though the actual physical stimulus remained extant.

Contemporary retinal stabilization techniques do not require devices placed on the eye, but rely on modern computer and eye-tracking technologies. A noninvasive eye tracker may be used to measure the subject's eye movements and transmit the eye-position data to a projection system that moves the image (presented on a standard computer monitor) with the eye. This transmission may entail a significant delay, however (as long as 7.5–10 ms in recent studies; see for example Poletti & Rucci, 2010), thereby resulting in unintended refresh of the retinal image (and the subsequent return of faded stimuli to visibility). This can pose difficulties for the replication and data interpretation of contemporary retinal stabilization studies.

In everyday life, constant eye motion usually prevents perceptual fading. But we can nevertheless experience a measure of fading during steady fixation, when we minimize our eye movements (Martinez-Conde et al., 2006). If you stare precisely at a specific point in visual space, while trying to hold your eyes still, you may notice that objects elsewhere become less visible—see Fig. 3.5. This perceptual phenomenon is known as Troxler fading, after the Swiss philosopher Ignaz Paul Vital Troxler. Blinking or making eye movements—or relaxing one's fixation—quickly restores the faded image.

Martinez-Conde et al. (2006) showed that the dynamics of Troxler fading are related to microsaccade production. Subjects made fewer microsaccades right before a visual target faded from perception, and more microsaccades right before the target reappeared—see Fig. 3.6. This pattern suggested a causal relationship between fixational eye movements and the prevention and reversal of perceptual fading. Subsequent studies showed that microsaccades restore the visibility of faded targets,

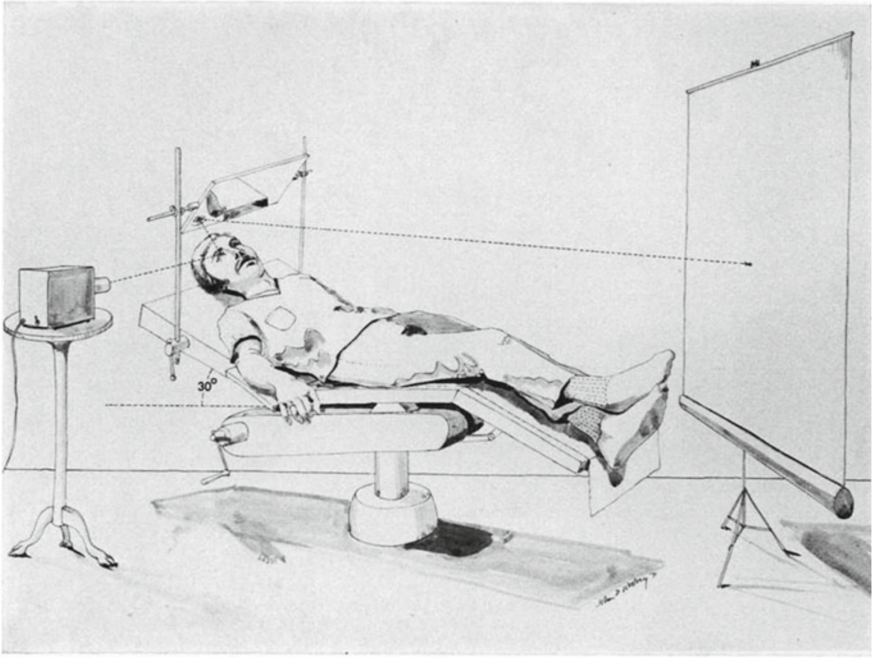
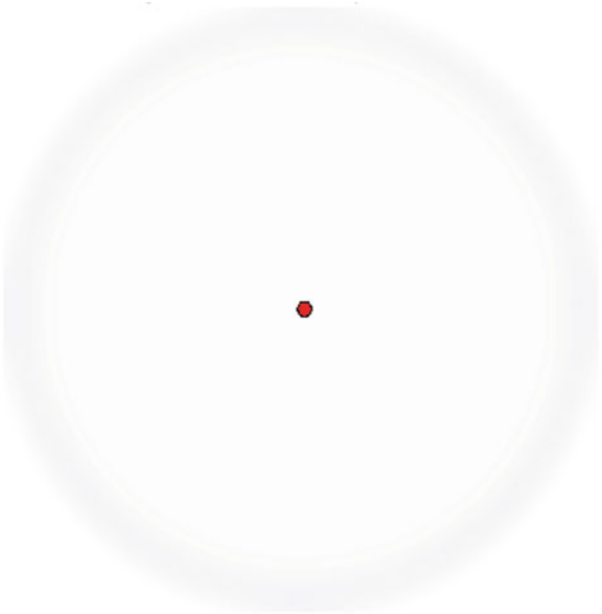
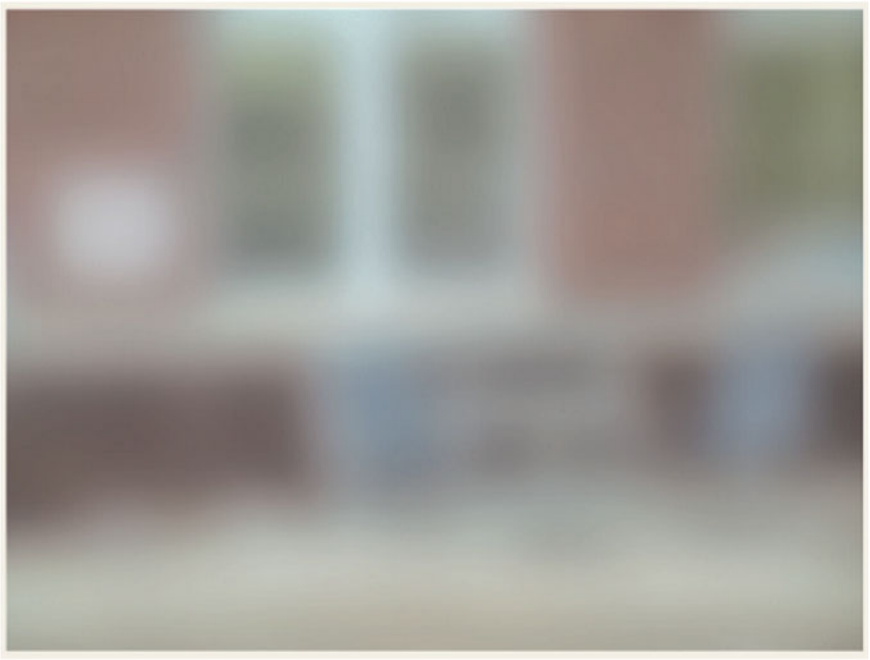


Fig. 3.4 A dramatic means of stopping the retinal image from moving is to paralyze the eye muscles. In the 1970s, the neuroscientist John K. Stevens underwent temporary paralysis, via the injection of curare (a paralytic drug). Pictured here is an artist's rendition of the experimental setup used to present visual stimuli and record Stevens's eye movements. Stevens's full-body paralysis prevented nearly all of his bodily motion, including eye movements. As a result, "image fading became a real problem" during the experiments. Reproduced from Stevens et al. (1976)

not only in the visual periphery but also in parafoveal and foveal regions (Carpenter, 1988; Costela, McCamy, Macknik, Otero-Millan, & Martinez-Conde, 2013; Martinez-Conde, 2006; Martinez-Conde et al., 2006; McCamy, Macknik et al., 2014; Troncoso, Macknik, & Martinez-Conde, 2008).

Whereas both microsaccades and drift may help prevent fading, only microsaccades (but not drift) can restore visibility after fading has occurred. This is likely due to microsaccades' higher speeds and the larger distances they cover compared to drift. Consistent with this proposal, larger microsaccades reverse fading more effectively than smaller ones do (McCamy, Macknik, et al., 2014; McCamy et al., 2012).





◀**Fig. 3.5** **Top** Fixate precisely on the red dot while attending to the ring around it. The ring will fade in a few seconds. Blinking or moving your eyes will bring the ring back to visibility. Reproduced from Martinez-Conde et al. (2004), see also Krauskopf (1967). **Bottom** Full-field Troxler fading shows that this phenomenon can affect not only the periphery, but the entire visual field. To experience it, fixate on the center of the image while attending to the entire scene. Maintaining your gaze carefully on a single point for several seconds will minimize the movement of your eyes, causing the image to fade to gray. If you move your eyes (or relax your fixation), the image will reappear. Reproduced from Simons et al. (2006)

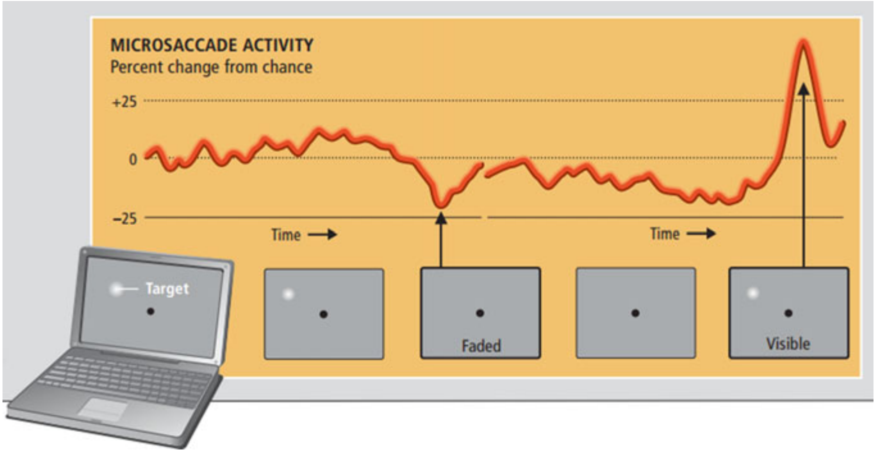


Fig. 3.6 Microsaccades become less frequent just before a static target vanishes from view. Conversely, microsaccades become more frequent just before the faded target reappears perceptually. Modified from Martinez-Conde and Macknik (2007)

Box 1: Measuring Fixational Eye Movements

To measure fixational eye movements, the first thing you will need is an eye-tracking device—see Table 3.1. The necessary equipment will depend on the specific fixational eye movement type(s) that you wish to study, particularly because most eye trackers are not sensitive enough to disambiguate tremor from noise. You should ensure that any system you use has sufficient temporal resolution, allowing you to capture enough data samples to adequately describe the movement (i.e. 250 Hz or greater may be adequate to characterize microsaccades, but not tremor). Similarly, the spatial resolution of the system—typically reported in RMS (root mean squared degrees of visual angle) and ideally measured against precisely controlled motion of an artificial eye—should be *significantly* higher than the movements that you wish to record. An eye tracker with a reported 0.5° RMS would be insufficient for measuring fixational eye movements, because it would only precisely record movements greater than 0.5° and only under the kinds of ideal circumstances that will rarely exist in the laboratory: There will usually be some other sources of noise, due (for example) to the differences between human participants and artificial eyes, to camera slippage, or other factors. If you are only interested in microsaccadic movements in human participants, chances are that you will want a high-resolution video-based eye tracker, a non-invasive and effective solution. Drift can also be explored on high-resolution video-based trackers. For measuring tremor, however, more specialized devices are required: Tremor measurements are typically made using piezoelectric probes or piezoelectric strain gauges. Both methods are somewhat invasive, requiring contact with the sclera and the application of a local anesthetic (Bengi & Thomas, 1968). Non-contact methods do exist, but are lower resolution and have not been used widely (Boyle, Coakley, & Malone, 2001; Ryle et al., 2009).

The next step is to acquire the actual measurements. Because fixational eye movements cannot be suppressed *completely* (Steinman, Cunitz, Timberlake, & Herman, 1967; Otero-Millan et al., 2008), and because they are made in virtually every task (Otero-Millan et al., 2008), any stimulus context can theoretically be used, from a blank screen to a fully-rendered 3d virtual reality display, to generate and record fixational eye movements. Your choice of stimulus will therefore depend primarily on the question you want to answer. Keep in mind that stimulus and task context will affect the size and frequency of fixational eye movements (McCamy, Najafian Jazi, Otero-Millan, Macknik, & Martinez-Conde, 2013; Thaler, Schutz, Goodale, & Gegenfurtner, 2013).

Importantly, some contexts will result in few fixational eye movements: For example, microsaccades are relatively rare in tasks where (larger) saccades are very frequent (Collewijn & Kowler, 2008; Otero-Millan et al., 2008). A common practice is to collect microsaccades while participants attempt to fixate a “fixation point”, such that any movements away from that point can be presumed to be involuntary (thus fulfilling the traditional definition of microsaccades as small, involuntary saccadic movements away from fixation).

Likewise, drift and tremor will be scarcer in conditions that produce few fixations than in conditions with a high prevalence of fixation periods. As in any other field of research, you may compare results across studies (such as in meta-analyses) when tasks and experimental contexts can be matched.

Once recordings have been acquired, you can examine the gaze positions over time and identify the changes in position that correspond to tremor, drift, and microsaccades. Different laboratories use various criteria to detect microsaccades (see Martinez-Conde et al., 2009 and the Historical Annotations in the current chapter), but the current convention is to accept microsaccade amplitudes up to one degree (Martinez-Conde et al., 2009, 2013; Otero-Millan et al., 2008). See Table 3.2 for a listing of microsaccade parameters recorded using different criteria and different eye tracking methodologies, through 2009. The contemporary convention of a one degree cut-off is consistent with the observed distribution of involuntary saccades produced during the attempt to fixate, as measured by contemporary eye tracking technologies and in most experimental contexts (e.g. Rolfs, 2009). Other criteria—such as labelling all saccadic movements produced during fixation as microsaccades—would result in similar data patterns, but because differences in criteria can lead to different results, it is important to report what criteria were used.

Engbert and Kliegl (2003a, 2004) wrote the most widely used contemporary algorithm for automatic microsaccade detection from eye position data. This algorithm takes advantage of the fact that microsaccades are faster than the other fixational eye movements, and thus bases microsaccadic detection on eye movement velocity (in the same way one may use velocity thresholds to detect larger saccades; see Chap. 2, "Saccades", in the current volume). Specifically, an eye movement is classified as a (micro)saccade when its velocity exceeds a multiple of the standard deviation of the velocity distribution found in the data, and a researcher-specified velocity factor. So long as the researcher specifies the velocity factor in accordance with the amount of noise present in the data, this algorithm will detect only velocities that are greater than the noise in the system—and greater than the velocity of drift and tremor. Note that this algorithm will detect both small and large saccades, and researchers that wish to impose an amplitude-based criterion on their analyses may do so after all (micro)saccadic movements are detected. Once microsaccades have been identified, any remaining small motion trajectories will correspond to drifts, tremor, and/or noise.

Drift is usually measured in terms of changes in the x,y position of the eye over time (when not due to blinks or other movements). When using eye tracking systems that are sensitive enough to disambiguate tremor from noise, tremor can be observed in the same x,y positions as drift, where researchers can count each peak of the movement in order to determine the frequency of tremor. Because tremor so closely resembles noise, it is important to then confirm the presence of tremor by demonstrating statistical differences from noise inherent in the eye tracker.

Note that because drifts and tremor overlap in time (Carpenter, 1988; Pritchard, 1961) and also overlap in their spectral density (that is, tremor can be as slow as 40 Hz and drifts can be as fast as 40 Hz), there is no clear-cut way to distinguish between these two kinds of movements. When exploring the effects of drifts or tremor, one can exclude any data from time periods surrounding the start or end of a detected (micro)saccade, thereby preventing saccadic motions from influencing drift and tremor detection. Disambiguating between drifts or tremor, however, cannot be done in this way, because they occur simultaneously. Thus, when relating neural responses to eye movements, it is therefore difficult (and often impossible) to clearly identify which eye movements are responsible for a given neural response. In cases where neural responses change over time in correspondence with either drifts or tremor (but not both), a correlational relationship between the signal and the motion can then be inferred. Some researchers have also tested for effects of either drifts or

Table 3.1 Comparison of contemporary eye tracking methods

	Scleral search coil	Electro-oculography	Dual Purkinje image	Video-oculography
Fixational eye movements measured (all methods record horizontal and vertical rotations)	All eye movements except for tremor. Torsion and point-of-gaze measurements possible	Fixational eye movement recordings not possible. Torsion and point-of-gaze measurements also not possible. Large saccades, smooth pursuit, and nystagmus measurements possible	All eye movements except for tremor. Torsion measurements not possible. Point of gaze is possible	Device dependent. Torsion and point-of-gaze measurements possible
Physical contact	Contact lens on one eye with anesthetic. Wire exits eye	Electrodes around eyes	No contact	Headband or no contact required
View obstruction	No obstruction	No obstruction	Device obstructs view	Little or no obstruction with headband No obstruction without headband
Physical restraint	Usually head and chin rest	None	Head and chin rest (bite bar for higher precision)	Usually head and/or chin rest but can be unrestrained, sacrificing accuracy and precision
Accuracy	–0.2°	~3° to 7°	~1 arcmin	~0.25° to 1°
Precision	~0.25 arcmin		~1 arcmin	
Resolution	~0.5–1 arcmin	~0.5°	~1 arcmin	~0.01° to 0.2°
Range	~30°	~70°	~10° to 20°	~25° to 60°
Bandwidth	~200 Hz	~100 Hz	~400 Hz	Up to ~400 Hz (device dependent)
Ability to measure eye movements with eyes closed	Yes, but not for long periods	Yes, even during sleep	No	No

Table modified from McCamy, Macknik and Martinez-Conde (2013, p. 8)

Table 3.2 Physical parameters of microsaccades, as reported by original research studies, until 2009

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
5–120 min (range monocular); ~15 min (median monocular); 6–120 min (range conjugate binocular); ~30 min (median conjugate binocular)(a)	0.51 Hz (mean monocular); 1.43 Hz (mean conjugate binocular)	D–	MS–	Human	Video-based	Engbert and Kliegl (2003b)
		V 6–120 deg s^{-1} (range monocular); ~18 deg s^{-1} (median monocular); 10–120 deg s^{-1} (range conjugate binocular); ~25 deg s^{-1} (median conjugate binocular)(a)	BR Yes			
–	0.8–1.9 Hz (among different subjects)	D–	MS–	Human	500 Hz EyeLink II	Engbert and Kliegl (2004)
		V–	BR Yes			
32 min (mean)	1.3 ± 0.7 Hz (among different subjects)	D–	MS–	Human	250 Hz EyeLink	Engbert and Kliegl (2003b)
		V–	BR–			
12.7–60 min (range); ~42 min (median)(a, b)	–	D–	MS–	Human	240 Hz video-based ISCAN tracker	Hafed and Clark (2002)
		V 11.3–127.3 deg s^{-1} (range); ~28.3 deg s^{-1} (median)(a, b)	BR–			
12.7–65.9 min	0.23–0.93 Hz (mean 0.61 Hz)	D–	MS 28–97 deg s^{-1}	Human	250 Hz EyeLink	Moller et al. (2002)
		V 16–40 deg s^{-1} (mean acceleration 2322–6440 deg s^{-2})	BR Yes			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
<12 min	–	D–	MS–	Human	The Maryland revolving field monitor	Malinov, Epelboim, Herst, and Steinman (2000)
		V–	BR Yes			
17 min (mean)(c)	–	D–	MS–	Human	500 Hz Purkinje eye tracker	Kingstone, Fendrich, Wessinger, and ReuterLorenz (1995)
		V–	BR–			
5–32 min	2.1 Hz	D–	MS ~4.2–55 deg s^{-1}(b)	Human	Contact lens optical lever	Riggs, Armington, and Ratliff (1954)
		V–	BR Yes			
4.2–53 min	1–3 Hz	D–	MS 21.3–36.5 deg s^{-1}	Human	Contact lens optical lever	Schulz (1984)
		V–	BR–			
20.8 min (mean)(b)	2.62 Hz (mean)	D–	MS–	Human	100 Hz infrared reflection recording	Srebro (1983)
		V–	BR–			
–	1.3–1.43 Hz (means for two subjects)	D–	MS–	Human	120 Hz SRI double Purkinje eye trackers	Sabrin and Kertecz (1980)
		V–	BR–			
~7 min (mean)	~2 Hz (mean)	D–	MS–	Human	Search coil	Winterson and Colwijn (1976)
		V–	BR–			
7.7 min (median)	0.6 s (median)	D–	MS–	Human	Search coil	Skavenski, Robinson, Steinman, and Timberlake (1975)
		V–	BR–			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
7.7–8.9 min (means for two subjects)	2.1–2.5 Hz (means for two subjects; range 1–4 Hz)	D–	MS–	Human	Contact lens optical lever	Sansbury, Skavenski, Haddad, andSteinman (1973)
		V–	BR–			
~40 min (mean)	–	D–	MS–	Human	Contact lens optical lever	Steinman et al. (1973)
		V–	BR–			
1.8–5.8 min (medians for two subjects)	0.24–0.44 s (mean range for two subjects)	D–	MS–	Human	Contact lens optical lever	West and Boyce (1968)
		V–	BR–			
8.18 min (mean)	1.44 Hz(e)	D–	MS–	Human	Contact lens optical lever	Steinman et al. (1967)
		V–	BR–			
2–50 min	3–5% fixation time	D 0.01–0.02 s	MS–	Human	Contact lens optical lever	Yarbus (1967)
		V Mean: 10 deg s^{-1}	BR Yes			
3.7–9 min (median values for different subjects)(d)	0.20–0.72 s (mean values for different subjects)(d)	D–	MS–	Human	Contact lens optical lever	Boyce (1967)
		V–	BR–			
–	1.5 s (mean for one subject)	D–	MS–	Human	Contact lens optical lever	Fiorentini and Ercoles (1966)
		V–	BR–			
–	1.45–2.2 Hz (means for two subjects)	D–	MS–	Human	Contact lens optical lever	Steinman (1965)
		V–	BR–			
2–13 min(a)	–	D–	MS–	Human	Infrared reflection recording	Zuber and Stark (1965)

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
		V Mean: 3–14 deg $s^{-1(a)}$	BR–			
1.7–4.8 min (means for two subjects); 1.2–3.4 min (medians for two subjects)$^{(d)}$	–	D–	MS–	Human	Contact lens optical lever	Nachmias (1961)
		V–	BR–			
2.5–5.6 min (median range for two subjects)$^{(d)}$		D–	MS–	Human	Contact lens optical lever	Krauskopf et al. (1960)
		V–	BR Yes			
2–3.5 min (medians for two subjects)$^{(d)}$	0.45–0.5 s (means for two subjects); 0.35–0.4 s medians for two subjects)	D–	MS–	Human	Contact lens optical lever	Nachmias (1959)
		V–	BR–			
~3.5 min^{a}	~0.4 Hz$^{(a)}$	D–	MS–	Human	Contact lens optical lever	Cornsweet (1956)
		V–	BR–			
1–20 min	0.03–5 s	D 0.025 s	MS Max. acceleration 1000 deg s^{-2}	Human	Contact lens optical lever	Ditchburn and Ginsborg (1953)
		V Mean: 10 deg s^{-1}	BR Yes$^{(f)}$			
2–25 min	0.76–10 s (means for different subjects)	D 0.02–0.03 s	MS–	Human	Ultra-violet corneal reflection/optical lever	Lord (1951)
		V–	BR Yes			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
2.2–25.8 min (rarely >10 min; mean 5.6 min; median 3.9 min(d))	0.2–4 s	D 0.02 s	MS–	Human	Contact lens optical lever	Ratliff and Riggs (1950)
		V–	BR–			
3–14 min	1.5–2 Hz (mean frequencies for two subjects)	D 0.02–0.03 s	MS–	Human	Ultra-violet corneal reflection/optical lever	Lord and Wright (1948)
		V–	BR–			
12.5–17.5 min (mean 8 min(d);median 5.6 min(d))	1 Hz (mean frequency)	D–	MS–	Human	Optical lever (mirror placed directly on eye after anesthetic)	Adler and Fliegelman (1934)
		V–	BR–			
8.4–16.2 min (means for two monkeys)	2.3–2.5 Hz (mean frequencies for two monkeys)	D At least 8 ms	MS–	Human	250 Hz search coil	Horwitz and Albright (2003)
		V–	BR–			
~40 min (mean)	–	D–	MS–	Monkey	100 Hz double Purkinje monocular	Snodderly, Kagan, and Gur (2001)
		V Mean: ~30 deg s^{-1}	BR–			
~20 min (mean)	~3 to 5 Hz	D 29 ms (mean)	MS–	Monkey	1 kHz search coil, monocular	Martinez-Conde et al. (2000)
		V Mean: ~30 deg s^{-1}	BR–			
48 min (mean)	0.3–1.4 Hz (means for two monkeys)	D 25 ms (mean)	MS 9–110 deg s^{-1} (median 40 deg s^{-1})	Monkey	500 Hz search coil	Bair and O'Keefe (1998)
		V Mean: ~30 deg s^{-1}	BR -			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
10.1 min (median)	0.597 s (median)	D 0 ms (mean)	MS–	Monkey	200 Hz search coil	Leopold and Logothetis (1998)
		V–	BR–			
9.9–40.3 min (medians for four monkeys)	0.8–7.4 s (medians for four monkeys)	D–	MS–	Monkey	1 kHz search coil, monocular	Skavenski et al. (1975)
		V–	BR–			
40 min (mean; minimum amplitude 23 min)	–	D–	MS–	Monkey	Search coil, monocular	Steinman et al. (1973)
		V–	BR–			
1.4° max	>2 Hz mean frequency	D–	MS–	Monkey	1 kHz scleral search coil	Bosman et al. (2009)
		V Mean: 4.43°/s	BR No			
1°	–	D Min: 16.8 ms	MS 2008/s	Human	238 Hz SMI iView X Hi-Speed eye tracker	Valsecchi and Turatto (2009)
		V Max: 200°/s	BR No			
Mean 24.6′	0.25 Hz mean frequency; 20 ms min interval	D Min: 10 ms; max: 300 ms; mean: 20.3 ms	MS 38.88°/s	Monkey	200 Hz scleral search coil experiment 1	Herrington et al. (2009)
		V Min: 8°/s; peak: 38.8°/s	BR -			
Mean: 22.6′	0.76 Hz mean frequency; 20 ms min interval	D Min: 10 ms; max: 300 ms; mean: 19.7 ms	MS 30.3°/s	Monkey	200 Hz scleral search coil experiment 2	Herrington et al. (2009)
		V Min: 10°/s; peak: 30.3°/s	BR -			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
Mean: 21.4′-29.0′	0.65–1.7 Hz mean frequency; 20 ms min interval	D Min: 10 ms; max: 300 ms; mean: 23.9 ms	MS 16.4–47.1°/s	Monkey	200 Hz ASL6000 and 1000 Hz Eyelink1000 experiment 3	Herrington et al. (2009)
		V Min: 8–10°/s; peak: 16.4–47.1°/s	BR–			
Max: 2^3 mean: 13.7′	–	D–	MS–	Monkey	Magnetic induction coil	Hafed et al. (2009)
		V Min: 8°/s	BR No			
Max: 1.5°	–	D Min: 6 ms	MS–	Human	500 Hz EyeLink II	Valsecchi and Turatto (2009)
		V–	BR No			
Min: 3′; max: 2°		D–	MS–	Human	250 Hz HS video eye tracker	Beer et al. (2008)
		V Min: 10°/s	BR No			
Range: 1′–36′	Mean: 1.08 Hz min: 100 ms	D Min: 8 ms	MS–	Monkey	1000 Hz eye coil monocular	Cui et al. (2009)
		V Min: 110°/s	BR No			
Min: 3′–6′	–	D–	MS–	Monkey	100 Hz DPI eye tracker and 200 Hz eye coil monocular	Kagan et al. (2008)
		V Min: 10°/s	BR No			
Max: 1°	–	D Min: 6 ms	MS–	Human	500 Hz EyeLink II binocular	Kliegl et al. (2009)
		V–	BR Yes			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
Max: 1°; mode: 24.9′		D–	MS–	Human	500 Hz EyeLink II binocular	Laubrock et al. (2008)
		V–	BR Yes			
Max: 1°		D–	MS–	Human	500 Hz EyeLink II binocular	Moshel et al. (2008)
		V–	BR–			
Max: 1°; mean: 0.39°	Mean: 0.8 Hz; Min: 20 ms	D Mean: 12 ms	MS 39°/s	Human	500 Hz EyeLink II binocular (during prolonged fixation)	Otero-Milan et al. (2008)
		V–	BR Yes			
Max: 1°; mean: 0.55°–0.64°	Mean: 0.2–1.3 Hz; Min: 20 ms	D–	MS–	Human	500 Hz EyeLink II binocular (during free-viewing)	Otero-Milan et al. (2008)
		V–	BR Yes			
Max: 1°; mean: 0.41°	Mean: 0.7 Hz Min: 20 ms	D Min: 6 ms; mean: 13 ms	MS 41°/s	Human	500 Hz EyeLink II binocular	Troncoso et al. (2008)
		V–	BR Yes			
Max: 2°; mean: 0.4°	Mean: 1.0 Hz; min: 20 ms	D Min: 6 ms; mean: 18 ms	MS 38°/s	Human	500 Hz EyeLink II binocular	Troncoso et al. (2008)
		V–	BR–			
Max: 1.66° (observed)	–	D Min: 6 ms	MS–	Human	500 Hz EyeLink II binocular	Yuval-Greenberg et al. (2008)
		V–	BR Yes			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
	Range: 0.57–2.54 Hz Mean: 1.4 Hz	D Min: 12 ms	MS–	Human	500 Hz eye coil monocular	Betta et al. (2007)
		V–	BR No			
Max: 1.5°	–	D Min: 15 ms	MS–	Human	200 Hz Fourward DPI v.6.3, monocular	Horowitz et al. (2007)
		V–	BR no			
–	–	D–	MS–	Human	500 Hz EyeLink II	Mergenthaler and Engbert (2007)
		V–	BR–			
–	–	D–	MS–	Human	500 Hz EyeLink II binocular	Turatto et al. (2007)
		V–	BR–			
Max: 1.5°	–	D Min: 12 ms	MS–	Human	500 Hz EyeLink II binocular	Valsecchi and Turatto (2007)
		V–	BR–			
Max: 1.5°	Mean: 1.25–1.8 Hz	D Min: 12 ms	MS–	Human	500 Hz EyeLink II monocular and binocular	Valsecchi and Turatto (2007)
		V Max: 300°/s	BR Yes			
–	–	D–	MS–	Monkey	Eye coil monocular	Goffart et al. (2006)
		V Min: 5°/s	BR No			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
Max: 2°	–	D Min: 12 ms	MS–	Human	500 Hz EyeLink II monocular	Betta and Turatto (2006)
		V–	BR no			
Max: 1°	–	D Min: 6 ms	MS–	Human	500 Hz EyeLink II binocular	Engbert and Mergenthaler (2006)
		V–	BR Yes			
Mean: 0.24°; Max: 2°	–	D–	MS–	Human	500 Hz EyeLink II binocular	Martinez-Conde et al. (2006)
		V Min: 3°/s	BR no			
Max: 1°	–	D Min: 8 ms	MS–	Human	500 Hz EyeLink II binocular	Rolfs et al. (2006)
		V–	BR Yes			
–	–	D Min: 12 ms	MS–	Human	250 Hz EyeLink II binocular	Van Dam and van Ee (2005)
		V–	BR Yes			
–	–	D–	MS–	–	1000 Hz infrared limbus eye tracker binocular (horiz. only)	Murakami et al. (2006)
		V Min: 10°/s	BR–			
Min: 2.8′; mean: 22.19′	–	D Min: 8 ms	MS–	Human	500 Hz EyeLink II binocular	Laubrock et al. (2005)
		V–	BR Yes			

(continued)

Table 3.2 (continued)

Amplitude	ISI/MSF	Duration (D)/max speed (MS)/velocity (V)/binocularity required (BR)		Species	Eye tracking system used	Reference
Max: 1.5°; max: 1°	–	D–	MS–	Human	500 Hz EyeLink II binocular	Liang et al. (2005)
		V–	BR–			
Max: 15′–18′	–	D Min: 8 ms	MS–	Human	500 Hz EyeLink II binocular	Rolfs et al. (2005)
		V–	BR–			
–	Mean: 0.8–1.9 Hz	D Min: 12 ms	MS–	Human	250 Hz EyeLink II binocular	Van Dam and van Ee (2005)
		V–	BR–			
–	Mean: 0.8–1.9 Hz	D Min: 12 ms	MS–	Human	500 Hz EyeLink II binocular	Engbert and Kiegl (2004)
		V–	BR Yes			
Max: 1°	Mean: 1.2 Hz	D Min: 12 ms	MS–	Human	500 Hz EyeLink II monocular	Galfano et al. (2004)
		V–	BR No			
Max: 1°		D–	MS–	Human	500 Hz EyeLink II	Rolfs et al. (2004)
		V–	BR–			

Note Unless otherwise indicated, the “minimum” and “maximum” values indicate thresholds imposed by the authors to identify microsaccades. “Mean” and “range” values are the observed values in the studies. From Martinez-Conde et al. (2004, 2009)

[a] Values estimated from plots

[b] Vector values from H-components—a conversion factor of $\sqrt{2}$ has been assumed

[c] Calculated from horizontal and vertical components

[d] Recalculated in/taken from Ditchburn and Foley-Fisher (1967)

[f] Microsaccade rates for ‘large’ and ‘small’ targets have been averaged

[e] Small differences in amplitude and direction were observed in one out of two subjects

tremor independently from other, via *simulated* motions (i.e. stimuli that move in drift- and tremor-like patterns) (Greschner, Bongard, Rujan, & Ammermuller, 2002). One significant limitation of such studies is that simulated eye movements may or may not be truthful reproductions of physiological ocular motion.

3.4 Neural Responses to Microsaccades in Visual Brain Areas

There are significant obstacles to studying the responses that fixational eye movements elicit in visual neurons. Some of these include the difficulty of measuring tremor and drifts and in deciding which movement type triggered a neural response (see Text Box: Measuring fixational eye movements). There are also technical difficulties involved in recording neural activity from a moving eye, and a long history of considering microsaccades to be irrelevant to vision (before the discovery of microsaccade-driven neural responses along the visual pathway). We now know that microsaccades trigger neural activity in the lateral geniculate nucleus (LGN), area V1 (Martinez-Conde, Macknik, & Hubel, 2000; Martinez-Conde, Macknik, & Hubel, 2002), and the extrastriate cortex (Dimigen, Valsecchi, Sommer, & Kliegl, 2009; Meirovithz, Ayzenshtat, Werner-Reiss, Shamir, & Slovin, 2012; Tse, Baumgartner, & Greenlee, 2010). Neural responses to microsaccades presumably start in the retina (Donner & Hemila, 2007).

Most neurophysiological recordings have shown that, in the presence of visual stimulation, microsaccades trigger increases in neural activity (Martinez-Conde et al., 2013). This finding is consistent with the concept of microsaccade-driven receptive field displacements across the visual scene (a visual neuron's receptive field is the visual field region where the presentation of a stimulus may cause a change in the neuron's response). When you first fixate your gaze on a given spatial location, the light entering your eye results in patterns of neural activity. As neural adaptation ensues, visual neurons becomes less sensitive to the constant stimulus, and the image begins to fade perceptually (unless the visual stimulus moves or otherwise changes). When you move your eyes, different parts of the image fall on different receptive fields, counteracting and preventing neural adaptation and perceptual fading. The interplay between the respective locations of receptive fields and visual stimuli largely determines the neural responses to microsaccades.

This same interplay may explain why fading is often more prevalent in the visual periphery than in the center of vision (i.e. peripheral receptive fields are larger than central ones) and why larger eye movements are more likely to reverse fading than smaller eye movements do. First, the greater the retinal distances traveled, the greater number of receptive fields will receive changing visual stimulation, thereby reversing adaptation in a larger extent of the retina. In other words, if only a small part of the field of vision changes, it will prevent and reverse fading less effectively than if the entire pattern of light changes. Second, larger eye movements increase the likelihood that the observer will look at something *different* from before the eye motion. The larger the eye movement, the more likely the retinal stimulation will change, with the result of preventing fading, or reversing it if already present.

Various studies have also explored the neural responses to microsaccades in the *absence* of visual stimulation (which often means presenting a blank screen over the receptive fields of the visual neurons recorded). Because the local retinal image does not change in this scenario, changes in neural responses, if present, must be due to the very act of making an eye movement (rather than to visual stimulation of the neurons' receptive fields). In such cases, researchers have often observed *decreases* in neural activity in connection with microsaccade production. The magnitude of such decreases is generally smaller than the magnitude of the increases triggered by microsaccades in the presence of visual stimulation (Martinez-Conde et al., 2013). The cause of these decreases in activity is not yet clear, though it might result from extraretinal brain activity associated with microsaccade generation.

3.5 Neural Generation of Microsaccades

Researchers have begun to explore the neural circuits involved in the generation of fixational eye movements (see Martinez-Conde et al., 2013 for a review). Of the fixational eye movements, microsaccade generation is the best understood, because of the relationship between microsaccades and larger saccades.

Microsaccades and saccades have comparable physical characteristics, which indicates a common oculomotor origin (see also Rolfs, Laubrock, & Kliegl, 2006, 2008). For example, both microsaccades and saccades follow the "main sequence," the parametric relationship between saccadic peak velocity and saccadic amplitude (Martinez-Conde, 2006; Mergenthaler & Engbert, 2010; Otero-Millan et al., 2008)—see Fig. 3.7. In fact, the characteristics of microsaccades and saccades overlap to such a degree that only a functional distinction is possible, with microsaccades being

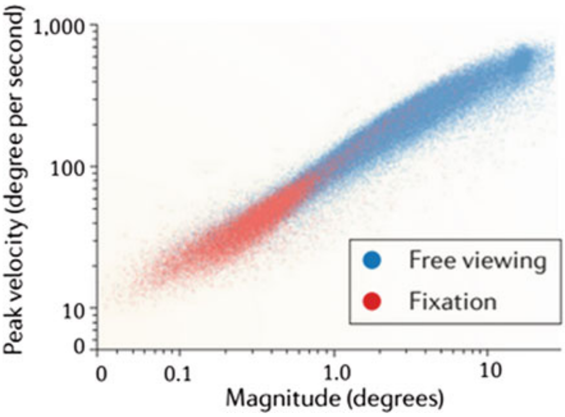


Fig. 3.7 The relationship between peak velocity and amplitude for (micro)saccadic eye movements, during both fixation and free viewing. Microsaccades are indistinguishable from larger saccades based on their physical characteristics. Reproduced from Martinez-Conde et al. (2013), modified from the original in Otero-Millan et al. (2008)

those *involuntary* saccades produced during the attempt to fixate (Otero-Millan et al., 2008). Since it is impossible to establish the intentionality (or lack thereof) of *every single eye movement*, the current convention is to label all saccadic movements which are under one degree in amplitude as "microsaccades."

Given their overlap in physical characteristics, we can reasonably assume that microsaccades and saccades share at least some of the same underlying neural circuitry. This assumption has provided a body of models—originally designed to explain saccadic generation—on which models of microsaccadic generation can be built (e.g. Moschovakis, Scudder, & Highstein, 1996; Munoz, 2002; Optican, 1995; Robinson, 1973; Scudder, Kaneko, & Fuchs, 2002; Sparks, 2002; Stanford, Freedman, & Sparks, 1996; see also Chap. 2 on saccades).

Saccade commands originate in the superior colliculus (SC) (Munoz & Wurtz, 1995; Schiller, 1984). The size and direction of the resulting saccadic movement is determined by *where* the activity is located in the SC (Schiller & Stryker, 1972). Large fluctuations in the caudal region of the SC give rise to saccades, while small shifts in the rostral region of the SC give rise to microsaccades (Hafed, Goffart, & Krauzlis, 2009; Rolfs, Kliegl, & Engbert, 2008). Omnipause neurons (OPNs) in the raphe nuclei inhibit excitatory burst neurons (EBNs) in the reticular formation, except when generating a saccade (e.g. Gandhi & Keller, 1999). The OPNs are in turn activated by signals from rostral SC (Guitton & Munoz, 1991). When microsaccades are initiated, inhibitory burst neurons cause the OPNs to decrease their firing rates, and therefore their inhibition of the EBNs (Brien, Corneil, Fecteau, Bell & Munoz, 2009). Activation in rostral SC thus causes a partial release of inhibition by OPNs on the EBNs. This allows the EBNs to generate small bursts, which relay the saccade commands to motor neurons, producing an eye movement—see Fig. 3.8.

3.6 Microsaccades and Neurological Disease

Despite recent progress in the characterization of fixational eye movements in various pathologies, their dynamics in many neurological and ophthalmic disorders remain largely unexplored.

Saccadic intrusions are saccades that interrupt (or "intrude on") accurate fixations, and are common in various neurological disorders (Leigh & Zee, 2015), including multiple system atrophy, Parkinson's disease, and corticobasal syndrome (Otero-Millan, Schneider, Leigh, Macknik, & Martinez-Conde, 2013, see chapter by Müri, Cazzoli, & Nyffeler (2019) in this volume).

Although (micro)saccadic intrusions are present in healthy individuals (Abadi & Gowen, 2004), and there is overlap in the amplitude and frequency of saccadic intrusions among healthy and clinical populations (Pinnock, McGivern, Forbes, & Gibson, 2010), there are differences in saccadic intrusion dynamics across pathologies that can aid diagnosis. For example, patients with progressive supranuclear palsy have a vertical saccadic palsy—a neurological abnormality that disrupts the patients' ability to make vertical (up-and-down) saccades—that is not characteristic of other

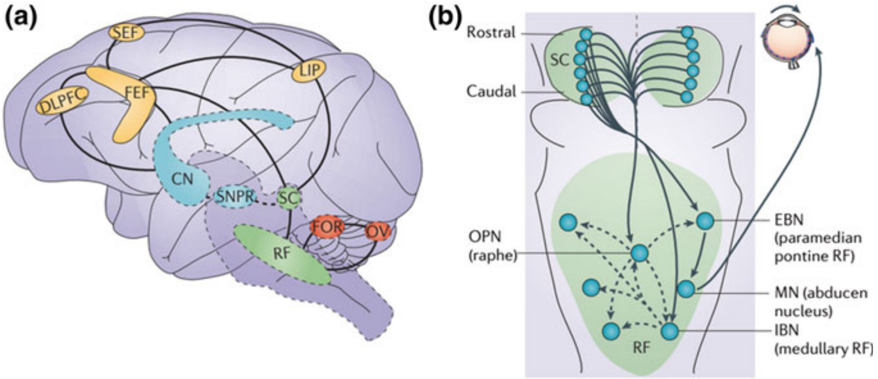


Fig. 3.8 The microsaccade generation circuit. **a** These are the brain areas implicated in the generation of saccades (also see Chap. 2). Cortical areas related to the voluntary control of gaze—shown in orange—include the lateral intraparietal area (LIP), frontal eye fields (FEF), supplementary eye fields (SEF) and the dorsolateral prefrontal cortex (DLPFC). The basal ganglia are shown in blue: The caudate nucleus (CN) and the substantia nigra pars reticulate (SNPR). The CN receives a projection from the FEF. The SNPR inhibits the superior colliculus (SC) in the brainstem, controlling saccade initiation. Brainstem areas–shown in green—include the SC and the reticular formation (RF). The SC projects to the RF to produce the final motor command. Areas of the cerebellum are shown in red: the oculomotor vermis (OV) and the fastigial oculomotor region (FOR). These regions control saccade accuracy via a feedback loop. **b** The brainstem circuit for producing a rightward microsaccade. SC neurons present two paths of connectivity: one that is strongest between rostral SC and omnipause neurons (OPNs), and one that is strongest between caudal SC and excitatory burst neurons (EBNs). During fixation, the SC rostral poles drive OPNs, which in turn inhibit IBNs and EBNs. Small shifts in rostral activity trigger microsaccades, whereas large shifts (to the caudal region) give rise to saccades. Shifts in SC activity increase the drive to burst neurons and/or decrease the drive to OPNs. Solid lines represent excitatory connections and dashed lines represent inhibitory connections. Reproduced from Martinez-Conde et al. (2013)

parkinsonian disorders. Saccades and saccadic intrusions in progressive supranuclear palsy moreover have slow speeds (Otero-Millan et al., 2013).

Fixational eye movements have also been used to distinguish hypoxia (decreased availability of oxygen in the body's tissues) from fatigue (Di Stasi, Cabestrero, et al., 2014). Various studies have shown that saccadic and microsaccadic velocities decrease as a function of fatigue in healthy subjects (Di Stasi, Catena, Cañas, Macknik, & Martinez-Conde, 2013; Di Stasi, McCamy, et al., 2013; Di Stasi, McCamy, et al., 2014; Di Stasi et al., 2015). Since one symptom of hypoxia *is* fatigue, being able to differentiate between fatigue and hypoxia is important. Thus, fixational eye movements can help distinguish between conditions that are difficult to differentiate by subjective report. The underlying reasoning is that two conditions with different neurological bases should result in different measurable eye movement dynamics.

Fixational eye movements occur with high frequency in natural vision, so alterations to their attributes may have a pervasive effect on daily life. Yet the potential for abnormal fixational eye movements to interfere with vision or underlie vision loss in pathology has been generally overlooked. As eye-tracking technology becomes more prevalent, fixational eye movements may increasingly contribute to the differential and early diagnosis of a variety of pathologies, as well as to the evaluation of therapy regimens (Alexander et al., 2018).

3.7 Gaze Correction, Control, and Acuity

Blinks, head movements, and other factors can result in differences between the actual and the desired gaze position. The oculomotor system triggers microsaccades in part as a response to deviations from the desired gaze position (Costela et al., 2014; Ko et al., 2010; Otero-Millan, Macknik, Serra, Leigh, & Martinez-Conde, 2011; Poletti, Listorti, & Rucci, 2013; Rucci et al., 2007). Thus, another role that microsaccades play is to correct for "fixation errors" (i.e. deviations from the desired fixation). Microsaccades and drifts both help correct undesired displacements in gaze position and bring the gaze back to the desired target (Cornsweet, 1956; Nachmias, 1959; Costela, Otero-Millan, et al., 2013; Ko et al., 2010; see Rolfs, 2009 for a review). In fact, drift is sometimes called "slow control" to emphasize its functional significance in maintaining fixation. However, both microsaccades and drift movements can move the gaze away from, and back to, the desired target, in healthy individuals and in clinical populations. Pathological abnormalities in fixational eye movements can lead to increases in movements away from fixation or decreases in the use of corrective movements to maintain the retinal image on the fovea (common in macular degeneration or with losses in oculomotor control). Decreases in corrective movements result in reduced visual acuity and reading speed (Falkenberg, Rubin, & Bex, 2007), highlighting the importance of gaze correction in daily life.

Individuals who are highly experienced in fixation tasks (i.e. as used in the lab) can suppress their microsaccades for a few seconds at a time when asked to "hold the eye still" (Steinman et al., 1967). Under such conditions, drifts can keep the observer's gaze position relatively stable. This does not necessarily mean that drifts are better than microsaccades at gaze correction in natural vision, however, especially in the correction of large fixation errors. Instead, drift may supplement the correcting role of microsaccades and/or correct gaze position in conditions where microsaccade production may be undesirable.

The ability to temporarily suppress microsaccades is important in that it demonstrates that we can exert a degree of control over our fixational eye movements. In the same way that you can control your breathing within certain limits (i.e. you cannot choose to stop breathing indefinitely or breathe more quickly than at a certain speed)—you can control microsaccade production to an extent. There appears to be a minimum intersaccadic interval (i.e. a refractory period of sorts) beyond which people cannot increase their microsaccade rates (Otero-Millan et al., 2008), and it is moreover likely that microsaccade generation cannot be decreased to fewer than one microsaccade every few seconds (Steinman et al., 1967). Our ability to lower microsaccade rates may help improve task performance in certain situations (i.e. in tasks where making many small eye movements would disrupt execution).

High-acuity tasks (such as threading a needle) can depend on information at such a small scale that a single microsaccade could take the subject's gaze away from all task-relevant information. Early studies concluded that microsaccades were detrimental to high-acuity task performance (see Collewijn & Kowler, 2008) and were suppressed during such tasks (Bridgeman & Palca, 1980; Winterson & Collewijn, 1976). These results may have resulted from expectancy effects, however, since anticipating a motor response can decrease microsaccade rate (Betta & Turatto, 2006). When task timing is randomized, reducing expectancy/anticipation of specific events (such as providing a motor response at the end of the task), this decrease in microsaccade rate is not found (Rolfs, 2009).

Current evidence suggests that microsaccades are beneficial for–rather than detrimental to–high-acuity tasks (Ko et al., 2010; Poletti et al., 2013; Rucci et al., 2007). Computational modelling efforts have also shown that microsaccades improve spatial resolution at the level of retinal responses, increasing sensitivity to edges and line patterns (Donner & Hemila, 2007). Microsaccade rates moreover increase during fixations on informative regions of images across a variety of tasks (McCamy, Otero-Millan, et al., 2014b; Otero-Millan et al., 2008). This suggests that the brain may actively generate microsaccades to sample meaningful information from the visual environment (McCamy, Otero-Millan et al., 2014b).

3.8 Microsaccades and Attention

There is a strong link between attention and large saccades (Yarbus, 1967; Findlay & Gilchrist, 2003; Gowen, Abadi, Poliakoff, Hansen, & Miall, 2007; see also the chapter by Foulsham in this volume). This relationship extends to microsaccades, due to biases in the SC activation map, which is involved in the generation of both microsaccades and large saccades. Specifically, microsaccadic directions tend to be biased towards visual field locations where covert (i.e. hidden) spatial attention is deployed (Engbert & Kliegl, 2003a; Gowen et al., 2007; Hafed & Clark, 2002; Laubrock, Engbert, & Kliegl, 2005; Laubrock, Kliegl, Rolfs, & Engbert, 2010; Turatto, Valsecchi, Tame, & Betta, 2007; Yuval-Greenberg, Merriam, & Heeger, 2014) even in the absence of visual information (Martinez-Conde & Alexander, 2019; Van Ede, Chekroud, & Nobre, 2019).

Some early studies presented mostly correlational evidence (Engbert & Kliegl, 2003a; Hafed & Clark, 2002; Laubrock et al., 2005) that visual attention modulated spatial and temporal microsaccade dynamics, and other research failed to find a link between microsaccades and attention altogether (Horowitz, Fine, Fencsik, Yurgenson, & Wolfe, 2007). However, more recent studies have convincingly shown that the directions of spontaneous microsaccades do reflect spatial shifts in covert attention (Yuval-Greenberg et al., 2014). Other work has moreover revealed a link between microsaccade-related brain potentials and the focusing of visuospatial attention (Meyberg, Werkle-Bergner, Sommer, & Dimigen, 2015). Furthermore, a computational model has been developed (Engbert, 2012; Engbert et al., 2011) that can explain microsaccade rate and direction effects as a function of neural activation, and which is modulated by perceptual input and visual attention. Thus, multiple lines of evidence support the proposal that attention and microsaccades are intimately related, and thus that microsaccades can serve as an index of covert attention (Laubrock, Engbert, Rolfs, & Kliegl, 2007; Laubrock et al., 2010).

Shifts in attention also cause changes in the *rate* of microsaccades. Fewer microsaccades are produced ~100 to 200 ms after an attentional cue onset. This transient decrease is followed by a momentary rebound in microsaccade production (Engbert & Kliegl, 2003a; Laubrock et al., 2010; Rolfs, Engbert, & Kliegl, 2004).

These combined studies suggest that task demands influence both microsaccadic direction and rate, within limits. No matter how hard we try to suppress microsaccades, we cannot eliminate them, or reduce their production below a certain rate. Similarly, only a fraction of all microsaccades is directed towards attended locations. Other microsaccades may help correct gaze position, or arise from neural fluctuations.

To sum up, fixational eye movements help prevent and reverse visual fading, move the eye to advantageous gaze positions, and correct fixation errors. Other roles have begun to be explored, but it seems clear that, far from being simple oculomotor noise, fixational eye movements have wide-ranging effects in vision and perception. Researchers have moreover started to use fixational eye movements to help facilitate clinical diagnoses and treatments of neural and ophthalmic diseases, to detect shifts in attention, and to develop a fuller understanding of our everyday visual experience. Over the course of the day, our gaze is most often engaged in fixation periods between large saccades. Thus, it is likely that—whatever your field of visual or perceptual research—fixational eye movements will have implications for the topics you are interested in. What is the perceptual experience of your experimental participants while they fixate their gaze during your study? Are fixational eye movements helping them perform the task you assigned them? If you think deeply about how fixational eye movements may impact you research question, you may be surprised about what you discover!

Box 2: Microsaccades and illusory motion Researchers may have only scratched the surface on the perceptual roles served by microsaccades. As an example, microsaccades have an influence on motion perception (Laubrock, Engbert, & Kliegl, 2008; Murakami, 2006; Murakami & Cavanagh, 1998, 2001), and several studies have moreover found that microsaccade production affects the perception of some types of illusory motion in static images. One study asked subjects to indicate when the speed of illusory motion increased or decreased in the “Enigma” illusion, and found more frequent microsaccades right before illusory motion gained speed. Conversely, there were less frequent microsaccades right before illusory motion slowed down or stopped completely (Troncoso, Macknik, Otero-Millan, & Martinez-Conde, 2008). Drift was initially linked to the perception of the “Rotating Snakes” illusion

(Murakami, Kitaoka, & Ashida, 2006), but a later study found that microsaccades, rather than drift, triggered the illusory perception of motion in the "Rotating Snakes"(Otero-Millan et al., 2012), see Fig. 3.9. This research proposed that quick image displacements due to microsaccades (and other transient ocular events such as saccades and blinks) may "trick" motion detectors in the visual cortex into responding as though there was actual motion in the world, thus detecting a motion signal in certain stationary, repetitive patterns. The perceived direction of such illusory motion is determined by the image pattern itself, rather than by the direction of the eye movement. The perceptual consequence for the "Rotating Snakes" illusion is that the "snakes" appear to move. Intriguingly, the "Rotating Snakes" illusion is more prevalent in younger than in older observers (Billino, Hamburger, & Gegenfurtner, 2009), but the reason for this age-dependency is not yet known (Macknik & Martinez-Conde, 2016).

3.9 Suggested Readings

Horowitz, T. S., Fine, E. M., Fencsik, D. E., Yurgenson, S., & Wolfe, J. M. (2007). Fixational eye movements are not an index of covert attention. *Psychological Science*, *18*, 356–363.

Laubrock, J., Engbert, R., Rolfs, M., & Kliegl, R. (2007). Microsaccades Are an Index of Covert Attention Commentary on Horowitz, Fine, Fencsik, Yurgenson, and Wolfe (2007). *Psychological Science*, *18*(4), 364–366.

Horowitz, T. S., Fencsik, D. E., Fine, E. M., Yurgenson, S., & Wolfe, J. M. (2007). Microsaccades and attention: does a weak correlation make an index? Reply to Laubrock, Engbert, Rolfs, and Kliegl (2007). *Psychological Science*, *18*(4), 367–368.

Laubrock, J., Kliegl, R., Rolfs, M., & Engbert, R. (2010). When do microsaccades follow spatial attention? Attention, *Perception, & Psychophysics*, *72*(3), 683–694.

- *This series of papers discusses whether there is a direct, causal relationship between microsaccades and the direction of covert spatial attention. Taken together, the above studies provide a glimpse into a complicated ongoing debate: What is the relationship between attention and microsaccades, and how should that relationship be characterized? The current state of the evidence suggests a tight link between spatial attention and microsaccades.*

Leigh, R. J., & Zee, D. S. (2015). *The neurology of eye movements* (5th ed.). Oxford: Oxford University Press.

Alexander, R. G., Macknik, S. L., & Martinez-Conde, S. (2018). Microsaccade characteristics in neurological and ophthalmic disease. *Frontiers in Neurology, 9*, 144. https://doi.org/10.3389/fneur.2018.00144. PMCID: PMC5859063.

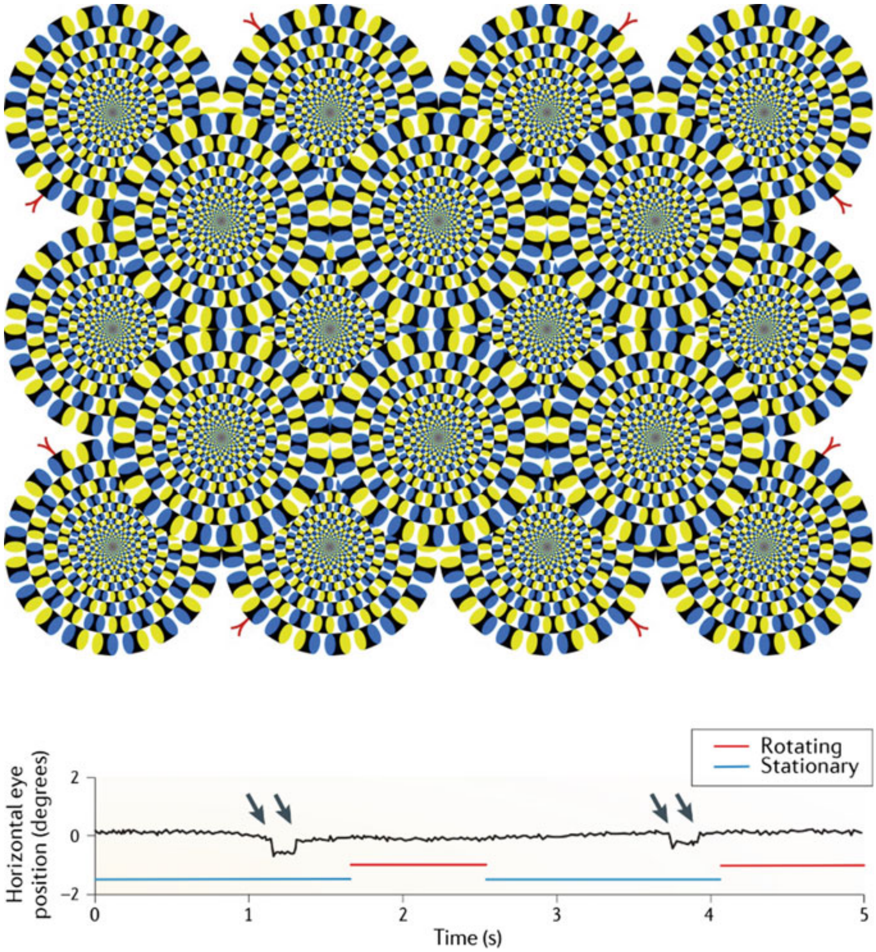


Fig. 3.9 Microsaccades trigger illusory rotation in the “Rotating Snakes” illusion (top panel). Look at the center of one of the “snakes” and the illusory motion will slow down or even stop. Move your eyes from one snake to another (or simply relax your fixation and allow more microsaccades to occur), and the snakes will begin to rotate again. The bottom panel shows eye positions and perceptual reports (with red lines indicating rotating snakes and blue lines indicating stationary snakes) from one subject from Otero-Millan, Macknik, and Martinez-Conde (2012). Arrows identify microsaccades. Top panel reproduced from Kitaoka (2003); bottom panel reproduced from (Martinez-Conde et al., 2013)

- *Provides some further (though non-exhaustive) discussion of abnormal fixational eye movements and their perceptual consequences, including ocular flutter, nystagmus, and the effects of ocular paralysis.*

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nature Reviews Neuroscience, 14*, 83–96.

- *Provides a comprehensive overview of the neural bases of microsaccades, including a detailed summary of the neural responses to microsaccades along the visual pathway.*

Rolfs, M. (2009). Microsaccades: small steps on a long way. *Vision Research, 49*, 2415–2441.

- *Provides a thorough review of research on the gaze-corrective functions of microsaccades.*

3.10 Questions Students Should Be Able to Answer

1. List drift, tremor, and microsaccades, from the smallest to the largest in average amplitude (in healthy individuals). What are the defining features of each kind of eye movement?
2. List and discuss some of the effects of microsaccades on visual perception.
3. Your friend John wants to run an experiment where he will unambiguously distinguish between microsaccades and saccades. Is this possible, and if so, how?
4. If you stare for many seconds at a single point in space, what will happen to the way you see some stationary objects in your visual periphery? Describe the perceptual phenomenon that will take place and its underlying basis, in terms of neural adaptation.
5. What is perceptual fading, and what fixational eye movement(s) can reverse it?
6. Why are larger eye movements more effective in reversing perceptual fading than smaller eye movements?
7. If you keep your experimental participants' heads as still as possible (for instance, using a bite bar), will they still make microsaccades?
8. How would you use microsaccadic features to determine whether participants paid attention to the peripheral target on Fig. 3.6, while maintaining fixation on the central dot?

Acknowledgements This work was supported by a challenge grant from Research to Prevent Blindness Inc. to the Department of Ophthalmology at SUNY Downstate, the Empire Innovation Program (Award to SMC), and the National Science Foundation (Award 1734887). We thank Max Dorfman for his comments and Daniel Cortes-Rastrollo for administrative assistance.

Bibliography

Abadi, R. V., & Gowen, E. (2004). Characteristics of saccadic intrusions. *Vision Research, 44*(23), 2675–2690.

Adler, F. H., & Fliegelman, M. (1934). Influence of fixation on the visual acuity. *Archives of Ophthalmology, 12,* 475–483.

Ahissar, E., & Arieli, A. (2012). Seeing via miniature eye movements: A dynamic hypothesis for vision. *Frontiers in computational neuroscience, 6*.

Alexander, R. G., Macknik, S. L., & Martinez-Conde, S. (2018). Microsaccade characteristics in neurological and ophthalmic disease. *Frontiers in Neurology, 9*(144), 1–9. https://doi.org/10.3389/fneur.2018.00144.

Aytekin, M., Victor, J. D., & Rucci, M. (2014). The visual input to the retina during natural head-free fixation. *The Journal of Neuroscience, 34*(38), 12701–12715.

Bair, W., & O'Keefe, L. P. (1998). The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Visual Neuroscience, 15*, 779–786.

Barlow, H. B. (1952). Eye movements during fixation. *Journal of Physiology, 116,* 290–306.

Beer, A. L. et al. (2008). A motion illusion reveals mechanisms of perceptual stabilization. *PLoS ONE, 3*, e2741.

Benedetto, S., Pedrotti, M., & Bridgeman, B. (2011). Microsaccades and exploratory saccades in a naturalistic environment. *Journal of Eye Movement Research, 4*(2), 1–10.

Bengi, H., & Thomas, J. (1968). Three electronic methods for recording ocular tremor. *Medical and Biological Engineering, 6*(2), 171–179.

Betta, E., & Turatto, M. (2006). Are you ready? I can tell by looking at your microsaccades. *NeuroReport, 17*(10), 1001–1004.

Betta, E. et al. (2007). Microsaccadic response during inhibition of return in a target-target paradigm. *Vision Research, 47*, 428–436.

Billino, J., Hamburger, K., & Gegenfurtner, K. R. (2009). Age effects on the perception of motion illusions. *Perception, 38*(4), 508–521.

Bosman, C. A. et al. (2009). A microsaccadic rhythm modulates gammaband synchronization and behavior. *Journal of Neuroscience, 29*, 9471–9480.

Boyce, P. R. (1967). Monocular fixation in human eye movement. *Proceedings of the Royal Society of London. B, 167*, 293–315.

Boyle, G., Coakley, D., & Malone, J. F. (2001). Interferometry for ocular microtremor measurement. *Applied Optics, 40*(1), 167–175.

Bridgeman, B., & Palca, J. (1980). The role of microsaccades in high acuity observational tasks. *Vision Research, 20*(9), 813–817.

Brien, D. C., Corneil, J. H., Fecteau, J. H., Bell, A. H., & Munoz, D. P. (2009). The behavioural and neurophysiological modulation of microsaccades in monkeys. *Journal of Eye Movement Research, 3*, 1–12.

Carpenter, R. H. S. (1988). *Movements of the eyes* (2nd ed.). London: Pion Ltd.

Cherici, C., Kuang, X., Poletti, M., & Rucci, M. (2012). Precision of sustained fixation in trained and untrained observers. *Journal of Vision, 12*(6).

Collewijn, H., & Kowler, E. (2008). The significance of microsaccades for vision and oculomotor control. *Journal of Vision, 8*(14). https://doi.org/10.1167/8.14.20.

Cornsweet, T. N. (1956). Determination of the stimuli for involuntary drifts and saccadic eye movements. *The Journal of the Optical Society of America, 46*(11), 987–993.

Costela, F. M., McCamy, M. B., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2013a). Microsaccades restore the visibility of minute foveal targets. *PeerJ, 1,* e119. https://doi.org/10.7717/peerj.119.

Costela, F. M., Otero-Millan, J., McCamy, M. B., Macknik, S. L., Troncoso, X. G., Jazi, A. N., & Martinez-Conde, S. (2014). Fixational eye movement correction of blink-induced gaze position errors. *PLoS ONE, 9*(10), e110889. https://doi.org/10.1371/journal.pone.0110889.

Costela, F. M., Otero-Millan, J., McCamy, M. B., Macknik, S. L., Troncoso, X., & Martinez-Conde, S. (2013). Microsaccades correct fixation errors due to blinks. *Journal of Vision, 13*(9), 1335–1335.

Cui, J. et al. (2009). Visibility states modulate microsaccade rate and direction. *Vision Research, 49*, 228–236.

Di Stasi, L. L., Cabestrero, R., McCamy, M. B., Ríos, F., Catena, A., Quirós, P., & Martinez-Conde, S. (2014a). Intersaccadic drift velocity is sensitive to short-term hypobaric hypoxia. *The European Journal of Neuroscience, 39*(8), 1384–1390. https://doi.org/10.1111/ejn.12482.

Di Stasi, L. L., Catena, A., Cañas, J. J., Macknik, S. L., & Martinez-Conde, S. (2013a). Saccadic velocity as an arousal index in naturalistic tasks. *Neuroscience and Biobehavioral Reviews, 37*(5), 968–975. https://doi.org/10.1016/j.neubiorev.2013.03.011.

Di Stasi, L. L., McCamy, M. B., Catena, A., Macknik, S. L., Cañas, J. J., & Martinez-Conde, S. (2013b). Microsaccade and drift dynamics reflect mental fatigue. *The European Journal of Neuroscience, 38*(3), 2389–2398. https://doi.org/10.1111/ejn.12248.

Di Stasi, L. L., McCamy, M. B., Macknik, S. L., Mankin, J. A., Hooft, N., Catena, A., & Martinez-Conde, S. (2014b). Saccadic eye movement metrics reflect surgical residents' fatigue. *Annals of Surgery, 259*(4), 824–829.

Di Stasi, L. L., McCamy, M. B., Pannasch, S., Renner, R., Catena, A., Cañas, J. J., & Martinez-Conde, S. (2015). Effects of driving time on microsaccadic dynamics. *Experimental Brain Research, 233*(2), 599–605.

Dimigen, O., Valsecchi, M., Sommer, W., & Kliegl, R. (2009). Human microsaccade-related visual brain responses. *Journal of Neuroscience, 29*(39), 12321–12331.

Ditchburn, R. W., & Foley-Fisher, J. A. (1967). Assembled data in eye movements. *Optica Acta (Lond), 14*(2), 113–118.

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature, 170,* 36–37.

Ditchburn, R. W., & Ginsborg, B. L. (1953). Involuntary eye movements during fixation. *Journal of Physiology, 119*(1), 1–17.

Donner, K., & Hemila, S. (2007). Modelling the effect of microsaccades on retinal responses to stationary contrast patterns. *Vision Research, 47*(9), 1166–1177.

Eizenman, M., Hallett, P. E., & Frecker, R. C. (1985). Power spectra for ocular drift and tremor. *Vision Research, 25*(11), 1635–1640.

Engbert, R. (2012). Computational modeling of collicular integration of perceptual responses and attention in microsaccades. *Journal of Neuroscience, 32*(23), 8035–8039. https://doi.org/10.1523/JNEUROSCI.0808-12.2012.

Engbert, R., & Kliegl, R. (2003a). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035–1045.

Engbert, R., & Kliegl, R. (2003b). The mind's eyes: Cognitive and applied aspects of eye movements. In: J. Hyona, R. Radach & H. Deubel (Eds.) (pp. 103–117), Oxford: Elsevier.

Engbert, R., & Kliegl, R. (2004). Microsaccades keep the eyes' balance during fixation. *Psychological Science, 15*(6), 431–436.

Engbert, R., & Mergenthaler, K. (2006). Microsaccades are triggered by low retinal image slip. *Proceedings of the National Academy of Sciences of the United States of America, 103*, 7192–7197.

Engbert, R., Mergenthaler, K., Sinn, P., & Pikovsky, A. (2011). An integrated model of fixational eye movements and microsaccades. *Proceedings of the National Academy of Sciences of the United States of America, 108*(39), E765–770.

Falkenberg, H. K., Rubin, G. S., & Bex, P. J. (2007). Acuity, crowding, reading and fixation stability. *Vision Research, 47*(1), 126–135.

Findlay, J., & Gilchrist, I. (2003). *Active vision: The psychology of seeing and looking*. Oxford: Oxford University Press.

Fiorentini, A., & Ercoles, A. M. (1966). Involuntary eye movements during attempted monocular fixation. *Atti della Fondazione Giorgio Ronchi, 21*, 199–217.

Galfano, G. et al. (2004). Inhibition of return in microsaccades. *Experimental Brain Research, 159*, 400–404.

Gandhi, N. J., & Keller, E. L. (1999). Activity of the brain stem omnipause neurons during saccades perturbed by stimulation of the primate superior colliculus. *Journal of Neurophysiology, 82*(6), 3254–3267.

Goffart, L. et al. (2006). Influence of background illumination on fixation and visually guided saccades in the rhesus monkey. *Vision Research, 46*, 149–162.

Gowen, E., Abadi, R. V., Poliakoff, E., Hansen, P. C., & Miall, R. C. (2007). Modulation of saccadic intrusions by exogenous and endogenous attention. *Brain Research, 1141,* 154–167. https://doi.org/10.1016/j.brainres.2007.01.047.

Greschner, M., Bongard, M., Rujan, P., & Ammermuller, J. (2002). Retinal ganglion cell synchronization by fixational eye movements improves feature stimation. *Nature Neuroscience, 5*(4), 341–347.

Guitton, D., & Munoz, D. P. (1991). Control of orienting gaze shifts by the tectoreticulospinal system in the head-free cat. I. Identification, localization, and effects of behavior on sensory responses. *Journal of Neurophysiology, 66*(5), 1605–1623.

Hafed, Z. M., & Clark, J. J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research, 42*(22), 2533–2545.

Hafed, Z. M., Goffart, L., & Krauzlis, R. (2009). A neural mechanism for microsaccade generation in the primate superior colliculus. *Science, 323*(5916), 940–943.

Herrington, T. M. et al. (2009). The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. *Journal of Neuroscience, 29*, 5793–5805.

Horwitz, G. D., & Albright, T. D. (2003). Short-latency fixational saccades induced by luminance increments. *Journal of Neurophysiology, 90*, 1333–1339.

Horowitz, T. S., Fine, E. M., Fencsik, D. E., Yurgenson, S., & Wolfe, J. M. (2007). Fixational eye movements are not an index of covert attention. *Psychological Science, 18*(4), 356–363.

Hubel, D. H. (1988). *Eye, brain, and vision* (Vol. 22). New York: Scientific American Library.

Jurin. (1738). Essay on distinct and indistintc vision. *Optics*.

Kagan, I. et al. (2008). Saccades and drifts differentially modulate neuronal activity in V1: Effects of retinal image motion, position, and extraretinal influences. *Journal of Vision, 8*(14), 19, 11–25.

Kingstone, A., Fendrich, R., Wessinger, C. M., & ReuterLorenz, P. A. (1995). Are microsaccades responsible for the gap effect? *Perception & Psychophysics, 57*, 796–801.

Kitaoka, A. (2003). *Akiyoshi's illusion pages*. Retrieved from http://www.ritsumei.ac.jp/~akitaoka/index-e.html.

Kliegl, R. et al. (2009). Microsaccadic modulation of response times in spatial attention tasks. *Psychological Research, 73*, 136–146.

Ko, H.-K., Poletti, M., & Rucci, M. (2010). Microsaccades precisely relocate gaze in a high visual acuity task. *Nature Neuroscience, 13*(12), 1549–1553. https://doi.org/10.1038/nn.2663.

Ko, H.-K., Snodderly, D. M., & Poletti, M. (2016). Eye movements between saccades: Measuring ocular drift and tremor. *Vision Research, 122,* 93–104.

Kowler, E., & Steinman, R. M. (1980). Small saccades serve no useful purpose: Reply to a letter by R. W. Ditchburn. *Vision Research, 20*(3), 273–276.

Krauskopf, J. (1967). Heterochromatic stabilized images: A classroom demonstration. *American Journal of Psychology, 80*(4), 634–637.

Krauskopf, J. et al. (1960). Analysis of eye movements during monocular and binocular fixation. *Journal of the Optical Society of America, 50*, 572–578.

Kuang, X., Poletti, M., Victor, J. D., & Rucci, M. (2012). Temporal encoding of spatial information during active visual fixation. *Current Biology, 22*(6), 510–514.

Laubrock, J., Engbert, R., & Kliegl, R. (2005). Microsaccade dynamics during covert attention. *Vision Research, 45*(6), 721–730. https://doi.org/10.1016/j.visres.2004.09.029.

Laubrock, J., Engbert, R., & Kliegl, R. (2008). Fixational eye movements predict the perceived direction of ambiguous apparent motion. *Journal of Vision, 8(14):13*, 1–17.

Laubrock, J., Engbert, R., Rolfs, M., & Kliegl, R. (2007). Microsaccades are an index of covert attention: Commentary on Horowitz, Fine, Fencsik, Yurgenson, and Wolfe (2007). *Psychological Science, 18*(4), 364–366; discussion 367–368.

Laubrock, J., Kliegl, R., Rolfs, M., & Engbert, R. (2010). When do microsaccades follow spatial attention? *Attention, Perception, & Psychophysics, 72*(3), 683–694. https://doi.org/10.3758/APP.72.3.683.

Leigh, R. J., & Zee, D. S. (2015). *The neurology of eye movements* (5th ed.). Oxford: Oxford University Press.

Leopold, D. A., & Logothetis, N. K. (1998). Microsaccades differentially modulate neural activity in the striate and extrastriate visual cortex. *Experimental Brain Research, 123*, 341–345.

Liang, J. R., Moshel, S., Zivotofsky, A. Z., Caspi, A., Engbert, R., Kliegl, R. et al. (2005). Scaling of horizontal and vertical fixational eye movements. *Physical Review E*, *71*, 031909.

Lord, M. P. (1951). Measurement of binocular eye movements of subjects in the sitting position. *British Journal of Ophthalmology, 35*, 21–30.

Lord, M. P., & Wright, W. D. (1948). Eye movements during monocular fixation. *Nature, 162*, 25–26.

Macknik, S. L., & Martinez-Conde, S. (2016). The age of illusion. *Scientific American Mind, 27*(1), 18–19.

Malinov, I. V., Epelboim, J., Herst, A. N., & Steinman, R. M. (2000). Characteristics of saccades and vergence in two types of sequential looking tasks. *Vision Research, 40*, 2083–2090.

Martinez-Conde, S. (2006). Fixational eye movements in normal and pathological vision. *Progress in Brain Research, 154,* 151–176. https://doi.org/10.1016/s0079-6123(06)54008-7.

Martinez-Conde, S., & Alexander, R. G. (2019). A gaze bias in the mind's eye. *Nature Human Behaviour, 3*(5), 424–425. https://doi.org/10.1038/s41562-019-0546-1.

Martinez-Conde, S., & Macknik, S. L. (2007). Windows on the mind. *Scientific American, 297*(2), 56–63.

Martinez-Conde, S., & Macknik, S. L. (2015). From exploration to fixation: An integrative view of Yarbus's vision. *Perception, 44*(8–9), 884–899. https://doi.org/10.1177/0301006615594963.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys [published erratum appears in Nature Neuroscience 2000 Apr; 3(4):409]. *Nature Neuroscience, 3*(3), 251–258. https://doi.org/10.1038/72961.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2002). The function of bursts of spikes during visual fixation in the awake primate lateral geniculate nucleus and primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 99*(21), 13920–13925. https://doi.org/10.1073/pnas.212500599.

Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2004). The role of fixational eye movements in visual perception. *Nature Reviews Neuroscience, 5*(3), 229–240. https://doi.org/10.1038/nrn1348.

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Dyar, T. A. (2006). Microsaccades counteract visual fading during fixation. *Neuron, 49*(2), 297–305. https://doi.org/10.1016/j.neuron.2005.11.033. S0896-6273(05)01056-1 [pii].

Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: A neurophysiological analysis. *Trends in Neurosciences, 32*(9), 463–475. https://doi.org/10.1016/j.tins.2009.05.006.

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience, 14*(2), 83–96. https://doi.org/10.1038/nrn3405.

McCamy, M. B., Collins, N., Otero-Millan, J., Al-Kalbani, M., Macknik, S. L., Coakley, D., & Wolf. T. R. (2013a). Simultaneous recordings of ocular microtremor and microsaccades with a piezoelectric sensor and a video-oculography system. *PeerJ, 1,* e14. https://doi.org/10.7717/peerj.14.

McCamy, M. B., Macknik, S. L., & Martinez-Conde, S. (2013b). Natural eye movements and vision. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences*. Cambridge, MA: The MIT Press.

McCamy, M. B., Najafian Jazi, A., Otero-Millan, J., Macknik, S. L., & Martinez-Conde, S. (2013c). The effects of fixation target size and luminance on microsaccades and square-wave jerks. *PeerJ, 1,* e9. https://doi.org/10.7717/peerj.9.

McCamy, M. B., Macknik, S. L., & Martinez-Conde, S. (2014a). Different fixational eye movements mediate the prevention and the reversal of visual fading. *The Journal of Physiology, 592*(19), 4381–4394. https://doi.org/10.1113/jphysiol.2014.279059.

McCamy, M. B., Otero-Millan, J., Di Stasi, L. L., Macknik, S. L., & Martinez-Conde, S. (2014b). Highly informative natural scene regions increase microsaccade production during visual scanning. *The Journal of Neuroscience, 34*(8), 2956–2966. https://doi.org/10.1523/jneurosci.4448-13.2014.

McCamy, M. B., Otero-Millan, J., Leigh, R. J., King, S. A., Schneider, R. M., Macknik, S. L., & Martinez-Conde, S. (2015). Simultaneous recordings of human microsaccades and drifts with a contemporary video eye tracker and the search coil technique. *PLoS ONE, 10*(6), e0128428. https://doi.org/10.1371/journal.pone.0128428.

McCamy, M. B., Otero-Millan, J., Macknik, S. L., Yang, Y., Troncoso, X. G., Baer, S. M., & Martinez-Conde, S. (2012). Microsaccadic efficacy and contribution to foveal and peripheral vision. *Journal of Neuroscience, 32*(27), 9194–9204. https://doi.org/10.1523/JNEUROSCI.0515-12.2012.

Meirovithz, E., Ayzenshtat, I., Werner-Reiss, U., Shamir, I., & Slovin, H. (2012). Spatiotemporal effects of microsaccades on population activity in the visual cortex of monkeys during fixation. *Cerebral Cortex, 22*(2), 294–307.

Mergenthaler, K., & Engbert, R. (2007). Modeling the control of fixational eye movements with neurophysiological delays. *Physical Review Letters, 98*, 138104.

Mergenthaler, K., & Engbert, R. (2010). Microsaccades are different from saccades in scene perception. *Experimental Brain Research, 203*(4), 753–757.

Meyberg, S., Werkle-Bergner, M., Sommer, W., & Dimigen, O. (2015). Microsaccade-related brain potentials signal the focus of visuospatial attention. *Neuroimage, 104,* 79–88.

Moller, F., Laursen, M. L., Tygesen, J., & Sjolie, A. K. (2002). Binocular quantification and characterization of microsaccades. *Graefes Archive for Clinical and Experimental Ophthalmology, 240*(9), 765–770. https://doi.org/10.1007/s00417-002-0519-2.

Moschovakis, A., Scudder, C., & Highstein, S. (1996). The microscopic anatomy and physiology of the mammalian saccadic system. *Progress in Neurobiology, 50*(2), 133–254.

Moshel, S. et al. (2008). Persistence and phase synchronization properties of fixational eye movements. *European Physical Journal Special Topics, 161*, 207–223.

Munoz, D. P. (2002). Commentary: Saccadic eye movements: Overview of neural circuitry. *Progress in Brain Research, 140,* 89–96.

Munoz, D. P., & Wurtz, R. H. (1995). Saccade-related activity in monkey superior colliculus. I. Characteristics of burst and buildup cells. *Journal of Neurophysiology, 73*(6), 2313–2333.

Murakami, I. (2006). Fixational eye movements and motion perception. *Progress in Brain Research, 154,* 193–209.

Murakami, I., & Cavanagh, P. (1998). A jitter after-effect reveals motion-based stabilization of vision. *Nature, 395,* 798–801.

Murakami, I., & Cavanagh, P. (2001). Visual jitter: Evidence for visual-motion-based compensation of retinal slip due to small eye movements. *Vision Research, 41*(2), 173–186.

Murakami, I., Kitaoka, A., & Ashida, H. (2006). A positive correlation between fixation instability and the strength of illusory motion in a static display. *Vision Research, 46*(15), 2421–2431.

Müri, R., Nyffeler, T., & Cazzoli, D. (2019). Neurology. In C. Klein & U. Ettinger (Eds.), *Eye movement research: An introduction to its scientific foundations and applications*. Berlin: Springer Publishers.

Nachmias, J. (1959). Two-dimensional motion of the retinal image during monocular fixation. *Journal of the Optical Society of America, 49,* 901–908.

Nachmias, J. (1961). Determiners of the drift of the eye during monocular fixation. *Journal of the Optical Society of America, 51*, 761–766.

Nyström, M., Hansen, D. W., Andersson, R., & Hooge, I. (2014). Why have microsaccades become larger? Investigating eye deformations and detection algorithms. *Vision Research*.

Optican, L. M. (1995). A field theory of saccade generation: Temporal-to-spatial transform in the superior colliculus. *Vision Research, 35*(23), 3313–3320.

Otero-Millan, J., Macknik, S. L., Langston, R. E., & Martinez-Conde, S. (2013a). An oculomotor continuum from exploration to fixation. *Proceedings of the National Academy of Sciences of the United States of America, 110*(15), 6175–6180. https://doi.org/10.1073/pnas.1222715110.

Otero-Millan, J., Troncoso, X. G., Macknik, S. L., Serrano-Pedraza, I., & Martinez-Conde, S. (2008). Saccades and microsaccades during visual fixation, exploration, and search: Foundations for a common saccadic generator. *Journal of Vision, 8*(14), 1–18, 21. https://doi.org/10.1167/8.14.21/8/14/21.

Otero-Millan, J., Macknik, S. L., & Martinez-Conde, S. (2012). Microsaccades and blinks trigger illusory rotation in the “rotating snakes” illusion. *Journal of Neuroscience, 32*(17), 6043–6051. https://doi.org/10.1523/JNEUROSCI.5823-11.2012.

Otero-Millan, J., Macknik, S. L., Serra, A., Leigh, R. J., & Martinez-Conde, S. (2011). Triggering mechanisms in microsaccade and saccade generation: A novel proposal. *Annals of the New York Academy of Sciences, 1233*(1), 107–116. https://doi.org/10.1111/j.1749-6632.2011.06177.x.

Otero-Millan, J., Schneider, R., Leigh, R. J., Macknik, S. L., & Martinez-Conde, S. (2013b). Saccades during Attempted Fixation in Parkinsonian Disorders and Recessive Ataxia: From microsaccades to square-wave jerks. *PLoS ONE, 8*(3), e58535. https://doi.org/10.1371/journal.pone.0058535.

Pinnock, R. A., McGivern, R. C., Forbes, R., & Gibson, J. M. (2010). An exploration of ocular fixation in Parkinson’s disease, multiple system atrophy and progressive supranuclear palsy. *Journal of Neurology.*

Poletti, M., Listorti, C., & Rucci, M. (2013). Microscopic eye movements compensate for nonhomogeneous vision within the fovea. *Current Biology, 23*(17), 1691–1695.

Poletti, M., & Rucci, M. (2010). Eye movements under various conditions of image fading. *Journal of Vision, 10*(3), 6 1–18.

Pritchard, R. M. (1961). Stabilized images on the retina. *Scientific American, 204,* 72–78.

Ratliff, F., & Riggs, L. A. (1950). Involuntary motions of the eye during monocular fixation. *Journal of Experimental Psychology, 40*, 687–701.

Riggs, L. A., & Ratliff, F. (1951). Visual acuity and the normal tremor of the eyes. *Science, 114*(2949), 17–18.

Riggs, L. A., & Ratliff, F. (1952). The effects of counteracting the normal movements of the eye. *Journal of the Optical Society of America, 42,* 872–873.

Riggs, L. A., Armington, J. C., & Ratliff, F. (1954). Motions of the retinal image during fixation. *Journal of the Optical Society of America, 44*, 315–321.

Robinson, D. A. (1973). Models of the saccadic eye movement control system. *Kybernetik, 14*(2), 71–83.

Rolfs, M. (2009). Microsaccades: Small steps on a long way. *Vision Research, 49*(20), 2415–2441. https://doi.org/10.1016/j.visres.2009.08.010. S0042-6989(09)00369-1 [pii].

Rolfs, M., Engbert, R., & Kliegl, R. (2004). Microsaccade orientation supports attentional enhancement opposite a peripheral cue: Commentary on Tse, Sheinberg, and Logothetis (2003). *Psychological Science, 15*(10), 705–707; author reply 708–710. https://doi.org/10.1111/j.0956-7976.2004.00744.x.

Rolfs, M. et al. (2005). Crossmodal coupling of oculomotor control and spatial attention in vision and audition. *Experimental Brain Research, 166*, 427–439.

Rolfs, M., Kliegl, R., & Engbert, R. (2008). Toward a model of microsaccade generation: The case of microsaccadic inhibition. *Journal of Vision, 8*(11), 5, 1–23. https://doi.org/10.1167/8.11.5.

Rolfs, M., Laubrock, J., & Kliegl, R. (2006). Shortening and prolongation of saccade latencies following microsaccades. *Experimental Brain Research, 169*(3), 369–376. https://doi.org/10.1007/s00221-005-0148-1.

Rolfs, M., Laubrock, J., & Kliegl, R. (2008). Microsaccade-induced prolongation of saccadic latencies depends on microsaccade amplitude. *Journal of Eye Movement Research, 1*(3), 1, 1–8.

Rucci, M., Iovin, R., Poletti, M., & Santini, F. (2007). Miniature eye movements enhance fine spatial detail. *Nature, 447*(7146), 851–854. https://doi.org/10.1038/nature05866.

Rucci, M., & Poletti, M. (2015). Control and functions of fixational eye movements. *Annual Review of Vision Science, 1,* 499–518.

Rucci, M., & Victor, J. D. (2015). The unsteady eye: An information-processing stage, not a bug. *Trends in Neurosciences, 38*(4), 195–206.

Ryle, J. P., Al-Kalbani, M., Collins, N., Gopinathan, U., Boyle, G., Coakley, D., & Sheridan, J. T. (2009). Compact portable ocular microtremor sensor: Design, development and calibration. *Journal of Biomedical Optics, 14*(1), 014021-014021-014012.

Sabrin, H. W., & Kertesz, A. E. (1980). Microsaccadic eye movements and binocular rivalry. *Perception & Psychophysics, 28*, 150–154.

Sansbury, R. V., Skavenski, A. A., Haddad, G. M., & Steinman, R. M. (1973). Normal fixation of eccentric targets. *Journal of the Opical Society of America, 63*, 612–614.

Schiller, P. H. (1984). The superior colliculus and visual function. In D.-S. I. (Ed.), *Handbook of physiology—The nervous system III. Sensory processes part I.* Bethesda MD: American Physiological Society.

Schulz, E. (1984). Binocular micromovements in normal persons. *Graefe's Archive for Clinical and Experimental Ophthalmology, 222*, 95–100.

Schiller, P. H., & Stryker, M. (1972). Single-unit recording and stimulation in superior colliculus of the alert rhesus monkey. *Journal of Neurophysiology, 35,* 915–924.

Scudder, C. A., Kaneko, C. R., & Fuchs, A. F. (2002). The brainstem burst generator for saccadic eye movements. *Experimental Brain Research, 142*(4), 439–462.

Simons, D., Lleras, A., Martinez-Conde, S., Slichter, D., Caddigan, E., & Nevarez, G. (2006). Induced visual fading of complex images. *Journal of Vision, 6,* 1093–1101.

Skavenski, A. A., Robinson, D. A., Steinman, R. M., & Timberlake, G. T. (1975). Miniature eye movements of fixation in rhesus monkey. *Vision Research, 15*(11), 1269–1273.

Snodderly, D. M., Kagan, I., & Gur, M. (2001). Selective activation of visual cortex neurons by fixational eye movements: Implications for neural coding. *Visual Neuroscience, 18*, 259–277.

Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience, 3*(12), 952–964.

Spauschus, A., Marsden, J., Halliday, D. M., Rosenberg, J. R., & Brown, P. (1999). The origin of ocular microtremor in man. *Experimental Brain Research, 126*(4), 556–562.

Srebro, R. (1983). Fixation of normal and amblyopic eyes. *Archives of Ophthalmology, 101*, 214–217.

Stanford, T. R., Freedman, E. G., & Sparks, D. L. (1996). Site and parameters of microstimulation: Evidence for independent effects on the properties of saccades evoked from the primate superior colliculus. *Journal of Neurophysiology, 76*(5), 3360–3381.

Steinman, R. M. (1965). Effect of target size, luminance, and color on monocular fixation. *Journal of the Optical Society of America, 55*, 1158–1165.

Steinman, R. M., Cunitz, R. J., Timberlake, G. T., & Herman, M. (1967). Voluntary control of microsaccades during maintained monocular fixation. *Science, 155*(769), 1577–1579.

Steinman, R. M. et al. (1973). Miniature eye movement. *Science, 181*, 810–819.

Stevens, J. K., Emerson, R. C., Gerstein, G. L., Kallos, T., Neufeld, G. R., Nichols, C. W., & Rosenquis, A. C. (1976). Paralysis of the awake human: Visual perceptions. *Vision Research, 16*(1), 93–98.

Thaler, L., Schutz, A. C., Goodale, M. A., & Gegenfurtner, K. R. (2013). What is the best fixation target? The effect of target shape on stability of fixational eye movements. *Vision Research, 76*, 31–42. https://doi.org/10.1016/j.visres.2012.10.012. S0042-6989(12)00338-0 [pii].

Troncoso, X. G., Macknik, S. L., & Martinez-Conde, S. (2008). Microsaccades counteract perceptual filling-in. *Journal of Vision, 8*(14), 15, 11–19. https://doi.org/10.1167/8.14.15.

Troncoso, X. G., Macknik, S. L., Otero-Millan, J., & Martinez-Conde, S. (2008). Microsaccades drive illusory motion in the Enigma illusion. *Proceedings of the National Academy of Sciences, 105*(41), 16033–16038. https://doi.org/10.1073/pnas.0709389105. 0709389105 [pii].

Tse, P. U., Baumgartner, F. J., & Greenlee, M. W. (2010). Event-related functional MRI of cortical activity evoked by microsaccades, small visually-guided saccades, and eyeblinks in human visual cortex. *Neuroimage, 49*(1), 805–816.

Turatto, M., Valsecchi, M., Tame, L., & Betta, E. (2007). Microsaccades distinguish between global and local visual processing. *NeuroReport, 18*(10), 1015–1018.

Valsecchi, M., & Turatto, M. (2007). Microsaccadic response to visual events that are invisible to the superior colliculus. *Behavioral Neuroscience, 121*, 786–793.

Valsecchi, M., & Turatto, M. (2009). Microsaccadic responses in a bimodal oddball task. *Psychological Research, 73*, 23–33.

van Dam, L. C., & van Ee, R. (2005). The role of (micro)saccades and blinks in perceptual bi-stability from slant rivalry. *Vision Research, 45*, 2417–2435.

Van der Geest, J., & Frens, M. (2002). Recording eye movements with video-oculography and scleral search coils: A direct comparison of two methods. *Journal of Neuroscience Methods, 114*(2), 185–195.

Van Ede, F., Chekroud, S. R. & Nobre, A. C. (2019). Human gaze tracks attentional focusing in memorized visual space. *Nature Human Behaviour*. https://doi.org/10.1038/s41562-019-0549-y.

Verheijen, F. (1961). A simple after image method demonstrating the involuntary multidirectional eye movements during fixation. *Journal of Modern Optics, 8*(4), 309–312.

West, D. C., & Boyce, P. R. (1968). The effect of flicker on eye movement. *Vision Research, 8*, 171–192.

Winterson, B. J., & Collewijn, H. (1976). Microsaccades during finely guided visuomotor tasks. *Vision Research, 16*(12), 1387–1390.

Yarbus, A. L. (1957). The perception of an image fixed with respect to the retina. *Biophysics, 2,* 683–690.

Yarbus, A. L. (1967). *Eye Movements and vision* (B. Haigh, Trans.). New York: Plenum Press.

Yuval-Greenberg, S. et al. (2008). Transient induced gamma-band response in EEG as a manifestation of miniature saccades. *Neuron, 58*, 429–441.

Yuval-Greenberg, S., Merriam, E. P., & Heeger, D. J. (2014). Spontaneous microsaccades reflect shifts in covert attention. *The Journal of Neuroscience, 34*(41), 13693–13700. https://doi.org/10.1523/jneurosci.0582-14.2014.

Zuber, B. L., & Stark, L. (1965). Microsaccades and the velocityamplitude relationship for saccadic eye movements. *Science, 150*, 1459–1460.

Chapter 4
Smooth Eye Movements in Humans: Smooth Pursuit, Optokinetic Nystagmus and Vestibular Ocular Reflex

Rebekka Lencer, Andreas Sprenger and Peter Trillenberg

Contents

R. Lencer (✉)
Department of Psychiatry and Psychotherapy, Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Münster, Münster, Germany
e-mail: rebekka.lencer@ukmuenster.de

A. Sprenger · P. Trillenberg
Department of Neurology, University of Lübeck, Lübeck, Germany

A. Sprenger
Institute of Psychology, University of Lübeck, Lübeck, Germany

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_4

Abstract Smooth pursuit eye movements (SPEM), optokinetic nystagmus (OKN) and the vestibular ocular reflex (VOR) enable us to focus our eyes constantly on objects that move relative to us. In real life situations, natural stimuli that elicit SPEM, OKN and the VOR are processed simultaneously, and do not always act synergistically.

Smooth Pursuit

Smooth pursuit eye movements provide optimal visual acuity of small moving objects by keeping the retinal image of a moving small object within the foveal area. Retinal slip velocity represents the input signal for pursuit. Besides visual motion, processing SPEM control requires sensorimotor transformation of the motion signal into an oculomotor command and its integration with so-called extraretinal mechanisms, i.e. prediction and the efference copy of the ocular motor command. Thus, SPEM are driven by a combination of the prediction of target velocity and visual feedback about performance quality that has to be integrated in various networks. Novel smooth pursuit tasks and quantitative measurement techniques can help unravel these different smooth pursuit components and complex neural systems involved in its control. Under laboratory conditions, pursuit tasks can be broken down into a number of different functions: (1) initiation of the eye movement, (2) pursuit maintenance and (3) using of knowledge of predictable movement of the target. These functions can be studied with appropriate stimuli independently from each other. Ramp or step-ramp stimuli are used to study pursuit initiation, i.e. pursuit latency and pursuit eye velocity or initial eye acceleration during the first 100 ms when internal feedback is yet not available. 100 ms after the onset of pursuit, retinal slip can be used as an input to control eye velocity, i.e. the system works "closed loop" receiving feedback of its own actions. This phase is referred to as pursuit maintenance. The system performance in this phase is quantified by pursuit gain, i.e. the ratio of eye to target velocity. To quantify the contribution of prediction to pursuit performance the target can be blanked off for short intervals. The analysis of smooth eye movement data requires the differentiation of the eye position signal. The result has to be compared to stimulus velocity quantitatively. This can be done by the calculation of latencies, gain (ratio of eye velocity to target velocity) and phase shift with respect to stimulus. Eye acceleration can be determined by a fit to the velocity signal if an additional differentiation must be avoided.

Different models of smooth pursuit with specific in- and output parameters have been developed for a better understanding of the underlying neurophysiological mechanisms and to make quantitative predictions that can be tested in experiments. Earlier models used control system approaches that address nonlinearities and delays in the system. More recent efforts concentrate on mechanisms that can resolve the problems associated with internal delays.

Functional brain imaging and neurophysiological studies have defined motion sensitive visual area V5, parietal (PEF), frontal (FEF) and supplementary (SEF) eye fields as core cortical smooth pursuit regions. In addition, a dense neural network is involved in the adjustment of an optimal smooth pursuit response by integrating also extraretinal information. These networks facilitate interaction of the smooth pursuit system with multiple other visual and non-visual sensorimotor systems on the cortical and subcortical level. Future studies with fMRI advanced techniques (e.g. event-related fMRI) promise to provide an insight into how smooth pursuit eye movements are linked to specific brain activation.

Recent genetic approaches aim at identifying the specific role of different neurotransmitters on the control of pursuit subfunctions, with yet preliminary findings. Most studies focus on polymorphisms involved in dopamine neurotransmission, e.g. the catechol-O-methyltransferase (COMT) Val158Met polymorphism, while others examined genetic polymorphisms of the Neuregulin 1 (NRG1) gene in relation to pursuit performance. Alterations of NRG1 will result in disturbed development of the neural system and may thus have an impact on smooth pursuit control.

Optokinetic Nystagmus

The purpose of OKN is to provide gaze stabilisation when large environmental patterns move relatively to the eyes thereby projecting large images moving over the entire retina. These patterns induce a slow eye movement to stabilize the retinal image of the pattern representing the slow phase of OKN. Both OKN and the smooth pursuit system are assumed to rely on widely overlapping neural pathways being volitionally controlled by a cortical neural network. Depending on several factors of which attention is one of the most important, an OKN response can be suppressed during smooth pursuit although this suppression might not be absolute. Since optokinetic nystagmus is elicited by coherent motion of the whole visual surround its stimulation requires projection of large patterns on a screen without details that stimulate the subjects to focus and pursue. The circuit for the early component of OKN is believed to overlap with smooth pursuit. The delayed component which is less important in humans relies on the nucleus of the optic tract (pretectum) and a velocity storage mechanism similar to the vestibular system but is not relayed via the cerebellum.

Vestibulo-Ocular Reflex

The VOR is characterized by precise compensation for head movements around all possible axes of rotation, short latency and potentially high eye velocity and acceleration. The VOR uses mainly the three-neurons-arc, the shortest connection between efference and afference—only three neurons—apart from stretch reflexes. The anatomical connections of the VOR are based on results from electro-physiological and tracer studies in animals. Signals from the sensory organ are transmitted via the vestibular nuclei and ascending brainstem connections to the oculomotor neurons and their related eye muscles. There are several ways to test the peripheral functions of the VOR: irrigation of the vestibular organ by cold and warm water or air. Cold irrigation induces nystagmus with slow phase velocities in the direction towards the treated ear whereas warm water results in slow phases away from the treated ear.

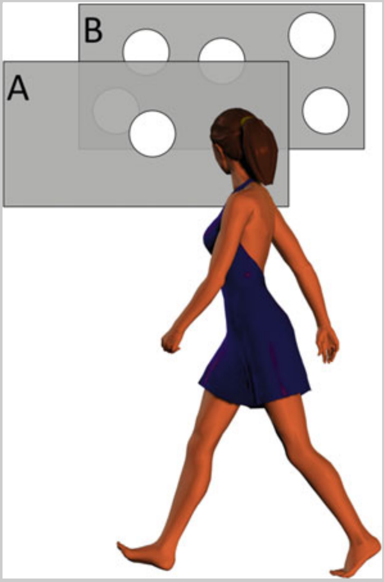
Semicircular canals can be tested separately by quick rotations of the head in the corresponding plane while the subject foveates a point in space. Whole body rotations with sudden stops at velocity saturation elicit nystagmus. Active head shaking may help to detect unilateral labyrinth lesions.

4.1 Introduction with Learning Objectives

Smooth pursuit eye movements (SPEM), optokinetic nystagmus (OKN) and the vestibular ocular reflex (VOR) enable us to focus our eyes constantly on objects which are moving relative to us. Smooth eye movements thereby provide clear vision and a stable percept of objects of interest while these objects are moving, we are moving ourselves and/or our environment is moving. SPEM is elicited by small moving objects, OKN by coherent motion of large visual patterns and VOR by head movements. As laid out in the chapter by Pierce et al. in this volume, saccades are generated in a ballistic manner, i.e. an oculomotor command for a saccade with an a priori defined amplitude is generated and then executed that brings the image of a distinct location in space quickly onto the central fovea. In contrast, the generation of smooth eye movements requires continuous analysis and adjustment of eye velocity to the velocity of the target, of the environment and/or motion of the head, respectively, even while the eyes are moving. A complex neural network involving visual motion processing, “online” feedback mechanisms, the integration of a copy of the oculomotor command, i.e. the efference copy, as well as cognitive mechanisms like prediction and anticipation are needed to drive smooth eye movements. SPEM and OKN are characterized by lower eye velocity and longer latency compared to VOR which is due to the need to compute target velocity by involving also higher-order cortical circuits. This latency causes a phase lag of the eye with respect to the target that can be compensated only by cognitive mechanisms, i.e. prediction. In contrast, the VOR relies mainly on a three-neurons-arc, a very short connection between efference and afference signalling enabling short reaction times and high smooth eye velocities.

In real life situations, natural stimuli eliciting SPEM, OKN and the VOR are processed simultaneously, and do not always act synergistically. As an example, consider a moving object against a background at a larger distance to be fixated, e.g. a bird in flight against the landscape. The moving bird will elicit a smooth pursuit movement in one direction while the landscape represents a large OKN pattern moving to the opposite direction. If additionally the subject also moves, the background will appear to move more slowly than the object while free head movement will provoke the VOR, for more details see Box 1.

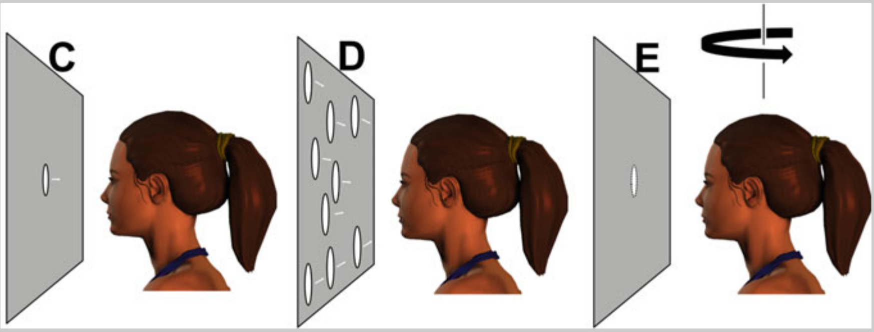
Box 1: Conflicting Stimuli to Evoke Smooth Eye Movements



Slow eye movements enable vision during movements. While we walk, our visual surround moves coherently relative to our eyes. To enable vision, the relative movement of the object around us must be compensated by an eye movement that constantly realigns the line of sight with the object.

In part, the apparent movement of objects can be predicted from the signal of our vestibular organs. They sense angular and translational acceleration and the brain calculates angular and translational velocity. The control of eye movement via vestibular signals is called "vestibular ocular reflex"(VOR). The exact movement of a visual target as a consequence of movements of the body depends on the distance of the target. Thus, a relevant object movement (the movement of the object on sheet A of the figure) must be distinguished from a large coherent movement of other objects (the movement of the objects on sheet B of the figure). The movement of the objects on sheet B can contribute to the drive for the compensatory eye movement if the difference between the movement of the relevant object and surrounding objects is small. If, however, objects on sheet A and B move in different directions (as an example: object A is a dog that runs with the same velocity as the viewer) the coherent motion of the surround (B) must be suppressed in order to pursue the target A effectively.

To study these components in isolation stimuli must be chosen appropriately. For pursuit (C) a single small bright moving object is presented in an otherwise dark room. If due to incomplete dimming of the room the projection screen and its borders are still visible these additional objects could diminish the drive for pursuit. Optokinetic following is stimulated with a large visual surround that moves coherently (D). The stimulus for the vestibular ocular reflex is a head movement. The VOR keeps the line of sight on an imaginary object, no real target is necessary (E).



Suppression of OKN during smooth pursuit will be particularly difficult when the slow phase of OKN goes in opposite direction to the smooth pursuit movement as it is in our example from above: the bird moves in one direction but the background pattern of the landscape potentially evokes a slow phase OKN in the opposite direction. These examples illustrate that different and even divergent retinal image motions must be processed in parallel during smooth pursuit. Integration of information from different retinal areas provides compensation for retinal localization errors and maintains position constancy for static objects during pursuit, which is required to guarantee a stable perception of the visual surrounding (van Beers, Wolpert & Haggard, 2001). When designing an experiment under laboratory conditions, in most cases only distinct aspects of smooth eye movements shall be investigated so that all kinds of confounds must be controlled. As an example, when investigating smooth pursuit of small objects all background patterns that could potentially elicit OKN should be eliminated and head movements should be avoided that could provoke a VOR.

In the following sections, we will lay out the neurophysiological characteristics separately for SPEM, OKN and the VOR before going into more detail on smooth eye movement testing, stimulus requirements and basic approaches of how to analyse recordings from smooth eye movements. We will then introduce basic models of smooth eye movements with a special emphasis on smooth pursuit. Generally, computational models are used for a better understanding of the underlying neurophysiological mechanisms and to make quantitative predictions that can be tested in experiments. Note that elements of models do not necessarily correspond to anatomical structures. However, with the increasing body of findings from structural and functional imaging studies hypotheses derived from models have also been related to anatomical and functional correlates of neural circuits subserving smooth eye movements. Therefore the section on models is followed by a section on current knowledge about neural networks driving and controlling smooth eye movements. The last section will give a comprehensive overview about first approaches to identify specific genetic contributions to smooth eye movement control. This aspect is highly relevant because disturbances of smooth eye movements can occur as first manifestation of genetically inherited neurological disease. In the case of severe mental illness such as schizophrenia and other psychotic disorders, a dysfunction of smooth pursuit is thought to represent a subtle neurophysiological deficit that is under genetic control and that contributes to the disease mechanisms of psychotic disorders. Smooth pursuit dysfunction is therefore regarded as the oldest endophenotype indicating an increased vulnerability to psychotic disorders (Diefendorf & Dodge, 1908; Gottesman & Gould, 2003). This underlines its importance as a quantifiable biological marker for mental disease.

This Chapter's Learning Objectives
You will learn about

1. The neurophysiological basis of smooth pursuit (SPEM), optokinetic nystagmus (OKN) and vestibular reflex (VOR) and how they interact.
2. Different stimuli for eliciting SPEM, OKN and VOR and what stimuli you need to test distinct oculomotor subfunctions. This knowledge should enable you to design your own experiment.
3. Different approaches to analyzing smooth eye movements and which approach is most appropriate for which stimulus.
4. The basics of models of smooth eye movements that can be used for testing quantitative predictions of neurophysiological mechanisms.
5. The neural circuits subserving smooth eye movements including key findings from functional imaging studies. This knowledge will enable you to develop hypotheses for your own study using brain imaging techniques together with smooth eye movement recordings.
6. Major genes that are thought to be involved in smooth eye movement control.

4.2 Historical Annotations

The evaluation of eye movements has been of interest since the renaissance (Laurentius, 1599). Most easily, the index finger of the examiner is used as target while the patient is instructed to follow the slowly moving finger with his/her eyes as precisely as possible. This procedure is still part of the standard neurological examination in clinical medicine. However, in this situation the evaluation of the quality of pursuit eye movements is based on the examiner's observation and experience. Therefore, techniques have been developed to measure eye movements quantitatively under controlled conditions. To this respect, major milestones in the research of smooth pursuit eye movements have been reached in the last 100 years.

At the beginning of the 20th century, Raymond Dodge, Professor of Psychology at Wesleyan University, US, invented the Dodge microphotograph to record quantifiable eye movements with a sampling rate of 100 Hz while subjects followed a pendulum with their eyes, Fig. 4.1 (Diefendorf & Dodge, 1908). Together with Allan Ross Diefendorf, Lecturer at Yale University, US, he assessed eye movements, especially smooth pursuit, in a range of patients with different psychiatric disorders to examine the biological basis for at that time modern neuropsychiatry. From their results they concluded that mental diseases like schizophrenia and mania can be differentiated based on specific patterns of smooth pursuit disturbances. This gave rise to their recommendation that eye movement recordings should become part of the regular laboratory assessment of mentally ill patients (Diefendorf & Dodge, 1908).

About 30 years later, first results from mapping studies of the visual fields in the striate and extrastriate cortex in monkeys were reported (Talbot & Marshall, 1941; Zeki, 1969), that led to Komatsu and Wurtz publishing their critical findings on the anatomical and functional properties of visual motion processing areas MT (middle temporal area) and MST (middle superior temporal area, Komatsu & Wurtz, 1988a, 1988b). In 1961, Rashbass introduced step-ramp stimuli to study the interaction between smooth pursuit (Rashbass, 1961) and saccadic eye movements which formed the basis for Carl and Gellman's work on the analysis of latency and acceleration during pursuit initiation in 1987 (Carl & Gellman, 1987). In 1965, Robinson published the first version of his key model on the pursuit system which mainly addresses pursuit generation to unpredictable target movement (Robinson, 1965).

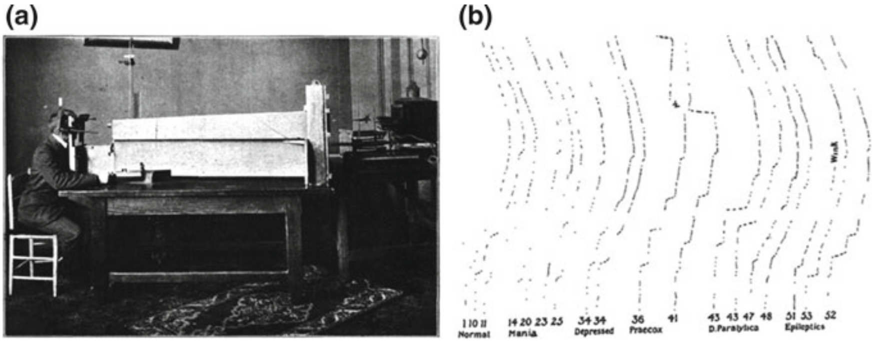


Fig. 4.1 The Dodge photochronograph. **a** While subjects followed a moving pendulum with their eyes, a light beam was projected through a hole onto their cornea unseen in the blind spot of the viewing eye. The reflection of the light was recorded at 100 Hz on a slowly moving photographic plate mounted within the photomicrograph. **b** Pursuit recording examples from healthy participants and patients with a range of different neuropsychiatric disease. Recordings are read from bottom up. Dashes represent flashes of light succeeding each other at 0.01 s. Fixations are indicated by vertical lines, horizontal steps to the right represent saccades occurring during pursuit, reprint by permission of Oxford University Press

In 1973, Holzman et al. rediscovered the findings of Diefendorf and Dodge from 1908 on specific smooth pursuit deficits being observed in patients with severe mental disorders, e.g. schizophrenia. He reported smooth pursuit dysfunction, i.e. eye tracking dysfunction (ETD), not only in patients but also their relatives suggesting that these deficits could indicate a genetic susceptibility to the disorder (Holzman, Proctor & Hughes, 1973). He further hypothesized that ETD is due to impaired motion signal integration in patients and their relatives.

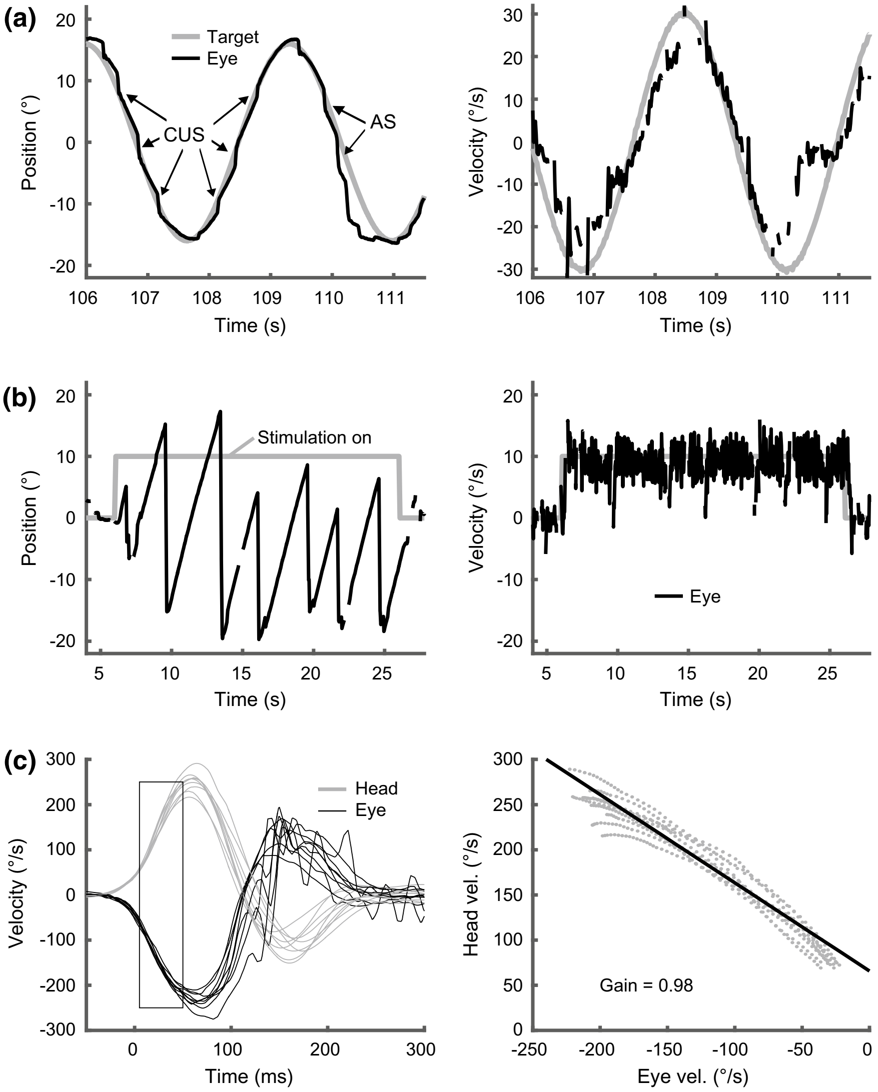
High precision recording of eye movements by scleral search coils (Robinson, 1963) resulted in increased knowledge about different components of eye movements which remained a domain of few university institutes. The developments in non-invasive high speed video based eye tracking from 1990 on enabled eye tracking studies on all aspects of smooth eye movements: even in MRI, MEG or during stance or walking (Holmquist, Andersson, Dewhurst, Jarodska, & van de Weijer, 2012).

During the last decade of the 20th century, first studies using novel brain imaging techniques such as functional magnetic resonance imaging (fMRI, Petit, Clark, Ingeholm & Haxby, 1997) and positron emission tomography (PET, O'Driscoll et al., 1998) were able to confirm and extend the knowledge from monkey studies and brain lesions studies in humans on the functional neural network for smooth eye movements in humans (for a review see Lencer & Trillenberg, 2008).

4.3 Neurophysiological Characteristics of Smooth Eye Movements

4.3.1 Smooth Pursuit—Keeping Small Moving Objects on the Fovea

The purpose of smooth pursuit may be best described as keeping the retinal image of a moving small object within the foveal area in order to enable optimal visual acuity during movement as opposed to fixating a stationary object (Leigh & Zee, 2015). As pointed out in other chapters of this volume (see also chapters Alexander & Martinez-Conde; Greenlee & Kimmig in this volume), the fovea represents the retinal area with the highest receptor density and the smallest receptive fields together enabling high visual resolution. A visual angle of 1° is mapped onto this area. Visual acuity decreases whenever the retinal image moves within or out of this area. The retinal slip velocity, i.e. the velocity of the moving object image on the retina (not the retinal position error as with saccades), represents the sensory input signal for pursuit. As a consequence, smooth eye movements will occur to compensate for the dislocation of the retinal image from the centre of the fovea. Ideally, smooth pursuit velocity matches the velocity of the moving object (see, Fig. 4.2a). In this the ratio of smooth pursuit eye velocity and object (i.e. target) velocity (the so-called "smooth pursuit gain") attains a value of one. Depending on whether predictable or unpredictable


(a)
Position (°)
20
10
0
-10
-20
Target
Eye
CUS
AS
106
107
108
109
110
111
Time (s)
Velocity (°/s)
30
20
10
0
-10
-20
-30
106
107
108
109
110
111
Time (s)
(b)
Position (°)
20
10
0
-10
-20
Stimulation on
5
10
15
20
25
Time (s)
Velocity (°/s)
20
10
0
-10
-20
Eye
5
10
15
20
25
Time (s)
(c)
Velocity (°/s)
300
200
100
0
-100
-200
-300
Head
Eye
0
100
200
300
Time (ms)
Head vel. (°/s)
300
250
200
150
100
50
0
Gain = 0.98
-250
-200
-150
-100
-50
0
Eye vel. (°/s)

◀**Fig. 4.2** Overview of smooth eye movements. **a** Horizontal sinusoidal smooth pursuit of ±16° amplitude and 0.3 Hz frequency, i.e. peak velocity = 30°/s. Left panel shows eye position data with catch-up (CUS) and anticipatory saccades (AS) occurring physiologically during pursuit. Right panel shows velocity data of the same dataset (saccades and blinks removed). **b** shows eye movements during optokinetic stimulation while black-and-white stripes move constantly in horizontal direction from left to right. The eyes follow the stripes to the right until they reach the edge of the orbit (slow phase of OKN), then a saccade brings the eyes back to the left (quick phase of OKN) to catch a new stripe. This results in so-called look OKN characterized by large amplitudes. Right panel shows eye velocity data of the same trial (again saccades and blinks removed). **c** shows eye movements during the vestibulo-ocular-reflex (VOR) evoked by a head-impulse test (Halmagyi-Test) in the left panel. The subject has to fixate a target in gaze straight ahead while the experimenter quickly moves the head to the right side. In order to keep the eye on the target the subject performs an eye movement with the same velocity as the head velocity but in opposite direction. Right panel shows the relation between eye and head velocity. Positive values on the y-axes indicate movements to the right, negative values indicate movements to the left

target movements are pursued, reliably high gain values have been reported for target speeds up to 100°/s at least for short times, although smooth pursuit performance is optimal for target speeds ranging between 15 and 30°/s (Meyer, Lasker & Robinson, 1985; Ettinger et al., 2003). Due to the dynamic properties of the smooth pursuit system, eye velocity will vary to some extent even during optimal smooth pursuit performance (Robinson, 1965). Note that vision would also fade if the target image would be perfectly fixed onto the fovea thus erasing the retinal stimulus to evoke smooth pursuit (see also chapter by Greenlee & Kimmig in this volume; Ditchburn & Ginsborg, 1952).

The neurophysiological mechanisms driving SPEM during the initiation phase rely on motion perception, which involves the sensory analysis of visual information to obtain retinal slip velocity, i.e. the difference between target and eye velocity. Next, the visual motion signal has to be transformed to an oculomotor command that is then sent to the oculomotor nuclei in the brainstem. The maintenance of pursuit is driven by a combination of retinal feedback about performance quality, i.e. retinal slip velocity and so-called extraretinal mechanisms, which are not directly derived from visual input. These extraretinal mechanisms, e.g. an efference copy of the oculomotor command, anticipation, and prediction, play an important role for pursuit generation and maintenance since simple feedback loops are too slow to enable perfect smooth pursuit. Furthermore, the more accurate the pursuit system matches eye velocity to target velocity the smaller the retinal slip velocity error signal becomes. This is a challenge for the system since a high output, i.e. high velocity, has to be maintained during pursuit despite the small error signal. To solve this problem, higher-order extraretinal or cognitive influences to the pursuit system have been postulated (Barnes, 2008). One extraretinal component is the efference copy, i.e. a copy of the oculomotor command, that codes information about current eye velocity and that is thought to be internally processed together with the retinal error signal (Yasui & Young, 1975). Prediction represents a cognitive component driving SPEM which can be understood as the extrapolation of the current eye velocity to the future whereas the term of anticipation refers more to the expectation of future events, e.g. the onset of a target movement (Kowler, 2011; Sprenger et al., 2011). Further cognitive components modifying SPEM include working memory, visuo-spatial attention as well as visuomotor control needed for coordination of smooth pursuit with other eye movements such as saccades, OKN and the VOR. This has been concluded from pursuit studies that observed activation of brain areas that are otherwise activated by saccades, OKN and VOR but which may not be considered core SPEM areas (Lencer et al., 2004). The relative contributions of retinal and extraretinal components during pursuit generation depend on the extent of experience with the pattern of target motion and the predictability of the stimulus (Becker & Fuchs, 1985; Burke & Barnes, 2008). Imagine situations when a pursuit target is temporarily hidden behind an object but the subject following the target is well aware that the target will reappear. Based on the experience with the target, the subject will predict target velocity from what was learned before the target disappeared and will anticipate its reappearance to continue smooth pursuit.

Pursuit eye movements can also be generated to follow auditory stimuli underlining the multimodal aspect of the smooth pursuit system. Furthermore, previous studies have demonstrated that subjects are capable of tracking their finger or limb in complete darkness by pursuit eye movements (Jordan, 1970; Levine & Lackner, 1979).

When examining smooth pursuit, two different phases, i.e. pursuit initiation and pursuit maintenance, have to be differentiated and analysed separately. **Pursuit initiation** comprises the first 50 to 100 ms after the onset of the pursuit movement. During this phase, pursuit gain is not controlled by internally generated feedback but is driven by visual motion information only; this is the so called “open loop” phase (Rashbass, 1961). Eye velocity and eye acceleration during this phase are the most specific measures for the use of sensory visual motion information by the pursuit system (see Fig. 4.4). Pursuit latency, i.e. the time needed to start the pursuit movement after onset of the target movement, further characterizes pursuit initiation ability. Pursuit latency in healthy subjects can vary from 120 to 180 ms depending on task conditions and whether subjects are trained or not. After another 100–150 ms, the **pursuit maintenance** phase integrating also extraretinal components as described above is established; this is referred to as the “closed loop” phase. Smooth pursuit velocity gain represents the parameter that is most often used to measure pursuit maintenance ability (see Fig. 4.6). Physiologically, small position errors due to reduced eye velocity are corrected by saccades consisting mostly of catch-up saccades, whereas anticipatory saccades will lead the eye ahead of the target, Fig. 4.2a (Keller & Johnsen, 1990; Van Gelder, Lebedev & Tsui, 1997). These saccades have to be detected and removed before smooth pursuit eye velocity gain is determined (see also Box 2).

4.3.2 *Optokinetic Nystagmus—Stabilizing Gaze to Coherently Moving Large Patterns*

The purpose of OKN is to stabilize gaze when large environmental patterns move relative to the eyes, thereby projecting large images moving over the entire retina. OKN comprises of a slow and a fast phase (saccade). Large moving patterns induce a slow eye movement to stabilize the retinal image of the pattern during the slow phase of OKN. When the eye approaches the edge of the orbit, a saccade (the quick phase of OKN) will occur that brings the eye orientation back to the primary position. This results in a new retinal image that will be stabilized by the next slow phase (see Fig. 4.2b). Depending on whether certain items (e.g. a single tree) of the background pattern are fixated or the pattern is viewed passively, look and stare OKN can be differentiated, respectively. Look OKN is characterized by large slow phase amplitudes rather resembling smooth pursuit and low frequency of fast phases. In stare OKN, by contrast, slow phases have smaller amplitudes and fast phases occur with higher frequency (Ter Braak, 1936). Both look OKN and the smooth pursuit system are assumed to rely on widely overlapping neural pathways being volitionally controlled by a cortical neural network, whereas stare OKN is usually characterized as being more reflexive in nature (Holm-Jensen, 1984; Valmaggia, Proudlock & Gottlob, 2005). Depending on which of two conflicting OKN and pursuit stimuli is attended to, an OKN response can be suppressed during smooth pursuit although this suppression might not be absolute (Collewijn & Tamminga, 1984; Lencer & Clarke, 1998). The consequences of a conflict between OKN and smooth pursuit stimuli are explained in Box 1 in more detail.

4.3.3 *Vestibular Ocular Reflex—The Antishake Mechanism for Human Vision*

The VOR responds to a head movement with an eye movement in the opposite direction with the same amplitude such that gaze direction is kept constant. The velocity of the head movement is detected by the vestibular organ: the semicircular canals provide information about angular acceleration and the otolith organs provide information about linear acceleration. This is done for head movements around all possible axes of rotation, with a short latency and potentially high eye velocity and acceleration, Fig. 4.2c. The VOR uses mainly a three-neurons-arc, i.e. a connection via three neurons between efference and afference. Nevertheless, recent research demonstrated that efference copy and further cognitive processes such as prediction are used and become obvious in case of pathological changes in the vestibular system. Changes in gain and direction of VOR can be induced by adaptation (Trillenberg, Shelhamer, Roberts & Zee, 2003; Sprenger et al., 2006; Sprenger, Wojak, Jandl, Hertel & Helmchen, 2014).

4.4 The Challenge—Defining the Right Parameters and Choosing the Right Stimuli and Recording Conditions for Measuring Specific Aspects of Smooth Eye Movements

4.4.1 Smooth Pursuit

In the real world, many potential visual targets appear or disappear, move unpredictably and change their speed and their direction rapidly. The oculomotor reactions to such complex stimuli can hopefully be understood as the result of limited number of archetypic eye movements each elicited by a particular aspect of those complex stimuli. To test a particular type of eye movement in isolation all confounding influences must be minimized. With respect to smooth pursuit, measurements are ideally done in a dark or only dimly lit room (luminance < 1 cd/m^2) with only the pursuit target being illuminated (because otherwise the stationary surround might stimulate fixations). During measurements, at least a chin rest should be used to avoid head movements that could provoke VOR. In some cases, an individual bite bar is used that supports head fixation. Alternatively, a tape stripe can be fixed between the testing subject’s forehead and the side mounting of the chinrest that provides subtle feedback whenever the testing subject moves the head, reminding him or her of keeping the head still. Despite all these efforts, changes in head position of 2° within five minutes have been demonstrated (Richter, Trillenberg, Schweikard & Schlaefer, 2013). With advanced eye tracking devices that use an additional camera, head movements can be registered separately and can be considered in offline eye movement analyses. Some systems calculate gaze positions by using head position data or a combination of pupil and corneal reflex positions. However, whenever head movements are not subject to experimental control every effort should be undertaken to avoid them.

In principle the target movement for a pursuit experiment could have any direction, most frequently a horizontal target movements is studied. Pursuit tasks are designed to measure different aspects of pursuit: (1) the initiation of the pursuit movement, (2) pursuit maintenance and (3) cognitive components such as prediction and anticipation using knowledge of predictable target movement. These functions can be studied with corresponding stimuli independently from each other. Basically, three types of stimuli are used: (1) single ramp or step-ramp tasks with intervals of constant target speeds, (2) continuous movements such as triangle waves, sinusoids or oscillating tasks, Fig. 4.3, and (3) tasks in which the visible target is blanked for a variable duration during an ongoing movement (see Fig. 4.8).

4.4.1.1 Pursuit Initiation—The Window into Visual Motion Processing

Pursuit initiation comprises the first 100 ms of pursuit that do not rely on visual feedback of the target movement (“open loop”). For the study of pursuit initiation,

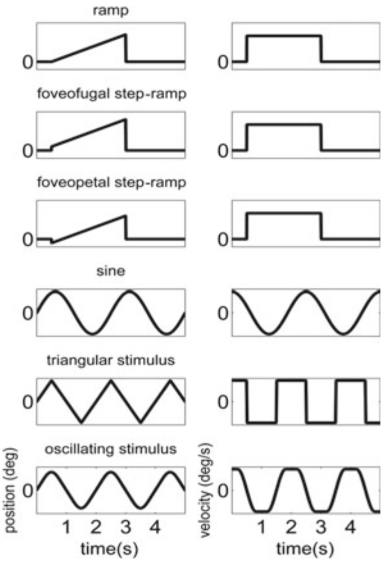


Fig. 4.3 Common stimuli for pursuit experiments (from top to bottom): **Ramp stimulus:** target starts at the center of the screen and moves with constant velocity (in this case to the right). **Foveofugal step-ramp:** a jump is followed by a smooth movement in the same direction. **Foveo-petal step-ramp:** a jump is followed by a smooth movement in the opposite direction. **Sine stimulus:** the target position varies sinusoidally with time. **Triangular stimulus:** the target moves with constant velocity between maximum eccentricities to the left and to the right, respectively. **Oscillating task:** very abrupt changes of target velocity as seen with the triangular stimulus are smoothed out

ramp or step-ramp tasks are used in which the target starts from rest, most often in primary position, and then moves with a constant velocity in one direction constituting a ramp (Rashbass, 1961). Movement direction and velocity can either vary randomly from trial to trial or can be the same making the task fully predictable. Since initially the eye movement lags behind the target, catch-up saccades will occur which interfere with the examination of the initiation phase. To avoid these catch-up saccades, the classical foveopetal step-ramp stimulus was designed, in which the target movement therefore starts with a "step" in the direction opposite to the intended pursuit movement before the target starts moving smoothly in the intended direction ("ramp", see Figs. 4.3 and 4.4). The step amplitude has to be chosen such that the target crosses midline after the average latency of pursuit onset so that healthy subjects will directly initiate a smooth pursuit movement in the intended direction without performing a catch-up saccade (compare also Fig. 4.6b, c). Note, that the step amplitude depends on target speed: in a 10°/s ramp task with an expected pursuit onset latency of 0.150 s the step amplitude should be of 1.5° in order to cross midline at pursuit onset ($10 * 0.15 = 1.5$). Catch-up saccades will occur only later during the pursuit movement but not during the initiation phase.

Carl and Gellman (1987) proposed the following procedure to determine pursuit initiation latency and initial eye acceleration (see Fig. 4.4; Carl & Gellman, 1987).

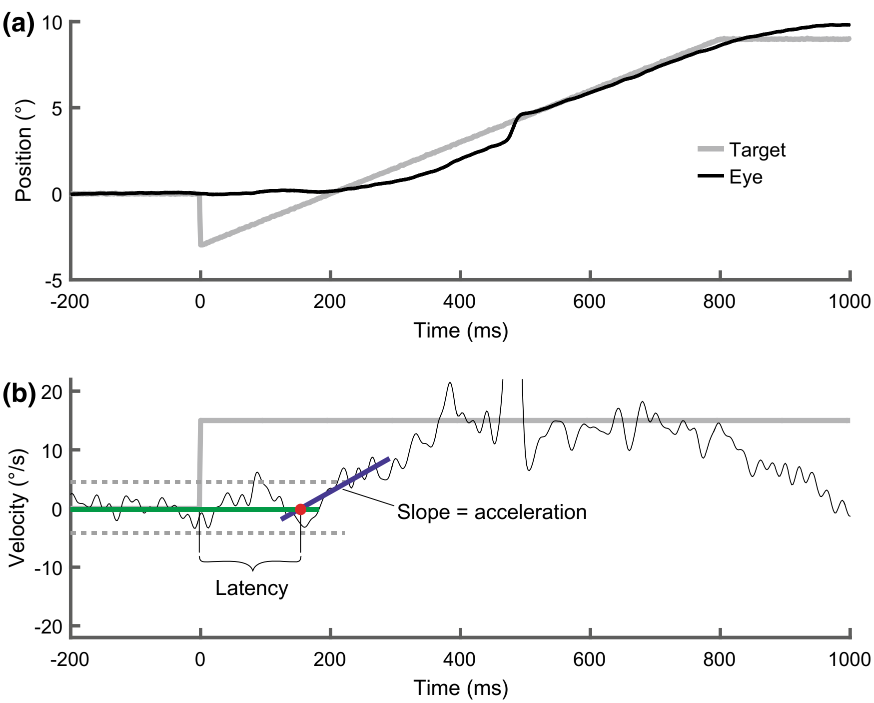


Fig. 4.4 Determining pursuit latency and eye acceleration in foveo-petal step-ramps. **a** shows a horizontal foveo-petal step-ramp pursuit of a healthy subject. The target steps to the left and starts immediately moving at constant velocity to the right. Usually there is no initial catch-up saccade so that initial acceleration can be determined by linear regression of eye velocity (blue line in **b**) in a 60 ms time window after a threshold is exceeded, e.g. 3.2-fold standard deviation of the last 200 ms before trial start (dotted line, see Carl & Gellman, 1987). The intersection (red dot) between the regression line and mean eye velocity before trial start (green line) indicates the start of the pursuit eye movement. The time between trial start and the intersection indicates pursuit latency

The eye position signal is differentiated to obtain eye velocity, and saccades are removed. In a fixed interval (recommendation by Carl and Gellman: 60 ms) after the eye velocity first exceeds 3.2 times the standard deviation of the baseline noise in the velocity signal, a least square fit with a linear function is calculated. The slope of the fit line represents initial eye acceleration. The interval of the intercept of the fit line with the abscissa represents the pursuit latency (Carl & Gellman, 1987). Alternatively, to initial acceleration, mean eye velocity can be determined during the 100 ms of the initial "open loop" pursuit to calculate mean initiation eye velocity gain. Others have defined pursuit latency as the interval between the onset of the ramp target movement and the moment when eye velocity reached 2°/s for at least 20 ms if that preceded the first catch-up saccade (Lencer et al., 2010).

Stimuli with step and ramp in the same direction are called foveo-fugal step-ramp and allow for measuring the saccade position error of the initial catch-up saccade during the pursuit initiation phase (see Figs. 4.3 and 4.5). Eye position error of the

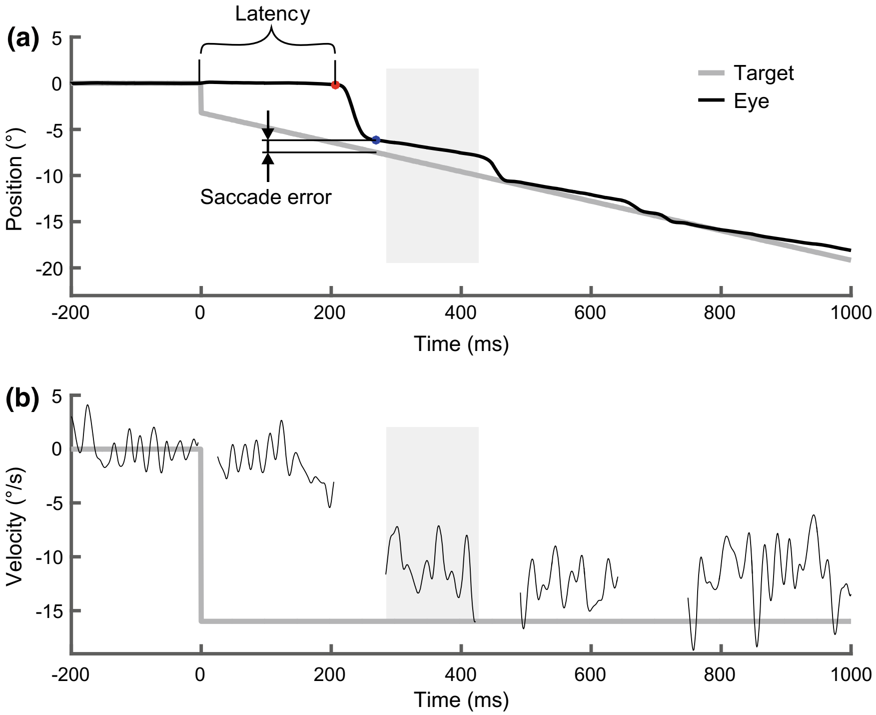


Fig. 4.5 Parameters derived from foveo-fugal step-ramps. The eye position error of the initial catch-up saccade elicited by the target step indicates how correctly the system takes into account target velocity when calculating the landing position of the catch-up saccade. Red and blue dotes indicate start and endpoint of the catch-up saccade, respectively. The lower trace shows eye velocity data of the same dataset (saccades and blinks removed). Gray areas refer to the intervals used for calculating post-saccadic initial gain

initial catch-up saccade after pursuit onset indicates how correctly the system takes into account target velocity when calculating the landing position of the catch-up saccade elicited by the target step. Thus, the magnitude of the eye position error indicates the integrity of visual motion processing during the open loop phase. The advantage of foveo-fugal step-ramp tasks as opposed to single ramp tasks is the fact that the initial catch-up saccade occurs under controlled conditions rather than being provoked by motion processing. That means, in foveo-fugal ramp the initial catch-up saccade is elicited by a target step of a predefined amplitude in the same direction as the following ramp movement. Furthermore, eye velocity in an interval directly after the initial catch-up saccade can be determined to calculate post-saccadic initial gain. This performance measure of a saccade with a subsequent smooth pursuit reflects how precisely the pursuit and the saccadic system are interacting.

4.4.1.2 Pursuit Maintenance—When Higher-Order Cortical Mechanisms Come into Play

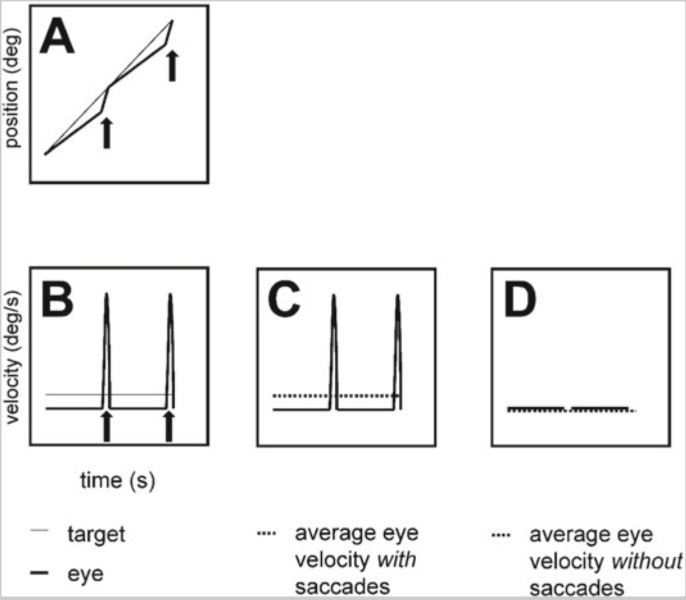
System performance during the maintenance phase when retinal slip velocity is used for feedback control of eye velocity is quantified by pursuit gain (see Sect. 5.1). Besides ramp tasks, most often sinusoids, triangular wave or oscillating tasks are used. Alternatively to gain, the root-mean-square error (RMSE) on eye position in an interval of interest can be determined to characterize global smooth pursuit maintenance. This rather global measure indicates to what extent eye position differs from target position.

Stimuli with a sinusoidal variation of the target position with time (see Fig. 4.3) have the advantage that they can be characterized by a single frequency in Fourier analysis. However, target velocity is constantly changing. Depending on the experiment's goals and the parameters of interest, different variations of sine wave tasks can be designed such as stimuli with varying amplitudes and/or frequencies to make the stimulus less predictable (Barnes, Donnelly & Eason, 1987). The so called Lissajou pursuit uses target traces which are the results of two harmonic sine waves that are projected at right angle to each other thus constructing a 2D target trace (Benson et al., 2012).

In triangular wave stimuli the target moves back and forth with constant velocity (see Fig. 4.3). The term "triangular" has been inspired by the shape of their time-position characteristic. The advantage of triangular stimuli is that target velocity is constant, which facilitates calculation of mean eye velocity gain. Their disadvantage are abrupt reversals of target movement at the end of each ramp. Although abrupt changes may occur under natural viewing conditions, e.g. when watching a tennis match, they represent a challenge for performing smooth pursuit since they require immediate pursuit initiation from zero directly after direction reversal of the target. Thus, abrupt changes induce many saccades around turning points which have to be omitted from analyses when calculating pursuit eye velocity gain. In this respect, sinusoids behave more physiologically since target velocity slows down smoothly before reaching the endpoint of the target trace and eye velocity will increase smoothly again after the turning point. Oscillating targets combine the advantages of triangular wave tasks and sinusoids (see Fig. 4.3; Lencer et al., 2008). They consist of constant velocity ramps in the middle of the trace, but target speed will gradually decelerate to reverse its direction at the endpoint of the trace at which point it will gradually accelerate again until it reaches a constant speed in the opposite direction.

Analysis of pursuit maintenance requires first identification and exclusion of saccades occurring during pursuit. For this, they have to be identified based on their high velocity, e.g. >40°/s, for more details see Box 2.

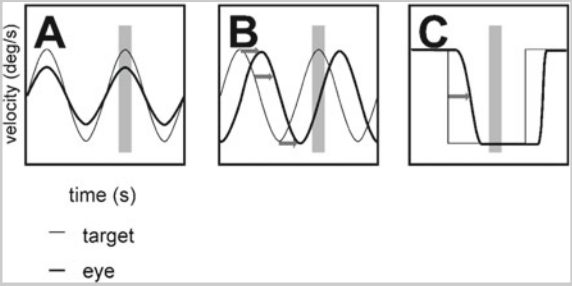
Box 2: Removal of Saccades from Position and Velocity Traces to Calculate Mean Smooth Pursuit Eye Velocity



Panel A shows the position of a target that moves with constant velocity to the right (thin line). The eye (bold line) follows the target but eye velocity is lower than target velocity. The increasing position error is corrected by two catch-up saccades indicated by arrows. Panel B shows the corresponding velocities of the target (thin line) and the eye (bold line). During pursuit eye velocity is lower than target velocity. The two saccades can be identified as episodes with high eye velocity (e.g. >40°/s). For the analysis of this hypothetical recording the average eye velocity in an interval is required. If the full interval is chosen (Panel C) the average eye velocity is identical to the target velocity (compare target velocity in Panel B to average eye velocity with saccades in Panel C). This can be seen from the fact that the eye and the target cover identical angular distances. If however, saccadic episodes are excluded from the analysis average eye velocity without saccades (Panel D) reflects correctly that smooth pursuit eye velocity is lower than target velocity (compare target velocity in Panel B to average eye velocity in Panel D), resulting in a gain value <1.

In sinusoidal stimuli, pursuit performance is characterized by gain and phase angle. The phase angle defines the temporal relationship between eye movement and target movement. Phase angles reflect internal delays in processing afferent and efferent signals. They range between −180° and 180° with positive values implying a phase lag, and negative values indicating a phase lead of the eye velocity. Note that the phase angle should not be calculated for the eye position since a phase lag can simply be caused by a gain reduction of eye velocity without any true phase lag. Gain and phase angle have to be determined simultaneously (at least if the phase angle is not small) because only for a phase angle of zero degrees identical intervals of target and eye movements, respectively, can be used for gain determination. If phase angle is not close to zero analyses can be accomplished by a Fourier analysis or by fitting a sine wave to the eye velocity (for more details, see Box 3). In cases, when the phase angle cannot be determined the ratio between peak eye velocity and peak target velocity can be used.

Box 3: Calculation of Gain: Problems



In Box 2 we showed that saccades interfere with the calculation of smooth pursuit eye velocity and gain. A lag of eye velocity behind a continuously changing target velocity can also interfere with the calculation of smooth pursuit gain. Consider a sinusoidal stimulus (thin line in Panel A). If an idealized smooth pursuit eye movement (bold line in Panel A) matches the temporal pattern of the stimulus exactly the peaks of eye and target velocity are attained simultaneously. However peak eye velocity does not reach peak target velocity. Therefore the smooth pursuit gain can be calculated as the ratio of eye and target velocity in the interval indicated by the grey rectangle.

If internal delays caused by processing of the stimulus cause a shift of the eye velocity with respect to the target velocity (Panel B) the ratio cannot be calculated from a single interval. In the idealized example in Panel B peak eye velocity matches peak target velocity therefore pursuit gain is 1 in contrast to the result obtained by using the interval indicated by the grey rectangle. The situation is easier for a triangular stimulus (Panel C). Even in the presence of phase lag (the reversal of eye movement direction lags behind the reversal of target movement direction as indicated by the grey arrow) an interval can be chosen to determine smooth pursuit gain as a ratio of average eye velocity in a single interval. For sinusoidal stimuli the phase lag of the eye behind the target is a function of stimulus frequency. For low frequency stimuli up to 0.5 Hz Phase lags are negligible and increase to 5° for a stimulus frequency of 1 Hz. Both Fourier transform of the eye velocity and a fit of a sine to the eye velocity (after removal of saccades) yield gains that are not confounded by phase lags and at the same time indicate phase lags if present.

In triangular wave and oscillating stimuli with ramps of constant target velocity, gain can be calculated directly. In most cases, an interval of about ±250 ms around straight ahead (primary position) is defined for calculating mean eye velocity so that artefacts around turning points are excluded from analysis (see Fig. 4.6a). The temporal relationship between target and eye and therefore anticipation of target direction reversal can be considered by measuring the latency of the reversal of eye movement direction with respect to the change of target movement direction (Heide, Koenig, Trillenberg, Kompf & Zee, 1999).

Pursuit maintenance can also be measured in ramp and step-ramp tasks (see Fig. 4.6b–d). With such stimuli, a stable maintenance of eye velocity is reached after 200 ms of pursuit onset, implying that the onset of the interval for determining mean maintenance gain should be after 300 ms of pursuit onset at the earliest. On the other hand, eye velocity will decrease towards the end of ramps due to anticipation of the ramp's ends. This deceleration phase should also be excluded from analysis. Depending on the study design, velocities and directions of ramps can be either identical as in studies investigating predictive mechanisms, or can vary from trial to trial.

Mean maintenance gain should be calculated across trials of same velocity and direction. To date, in healthy subjects, no differences have been reported between mean gains of rightward and leftward movements so that in most studies gain values of left- and rightward trials are merged. Side differences with considerably higher gains for one compared to the opposite direction are observed in, e.g. neurological patients with unilateral lesions of frontal eye fields (Heide, Kurzidim & Kompf, 1996) or patients with autism (Takarae, Minshew, Luna, Krisky & Sweeney, 2004). Generally, maintenance gains for velocities in a range of 10°/s to 30°/s are more stable than with very slow or very fast target movements. Pursuit performance has been shown to be highly reliable over time intervals of up to 52 weeks (see Fig. 4.7, (Ettinger et al., 2003; Lencer et al., 2008)). Ongoing stimuli such as oscillating tasks are performed better (see Fig. 4.7a) compared to ramp or step-ramp tasks of same velocity (see Fig. 4.7b): the predictability of target movement is used by the smooth pursuit system in order to yield high performance. It even helps to compensate for age related decline in smooth pursuit performance (Sprenger et al., 2011).

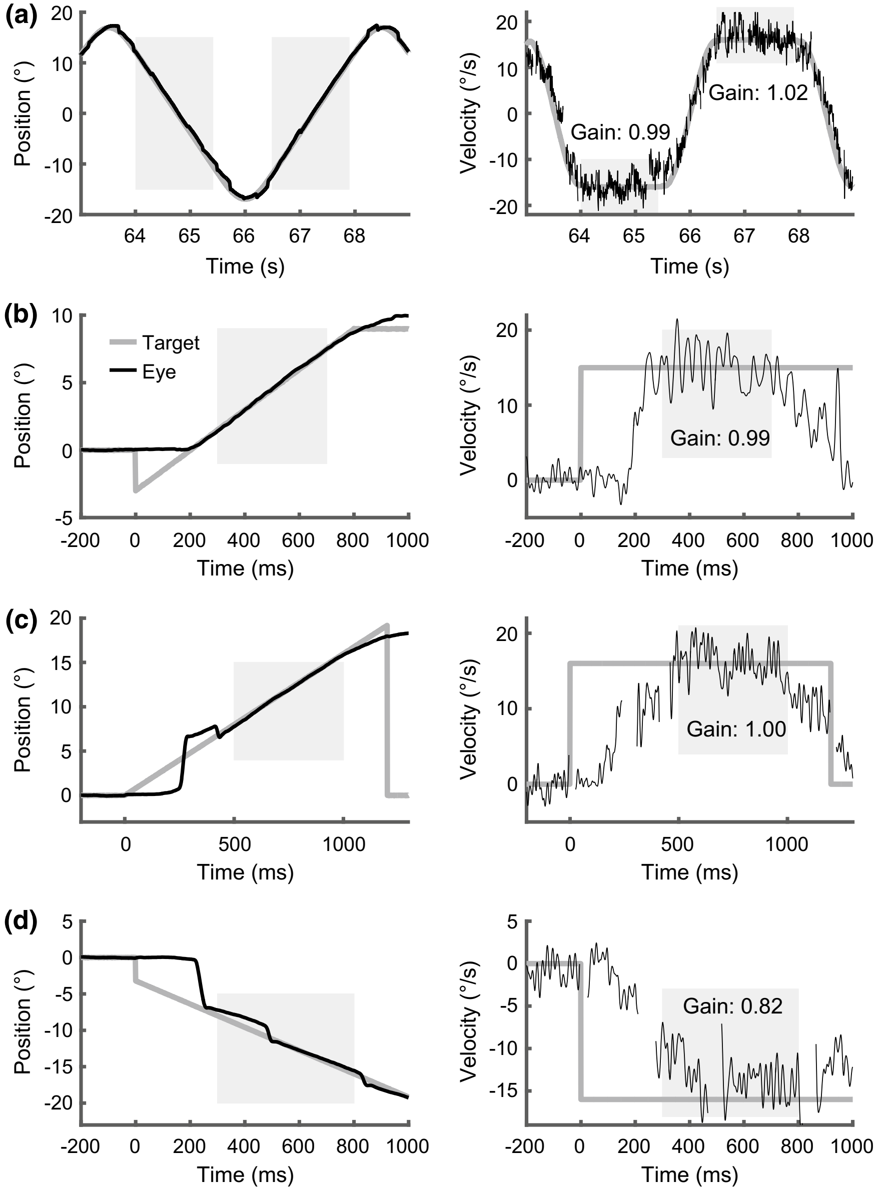


Fig. 4.6 Determining mean maintenance gain in oscillating tasks (**a**), foveo-petal step-ramps (**b**), single ramps (**c**) and foveo-fugal step-ramps (**d**). Grey areas refer to the intervals used for calculating mean eye velocity after saccades were removed from velocity traces. Target velocity in all tasks was 16°/s. For differential effects of target velocity on mean gain depending on stimulus characteristics, see Fig. 4.7

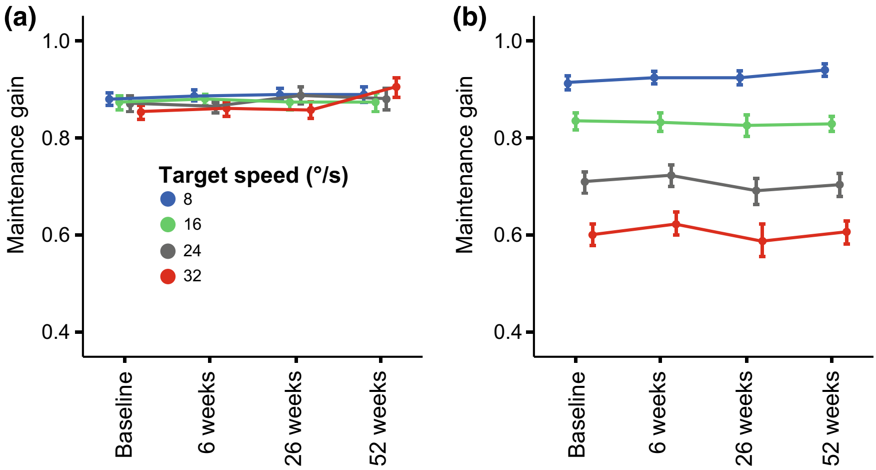


Fig. 4.7 Velocity effects on smooth pursuit maintenance gain. Mean maintenance gain with standard error measured in an oscillating task (**a**) and a ramp task (**b**). Thirty nine healthy volunteers were assessed at baseline and three follow-up sessions showing high temporal stability of pursuit performance. Note that gain decreased with increasing target velocity in ramp tasks but not in oscillating tasks

4.4.1.3 Pursuit Eye Movements During Target Blanking – a Way to Study Extraretinal Mechanisms

One experimental approach for studying extraretinal or cognitive input to pursuit control is to blank the visual target during ongoing pursuit movements (Becker & Fuchs, 1985; Barnes, 2008; Sprenger et al., 2011). Depending on the point in time when a blanking interval occurs during the pursuit movement, different aspects of extraretinal pursuit generation can be studied. First, if the blanking interval is implemented at the beginning of a ramp and the target becomes visible only later in the trial, e.g. after 500 ms (*early blanking* task, see Fig. 4.8), the eyes will begin to move in the direction of the expected target movement prior to target appearance. The velocity of this initial eye drift can be measured reflecting extraretinal anticipatory pursuit drive. Second, if the target is visible during pursuit initiation, is then blanked off and does not reappear until the end of the ramp, any eye movement that persists after a latency of about 300 ms after disappearance of the stimulus can be regarded a consequence of prediction. In this situation, predictive drive is based on the extrapolation of recent pursuit eye velocity to the future. However, after 300 ms eye velocity will decay towards zero (*short-ramp* task, see Fig. 4.8). Eye deceleration can be measured by the slope of the regression line of eye velocity after target extinction. The latency after which the eye starts decelerating after target extinction is indicated by the time interval between time of target extinction and the intercept of the regression line with the eye velocity given by peak gain. It measures how well predictive mechanisms can drive pursuit after visual input was removed from

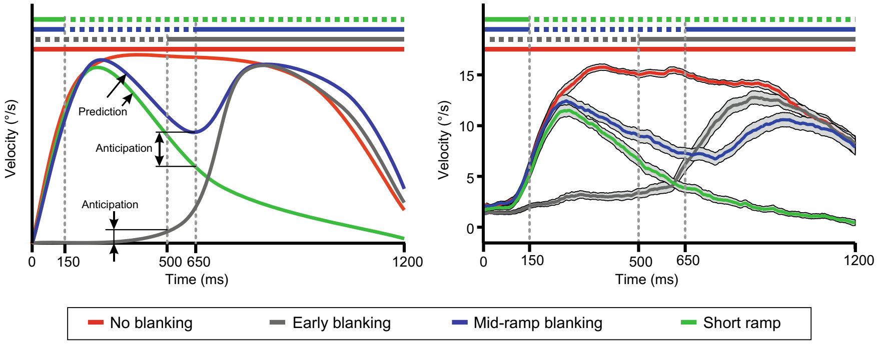


Fig. 4.8 Set of tasks with target blanking to assess different aspects of prediction and anticipation of smooth pursuit. Left panel shows the theoretical scheme of eye velocity traces depending on when the target is blanked during a pursuit movement. The right panel shows grand means with standard error of a group of 24 healthy volunteers naïve to pursuit experiments. Horizontal lines on top indicate whether the target was visible (solid lines) or invisible (dashed lines). The slope of deceleration after target extinction reflects eye velocity driven by prediction from what was learned before. Note, that peak gain is reached after about 200 ms even if the target is extinguished after 150 ms. The difference between eye velocity in mid-blanking tasks when reappearance of the target is expected (blue trace) and eye velocity in the short ramp tasks (green race) indicates the effect of anticipation

the system (Table 1).Third, in tasks with reappearance of the target at the end of the blanking interval (*mid-blank* task, see Fig. 4.8) there is an initial decay of eye velocity similar to *short-ramp* tasks but eye velocity does not decrease to zero. Instead, some non-zero "residual eye velocity" is attained (Becker & Fuchs, 1985), which can be interpreted as a consequence of anticipation driving pursuit. Accordingly, residual gain is defined as mean eye velocity after the end of deceleration related to original target velocity. The magnitude of this residual velocity highly depends on whether subjects are trained and well experienced with this kind of task or not. Well trained subjects are able to reach 80 to 100% of the eye velocity achieved with a visible target (Becker & Fuchs, 1985; Barnes & Collins, 2008) while in untrained, naïve subjects residual eye velocity will decrease to about 20% of the eye velocity achieved with a visible target (Sprenger et al., 2011).

To fully study the different components of cognitive contributions that drive pursuit, a set of tasks including *early blanking*, *short-ramp* and *mid-blank* tasks is necessary so that critical parameters, e.g. initial eye acceleration, can be compared between tasks (see Table 1). Investigations of the effects of aging on retinal and extraretinal contributions to pursuit drive have shown that parameters measured during visually driven pursuit, e.g. mean gain, decline with increasing age whereas parameters indicating cognitive input such as anticipation and prediction of target motion do not change with age (Sprenger et al., 2011). These findings suggest that the elderly maintain their capability of using cognitive input in the immediate control of motor behaviour, which might be a way of compensating for age-related delays in sensorimotor transformation.

Table 1 Parameters to define the integrity of smooth pursuit performance

Parameter	Stimulus	Interpretation
Pursuit initiation (first 100 ms)		
Pursuit latency	Foveopetal step-ramp or ramp	Time needed to compute eye velocity from target velocity
Initial acceleration during first 70–150 ms	Foveopetal step-ramp or ramp	Most direct measure of visual information driving pursuit eye velocity without feedback control
Mean initial eye velocity gain during first 100 ms	Foveopetal step-ramp or ramp	Indicates visual information processing during pursuit initiation without visual feedback
Eye position error of first catch-up saccade	Foveofugal step-ramp	Integrity of visual motion processing for saccade planning during ongoing pursuit
Pursuit maintenance		
Mean eye velocity gain in an interval of at least 100 ms starting >300 ms after pursuit onset	Step-ramp and ramp tasks, triangular wave and oscillating tasks	System performance under feedback control using retinal slip velocity
Latency of direction reversal	Triangular wave	Anticipation of target direction reversal
Mean eye velocity gain under consideration of phase angle	Sine wave	System performance under feedback control using retinal slip velocity and anticipation of target direction reversal
Peak eye velocity gain	Sine wave (when phase angle is not available)	Optimal system performance under feedback control using retinal slip velocity
Root mean square error (RMSE)	Triangular, oscillating and sine wave tasks	Global measure of pursuit accuracy
Catch-up saccade frequency during pursuit	Triangular, oscillating and sine wave tasks	Compensating for low eye velocity when eyes lack behind the target and a certain position error threshold is reached
Anticipatory saccades during pursuit	Triangular, oscillating and sine wave tasks	Leading the eye ahead of the target, can be followed by a back-up saccade

(continued)

Table 1 (continued)

Parameter	Stimulus	Interpretation
Square wave jerks during pursuit	Triangular, oscillating and sine wave tasks	Pairs of saccades with same amplitude but in opposite direction, between saccades the target is fixated parafoveally so that pursuit can be continued without a significant velocity reduction
Pursuit responses to target blanking		
Initial drift (average eye velocity prior to target appearance)	Early blanking tasks	Reflects anticipatory mechanisms driving pursuit when the target is yet not visible
Deceleration and deceleration latency after target extinction	Short-ramp and mid-ramp blanking tasks	Indicates how long and to what extent extraretinal mechanisms such as prediction can drive pursuit after removal of visual motion input
Residual gain after end of deceleration phase	Mid-ramp blanking tasks	Indicates predictive and anticipatory mechanisms driving pursuit when reappearance of the target is expected

As an alternative to tasks with target blanking, the target can change its movement abruptly after a longer period of predictable movement (van den Berg, 1988). In this case, the eye movement is a mixture of the reaction to the actual target movement and the prediction from the previous movement. Technically ambitious is the stabilization of the target on the retina (van den Berg & Collewijn, 1987). For that purpose, the eye position is measured and the target position is adjusted online. Thus, the retinal slip velocity is nullified and any change of eye velocity is a consequence of internal actions including prediction.

4.4.2 *Optokinetic Nystagmus*

Since optokinetic nystagmus is elicited by coherent motion of the whole visual surround its stimulation requires projection of large patterns on a screen (see Fig. 4.2). Different velocities could occur in different parts of the screen if the surface of the projection screen is not perpendicular to the projection axis in the peripheral parts of the stimulus. Therefore, a curved projection screen can be necessary. As an alternative, the stimulation can be performed with a headset (if the recording system can still be applied in this situation). In order to avoid contributions of pursuit, the projected pattern should be composed of large objects without details that stimulate the subjects to focus onto. The optokinetic response comprises an early (immediately beginning) component and a delayed component (with a larger latency). The delayed component persists after switching off the visual stimulation in complete darkness ("OKAN", optokinetic after-nystagmus) and can be studied in isolation under this condition.

4.4.3 Vestibular Ocular Reflex

There are several ways to test the peripheral functions of the VOR: (1) irrigation of the vestibular organ by cold and warm water or air. Cold irrigation induces nystagmus with slow phase velocities in the direction towards the treated ear whereas warm water evokes slow phase velocities in the direction towards the untreated ear. (2) Semicircular canals can be tested separately by quick rotations of the head in the corresponding plane while the subject foveates a point in space. (3) Whole body rotations with sudden stops at velocity saturation elicit nystagmus as well. (4) Active head shaking may help to detect unilateral labyrinth lesions.

4.5 Models of the Smooth Eye Movements

4.5.1 What Models Can Tell About Physiological Systems

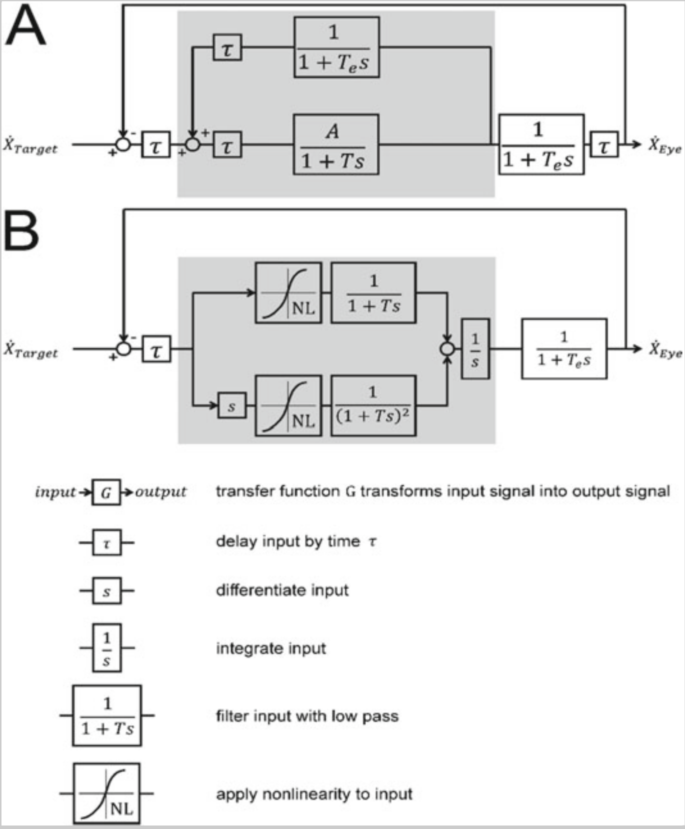
In general, a model is a computational reproduction of a physiological system. The creation of a model can be an important step towards understanding physiology. A model requires the specification of the input and output parameters of a system and of any calculation that transforms input into output. Parameters that influence the transform have to be specified. Ideally, a model can make quantitative predictions that can be tested in experiments.

On the other hand, a model reflects only calculations that are necessary in a physiological system and does not make a statement whether these calculations are carried out in specific brain regions. If a large number of parameters is incorporated in a model any experimental data can be met. The agreement of experiment and simulation then is not a very strong argument for the correctness of the model. For example, it has been shown that the two main models for smooth pursuit lead to almost identical results in simulating the response to a step-ramp stimulus although they have a completely different architecture (Krauzlis & Lisberger, 1994). Experimental data often do not provide unambiguous constraints for model elements: For example, it can be assumed that eye velocity during smooth pursuit does not exceed 100°/s (Meyer, Lasker & Robinson, 1985). Thus eye velocity as a function of stimulus velocity saturates at this value, which implies that a model of pursuit is necessarily nonlinear.

However, simulations show that introduction of a velocity saturation element leads to acceleration saturation in effect (Buizza & Ramat, 2005). Thus, from experimental data it is difficult to decide whether a model should have an acceleration saturation element or a velocity saturation element or both.

Models can rely on control systems theory and combine various elements to a transfer function that determines the output for a given input. An adaptive reaction to the input can be realized with neural networks. A “neuron” in a neural network is a variable. The value of this variable is calculated as the weighted sum of the values of other “neurons” (variables) in the network. The weights in the sum are changed according to a rule that makes sure that weights that lead to unwanted output are diminished and weights that lead to a desired output are enhanced. For a review on modelling in the oculomotor system see Glasauer (2007).

Box 4: Control Systems Models of the Smooth Pursuit



A: simplified version of the Robinson model (Robinson, Gordon & Gordon, 1986). B: simplified version of the Lisberger model (Krauzlis & Lisberger, 1989). Both models differ in the elements in the shaded area. The models have in common an outer feedback loop that guarantees that only retinal slip velocity is available as input. In the Robinson model (A) a second feedback loop adds the eye velocity as sent as a command to the eye muscles to the retinal slip velocity thus restoring target velocity. The inner low pass element (A/(1 + Ts)) effectively adjusts the output of the system to target velocity. In the final version as presented by Robinson (Robinson et al., 1986), this low pass element is replaced by a more complicated element that can emulate saturation of velocity and acceleration. In the Lisberger model (B) there is no internal representation of target velocity. As an alternative, the derivative of retinal slip is used to generate a second contribution to the system output. The figure is simplified by showing only one branch for retinal slip acceleration. In the original version there are two branches with different nonlinearities. C: glossary of typical elements of control systems models. The output of delay element is the input signal shifted in time. For s and 1/s the input is differentiated or integrated, respectively, to obtain the output. A low pass element 1/(1 + Ts) attenuates components of the input with frequency $\gg$ 1/T and leaves components with frequency $\ll$ 1/T unchanged.

4.5.2 Control Systems Models of Smooth Pursuit—Calculate the Eye Command from Moment to Moment

The key models for understanding the pursuit system have been suggested by Robinson (Robinson et al., 1986) and by Krauzlis and Lisberger (1989). Both models address pursuit of an unpredictably moving target (for more details, see Box 4). Since information about the target movement can enter the brain only via the retina, only target movement relative to the movement of the eye can be the input to the pursuit system. In control systems models this is reflected by a trivial outer feedback loop to guarantee that only retinal slip velocity (the difference of target velocity and eye velocity) enters processing by the models. All models in Box 4 have this feedback loop in common.

The main problem in pursuit of non-predictable targets is a combination of internal delays and high gains that render control systems unstable. Ideally, retinal slip velocity is small if eye velocity matches target velocity. This small input has to generate a high output (eye velocity). Therefore, the system has to operate at a high gain. Internal delays arise due to computational effort to define the input (note that the velocity of an image of the retina has to be inferred) and to create the desired output. The Robinson and the Lisberger models differ in the way they try to solve this problem.

Robinson refined an idea by Young and colleagues that a copy ("efference copy") of the command to the eye is added to retinal slip velocity which ideally (e.g. if delays can be ignored) reconstructs target velocity. If this signal is used as the input to the system a gain close to unity is required (Robinson et al., 1986; Young, Forster & van Houtte, 1968). Alternatively, it has been suggested to use the acceleration of retinal slip in addition to retinal slip itself an input signal (Churchland & Lisberger, 2001). This would stabilize eye velocity at target velocity, because if the eye velocity should decrease this would create an acceleration of retinal slip that would in turn create an additional input to the system.

If a parameter (for example, target velocity or acceleration of retinal slip velocity) is thought to play a role in a model this implies that there are neurons with a firing rate proportional to this particular parameter. Thus, the Robinson model predicts the existence of neurons with firing rate proportional to the target speed whereas the Lisberger model predicts the neurons with retinal slip acceleration. With respect to target velocity, recordings from such neurons have in fact been made and therefore provide evidence for the existence of such neurons (Thier & Ilg, 2005).

4.5.3 State Space Models of Smooth Pursuit—Calculate the Eye Command Based on the Whole Pattern

A detailed account of models for pursuit of predictably moving targets is beyond the scope of this chapter. We will outline the general idea and refer the reader to the references cited for details.

Experiments show that targets with a predictable path can be pursued with higher gains and smaller delays than targets with unpredictable paths (van den Berg, 1988). Therefore, models of pursuit should comprise an option to represent the expected motion of the target. This is very difficult for control systems models but much easier for "state space models". In these models, a number of variables (position, velocity, higher derivatives if necessary) characterizes the "state" of the target. A small number of parameters that enter the equations of the model controls the variation of these variables with time. The values for the parameters are continuously adjusted such that the observed behaviour of the target (characterized by the state variables) matches the behaviour that is predicted by the model equations. Then—as long as the target still follows the pattern that is represented by the set of parameters—the model can anticipate the movement of the target and thus compensate for internal delays. As an example the adjustment of parameters can be implemented with a neural network (Shibata, Tabata, Schaal & Kawato, 2005). This model could predict target position during pursuit of a sinusoidal target movement well before a full cycle had been completed but oscillations that typically occur during pursuit initiation were not obtained with the correct frequency.

4.5.4 Models of Optokinetic Nystagmus

Although the stimulus for optokinetic nystagmus has a constant velocity, experiments showed that the eye movements vary from slow phase to slow phase and from fast phase to fast phase (Waddington & Harris, 2012). Thus statistical models are required to explain random fluctuations of the relative timing of slow phases and fast phases (see Sect. 5.2) and statistical properties of fast phase timing (Anastasio, 1997; Balaban & Ariel, 1992). As a simple approach, some rate parameter can be assumed to be integrated over time until a threshold is reached to elicit a fast phase. Different assumptions on variations of the threshold or variations of the rate result in different distributions for fast phase intervals. Unfortunately, it turns out to be very difficult to differentiate between these distributions experimentally (Trillenberg, Zee & Shelhamer, 2002), even more if larger numbers of parameters are introduced (Waddington & Harris, 2013).

4.5.5 Vestibular Ocular Reflex Models

Three aspects of VOR have been in the focus of modelling. First, three pairs of semicircular canals serve as input for the VOR. From their different inputs the correct rotation direction of the eye has to be calculated. Thus, a three dimensional input has to be transformed into a three dimensional output. This aspect has been addressed early on (Robinson, 1975) and must be included if more complicated phenomena like anticipation or eye-head coordination are modelled. Second, the anticipation of signals expected from self-generated head movements has been addressed (King, 2013) and third, the interaction with saccades and smooth pursuit has been described with models (Daemi & Crawford, 2015). For details we refer to Glasauer (2007).

4.6 Neural Networks for Smooth Eye Movements

4.6.1 Smooth Pursuit

Based on the knowledge from neurophysiological studies in non-human primates and behavioural observations in neurological patients with distinct cortical lesions (Ilg & Thier, 2008; Sharpe, 2008), hypothetical schemes for neural networks involved in the generation of smooth pursuit eye movements in humans have been developed (see Fig. 4.9). In the last decades, advances in technology have opened oculomotor research to functional brain imaging. Online eye movement recording during MRI-scanning not only allows for monitoring task performance, it also provides an insight into how behaviour is associated with specific brain activation when quantitative smooth pursuit data is integrated into the model for imaging data analysis.

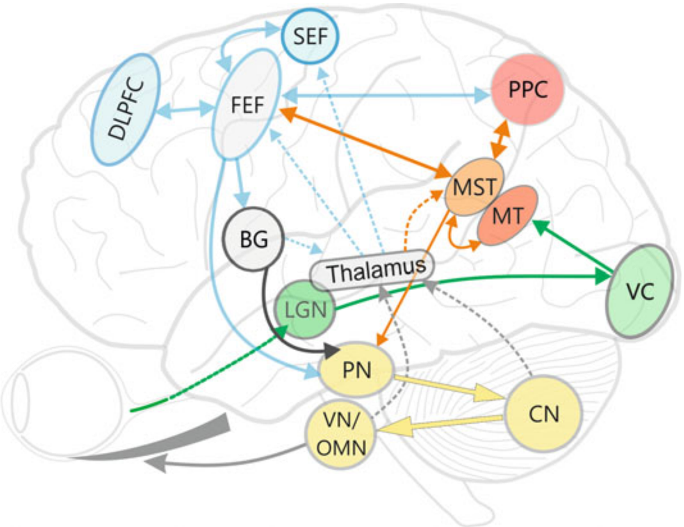


Fig. 4.9 Hypothetical scheme for the smooth pursuit network. Retinal image signals reach the primary visual cortex (VC) via the lateral geniculate nucleus (LGN). VC projects to extrastriatal visual area V5 consisting of the middle temporal visual area (MT) and the medial superior temporal visual area (MST). From there, signals are sent directly or via the posterior parietal cortex (PPC) to frontal areas, one of these being the frontal eye field (FEF) where the oculomotor command for smooth pursuit is generated. Further frontal smooth pursuit regions in humans involve the supplementary eye field (SEF) and dorsolateral prefrontal cortex (DLPFC) which together are involved in higher cognitive control of pursuit movements. Extrastriatal and frontal visual areas project directly and via the basal ganglia (BG) to the pontine nuclei (PN) encoding both eye movement and visual information signals. From there, information is sent to the cerebellar nuclei (CN), specifically the paraflocculus, the flocculus and the vermis. Their output is then projected via the vestibular nuclei (VN) to the oculomotor nuclei (OMN). Feedback projections are assumed to link subcortical areas via the thalamus to cortical eye fields, e.g. FEF and SEF

Using these novel techniques, the quantitative evaluation of pursuit performance on different pursuit tasks has helped to unravel the complex interactions in sensorimotor and cognitive systems involved in smooth pursuit control not only in healthy individuals but also in neurological and mental disorders, e.g. schizophrenia. While most studies using functional magnetic resonance imaging (fMRI) and positron emission tomography (PET) focussed on saccade related activation maps, some functional imaging studies have revealed activation maps of the entire cortical smooth pursuit network. The comparison of pursuit related activations to those observed during visually-guided saccade tasks showed overlapping networks for both smooth pursuit and saccades. These networks involve the fontal eye fields (FEF), supplementary eye fields (SEF), the dorsolateral prefrontal cortex (DLPFC), the parietal areas located in the intraparietal cortex, the precuneus and extrastriate visual area V5, but also the basal ganglia and cerebellar areas (for a review see Lencer & Trillenberg, 2008). It was pointed out that pursuit-related areas differ from those of saccades in terms of spatial extent and location, being somewhat smaller and lying more inferior or posterior. Some authors found a greater activation of the caudate (O’Driscoll et al., 2000) or the DLPFC (Burke & Barnes, 2008) during pursuit than during saccades. In the following section we will describe the key brain areas involved in pursuit eye movement control.

4.6.1.1 Area V5—The Motion Processing Centre for Pursuit

Retinal image motion information is encoded in retinal cells which are projected through the lateral geniculate nucleus (LGN) to neurons in the primary visual cortex (V1) also referred to as striate cortex. Striate neurons respond to motion signals from moving objects. However, these neurons have small receptive fields, and only a limited range of target directions can be coded. V1 projects directly and indirectly to human visual area V5, an extrastriate area that corresponds to the middle temporal visual area (MT) and the medial superior temporal visual area (MST) described in monkeys. V1 projections to MT are retinotopic and ipsilaterally organized. MT and MST neurons differ with respect to their physiological properties. While MT neurons in monkeys have been shown to be direction-selective and selective to moving targets, MST neurons also respond during target blanking, implying that MST neurons receive extraretinal input from higher-order cortical areas, e.g. the FEF. From monkey studies it has been concluded that MST neurons are also influenced by visual stimuli presented in the region surrounding their receptive field. These neurons play an important role for the interaction of different, even conflicting retinal stimuli. The parallel processing of conflicting information is required in daily life during smooth pursuit on background textures. Furthermore, MST neurons apparently receive vestibular information thus being able to encode moving stimuli in head-centred coordinates (Page & Duffy, 2003).

The human homologue of MT is assumed to be located posterior to the superior temporal sulcus junction of Brodmann areas 19, 37 and 39. Area MST in humans apparently lies adjacent to MT in the occipital-temporal-parietal junction (Dukelow et al., 2001; Ohlendorf, Sprenger, Speck, Haller & Kimmig, 2008). V5 subregions in humans appear to be arranged in a similar manner to those seen in monkeys. MT and MST respond differently to the stimulation of the full field (large MT and MST activation) or ipsilateral peripheral hemifield (small dorsolateral MST activation), and visible pursuit targets (large MT and MST activation) or non-visual pursuit targets (small anterolateral MST activation) (Dukelow et al., 2001). There is evidence that signal intensity in V5 is greater during active smooth pursuit of small targets than during motion perception when passively viewing moving gratings (Barton et al., 1996). Furthermore, neural activity in the V5 complex was demonstrated to be linearly correlated to stimulus speed (Chawla et al., 1999; Nagel, Sprenger, Hohagen, Binkofski & Lencer, 2008) and eye velocity (Nagel et al., 2006). MT neurons in humans are sensitive to speed, acceleration and direction of moving stimuli with a preference for so-called first-order motion stimuli (e.g. moving objects), but second-order motion stimuli relying on contrast or flicker modulations can also be coded (Maunsell & Van Essen, 1983). Neurons in area MST, especially its ventrolateral part, are activated when subjects smoothly pursue small targets, but no activation was found when they looked at large moving patterns while keeping their eyes still (Barton et al., 1996). These findings support the observation from monkey studies that area V5 receives extraretinal signals coding for attentional input, corollary eye movement information and a pursuit command such as the efference copy signal (Barton et al., 1996; Dukelow et al., 2001). In conclusion, these properties demonstrate that area V5 must be regarded as a core region not only for motion perception but also for smooth pursuit control.

4.6.1.2 Parietal Cortex—Core Region for Multimodal Sensorimotor Transformation and Attentional Modulation

V5 neurons project to the intraparietal sulcus (IPS) and other parietal and temporal areas, transferring visual motion information for perceptual analysis. Generally, the posterior parietal cortex (PPC) is involved in shifting attention during eye movement tasks as it is required during smooth pursuit to maintain foveal fixation. The PPC encodes information in gaze-centered, eye-fixed frames of reference which are updated across eye movements to remain accurate and maintain spatial constancy (Ilg, Schumann & Thier, 2004). Depending on the attentional load during attentional visual tracking, activations are seen in the superior parietal lobe, intraparietal sulcus, precuneus, besides activations in FEF and precentral sulcus. In a recent study in humans, the posterior parietal cortex was the only area where a specific dissociation of attention from gaze was found (Ohlendorf, Kimmig, Glauche & Haller, 2007).

Activation in the IPS has been preferentially related to the planning of saccades during saccade tasks coding for ‘preparatory set’ activity, but seems also to be related to the suppression of saccades during smooth pursuit (Nagel et al., 2006). In monkeys, neural activity in the parietal cortex was found to be related to eye position during SPEM (Bremmer, Distler & Hoffmann, 1997). More specifically, a direction-specific pursuit-related activity was observed in the lateral intraparietal area (LIP). In humans, the ventral intraparietal area (VIP) seems to process eye movement information by relating the different components of an entire action to each other in head-centred coordinates, a function that seems to work over different modalities (Bremmer, 2005). Neurons in this area are also believed to process visual motion information for monitoring pursuit responses by coding target velocity during pursuit. Furthermore, these neurons are known to respond during passive viewing of optic flow patterns (Konen & Kastner, 2008; Nagel et al., 2008).

4.6.1.3 Frontal Oculomotor Areas—Generating the Oculomotor Command and Adding Cognitive Input

In the frontal eye fields (FEF), the oculomotor command for smooth pursuit is generated, thereby contributing not only to smooth pursuit maintenance but especially to smooth pursuit initiation and prediction (MacAvoy, Gottlieb & Bruce, 1991;

Fukushima et al., 2002). Results from single cell recordings in monkeys underline the importance of FEF neurons for prediction during pursuit, showing increased firing rates in neurons coding for a preferred direction even if the target was temporarily blanked (Ilg & Thier, 2008). Notably, neurons started firing even before the target started moving underlining the role of the FEF for anticipation.

In humans, activation of the FEF during smooth pursuit was one of the first results yielded by functional imaging studies. Evaluation of specific pursuit related FEF activation revealed a smooth pursuit location deeper along the anterior wall of the precentral sulcus but with less intensity than that seen during saccades (Rosano et al., 2002). Further frontal smooth pursuit regions in humans involve the supplementary eye field (SEF) and the anterior cingulate which are suggested to be particularly involved in immediate motor planning, learning of timing and trajectory, and in prediction (Schmid, Rees, Frith & Barnes, 2001; Lencer et al., 2004). More recently it has been shown by transcranial magnet stimulation during predictive sinusoidal pursuit that the gain of the transformation of predictive signals into motor commands was controlled by frontal pursuit areas (i.e. FEF), whereas SEF-stimulation was suggested to have a distinct effect on target direction reversal (Gagnon, Paus, Grosbras, Pike & O'Driscoll, 2006). The DLPFC together with the parietal eye field (PEF) has been shown to reflect attentive response selection and monitoring aspects (Schmid et al., 2001; Nagel et al., 2006). The DLPFC is also thought to be involved in prediction during pursuit coding working memory information (Schmid et al., 2001; Burke & Barnes, 2008) being up-regulated in response to reduced pursuit gain (Nagel et al., 2006).

4.6.1.4 Brainstem and Cerebellar Regions Involved in Pursuit

Both extrastriatal and frontal visual areas project to the pontine nuclei where the dorsolateral pontine nuclei and the nucleus reticularis tegmenti pontis encode both eye movement and visual information signals. From there, information is sent to the cerebellum, specifically the paraflocculus, the flocculus and the vermis. In general, the cerebellum is involved in the immediate guidance of SPEM modulating both the initiation and maintenance of smooth pursuit (Thier & Ilg, 2005). Cerebellar output is then projected via the vestibular nuclei to the oculomotor nuclei. There are tight reciprocal interconnections between the cerebellum and cortical areas. It has been shown in primates as well as in humans that cerebellar output also projects to the prefrontal cortex and thereby may modulate cognitive functions like planning, working memory and rule-based learning (Middleton & Strick, 2000). Areas within the posterior cerebellar lobe are suggested to support mechanisms of prediction, visual attention and shifting visual attention. Furthermore, activation in the cerebellar lobule VII (Crus II) has been shown to be involved in visuospatial working memory including the cueing of visuospatial attention and the suppression of visual reflexive saccades (Nitschke et al., 2004; Stoodley, Valera & Schmahmann, 2012).

4.6.1.5 Differentiating Cognitive Components of Pursuit

Within the last years, the research interest using imaging techniques has shifted towards the impact of specific cognitive mechanisms, such as prediction and anticipation, on the control of smooth pursuit eye movements in humans. During purely predictive pursuit, smooth pursuit velocity was shown to be positively correlated with the neural activity in the V5 complex underlining its essential role as a core region whereas smooth pursuit velocity was negatively correlated with the activity in the DLPFC implying a compensatory function of working memory processing during smooth pursuit (Nagel et al., 2006). When contrasting predictable to unpredictable smooth pursuit, differential increases were seen in the anterior cingulate and preSEF in predictable conditions. In contrast, during non-predictable conditions, the DLPFC and intraparietal sulcus showed high levels of activity that further increased throughout the condition (Schmid et al., 2001). These changes in neural activity reflected the time course of the short-term learning of eye movements seen behaviourally, and were congruent with learning-related changes that have been reported for other motor paradigms. Another study found higher activity in the SEF and the superior parietal lobule that was associated with predictable but not unpredictable eye movements (Burke & Barnes, 2008). The authors further assume that a higher sustained level of FEF activation is necessary during visually guided unpredictable eye movements as compared to memory-driven predictable ones. In addition, the authors found learning-related activation as reflected by an increase over time during predictable eye movements in visual area V5, the DLPFC, and the cerebellum (Burke & Barnes, 2008).

Using a task with mid-ramp target blanking yielded an increase of neural activation in regions seen activated during retinal pursuit like the posterior parietal cortex, the FEF and cerebellar regions and led to additional activation of the SEF, preSEF, the DLPFC, the premotor cortex, the anterior cingulate and the basal ganglia implying that these regions receive extraretinal input (Lencer et al., 2004; Ding, Powell & Jiang, 2009, see Fig. 4.10). More specifically, during non-visual pursuit neural activity in the FEF, the parieto-insular vestibular cortex and the angular gyrus increased with decreasing smooth pursuit velocity indicating their role for reconstitution and prediction whenever smooth pursuit velocity decreased (Nagel et al., 2006). Investigating the neural correlates of smooth pursuit velocity adjustment during retinal and extraretinal pursuit of unpredictable target ramps by event-related fMRI revealed that within a common oculomotor network involving V5, FEF, SEF and a lateral intraparietal area, neural activity increases as target velocity increases (Nagel et al., 2008). Additional engagement of DLPFC, angular gyrus, parahippocampal gyrus and superior temporal gyrus (STG) was found during extraretinal pursuit, whereas the inferior parietal cortex seemed to be important for task switching and sensorimotor transformation.

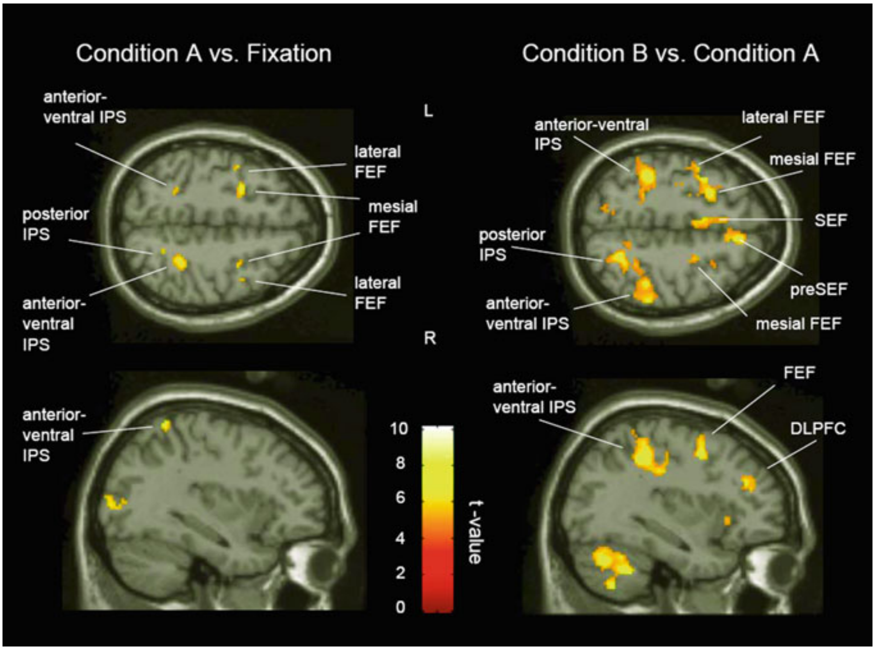


Fig. 4.10 Activation maps from a functional imaging study (1.5T) examining smooth pursuit with continuously visible target movement (condition A), and pursuit with target blanking (condition B–A) obtained by random effects analysis of 16 healthy subjects. Sagittal sections were centred at $x = -34$, and transversal sections at $z = 52$. Thresholds were set at $p = 0.003$ corrected for multiple comparisons. In both contrasts, activation was related to the frontal eye field (mesial and lateral FEF) and the anterior-ventral part of the intraparietal sulcus (IPS) as well as the posterior IPS (right hemisphere). During pursuit with target blanking (contrast B–A), additional activation was found in the supplementary eye field (SEF), the pre-supplementary eye field (preSEF), and the dorsolateral prefrontal cortex (DLPFC) (Lencer et al., 2004)

4.6.2 Interaction of the Pursuit Network with Neural Networks for OKN and VOR

Determining the neural basis of OKN and smooth pursuit by functional imaging yielded a largely overlapping neural circuitry (Konen, Kleiser, Seitz & Bremmer, 2005). For the generation of the slow phase of look OKN, visual information from the retina is projected through V1 and V5 to the nuclei of the optic tract in the pretectum and to oculomotor nuclei. Accordingly, look OKN has been shown to evoke activity in cortical oculomotor centers but stare OKN failed to elicit significant signal changes in these same regions. Furthermore, smooth pursuit evoked a bilateral deactivation of the human equivalent of the parieto-insular vestibular cortex suggesting evidence that the reciprocally inhibitory visual–vestibular interaction involves not only OKN but also smooth pursuit, both being linked with the encoding of object-motion and self-motion. Investigating retinal image motion of stable textured backgrounds during smooth pursuit revealed that temporal regions as well as the posterior parietal cortex (PPC) could represent core regions for the internal referential calculation thereby enabling sustained visual perceptual stability during eye movements (Trenner et al., 2007). The specific role of cerebro-cerebellar circuits for the perceptual cancellation of retinal image motion induced during smooth pursuit was focused in another study (Lindner, Haarmeier, Erb, Grodd & Thier, 2006). Perceptual cancellation mechanisms are assumed to rely on comparing retinal image motion with an internal reference signal predicting the sensory consequences of the eye movement. Notably, neural activity in the lateral cerebellar hemispheres (lobule VI) was found to be correlated with the amount of motion predicted by the reference signal (Lindner et al., 2006). The fact that these cerebellar regions are functionally coupled with the left parieto-insular cortex and the SEF points to these cortical areas as the sites of interaction between predicted and experienced sensory events thus giving rise to the perception of a stable environment.

In contrast to these tight interactions between the direct OKN component and smooth pursuit, the delayed component of OKN seems of minor importance in humans. This OKN component rather relies on the nucleus of the optic tract (pretectum) and a velocity storage mechanism similar to the vestibular system but is not relayed via the cerebellum. The anatomical connections of the VOR are based on results from electrophysiological and tracer studies in animals. Signals from the sensory organ are transmitted via the vestibular nuclei and ascending brainstem connections to the oculomotor neurons and their related eye muscles. Although this pathway has strong connections the vestibular nucleus receives feedback signals such as eye position or cerebellar control signals which enable adaptive learning in case of dysfunctions.

4.7 Possible Genetic Contributions to Pursuit Subfunctions

Interest on the genetic control of smooth pursuit eye movements mostly comes from studies using smooth pursuit as an indicator, i.e. endophenotype, for increased vulnerability to psychotic disorders like schizophrenia (Calkins, Iacono & Ones, 2008). Following the endophenotype concept, a dysfunction of smooth pursuit seems to represent a subtle neurophysiological deficit that is under genetic control and that contributes to the disease mechanism of the disorder (Gottesman & Gould, 2003).

4.7.1 Genes Modulating Dopamine Transmission

Most genetic association studies using smooth pursuit dysfunction as phenotype have focused on polymorphisms involved in dopamine neurotransmission, specifically the gene coding for the catechol-O-methyltransferase (COMT). COMT is involved in dopamine transmission control by degrading dopamine and is highly expressed in prefrontal neural circuitry. Here, a Val158Met polymorphism represented by an exchange of the amino acid valine by methionine at position 158 of the COMT gene is of high interest. This amino acid exchange results in a four times lower degradation of dopamine in the Met variant compared to the Val variant. Therefore, higher dopamine levels in the brain of Met carriers can be expected compared to Val carriers. In healthy subjects, genetic association studies using smooth pursuit have reported higher predictive pursuit gain in carriers of the Met/Met genotype compared to carriers of the Val/Val genotype implying that increased levels of dopamine in prefrontal cortex may enhance predictive pursuit (Thaker, Wonodi, Avila, Hong & Stine, 2004). In the predictive pursuit task the target is intermittently blanked in order to specifically test extraretinal processes involved in SPEM (see Sect. 6.1.3 of this chapter). Thaker's findings are supported by a more recent study using *early blanking* tasks in which the target becomes visible only later in the trial, e.g. after 500 ms so that the eyes will begin to move in direction of the expected target movement prior to target appearance driven by anticipation (Billino, Henning & Gegenfurtner, 2012). In that study, Met/Met carriers showed higher anticipatory eye velocity before the target became visible than Val/Val carriers. Together, these findings suggest that extraretinal SPEM processes may be more sensitive to differences in prefrontal dopamine levels while no association has been found between the Val158Met polymorphism and mean eye velocity during visually guided pursuit (Haraldsson et al., 2009; Billino et al., 2012).

The association between the gene coding for the dopamine transporter 1 (DAT1) and predictive pursuit during target blanking has also been studied (Wonodi et al., 2009). DAT1 is involved in regulating the duration of extracellular dopamine activity. A functional effect of the DAT1 10-repeat variant (10/10 genotype) has been suggested to modulate cortical dopamine resulting in decreased DAT expression and putatively, increased synaptic dopamine levels in carriers of the 10/10 genotype compared to carriers of non-10/10 genotypes, but other reports have demonstrated opposite genotype effects. Healthy carriers of the 10/10 genotype had better predictive smooth pursuit performance than those who belonged to the non-10/10 genotype group (Wonodi et al., 2009). This finding underlies the notion of cortical dopamine transmission being involved in smooth pursuit control.

Other genes involved in dopamine neurotransmission include the gene coding for the dopamine DR3 receptor (DRD3). Here, a polymorphism containing either the amino acids serine or glycine at position 9 (Ser9Gyl polymorphism) seems to be associated with smooth pursuit performance during visually guided pursuit. Carriers of the Ser/Ser variant performed worse than carriers of the Gly/Gly variant, independently whether they belonged to a group of healthy subjects or patients with schizophrenia (Rybakowski, Borkowska, Czerski & Hauser, 2001).

4.7.2 Neuregulin 1 Gene and Smooth Pursuit Performance

Neuregulin 1 (NRG1) is a protein that acts on tyrosine kinase transmembrane receptors in a variety of human tissue including the brain. Alterations of NRG1 will result in disturbed development of the neural system and may thus have an impact on smooth pursuit control. A recent genetic study investigated the association between polymorphisms of the NRG1 gene and smooth pursuit performance in both healthy subjects and patients with schizophrenia. The results are ambiguous showing a tendency for worse performance on several smooth pursuit variables (eye velocity gain and saccade frequency at three different target velocities) in carriers of two NRG1 risk alleles for schizophrenia (the G allele in the single nucleotide polymorphism SNP8NRG222662 and the T allele in SNP8NRG243177) (Haraldsson et al., 2010). These findings were independent of whether subjects belonged to a patient or a control group.

A further genetic association study on the NRG1 gene and smooth pursuit performance included 2,243 young male military conscripts in Greece. Results show that a deficit in global smooth eye pursuit performance measured by the root-mean-square error (RMSE) was related to the T risk allele of SNP8NRG243177 for schizophrenia, whereas increased saccade frequency during pursuit could not be confirmed to be related with the risk allele of SNP8NRG433E1006 (Smyrnis et al., 2011). In all, NRG1 genotype variations were related to smooth eye pursuit variations both at the SNP level and at the haplotype level adding to the validation of this gene as an indicator for increased susceptibility to schizophrenia related disorders. However, another study investigated the NRG1 polymorphism in SNP rs3924999 in 114 Caucasian healthy subjects but could not find an association to visually guided smooth pursuit performance using triangular wave tasks (Schmechtig et al., 2010).

4.8 Future Perspectives

Future studies using advanced techniques, e.g. event-related fMRI, promise to reveal more specific information about distinct mechanisms and neural networks involved in the control of smooth eye movements such as subcortical systems, e.g. basal ganglia, thalamus and cerebellum. Online eye movement recording during MRI-scanning not only allows for monitoring task performance. It also provides an insight into how behaviour is linked to specific brain activation when quantitative eye movement data is integrated into a model for imaging data analysis. Using these novel techniques, the quantitative evaluation of smooth eye movements can help unravel the complex sensorimotor, and cognitive systems involved in their control not only in healthy individuals but also in neurological and mental disorders, e.g. schizophrenia. This may be done in medicated naïve or medication-free patients, independent from potential medication confounds. Further, assessing the impact of psychotropic medication on smooth eye movement systems in previously untreated patients may be informative about drug effects on multiple brain systems. For psychotic disorders, i.e. schizophrenia, schizoaffective and psychotic bipolar disorder genetic linkage studies further indicate that smooth pursuit deficits may represent shared intermediate phenotypes. These findings promise support for uncovering neural network disturbances contributing to disease mechanisms of severe mental disorders. Thus, investigating smooth eye movement performance provides a sophisticated research tool in translational basic neuroscience and neuropsychiatry.

4.9 Suggested Readings

1. A "must" for all who are interested in eye movements: John R. Leigh, David Zee (2015), The Neurology of Eye Movements, 5th edition, New York: Oxford University Press.
2. Comprehensive chapters on smooth pursuit—neurophysiological basis and cognitive processes: Brain & Cognition—Special Issue 68 (2008), Eye Movements in Psychiatric Research, Edited by Christoph Klein & Ulrich Ettinger, specifically see chapters on
   - The neural basis of smooth pursuit eye movements in the rhesus monkey brain (Ilg & Thier, 2008)
   - Neurophysiology and neuroanatomy of smooth pursuit: lesion studies (Sharpe, 2008)
   - Neurophysiology and neuroanatomy of smooth pursuit in humans (Lencer & Trillenberg, 2008)
   - Cognitive processes involved in smooth pursuit eye movements (Barnes, 2008).
3. **Web courses** by Tutis Vilis with excellent visualization and animated models http://www.tutis.ca/courses.htm.

4.10 Questions Students Should Be Able to Answer

1. What are the main purposes of smooth eye movements during (a) smooth pursuit, (b) optokinetic nystagmus and (c) the vestibular ocular reflex?
2. What is the appropriate visual stimulus for smooth pursuit eye movements?
3. What is the range of eye velocity during smooth pursuit eye movements?
4. What would happen if a moving target would be perfectly stabilized on the fovea?
5. How fast can the SPEM system react to changes of target movement?

6. How does the brain react if the stimulus moves too fast for the SPEM system? Try to observe this in a healthy person by presenting a fast stimulus movement with your finger!
7. Suggest criteria that could be used in a computer program to identify saccades during SPEM!
8. Suggest a method to measure the latency of smooth pursuit.
9. What is the definition of "gain" for smooth pursuit eye movements?
10. Describe patterns of target movement that are used to study SPEM!
11. Discuss advantages and disadvantages of these patterns!
12. Explain why a ramp stimulus can be "foveopetal" or "foveofugal.
13. In the "antisaccade paradigm" the subject observes the target to jump to one side and has to make a saccade to the opposite site. Does it make sense to study an "antipursuit paradigm"?
14. Suggest an experiment that proves that nonlinear elements are necessary in pursuit models.
15. What are the core brain areas involved in smooth pursuit control?
16. What is meant by extraretinal factors driving pursuit movements? Why are they essential for smooth pursuit drive?
17. Design an experiment for investigating the predictive components of pursuit.

References

Anastasio, T. J. (1997). A burst-feedback model of fast-phase burst generation during nystagmus. *Biol Cybern Biological Cybernetics, 76,* 139–152.

Balaban, C. D., & Ariel, M. (1992). A beat-to-beat interval generator for optokinetic nystagmus. *Biol. Cybern. Biological Cybernetics, 66,* 203–216.

Barnes, G. R. (2008). Cognitive processes involved in smooth pursuit eye movements. *Brain and Cognition, 68,* 309–326.

Barnes, G. R., & Collins, C. J. (2008). Evidence for a link between the extra-retinal component of random-onset pursuit and the anticipatory pursuit of predictable object motion. *Journal of Neurophysiology, 100,* 1135–1146.

Barnes, G. R., Donnelly, S. F., & Eason, R. D. (1987). Predictive velocity estimation in the pursuit reflex response to pseudo-random and step displacement stimuli in man. *Journal of Physiology, 389,* 111–136.

Barton, J. J., Simpson, T., Kiriakopoulos, E., et al. (1996). Functional MRI of lateral occipitotemporal cortex during pursuit and motion perception. *Annals of Neurology, 40,* 387–398.

Becker, W., & Fuchs, A. F. (1985). Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Experimental Brain Research, 57,* 562–575.

Benson, P. J., Beedie, S. A., Shephard, E., Giegling, I., Rujescu, D., & St Clair, D. (2012). Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. *Biological Psychiatry, 72,* 716–724. https://doi.org/10.1016/j.biopsych.2012.04.019.

Billino, J., Henning, J., & Gegenfurtner, K. R. (2012). Dopaminerge Modulation der okulomotorischen Kontrolle bei Gesunden. *Psychologie und Gehirn* (pp. 155–156). Germany: Jena.

Bremmer, F. (2005). Navigation in space–the role of the macaque ventral intraparietal area. *Journal of Physiology, 566,* 29–35.

Bremmer, F., Distler, C., & Hoffmann, K. P. (1997). Eye position effects in monkey cortex. II. Pursuit- and fixation-related activity in posterior parietal areas LIP and 7A. *Journal of Neurophysiology, 77,* 962–977.

Buizza, A., & Ramat, S. (2005). About the effects of velocity saturation on smooth pursuit. *Annals of the New York Academy of Sciences, 1039,* 459–462.

Burke, M. R., & Barnes, G. R. (2008). Brain and behavior: a task-dependent eye movement study. *Cerebral Cortex, 18,* 126–135.

Calkins, M. E., Iacono, W. G., & Ones, D. S. (2008). Eye movement dysfunction in first-degree relatives of patients with schizophrenia: a meta-analytic evaluation of candidate endophenotypes. *Brain and Cognition, 68,* 436–461.

Carl, J. R., & Gellman, R. S. (1987). Human smooth pursuit: stimulus-dependent responses. *Journal of Neurophysiology, 57,* 1446–1463.

Chawla, D., Buechel, C., Edwards, R., Howseman, A., Josephs, O., Ashburner, J., Friston, K. J. (1999). Speed-dependent responses in V5: A replication study. *Neuroimage, 9,* 508–515.

Churchland, M. M., & Lisberger, S. G. (2001). Experimental and computational analysis of monkey smooth pursuit eye movements. *Journal of Neurophysiology, 86,* 741–759.

Collewijn, H., & Tamminga, E. P. (1984). Human smooth and saccadic eye movements during voluntary pursuit of different target motions on different backgrounds. *Journal of Physiology, 351,* 217–250.

Daemi, M., & Crawford, J. D. (2015). A kinematic model for 3-D head-free gaze-shifts. *Front Comput Neurosci, 9,* 72. https://doi.org/10.3389/fncom.2015.00072.

Diefendorf, A. R., & Dodge, R. (1908). An experimental study of the ocular reactions on the insane from photographic records. *Brain, 31,* 451–489.

Ding, J., Powell, D., & Jiang, Y. (2009). Dissociable frontal controls during visible and memory-guided eye-tracking of moving targets. *Human Brain Mapping, 30,* 3541–3552. https://doi.org/10.1002/hbm.20777.

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature, 170,* 36–37.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT + complex using visual fields and pursuit eye movements. *Journal of Neurophysiology, 86,* 1991–2000.

Ettinger, U., Kumari, V., Crawford, T. J., Davis, R. E., Sharma, T., & Corr, P. J. (2003). Reliability of smooth pursuit, fixation, and saccadic eye movements. *Psychophysiology, 40,* 620–628.

Fukushima, K., Yamanobe, T., Shinmei, Y., Fukushima, J., Kurkin, S., & Peterson, B. W. (2002). Coding of smooth eye movements in three-dimensional space by frontal cortex. *Nature, 419,* 157–162.

Gagnon, D., Paus, T., Grosbras, M. H., Pike, G. B., & O'Driscoll, G. A. (2006). Transcranial magnetic stimulation of frontal oculomotor regions during smooth pursuit. *Journal of Neuroscience, 26,* 458–466.

Glasauer, S. (2007). Current models of the ocular motor system. *Developments in Ophthalmology, 40,* 158–174.

Gottesman, I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry, 160,* 636–645. https://doi.org/10.1176/appi.ajp.160.4.636.

Haraldsson, H. M., Ettinger, U., Magnusdottir, B. B., et al. (2010). Neuregulin-1 genotypes and eye movements in schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience, 260,* 77–85. https://doi.org/10.1007/s00406-009-0032-2.

Haraldsson, H. M., Ettinger, U., Magnusdottir, B. B., Sigmundsson, T., Sigurdsson, E., Ingason, A., Petursson, H. (2009). COMT val(158)met genotype and smooth pursuit eye movements in schizophrenia. *Psychiatry Research, 169,* 173–175. https://doi.org/10.1016/j.psychres.2008.10.003.

Heide, W., Koenig, E., Trillenberg, P., Kompf, D., & Zee, D. S. (1999). Electrooculography: technical standards and applications. The International Federation of Clinical Neurophysiology. *Electroencephalography and Clinical Neurophysiology. Supplement, 52,* 223–240.

Heide, W., Kurzidim, K., & Kompf, D. (1996). Deficits of smooth pursuit eye movements after frontal and parietal lesions. *Brain, 119*(Pt 6), 1951–1969.

Holm-Jensen, S. (1984). The significance of attention and duration of the stimulation in optokinetic nystagmus. *Acta Oto-Laryngologica, 98,* 21–29.

Holmquist KN, M.; Andersson, R.; Dewhurst, R.; Jarodska, H.; van de Weijer J. (2012) Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press London.

Holzman, P. S., Proctor, L. R., & Hughes, D. W. (1973). Eye-tracking patterns in schizophrenia. *Science, 181,* 179–181.

Ilg, U. J., Schumann, S., & Thier, P. (2004). Posterior parietal cortex neurons encode target motion in world-centered coordinates. *Neuron, 43,* 145–151.

Ilg, U. J., & Thier, P. (2008). The neural basis of smooth pursuit eye movements in the rhesus monkey brain. *Brain and Cognition, 68,* 229–240.

Jordan, S. (1970). Ocular pursuit movement as a function of visual and proprioceptive stimulation. *Vision Research, 10,* 775–780.

Keller, E., & Johnsen, S. D. (1990). Velocity prediction in corrective saccades during smooth-pursuit eye movements in monkey. *Experimental Brain Research, 80,* 525–531.

King, W. M. (2013). Getting ahead of oneself: anticipation and the vestibulo-ocular reflex. *Neuroscience, 236,* 210–219. https://doi.org/10.1016/j.neuroscience.2012.12.032.

Komatsu, H., & Wurtz, R. H. (1988a). Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *Journal of Neurophysiology, 60,* 580–603.

Komatsu, H., & Wurtz, R. H. (1988b). Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *Journal of Neurophysiology, 60,* 621–644.

Konen, C. S., & Kastner, S. (2008). Representation of eye movements and stimulus motion in topographically organized areas of human posterior parietal cortex. *Journal of Neuroscience, 28,* 8361–8375.

Konen, C. S., Kleiser, R., Seitz, R. J., & Bremmer, F. (2005). An fMRI study of optokinetic nystagmus and smooth-pursuit eye movements in humans. *Experimental Brain Research, 165,* 203–216.

Kowler, E. (2011). Eye movements: the past 25 years. *Vision Research, 51,* 1457–1483. https://doi.org/10.1016/j.visres.2010.12.014.

Krauzlis, R. J., & Lisberger, S. G. (1989). A control systems model of smooth pursuit eye movements with realistic emergent properties. *Neural Computation, 1,* 116–122.

Krauzlis, R. J., & Lisberger, S. G. (1994). A model of visually-guided smooth pursuit eye movements based on behavioral observations. *Journal of Computational Neuroscience, 1,* 265–283.

Laurentius A (1599) A discourse of the Preservation of the Sight: of Melancholike Diseases; of Rheumes, and of Old Age. Translated by Surphlet, R. (facsimile edition 1938). Oxford University Press, London.

Leigh, R. J., & Zee, D. S. (2015). *The neurology of eye movements*. New York: Oxford University Press.

Lencer, R., Nagel, M., Sprenger, A., Zapf, S., Erdmann, C., Heide, W., Binkofski, F. (2004). Cortical mechanisms of smooth pursuit eye movements with target blanking. *An fMRI study. Eur J Neurosci, 19,* 1430–1436.

Lencer, R., Reilly, J. L., Harris, M. S., Sprenger, A., Keshavan, M. S., & Sweeney, J. A. (2010). Sensorimotor transformation deficits for smooth pursuit in first-episode affective psychoses and schizophrenia. *Biological Psychiatry, 67,* 217–223. https://doi.org/10.1016/j.biopsych.2009.08.005.

Lencer, R., Sprenger, A., Harris, M. S., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2008). Effects of second-generation antipsychotic medication on smooth pursuit performance in antipsychotic-naive schizophrenia. *Archives of General Psychiatry, 65,* 1146–1154.

Lencer, R., & Trillenberg, P. (2008). Neurophysiology and neuroanatomy of smooth pursuit in humans. *Brain and Cognition, 68,* 219–228.

Lencer, R. M., & Clarke, A. H. (1998). Influence of optokinetic and vestibular stimuli on the performance of smooth pursuit eye movements: implications for a clinical test. *Acta Oto-Laryngologica, 118,* 161–169.

Levine, M. S., & Lackner, J. R. (1979). Some sensory and motor factors influencing the control and appreciation of eye and limb position. *Experimental Brain Research, 36,* 275–283.

Lindner, A., Haarmeier, T., Erb, M., Grodd, W., & Thier, P. (2006). Cerebrocerebellar circuits for the perceptual cancellation of eye-movement-induced retinal image motion. *Journal of Cognitive Neuroscience, 18,* 1899–1912.

MacAvoy, M. G., Gottlieb, J. P., & Bruce, C. J. (1991). Smooth-pursuit eye movement representation in the primate frontal eye field. *Cerebral Cortex, 1,* 95–102.

Maunsell, J. H., & Van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology, 49,* 1127–1147.

Meyer, C. H., Lasker, A. G., & Robinson, D. A. (1985). The upper limit of human smooth pursuit velocity. *Vision Research, 25,* 561–563.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and cognitive circuits. *Brain Research. Brain Research Reviews, 31,* 236–250.

Nagel, M., Sprenger, A., Hohagen, F., Binkofski, F., & Lencer, R. (2008). Cortical mechanisms of retinal and extraretinal smooth pursuit eye movements to different target velocities. *Neuroimage, 41,* 483–492.

Nagel, M., Sprenger, A., Zapf, S., et al. (2006). Parametric modulation of cortical activation during smooth pursuit with and without target blanking. an fMRI study. *Neuroimage, 29,* 1319–1325.

Nitschke, M. F., Binkofski, F., Buccino, G., et al. (2004). Activation of cerebellar hemispheres in spatial memorization of saccadic eye movements: an fMRI study. *Human Brain Mapping, 22,* 155–164. https://doi.org/10.1002/hbm.20025.

O'Driscoll, G. A., Strakowski, S. M., Alpert, N. M., Matthysse, S. W., Rauch, S. L., Levy, D. L., Holzman, P. S. (1998). Differences in cerebral activation during smooth pursuit and saccadic eye movements using positron-emission tomography. *Biological Psychiatry, 44,* 685–689.

O'Driscoll, G. A., Wolff, A. L., Benkelfat, C., Florencio, P. S., Lal, S., & Evans, A. C. (2000). Functional neuroanatomy of smooth pursuit and predictive saccades. *NeuroReport, 11,* 1335–1340.

Ohlendorf, S., Kimmig, H., Glauche, V., & Haller, S. (2007). Gaze pursuit, 'attention pursuit' and their effects on cortical activations. *European Journal of Neuroscience, 26,* 2096–2108.

Ohlendorf, S., Sprenger, A., Speck, O., Haller, S., & Kimmig, H. (2008). Optic flow stimuli in and near the visual field centre: a group FMRI study of motion sensitive regions. *PLoS ONE, 3,* e4043.

Page, W. K., & Duffy, C. J. (2003). Heading representation in MST: sensory interactions and population encoding. *Journal of Neurophysiology, 89,* 1994–2013.

Petit, L., Clark, V. P., Ingeholm, J., & Haxby, J. V. (1997). Dissociation of saccade-related and pursuit-related activation in human frontal eye fields as revealed by fMRI. *Journal of Neurophysiology, 77,* 3386–3390.

Rashbass, C. (1961). The relationship between saccadic and smooth tracking eye movements. *Journal of Physiology, 159,* 326–338.

Richter, L., Trillenberg, P., Schweikard, A., & Schlaefer, A. (2013). Stimulus intensity for hand held and robotic transcranial magnetic stimulation. *Brain Stimul, 6,* 315–321. https://doi.org/10.1016/j.brs.2012.06.002.

Robinson, D. A. (1963). A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Transactions on Biomedical Engineering, 10,* 137–145.

Robinson, D. A. (1965). The mechanics of human smooth pursuit eye movement. *Journal of Physiology, 180,* 569–591.

Robinson, D. A. (1975). A quantitative analysis of extraocular muscle cooperation and squint. *Invest Ophthalmol, 14,* 801–825.

Robinson, D. A., Gordon, J. L., & Gordon, S. E. (1986). A model of the smooth pursuit eye movement system. *Biological Cybernetics, 55,* 43–57.

Rosano, C., Krisky, C. M., Welling, J. S., Eddy, W. F., Luna, B., Thulborn, K. R., & Sweeney, J. E. (2002). Pursuit and saccadic eye movement subregions in human frontal eye field: a high-resolution fMRI investigation. *Cerebral Cortex, 12,* 107–115.

Rybakowski, J. K., Borkowska, A., Czerski, P. M., & Hauser, J. (2001). Dopamine D3 receptor (DRD3) gene polymorphism is associated with the intensity of eye movement disturbances in schizophrenic patients and healthy subjects. *Mol Psychiatry, 6,* 718–724.

Schmechtig, A., Vassos, E., Kumari, V., et al. (2010). Association of Neuregulin 1 rs3924999 genotype with antisaccades and smooth pursuit eye movements. *Genes Brain Behav, 9,* 621–627. https://doi.org/10.1111/j.1601-183X.2010.00594.x.

Schmid, A., Rees, G., Frith, C., & Barnes, G. (2001). An fMRI study of anticipation and learning of smooth pursuit eye movements in humans. *NeuroReport, 12,* 1409–1414.

Sharpe, J. A. (2008). Neurophysiology and neuroanatomy of smooth pursuit: lesion studies. *Brain and Cognition, 68,* 241–254.

Shibata, T., Tabata, H., Schaal, S., & Kawato, M. (2005). A model of smooth pursuit in primates based on learning the target dynamics. *Neural Netw, 18,* 213–224.

Smyrnis, N., Kattoulas, E., Stefanis, N. C., Avramopoulos, D., Stefanis, C. N., & Evdokimidis, I. (2011). Schizophrenia-related neuregulin-1 single-nucleotide polymorphisms lead to deficient smooth eye pursuit in a large sample of young men. *Schizophrenia Bulletin, 37,* 822–831. https://doi.org/10.1093/schbul/sbp150.

Sprenger, A., Trillenberg, P., Pohlmann, J., Herold, K., Lencer, R., & Helmchen, C. (2011). The role of prediction and anticipation on age-related effects on smooth pursuit eye movements. *Annals of the New York Academy of Sciences, 1233,* 168–176. https://doi.org/10.1111/j.1749-6632.2011.06114.x.

Sprenger, A., Wojak, J. F., Jandl, N. M., Hertel, S., & Helmchen, C. (2014). Predictive mechanisms improve the vestibulo-ocular reflex in patients with bilateral vestibular failure. *Journal of Neurology, 261,* 628–631. https://doi.org/10.1007/s00415-014-7276-0.

Sprenger, A., Zils, E., Stritzke, G., Kruger, A., Rambold, H., & Helmchen, C. (2006). Do predictive mechanisms improve the angular vestibulo-ocular reflex in vestibular neuritis? *Audiol Neurootol, 11,* 53–58. https://doi.org/10.1159/000088926.

Stoodley, C. J., Valera, E. M., & Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: an fMRI study. *Neuroimage, 59,* 1560–1570. https://doi.org/10.1016/j.neuroimage.2011.08.065.

Takarae, Y., Minshew, N. J., Luna, B., Krisky, C. M., & Sweeney, J. A. (2004). Pursuit eye movement deficits in autism. *Brain, 127,* 2584–2594.

Talbot, S. A., & Marshall, W. H. (1941). Physiological Studies on Neural Mechanisms of Visual Localization and Discrimination*. *American Journal of Ophthalmology American Journal of Ophthalmology, 24,* 1255–1264.

Ter Braak, J. W. G. (1936). Untersuchungen über optokinetischen Nystagmus. *Archives Neerlandaises de Physiologie de l'homme et des animaux, 21,* 309–376.

Thaker, G. K., Wonodi, I., Avila, M. T., Hong, L. E., & Stine, O. C. (2004). Catechol O-methyltransferase polymorphism and eye tracking in schizophrenia: a preliminary report. *American Journal of Psychiatry, 161,* 2320–2322.

Thier, P., & Ilg, U. J. (2005). The neural basis of smooth-pursuit eye movements. *Current Opinion in Neurobiology, 15,* 645–652.

Trenner MU, Fahle M, Fasold O, Heekeren HR, Villringer A, Wenzel R (2007) Human cortical areas involved in sustaining perceptual stability during smooth pursuit eye movements. Hum Brain Mapp.

Trillenberg, P., Shelhamer, M., Roberts, D. C., & Zee, D. S. (2003). Cross-axis adaptation of torsional components in the yaw-axis vestibulo-ocular reflex. *Experimental Brain Research, 148,* 158–165. https://doi.org/10.1007/s00221-002-1285-4.

Trillenberg, P., Zee, D. S., & Shelhamer, M. (2002). On the distribution of fast-phase intervals in optokinetic and vestibular nystagmus. *Biol Cybern Biological Cybernetics, 87,* 68–78.

Valmaggia, C., Proudlock, F., & Gottlob, I. (2005). Look and stare optokinetic nystagmus in healthy subjects and in patients with no measurable binocularity. *A prospective study. Klin Monatsbl Augenheilkd, 222,* 196–201.

van Beers, R. J., Wolpert, D. M., & Haggard, P. (2001). Sensorimotor integration compensates for visual localization errors during smooth pursuit eye movements. *Journal of Neurophysiology, 85,* 1914–1922.

van den Berg, A. V. (1988). Human smooth pursuit during transient perturbations of predictable and unpredictable target movement. *Experimental Brain Research, 72,* 95–108.

van den Berg, A. V., & Collewijn, H. (1987). Voluntary smooth eye movements with foveally stabilized targets. *Experimental Brain Research, 68,* 195–204.

Van Gelder, P., Lebedev, S., & Tsui, W. H. (1997). Peak velocities of visually and nonvisually guided saccades in smooth-pursuit and saccadic tasks. *Experimental Brain Research, 116,* 201–215.

Waddington, J., & Harris, C. M. (2012). Human optokinetic nystagmus: a stochastic analysis. *J Vis, 12,* 5. https://doi.org/10.1167/12.12.5.

Waddington, J., & Harris, C. M. (2013). The distribution of quick phase interval durations in human optokinetic nystagmus. *Exp Brain Res Experimental Brain Research, 224,* 179–187.

Wonodi, I., Hong, L. E., Stine, O. C., et al. (2009). Dopamine transporter polymorphism modulates oculomotor function and DAT1 mRNA expression in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet, 150B,* 282–289.

Yasui S, Young LR (1975) Perceived visual motion as effective stimulus to pursuit eye movement system. Science (New York, N.Y.) 190:906–908.

Young LR, Forster JD, van Houtte N (1968) A revised stochastic sampled data model for eye tracking movements. In: University Conference on Manual Control, vol Fourth Ann NASA. NASA, University of Michigan, Ann Arbor, Michigan.

Zeki, S. M. (1969). Representation of central visual fields in prestriate cortex of monkey. *Brain Research, 14,* 271–291.

Chapter 5
Visual Perception and Eye Movements

Mark W. Greenlee and Hubert Kimmig

Contents

Abstract Our perception of the world appears to be steady and focused, despite the fact that our eyes are constantly moving. In this chapter, we review studies on the neural mechanisms and visual phenomena that endow us with stable visual perception despite frequent eye movements and gaze shifts. We describe how sensitivity to stationary and moving stimuli is suppressed just before and during a saccadic eye movement, a phenomenon referred to as *saccadic suppression*. We also depict

---

M. W. Greenlee (✉)
Institute of Experimental Psychology, University of Regensburg, Regensburg, Germany
e-mail: mark.greenlee@ur.de

H. Kimmig
Schwarzwald-Baar Klinikum, University of Freiburg, Freiburg, Germany

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_5

the neural correlates of saccadic suppression based on studies conducted in alert monkeys during single-unit recordings and in human subjects using functional MRI. In addition to saccadic suppression, the phenomena of *saccadic suppression of displacement* and *saccadic mislocalization* suggest that motion sensitivity and perceived location of objects are altered during eye movements. For example, target motion during saccades goes largely unnoticed. Clearly visible targets that are flashed during an eye movement are apparently displaced towards the location of the end point of the saccade. These findings suggest that visual perception is based on the spatio-temporal integration of information gathered during sequential fixation periods. To ensure stability of our percepts during eye movements, some form of compensation or remapping must take place to compensate for retinal displacements of stimuli in a viewed scene. We also examine the interactions that take place between motion perception and pursuit eye movement. Furthermore, we review findings from behavioral, psychophysical and imaging studies in human observers and single-unit recordings in non-human primates that are relevant to perceptual phenomena arising during saccadic and pursuit eye movements. These studies suggest that conscious vision results from the dynamic interplay between sensory and motor processes. Specifically, we propose that our perception of the visual world is built up over time during consecutive fixations.

5.1 Introduction and Learning Objectives

Foveal vision is one of the outstanding achievements of evolution, beginning with the tiniest primates (e.g., the mouse lemur *Microcebus berthae* weighs only 30 g) and ending with Homo sapiens. In humans, the central rod-free area of our visual field is subserved by 150,000–200,000 photoreceptors densely packed in each squared millimeter with the foveal pit or foveola having the highest receptor density (Curcio, Sloan, Packer, Hendrickson, & Kalina, 1987; Østerberg, 1935). Of the three different cone photoreceptor classes only the middle and long-wavelength sensitive cones are found in this area of 10–20 min arc diameter. Outside this “pit” the parafovea contains a mixture of all three cone types and rods, the packing density (or receptor spacing) of which varies with eccentricity. Cone density peaks in the fovea and falls off rapidly towards 5° of eccentricity, as does visual acuity (Hirsch & Curcio, 1989). Further in the periphery, the number of rods exceeds that of the cones with a maximum at approx. 20° of eccentricity. This distribution has important consequences for primate vision. During daylight we see best in central vision, while objects presented to the peripheral retina are resolved poorly. Often, we only see that there is something, but can only guess what it is. To scrutinize the object, we need to perform a saccade, a fast ballistic movement of the eye from the present focus to the peripheral target (see chapter by Pierce et al. in this volume). Hereafter the object can be closely inspected by the fovea and high-resolution processing can take place at this new location. Vision thus requires the interplay between sensory and motor processes:

visual processing that is centered at the fovea and oculomotor processing that guide the fovea toward the area of interest in the peripheral visual field.

We typically move our eyes three to four times per second. Periods of relative gaze[1] stability during fixation last only 150–250 ms and are interrupted by ballistic shifts of gaze to new locations in the visual field. Even during fixations tiny saccades, referred to as microsaccades and drifts, keep the eye in more or less constant motion (see chapter by Alexander & Martinez-Conde in this volume). Indeed if these micromovements of the eye are suppressed by artificial image stabilization (Ditchburn & Ginsborg, 1952; Riggs, Armington, & Ratliff, 1954; Riggs & Schick, 1968), vision rapidly fades owing to local adaptation at the level of the retina.

In this chapter we will focus on the interplay between visual perception during fixations and eye movements. Perceptual phenomena that are associated with the transitions between fixations and eye movements, such as *saccadic suppression* for flashed targets and *saccadic suppression of spatial displacement* are considered. Conjugate eye movements (i.e., binocular eye movements in the same direction) also take place during the tracking of a moving object and these smooth eye movements are referred to as pursuit (see chapter by Lencer et al. in this volume). Pursuit eye movements are much slower than saccades and are under voluntary control and receive visual feedback. In Sect. 5.3.9 of this chapter, we will examine the interplay between motion perception and pursuit eye movement. Finally we will review studies on the role of attention and memory in the pattern of eye movements observed during visual search and visual memory tasks. These studies point to a close interaction between the cognitive demands of visual tasks and the resulting pattern of eye movements. We conclude that neither visual processing nor oculomotor control should be studied in isolation. Rather *active vision* represents a dynamic interplay between these sensory and motor processes (Findlay & Gilchrist, 2003). Understanding how these sensorimotor mechanisms interact poses a significant challenge to contemporary systems neuroscience.

In this chapter we will focus on the following topics and describe, at the onset, our learning objectives:

(1) The anatomical structure of the human retina poses important constraints of visual perception. We first describe the factors that influence visual acuity and contrast sensitivity in central and peripheral vision. Further limitations in peripheral vision concern the phenomenon known as visual crowding, i.e., the inability to discriminate visual objects viewed in the peripheral visual field owing to the presence of flanking stimuli. Our goal here will be to obtain a better understanding of the anatomical and functional constraints imposed on human vision thereby requiring eye movements.
(2) The decline in contrast sensitivity from foveal to eccentric vision and the incline in crowding effects imply that eye movements represent an essential aspect of

---

[1] Gaze shifts arise from the combination of head and eye movements. In the laboratory the subject's head is often rendered stationary by the use of a chin rest or bite bar. In the real world, saccades exceeding 20° are usually accompanied by a shift in head position in the direction of the eye movement.

everyday vision. Peripheral vision provides the gist for scene analysis, whereas foveal vision provides the requisite acuity for stimulus discrimination and object recognition. Dynamic shifts in gaze are needed to move the fovea to the area of interest within a scene for detailed visual analysis. Our goal here will be to obtain a better understanding of how central and peripheral vision interact to stabilize visual perception.

(3) Fast eye movements induce retinal image shifts. The brain needs to discount these retinal image shifts from true object motion. A corollary discharge arises from the oculomotor command to cancel image processing while the eyes are in motion. These phenomena are known as saccadic suppression and saccadic suppression of stimulus displacement. A better understanding of the brain mechanisms underlying this gating of information flow will be our learning goal.

(4) Saccadic mislocalization occurs when brief stimuli are flashed during the subject's own eye movement. The location of these peri-saccadic flashed stimuli appears to be shifted toward the focus of attention at the saccade target location. These observations imply that the representation of peri-personal space is not constant, but rather is a dynamic process that varies with the focus of attention. Our goal here will be to obtain a better understanding of the underlying neural mechanisms of vision-based spatial representations.

(5) In addition to the above-described suppressive effects, a peri-saccadic shift of a target can be compensated for by an adaptive gain change. Saccadic adaptation represents a gain adjustment that takes place when recurrent undershoots or overshoots occur in saccade tasks. We describe how gain changes take place and which neural structures appear to be involved in this form of adaptation. A better understanding of the underlying neural mechanisms of saccadic adaptation is our learning goal in this section. The visual analysis of object motion is an essential prerequisite for smooth pursuit of a moving target. We describe the results of studies that focused on the interplay between motion perception and pursuit eye movements. Such phenomena like the Filehne illusion, an illusory motion of the background during eye movements, are normally cancelled by a similar corollary discharge as that found for saccades. We will try to obtain more insight into the role of pursuit eye movements in visual stabilization.

(6) The differentiation between self and object motion is a further task that needs to be solved by the brain. Here additional signals from the vestibular systems for head motion and from the oculomotor system for eye movements are compared in cases where the source of retinal motion is ambiguous. We demonstrate that, under certain circumstances, these corollary signals can even enhance motion perception, as is the case for coherent dot motion detection during smooth pursuit eye movements. Our goal is the understand how self and object motion are differentiated by extra-retinal sources of information.

(7) Despite on-going movement of the eyes in the orbit and our head in space, we can accurately locate objects in space. This ability implies the existence of coordinate transformations from retinal to head and space coordinates, leading to a dynamic map of objects in space around the observer. Our goal will be to better understand the neural mechanisms involved in such coordinate transformations.

(8) Visual attention allows us to focus on certain stimuli in a visual scene while ignoring others. These objects of interest can become the target for a saccade. We review studies that explore the close association between shifts in spatial attention and goal-directed saccadic eye movements. Our learning goal is to better understand how cognitive processes like visual/spatial attention interact with motor processes to guide our gaze to specific locations in peri-personal space.

(9) Visual working memory reflects our ability to store and recall information about visual objects to help guide our interaction with these objects. Finally, our goal will be to understand how the limited capacity of visual memory affects how we move our eyes during visual tasks.

5.2 Historical Annotations

Pioneering studies on human vision began with visual psychophysics, the systematic study of the relationship between the physical stimulus and the perceptual report of the observer (Fechner, 1889). Electrophysiological recordings in the visual pathway of cats and monkeys substantiated the psychophysical results (Hubel & Wiesel, 1962, 1968) by providing support for mechanisms in early vision related to stimulus processing. Most of our understanding about the neural basis of visual perception was derived by exact recordings of the sensitivity of single neurons in electrophysiological experiments in primates or by psychophysical observations in the human observer during strict central fixation. By introducing oculographic measures (Dodge, 1900) researchers began to allow their participants to move their eyes during stimulus presentation, thereby promoting the study of the interactions between perception and eye movements. New phenomena were discovered, some of which we report below, concerning the effects of eye movements on visual perception. With the advent of brain imaging techniques, researchers could begin to investigate the neural basis for eye movements and their effects on human perception. A number of excellent reviews on this topic have been published (Bremmer, 2011; Britten, 2008; Findlay & Gilchrist, 2003; Morrone, 2014; Pelisson, Alahyane, Panouilleres, & Tilikete, 2010).

5.3 Psychophysical and Oculomotor Studies of Human Vision

Below we review the most important studies related to our topic of visual perception during eye movements. We begin with psychological studies of central and peripheral vision. We then describe some of the known phenomena related to visual crowding in eccentric vision. When the eyes move, the brain creates a corollary discharge (also known as an efference copy, see below) that can be incorporated into visual processing in the moment before and during an eye movement. Several suppressive effects are described that appear to be related to this corollary discharge. Distortions in spatial representations at the time point of a saccade can be studied by flashing additional stimuli near the saccade target and asking subjects to estimate their location in space. Further studies explore the interplay between pursuit eye movements and motion perception. Finally, we review studies on the role of visual attention and visual working memory on saccadic eye movements and perception.

5.3.1 Contrast Sensitivity of Foveal and Peripheral Vision

Contrast sensitivity reflects the ability of the visual system to resolve spatial variations in the luminance of visual stimuli at very low levels of stimulus contrast. Sinusoidal gratings are often used to test contrast sensitivity, since stimulus energy is restricted to a narrow band of spatial frequency. Contrast sensitivity (the inverse of contrast threshold) is plotted as a function of the spatial frequency in cycles per degree (c/deg). In a typical experiment, the participant fixates a central point on a monitor. On a given trial, a temporal interval is defined (often by auditory tones) and a weak, low contrast stimulus is randomly presented or nothing is presented. The participant's task to be "detect" whether a stimulus was presented or not. Depending on the computer algorithm used, the stimulus contrast is either reduced or increased until a threshold level of performance is achieved. Typically, a performance level of 75% correct is selected. The stimulus contrast level that is associated with this performance level is referred to as the "threshold contrast" for that stimulus. The reciprocal value of the threshold (i.e., 1/threshold) is referred to as "contrast sensitivity". Figure 5.1a demonstrates how contrast sensitivity has a maximum between 3 and 5 c/deg falling off on either side. The contrast sensitivity function of the human visual system with foveal viewing is shown in Fig. 5.1b. Note that both axes are logarithmic, reflecting the way luminance is transduced by the visual system, i.e., the neural response increases linearly with the logarithm of luminance contrast. Campbell and Robson (1968) were the first to describe the contrast sensitivity function in the fovea. For these spatial frequencies and light levels, the Michelson contrast[2] at threshold is approximately 0.002 (corresponding to a contrast sensitivity of 500). This value reflects our high sensitivity for periodic luminance variations at those spatial frequencies with foveal viewing. At maximal contrast levels approaching 1.0, the upper cutoff for adult foveal vision varies from 30 to 60 c/deg, corresponding to a visual acuity value of 1.0–2.0. Sensitivity to low spatial frequencies is limited by lateral inhibition arising from the responsive neurons receptive field surrounds, i.e., light falling within the neuron's receptive field inhibitory surround leads to a weaker response in that neuron.

---

[2] Michelson contrast is defined for periodic patterns as $(L_{max} - L_{min})/(L_{max} + L_{min})$, where L corresponds to the luminance level at any given spatial location, L_{max} corresponds to the highest level and L_{min} to the lowest level. Michelson contrast varies between 0 and 1.

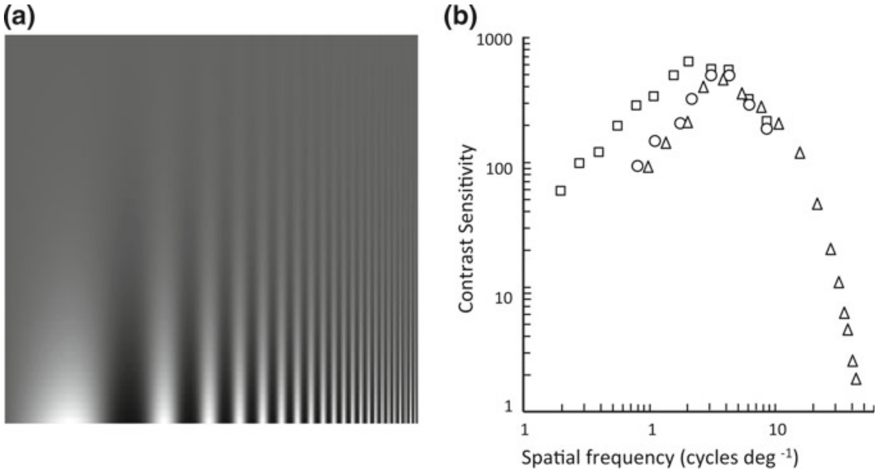


Fig. 5.1 Demonstration of contrast sensitivity function (adapted from Campbell & Robson, 1968). **a** This image shows a continuously increasing spatial frequency grating (from left to right) where the contrast is maximal at the lower edge and contrast decreases from bottom to top. At the appropriate viewing distance the shape of the contrast sensitivity function is revealed by the transition from visible to invisible (i.e., uniform grey background). **b** Contrast sensitivity as a function of spatial frequency (log-log axes) for observer FWC for photopic luminance levels (500 cd/m^2). The different symbols correspond to three binocular viewing conditions (squares: stimulus size 10° × 10° at viewing distance of 57 cm; circles stimulus size 2° × 2° at 57 cm; triangles: 2° × 2° at 285 cm). Reproduced with permission of the publisher and Dr. J. G. Robson

Sensitivity to high spatial frequencies is limited by the optics of the eye and the resolving power of the retina.

Several studies have investigated the changes of the contrast sensitivity function in central and peripheral vision. The results of one study (Rovamo, Virsu, & Nätänen, 1978) are shown in Fig. 5.2. Here the contrast sensitivity function is shown to shift downwards and to the left as the stimuli are viewed by increasingly more eccentric parts of the visual field. A comparison between the contrast sensitivity functions presented in Figs. 5.1 and 5.2 indicates the extent to which sensitivity to medium-to-high spatial frequencies is lost in eccentric vision.

5.3.2 *Visual Crowding in Eccentric Vision*

Suprathreshold vision (i.e., involving stimuli with contrast levels clearly above the contrast threshold) has also been shown to decline with eccentricity. These changes in performance occur for digit identification (Strasburger, Harvey, & Rentschler, 1991), letter identification (Chung, Legge, & Tjan, 2002), relative-phase discrimination (Renschler & Treutwein, 1985) and face recognition (Melmoth, Kukkonen, Mäkelä, & Rovamo, 2000). A further factor limiting peripheral vision is related to

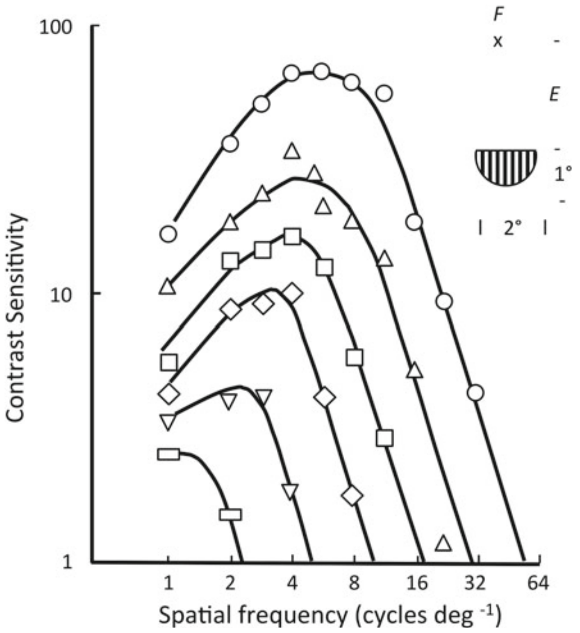


Fig. 5.2 Contrast sensitivity function for different eccentricities in the lower visual field. The different symbols signify the different eccentricities: circles: foveal viewing, triangles: 1.5°, squares: 4°; diamonds, 7.5°; inverted triangles, 14°; rectangles: 30°. The inset depicts the fixation point, F, and the eccentricity, E. The grating dimensions were 2° in width and 1° in height. Note that the overall foveal contrast sensitivity (circles) is reduced compared to that illustrated in Fig. 5.1. This reduction is most likely related to monocular viewing and the smaller size of the grating stimulus. Reprinted with permission of the publisher

the phenomenon of *visual crowding* (Pelli, 2008; Whitney & Levi, 2011), which reflects the negative effect of clutter (i.e., effect of closely spaced flanker stimuli) on stimulus recognition. The effect of visual crowding on discrimination performance is presented in Fig. 5.3. Stimulus spacing thus becomes critical for eccentrically viewed stimuli, such that stimuli viewed at a given eccentricity x should be spaced apart by at least $x/2$ to be recognizable (Fig. 5.3; Bouma, 1970). These observations suggest that parafoveal vision is not only characterized by a loss of contrast sensitivity for high spatial frequencies, but also by a coarser neural processing of eccentrically presented stimuli.

5.3.3 The Need for Eye Movements in Primate Vision

Because of the decrease in contrast sensitivity with retinal eccentricity and the detrimental effect of crowding as illustrated in Figs. 5.2 and 5.3, respectively, we need to actively inspect scenes by shifting our fovea to objects of interest in our visual environment. This is achieved by goal-directed eye movements. Owing to the limitations of peripheral vision discussed above, we learn, during the postnatal development of the retina (Ordy, Massopust, & Wolin, 1962), to direct our fovea to peripheral targets of interest such as objects, faces and scenes. High-acuity vision thus depends on how well objects of interest are kept in our fovea. In a world full of moving objects in which we ourselves move almost continuously (to reach different places, to turn

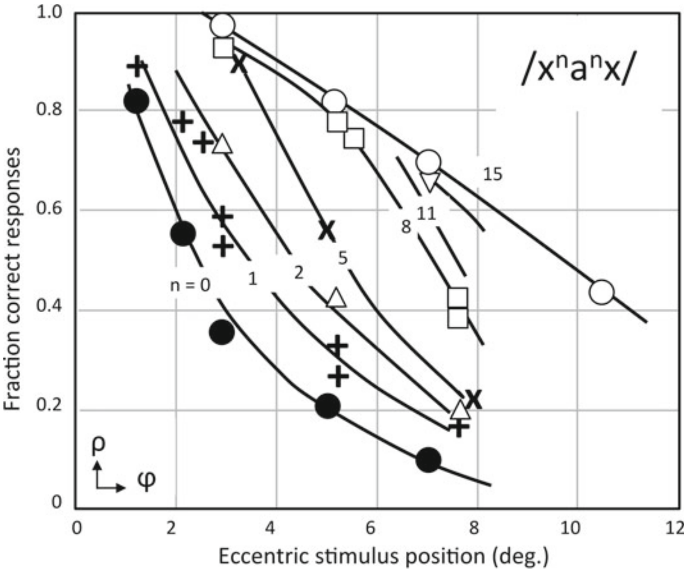


Fig. 5.3 Effect of crowding on letter discrimination in parafoveal vision. The proportion of correct responses is plotted as a function of stimulus eccentricity (degrees). The parameter reflects the relative spacing between target and flanking stimuli varying from 0° to 4.35° in n = 15 steps of 0.29°. The greater the spacing between the letters presented in eccentric vision, the greater was the likelihood that the target letter would be correctly recognized. Reprinted with permission of the publisher

and orient to new stimuli, to avoid obstacles, etc.), several types of eye movements have evolved to ensure that the targets remain on the fovea despite the observer's self motion. These include saccades, ocular pursuit, vestibular nystagmus and optokinetic movements.

5.3.4 Vision During Eye Movements

One obvious problem that arises when we move our eyes is that the otherwise stationary scene is rapidly moved across the retina in a direction opposite to the eye movement. Despite these almost constant retinal image shifts, we do not normally perceive the motion evoked by our own eye movements. Why is this so? Current theory suggests that the visual cortex receives an *efference copy* (von Holst & Mittelstaedt, 1950) or *corollary discharge* (Guthrie, Porter, & Sparks, 1983; Sperry, 1950) from the oculomotor system that informs the visual cortex that an oculomotor-induced retinal shift is about to take place. This efference copy is matched against the reafference resulting from the image shift on the retina. If cancellation occurs, the percept remains stable. If there is a difference between these signals, the object is perceived to move (exafference). In this section we will review studies that focus on visual phenomena that take place at the moment a saccade occurs.

5.3.5 Saccadic Suppression

Saccadic suppression is a form of visual desensitization for targets presented just prior to the onset of a saccade (Volkmann, Schick, & Riggs, 1968; Zuber & Stark, 1966; see Ross, Morrone, Goldberg, & Burr, 2001 for a review). The visual system is temporarily desensitized to stimulus contrast about 50 ms prior to and during the saccadic eye movement. This desensitization is not due to vitreous shearing nor to the Stiles-Crawford effect of the first kind[3] (Richards, 1969), but rather is attributed to a neural signal arising from the oculomotor system. One model of saccadic suppression proposes that a corollary discharge from the oculomotor system is added to the stimulus-driven signal that arises from the retinal image motion as a consequence of a saccade. This transient corollary discharge saturates the function describing the relationship between response and stimulus intensity (see, Naka & Rushton, 1966) thereby reducing sensitivity to stimuli flashed during the saccade (Diamond, Ross and Morrone, 2000). As such, the stimulus-evoked activity is masked by the effects arising from the eye movement. The putative corollary discharge could arise from the oculomotor nuclei in the brainstem that project into the lateral geniculate nucleus (LGN) of the thalamus (Reppas, Usrey, & Reid, 2002).

Interestingly, in psychophysical experiments, the processing of stimuli preferred by the magnocellular pathway has been found to be more strongly affected than the processing of stimuli preferred by the parvocellular pathway leading to the ventral stream (Burr, Morrone, & Ross, 1994). This finding suggests that neural activity in the dorsal visual stream (magnocellular pathway) is selectively suppressed. Consistent with this assumption, single-unit recordings in macaque dorsal areas MT (medial temporal area), MST (medial superior temporal area), VIP (ventral intraparietal area) show a significant saccadic suppression of neural responses to a flashed bar of light, presented ±60 ms with respect to saccade onset (Bremmer, Kubischik, Hoffmann, & Krekelberg, 2009; see their Fig. 5.5).

At which level of the visual process does this suppression take place? At the level of the LGN, relay neurons exhibit a bi-phasic response pattern with a brief, low amplitude, inhibitory phase (from approx. 50 ms prior to saccade onset up to 50 ms post saccadic onset) followed by a pronounced response enhancement (Reppas et al., 2002). In addition, the mediodorsal (MD) nucleus of the thalamus connects the frontal eye fields with the superior colliculus. Presaccadic activity in these neurons could represent an important source of corollary discharge. Using a double-step

---

[3] The Stiles-Crawford effect of the first kind (Stiles & Crawford, 1933) describes the directional selectivity of the light response of the photoreceptors. Light entering the eye at the rim of the pupil is less effective for vision than for light of the same intensity that enters the eye from the center of the pupil.

saccade task, Sommer and Wurtz (2002) showed that selective inhibition of neural activity in MD relay neurons led to systematic errors in the end points of the second saccade, without affecting the amplitude and velocity of the initial saccade (see, their Fig. 5.2). In humans, the response to flickering Ganzfeld stimuli (i.e., homogeneous illumination of the entire visual field) is suppressed by saccades at the level of the lateral geniculate nucleus and the primary visual cortex in fMRI (Sylvester, Haynes, & Rees, 2005). Interestingly, in total darkness, saccades lead to activation in LGN and V1. The retinotopic-specific fMRI-BOLD response to Gabor patches (i.e., sinusoidal gratings with a two-dimensional Gaussian contrast envelope) flashed just prior to the onset of a visually guided saccade is significantly reduced just before saccade onset (Fig. 5.4, Vallines & Greenlee, 2006). These findings suggest that saccadic suppression is an active process, where stimulus-driven responses in early areas of the visual pathway are selectively inhibited. As such saccadic suppression would contribute to visual stability by blocking out unwanted motion signals and blur that would otherwise arise during eye movements (see also chapter by Pierce et al. in this volume).

5.3.6 Saccadic Suppression of Displacement

In addition to saccadic suppression there is also an effect related at the inability to see stimulus displacements during eye movements. This effect is called saccadic suppression of displacement (Bridgeman, Hendry, & Stark, 1975; Deubel, Schneider, & Bridgeman, 1996). This inability to detect spatial shifts in target locations occurs when targets are displaced while the eyes are in motion (Bridgeman et al., 1975). It suggests that motion-sensitive cortex is also affected by the corollary discharge arising from the oculomotor system (Fig. 5.5). The larger the amplitude of the saccade, the greater the displacement must be in order for the spatial displacement to be detected by the observer (Li & Matin, 1990). The strengths of the effects are comparable for horizontal and vertical stimulus displacements (Stark, Kong, Schwartz, Hendry, & Bridgeman, 1976). The neural basis for this selective loss in motion sensitivity remains largely unknown. Single-unit recordings in behaving macaque monkeys indicate that the preferred motion direction of MST cells can reverse during saccades (Thiele, Henning, Kubischik, & Hoffmann, 2002), suggesting that stimulus motion and retinal slip due to eye movements are encoded by different mechanisms. The corollary discharge arising from the oculomotor system might not only suppress sensitivity to targets presented during the eye movement, but also alters neural responses to the motion of suprathreshold stimuli.

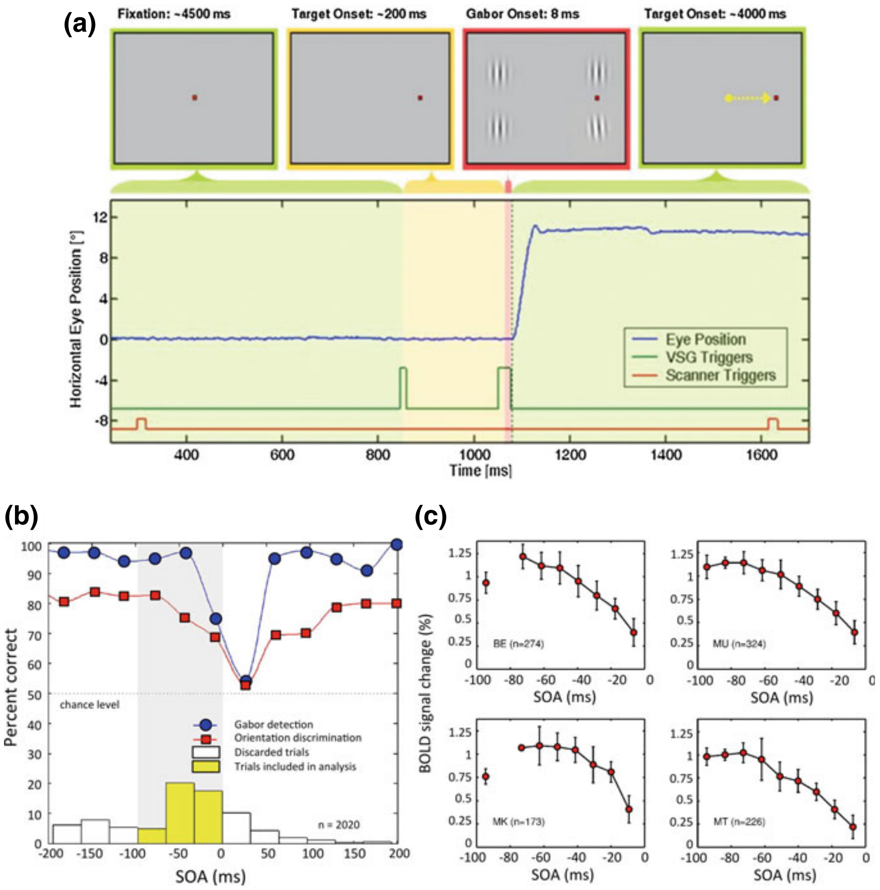


Fig. 5.4 Saccadic suppression in human parafoveal vision: **a Paradigm** Schematic illustration of the experimental design and stimuli used. A typical trial began with central fixation followed by the target shifted to an eccentric location. During the pre-saccadic period, 4 Gabor stimuli were flashed briefly (8 ms) and the subject had to judge whether a stimulus was presented and if so, whether one of the Gabors was tilted to the left or right. In the example shown, the lower Gabor is slightly tilted to the left. The lower blue trace shows the horizontal eye position on a trial where the subject made a 10-degree saccade to the right. The green trace depicts the time of onsets of saccade target and Gabors, respectively. **b Behavioral results** Percent correct on trials where the observers executed a saccade is shown as a function of the peri-saccadic time (stimulus onset asynchrony between stimulus onset and the initiation of the saccade—SOA in ms). Observers performed two tasks: a detection task (filled circles) whether Gabor stimuli were presented on that trial and a discrimination task, where the orientation of one of the Gabor stimuli was tilted to the left or right. These psychophysical data were collected in the MR scanner. Saccade onsets were sorted post hoc into time bins according to SOA. For the fMRI analysis, only trials on which the Gabor stimuli were presented prior to saccade onset were used (100 ms or less presaccadic, shaded region). **c Imaging results** The results of the region-of-interest (ROI) analysis for the retinotopic locations in V1 in four observers. Response magnitude (in percent signal change) is shown as a function of the SOA between Gabor stimuli onset and saccade onset. The closer the stimulus onset is to saccade initiation (SOA = 0), the lower the BOLD response becomes. This drop in cortical response reflects saccadic suppression in early visual cortex. From Vallines and Greenlee (2006) with permission of the publisher

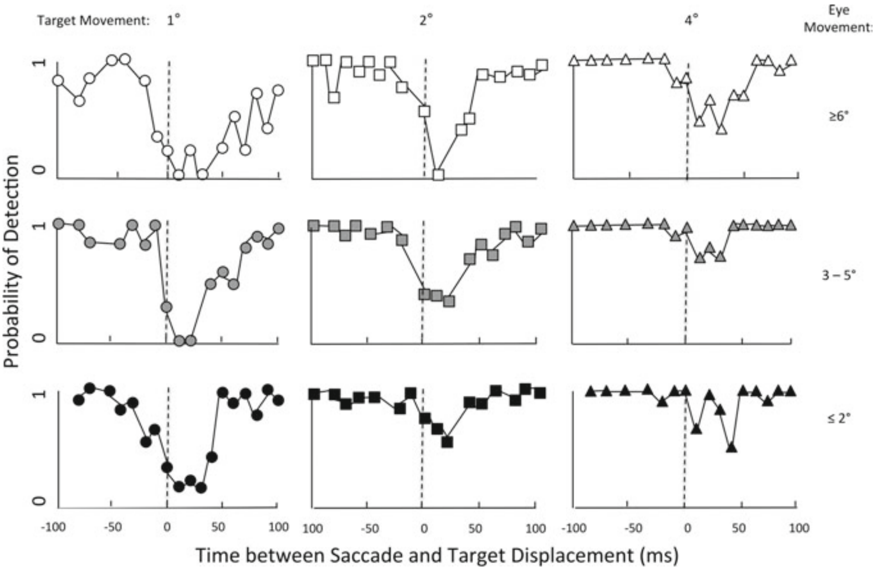


Fig. 5.5 Effect of saccadic eye movements on the ability of observers to detect stimulus movement. The proportion of correct responses is plotted as a function of pre- and postsaccadic onset time (ms) for 1 (circles), 2 (squares) and 4 (triangles) degrees of stimulus displacement (leftmost to rightmost columns, respectively). The decrease in motion sensitivity begins approx. 50 ms prior to saccade onset and lasts until the eye has reached the saccadic target. The mean results from 4 subjects are shown for saccades larger than 6° (upper row), for saccades of medium amplitude (3°–5°; middle row) and for small (<2°) saccades (lower row). Replotted from Bridgeman, Hendry, and Stark (1975) with permission of the publisher

5.3.7 Saccadic Mislocalizations

It has been known for some time, that visual targets flashed onto the display at the time of saccade initiation are perceptually mislocalized (Bischof & Kramer, 1968; Helmholtz, 1867; Matin & Pierce, 1965). In this paradigm a clearly visible stimulus bar is flashed while the participant executes a saccade to a target. After the saccade a pointer is presented on the display and the participant is asked to place the pointer at the perceived location of the flashed bar. The localization errors are most pronounced in the direction of the saccade itself, either along the horizontal (Honda, 1989) or vertical (Honda, 1991) visual field meridian. Similar to saccadic suppression (see above), saccadic mislocalization is most pronounced 100 ms or less before saccade initiation (Awater & Lappe, 2006). The perceptual distance between the saccadic target and the location of briefly flashed targets also appears to be compressed at the time of the saccade (Morrone, Ross, & Burr, 1997). The model suggested by Morrone et al. (1997, their Fig. 15) involves an anticipatory shift in location of the flashed target coupled with a transitory compression of space around the site of the upcoming saccade end point. These findings suggest that during saccades the spatial uncertainty of target locations increases. This increased spatial uncertainty is indicated by a shift in the apparent location of these targets towards the saccade end point. Saccadic mislocalizations would accordingly result from the transient increase in spatial uncertainty and the perceptual shrinking of space around the saccade target. A computational model of saccadic compression and mislocalization has been presented (Hamker, Zirnsak, Ziesche, & Lappe, 2011).

5.3.8 Saccadic Adaptation

As described above, the resolution of foveal vision largely exceeds that of peripheral vision. In order to make efficient use of foveal vision we need to be able to accurately direct our gaze to objects in peripheral parts of the visual field. This requires that the location of peripheral targets be precisely encoded to ensure as small an error as possible between the saccadic landing site and the target under consideration (usually under 0.5°). This computation appears to be optimized by adaptation.

Saccadic adaptation is a form of oculomotor gain adjustment that takes place when recurrent saccadic errors (i.e., under- or overshoots) occur, particularly in cases where the otherwise stationary saccade target is displaced either towards or away from its original position during the eye movement (Hopp & Fuchs, 2004). Saccadic errors can be induced by a peri-saccadic target step in a so-called double-step paradigm (McLaughlin, 1967). During the eye movement the saccade target is either systematically shifted inwards or outwards, leading to repetitive overshoots (after an inward target step) or undershoots (after an outward target step). Since these target displacements take place during the eye movement, the observer is unaware of any change in target location (owing to saccadic suppression of displacement, see above). Despite this lack of awareness, the oculomotor system adjusts its gain to reduce the post-step saccadic error on repetitive adaptation trials.

The neural mechanisms of saccadic adaptation have been studied in primates and humans (Hopp & Fuchs, 2004; Noto, Watanabe, & Fuchs, 1999). Noto et al. (1999) examined the extent to which gain changes carried over to saccades with slightly different directions and sizes, leading to the concept of gain adaptation field, which has its maximum for the adapted saccade direction and amplitude and falls off rapidly for other directions and sizes (see Fig. 5.2 in Noto et al., 1999). Steenrod, Phillips and Goldberg (2013) used saccadic adaptation to dissociate the coding of saccadic targets from saccadic amplitudes in the responses of LIP neurons. Following gain-reducing saccadic adaptation, LIP neurons continued to code the spatial location of the principal saccadic target. No evidence was found for a remapping of activity in the macaque superior colliculus after saccadic adaptation (Quessy, Quinet, & Freedman, 2010). Thus, the exact location of gain changes related to saccadic adaptation remains largely unknown.

In humans, Desmurget et al. (1998) reported a positive correlation between regional cerebral blood flow in the medioposterior cerebellum and the amount of saccadic adaptation. Using gaze-contingent displays in the fMRI scanner, we demonstrated that a reduction in activity in the left supplementary eye field is correlated with the reduction in saccadic gain across subjects (Blurton, Raabe, & Greenlee, 2011). A reduction in saccadic gain compensates for inward displacements of the saccadic target for rightward-directed saccades (Fig. 5.6, taken from Blurton et al., 2011). While subjects performed the double-step task, the fMRI percent signal change decreased in supplementary eye fields (SEF) with increasing adaptation. Significantly, the magnitude of this BOLD signal change is positively correlated with the gain reduction following adaptation (Fig. 5.6b), accounting for 45% of the variance in the adaptation-induced BOLD signal change in SEF. We found saccade-induced activations in other saccade-related ROIs (e.g., in the frontal and parietal eye fields, as well as in the oculomotor vermis of the cerebellum), but none of these regions exhibited a significant correlation between a change in BOLD response and the magnitude of saccadic adaptation across subjects.

In addition to the adaptation-induced reduction of BOLD response in the supplementary eye fields, Blurton et al. (2011) found a reduced negative BOLD in the posterior insular vestibular cortex (PIVC). Negative BOLD signals, i.e., a decrease in fMRI signal on trial onset (Smith, Singh, & Greenlee 2000; Smith, Williams, & Singh, 2004), are thought to represent deactivation or neural inhibition. Although the amplitude of this BOLD response reduction in PIVC was not correlated with the amount of gain reduction, we speculate that the supplementary eye fields might use these vestibular-derived cortical signals to create a gaze-centered frame of reference. This shift from an eye-centered to a head-centered reference frame could be useful when correcting saccadic amplitudes for the small head rotations in the direction of the saccade end point occurring during the execution of saccades. We would like to

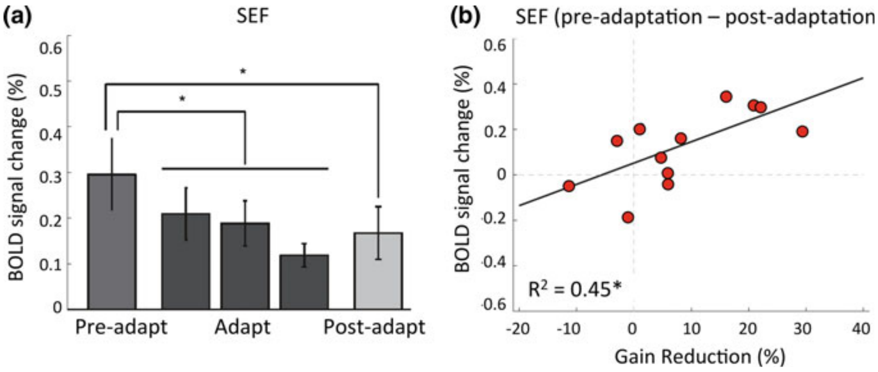


Fig. 5.6 BOLD signal (percent signal change) is presented for a region-of-interest (ROI) located in the left supplementary eye field. **a** BOLD percent signal change for pre-adaptation level (Pre-Adapt) and three blocks of 12 trials each during adaptation (Adapt). After adaptation the BOLD response begins to recover to pre-adaptation baseline levels (Post-Adapt). **b** The across-subjects correlation between the amount of BOLD signal change during adaptation and the relative amount of gain reduction (in percent) for the 13 subjects tested. There is a significant positive correlation between the amount of adaptation-induced BOLD response and the amount of gain reduction during saccadic adaptation. From Blurton, Raabe, and Greenlee (2011) with permission of the publisher

emphasize this result was unexpected, so the role of the vestibular cortex in saccadic gain changes remains speculative. It should be emphasized that the role of the cerebellum in gain changes resulting from saccadic adaptation has been supported by the studies of Desmurget et al. (1998) and Pelisson et al. (2010).

5.3.9 Motion Perception and Pursuit Eye Movements

Primate vision is characterized by a superb ability to detect and discriminate the velocity of moving objects. Extensive research has been conducted to enhance our understanding of the visual processing of object motion (for a review see, Albright & Stoner, 1995; Burr, 2014; Nakayama, 1985). The neuronal basis of motion perception in primates has been explored with a number of methods, including microelectrode recording techniques (Dubner & Zeki, 1971; Newsome, Britten, Salzman, & Movshon, 1990; van Essen & Gallant, 1994; Zeki, 1978) and functional magnetic resonance imaging (Smith, Greenlee, Singh, Kraemer, & Hennig, 1998; Sunaert, Van Hecke, Marchal, & Orban, 1999; Tootell et al., 1995). This section focuses on how the perception of object motion is affected by smooth pursuit eye movements and vice versa how motion perception impacts pursuit. These studies suggest that the pursuit system is closely controlled by perceptual information about target velocity (see also chapter by Lencer et al. in this volume).

It has been known for some time that pursuit eye movements affect visual motion perception (for a recent review see Schütz, Braun, & Gegenfurtner, 2011). One type of illusory motion induced by pursuit is referred to as the Filehne illusion (Bridgeman, Kirch, & Sperling, 1981; Mack & Herman, 1973), where an otherwise stationary background appears to move in a direction opposite to the smooth pursuit eye movements. Haarmeier, Thier, Repnow and Petersen (1997) investigated a patient with bilateral lesions in occipitotemporal extrastriate visual cortex. The patient complained about vertigo. Upon more careful examination, the patient reported strong Filehne illusions when he tracked a moving dot on a stationary background. Further tests of motion perception in this patient resulted in normal visual threshold values. MRI revealed that the patient had a lesion in the MT+ complex (consisting of areas MT and MST). These results imply that regions in the dorsal pathway, most likely area MT and MST, process a corollary discharge from the pursuit system that normally cancels illusory motion (thereby reducing the Filehne illusion). In the absence of feedback from MT/MST, the patient studied by Haarmeier et al. (1997) suffered from illusory scene motion during eye movements leading to vertigo. Taken together the findings suggest that the perception of background motion evoked by eye movements is actively suppressed during pursuit.

Several experiments have been conducted to determine whether the pursuit system follows the retinal signal or whether is it related more to the percept of target motion (Stone, Beutter, & Lorenceau, 2000). The relationship between stimulus speed and perceived velocity of moving stimuli with varying contrast levels has been studied in human subjects psychophysically. It was found that the perceived motion of a moving target declines with decreasing contrast levels (Stone & Thompson, 1992; Thompson, 1983), as does the gain of pursuit (Spering, Kerzel, Braun, Hawken, & Gegenfurtner, 2005). In the latter study, the authors used a step-ramp paradigm (Rashbass, 1961), where the subject first executed a saccade to a peripheral Gabor patch and began to pursue the target with their eyes as it moved towards the center of the screen. At the end of the trial, subjects were asked to judge the direction and speed of the Gabor stimulus. Their results indicate that the contrast sensitivity for the Gabor patch depends not only on its spatial frequency but also on the velocity of pursuit. Interestingly, the gain of pursuit is reduced for low contrast stimuli, suggesting again that pursuit follows perception (Steinbach, 1976).

5.3.10 Differentiation of Self- and Object Motion at Threshold

If subjects pursue a small moving dot in complete darkness or view a drifting grating or plaid on a homogenous background, the lower threshold for motion perception is approximately 0.1°/s or even less (Campbell & Maffei, 1981; Müller & Greenlee, 1994; Wright & Gurney, 1992). At these low speeds subjects cannot differentiate whether the eyes are moving (and thus pursuing a moving dot) or whether they themselves are moving. This phenomenon is referred to as dot-motion induced vection. Only in cases where stimulus speed exceeds the vestibular motion perception threshold of 0.2°/s can subjects differentiate between object (i.e., eye) motion and self motion (Kolev et al., 1996). Extensive research has been conducted in humans (Brandt et al., 1998; Eickhoff et al., 2006; Frank et al., 2014; Lopez & Blanke, 2011) and primates (Britten, 2008; Bremmer, 2011; DeAngelis & Angelaki, 2012; Guldin & Grüsser, 1998) indicating that regions in the posterior insula and occipito-parietal cortex are involved in the integration of visual and vestibular signals related to object and self motion. This information could be used to locate an object in space with respect to the position of the head (craniocentric) or body (egocentric).

5.3.11 Coherent Dot Motion Perception During Stable Fixation and Smooth Pursuit

Global motion in noisy stimulus displays can be extracted by human observers (Morgan & Ward, 1980; Williams & Sekuler, 1984). Studies using random-dot motion displays have revealed that such stimuli can be used to assess the gain of the pursuit system in the open-loop component (i.e., within the first 130 ms from stimulus onset; see, Watamaniuk and Heinen, 1999). Here it was found that the pursuit gain is reduced for increasing levels of directional noise in random-dot motion displays.

Pursuit eye movements normally cancel or slow down the perception of object motion, but they can, under specific circumstances, also enhance motion perception. In a study from our group (Greenlee et al., 2002), participants either fixated a stationary red dot or pursued the same dot with their eyes while it moved horizontally with a maximum speed of 12.6 deg/s. During fixation or pursuit, 400 moving random dots were displayed and subjects were requested to detect whether any of those dots moved left or right. The derived psychometric functions indicated a significant effect of pursuit on coherent motion detection thresholds (Fig. 5.7, left). When the coherent motion was in-phase with the subject's eye movements, detection thresholds declined to a value as low as 2%. The subjects indicated that the coherently moving dots appeared to "pop-out" from the randomly moving background dots. This perceptual pop-out depended on the relative direction of the pursuit eye movements and the direction of coherent motion: coherent random-dot motion in the opposite direction of the eye movements (out-of-phase) was more difficult to detect than when the eyes remained stationary. Interestingly, coherent dot motion also enhanced the gain of pursuit (Fig. 5.7, right side), but only when the coherent motion was in-phase with the pursuit eye movements. These findings suggest that pursuit eye movements and motion perception mutually influence each other, perhaps involving a comparison between a corollary discharge and the visual input.

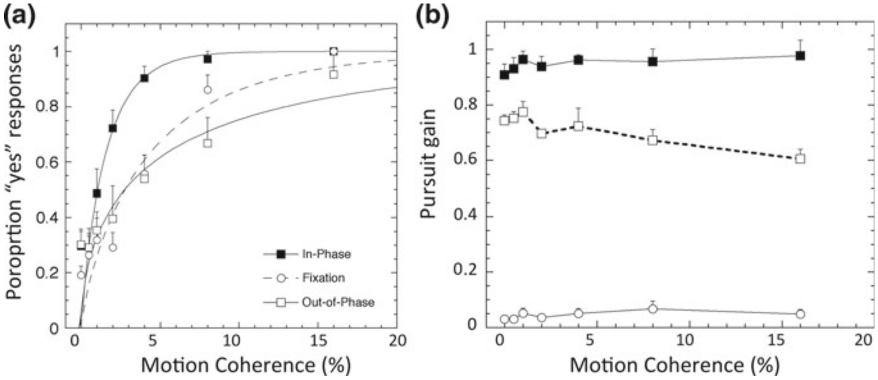


Fig. 5.7 Psychometric functions for coherent motion detection: **a** the proportion "yes" responses (i.e., trials on which the subjects reported the presence of coherent motion) are shown as a function of the motion coherence level (%) for the three viewing conditions: pursuit in-phase (filled squares) with coherent motion, stationary fixation (open circles) and pursuit out-of-phase (open squares) with coherent motion. In-phase pursuit is shown to significantly enhance performance on the coherent motion detection task. **b** Pursuit gain is shown as a function of motion coherence level (%) for the conditions of pursuit in-phase with coherent dot motion or out-of-phase with coherent motion, as well as for the stationary fixation condition. In-phase coherent motion is shown to enhance smooth pursuit. From Greenlee, Schira, and Kimmig (2002) with permission of the publisher

5.3.12 Sensorimotor Transformations During Pursuit

Smooth pursuit eye movements are driven by moving targets (see Lencer et al. in this volume). The pursuit system processes the visual input signals and transforms this information into an oculomotor output signal. Areas MT and MST in the dorsal visual pathway play a critical role in the visual motion processing required for smooth pursuit (Komatsu & Wurtz, 1988a, 1988b). Despite the object movement on the retina and the eyes' movement in the orbit, we are able to locate the object in space, implying coordinate transformations from retinal to head-in-space coordinates (Bremmer, 2011).

In functional magnetic resonance imaging in humans cortical activations during pursuit are found in a network of occipito-parietal and frontal brain regions. Visual motion processing during pursuit takes place in the motion-sensitive, temporo-parieto-occipital region MT+ and the right posterior parietal cortex (PPC). Motor components comprise more widespread activation in these regions and additional activations in the frontal and supplementary eye fields (FEF, SEF), the cingulate gyrus and precuneus. Possible transformation sites are found in MT+ and PPC. We found that the MT+ activation evoked by the motion of a single visual dot was highly localized, while the activation of the same single dot motion driving the eye movement was rather extended across MT+. We concluded that the eye movement information is dispersed across the visual map of MT+. This could be interpreted as a transfer of the one-dimensional eye movement information onto the two-dimensional visual map. Potentially, the dispersed information could be used to remap MT+ to space coordinates rather than retinal coordinates and to provide the basis for a motor output control. A similar interpretation was put forth for the PPC region (Kimmig et al., 2008). Eye position signals have been demonstrated in human MT+ (d'Avossa et al., 2007) and these appear to depend on spatial attention (Crespi et al., 2011).

By varying the amount of background dots during a pursuit task Ohlendorf et al. (2010) showed that the visual component activity in the pursuit cortical network increased with the number of dots in the background (summation effect) but broke down at a certain amount of background motion revealing another not yet identified physiological process. Early visual areas like V1 appeared to receive eye-movement motor information, which could be used for spatial remapping, in analogy to a saccade task during which V1 activation was interpreted as spatial updating (Merriam et al., 2007).

It is known that area MT+ in monkey receives retinal and extraretinal input and can process motion in world coordinates (Bremmer et al., 1997, Ilg et al., 2004). Ohlendorf et al. (2010) described two subregions in human MT+, one that is strongly modulated by the quantity of retinal stimulation, and another that is multimodally stimulated by eye pursuit and visual motion, but independent of the amount of retinal stimulation. The latter region was identified as subarea MST (which responds to ipsilateral optic flow; Dukelow et al., 2001; Huk et al., 2002; Ohlendorf et al., 2008). MST was therefore thought to be a likely candidate for visual-to-oculomotor transformation during pursuit. Another region activated by visual motion and eye motion was V7/LOC (lateral occipital cortex), but only when higher amounts of background dots appeared to represent a form or shape. Indeed, Area LO (including a subregion we have termed lateral occipito-parietal, LOP) has been shown to be activated preferentially by shape in comparison to scrambled objects (Grill-Spector et al., 1998) and has been shown to be involved in extracting information about object structure from different image cues (Kourtzi & Kanwisher, 2000). Accordingly, MST could be more involved in the spatial remapping of the eye position in space whereas V7/LOC appears to perform visuo-oculomotor transformations in the presence of a structured background.

The posterior parietal cortex integrates information from different senses to form spatial representations. Parts of the posterior parietal cortex (PPC) in monkeys have been described to be multimodal and to code the spatial location of visual objects (Andersen, Snyder, Bradley, & Xing, 1997; Bremmer et al., 2002). Furthermore, PPC has been shown to process motion in space (Andersen et al., 1997; Andersen, Snyder, Batista, Buneo, & Cohen, 1998; Grefkes & Fink, 2005). Ohlendorf et al. (2010) showed that PPC activation in their experiment cannot represent eye motion in head coordinates nor motion in space alone, but codes for the differential motion between eye and background (qualitative reference frame information).

5.3.13 The Role of Attention in Vision and Eye Movements

Visual attention reflects our ability to focus on certain aspects of a visual scene while we ignore all other aspects. William James (1890) emphasized the volitional aspect of spatial attention. In contrast, almost a century later Posner (1980) identified exogenous and endogenous control processes, using reaction time measurements in cueing tasks. Indeed, the control of covert shifts of attention has many similarities to the mechanisms underlying overt shifts of attention in the form of saccadic eye movements. Cells in the superficial layers of the superior colliculus respond more robustly, when a stimulus that falls within its receptive field becomes the target for a subsequent saccade (Goldberg & Wurtz, 1972). Response enhancements have also been found in neurons in the extrastriate (prelunate) visual cortex (Fischer & Boch, 1981) and in the lateral intraparietal cortex (Goldberg & Bruce, 1985) in alert monkeys, when they execute saccades to a target falling within their receptive field or when they just shifted their attention covertly to that location. Indeed, in humans fMRI activity evoked by covert shifts in spatial attention is comparable to that observed when the same subjects performed overt shifts of gaze (Perry & Zeki, 2000). The fMRI activations in prefrontal and parietal cortex for covert and overt shifts of attention exhibit considerable overlap (Corbetta et al., 1998; Corbetta & Shulman, 2002).

Saccadic reaction times (SRTs) can be used as an index of neural processing speed and as such can reflect the effects of attention. Fischer and colleagues (e.g., Fischer, 1986: Fischer et al., 1997) compared SRTs in the so-called "gap" and "overlap" tasks (see also the chapter by Pierce et al., in this volume). In these tasks the participant fixates a central stimulus (usually a small dot on a dark background). In the gap task, the fixation point is switched off and, after some delay, an eccentric stimulus is switched on. The participant executes a saccade to the eccentric target. In the overlap task, the central fixation point remains on, that is it temporally “overlaps” with the onset of the eccentric target. The behavioral results indicate that the gap task evokes bimodal saccadic reaction time distributions, with a maximum in the express saccade range of 100–120 ms and a second maximum in the “fast regular” range below 200 ms. In comparison, the overlap task leads to saccades with significantly longer latencies (in the “slow regular” range of 240 ms and longer). These results imply that in the overlap paradigm attention needs to be disengaged from the central fixation point before it can be re-engaged at an eccentric location. The gap task leads to shorter SRTs because the blank period between fixation point offset and saccadic target onset allows the brain to disengage attention from the central fixation point. The presence of additional processes has been discussed (see Reuter-Lorenz et al., 1991).

Attention not only affects the latency of saccadic eye movements, but also affects perception at the saccade end point. To investigate this supposed interaction between spatial attention and saccadic eye movements, dual tasks have been conducted, where the primary task is to make a visually guided saccade to a target and the secondary task is to recognize a letter at or near the target location (Deubel & Schneider, 1996; Kowler et al., 1995). Letter discrimination is best, when the location, to which attention is newly directed, is identical to the location of the letter target to be discriminated (Fig. 5.8). Performance falls off rapidly when the letter to be discriminated is spatially offset from the saccade target. These findings suggest that the control of spatial attention is closely linked to the control of saccadic eye movements. Although it is generally assumed that spatial attention and gaze location overlap, under certain

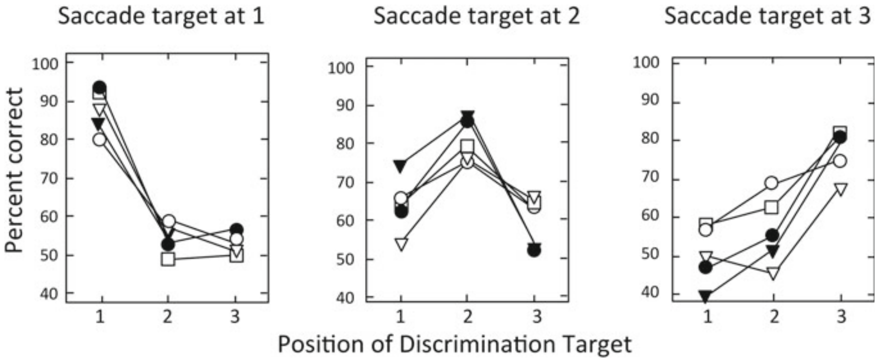


Fig. 5.8 Letter discrimination performance as a function of location for the conditions where the saccade target was located at the first (left panel), second (middle panel) or third (right panel) position. The different symbols present the results from each of the five subjects. Letter discrimination performance is greatest when the location of the letter to be discriminated corresponds with the location of the saccade target. Results are from Deubel and Schneider (1996) and are reprinted with permission from the publisher

circumstances the two can become dissociated, as is the case for the anti-saccade task, where the target cue appears on one side of fixation and the subject saccades to its mirror location.

5.3.14 Visual Memory and Eye Movements

Another important factor underlying our need to make several eye movements while scanning a scene is related to the limited capacity of working memory. Visual working memory denotes our ability to store and later recall information about previously viewed stimuli to help guide behavior (Baddeley, 2000, 2003; Pasternak & Greenlee, 2005; Ungerleider et al., 1998). The capacity of visual working memory is thought to be limited, with a capacity limit of 3 to 4 items (Cowan, 2004; Luck & Vogel, 1997). Working memory is essential for visual stability during eye movements. Since we move our eyes three to four times per second, it is essential that the visual information processed during each fixation should be integrated into a uniform percept of visual space (Irwin, 1991). Visual working memory assists us to locate previously fixated objects in space (Melcher & Kowler, 2001). Memory enhancement for stimuli that become the target for a saccade has been demonstrated (Bays & Husain, 2008). Functional MRI can be used to locate brain activity associated with visual working memory (Todd & Marois, 2004). In a typical experiment, the subjects are presented sequentially two grating patterns that differ either in orientation or color. Depending on the stimulus dimension to be remembered, the pattern of fMRI responses differs. Based on these fMRI data, multivariate pattern analysis (MVPA) has been used to determine the extent to which information about the stored stimulus dimension (i.e., orientation or color) can be extracted from the pattern of fMRI activity (Harrison & Tong, 2009; Kamitani & Tong, 2005; Serences et al., 2009; Sneve et al., 2012). In a study from our group (Raabe et al., 2013), pattern analysis of fMRI activity revealed four clusters in occipital and fusiform cortex and in the frontal eye fields that contained significant information about the color and orientation of the saccade target in a memory-guided task. The results of these studies suggest that visual working memory recruits the same brain areas involved in the original sensory analysis of the to be remembered stimuli.

Tasks involving sequences of saccades point to a prioritization of information in working memory for saccade targets (Khayat et al., 2004). Activation in the dorsolateral prefrontal cortex and the supplementary eye field is associated with the manipulation of information in working memory required to complete a triple-step, saccade-sequence task (Baumann et al., 2007). The early experiments by Irwin (1991) suggest that very short-term iconic memory supports visual pattern memory up to 300 ms post-stimulus, whereas short-term working memory retains information up to several seconds (see, Cornelissen & Greenlee, 2000). This former type of storage would underlie "transsaccadic" memory, since it is less reliant on absolute spatial positions, but rather stores the relative spatial position of elements thereby forming patterns. Indeed, when line drawings of real-world objects were used as stimuli, specific contour information (i.e., relative position of the line-drawing with respect to a grating mask) was not retained across saccades (Henderson, 1997). These latter results rather point to an abstract representation of the object in working memory.

In summary, there is no strong evidence for a specific form of "transsaccadic" memory. For varying stimuli, stimulus durations and inter-stimulus intervals (ISI), visual memory performance for post-saccadic stimuli is similar to that found for experiments without saccades. The results of memory tests with very short ISIs can be influenced by sensory (iconic) memory, by apparent motion or similar sensory-related phenomena. The results for long ISIs (over a few seconds) are similar to those expected based on a limited-capacity storage of more abstract object information.

5.4 Conclusions

Human foveal vision is one of the outstanding achievements of evolution. In order to fully exploit the advantages of central vision, we need to move our eyes to regions of interest contained within real-world scenes. Research over the last five decades has revealed the neural mechanisms underlying the control of visually guided eye movements. A significant challenge to the primate brain is to integrate the information that is obtained by a series of saccadic eye movements to yield stable, on-going and continuous perception. We find that certain aspects of visual perception are altered in the brief moments before and during saccades. We have reviewed selected studies that have explored the effects of saccadic eye movements on perception. Although in everyday life, for the most part, going unnoticed, sensitivity to briefly flashed stimuli is reduced when stimulus onset is coincident with the onset of a saccade. Our sensitivity for stimulus displacements is also considerably lowered when these displacements occur prior to or during a saccade. In addition, stimuli flashed during the preparation and execution of saccades are mislocalized in space. These stimuli appear to be shifted towards the location of the future saccade target as if space around the target were compressed. Because of this peri-saccadic reduced sensitivity to stimulus displacements, repeated peri-saccadic shifts in target location can go completely unnoticed by the subject. Despite this lack of awareness for the shift in the saccade target location, saccade adaptation leads to a reduction in oculomotor gain to match the new target location in the inward, double-step paradigm, as first described by McLaughlin (1967). Such gain reductions have been associated with changes in fMRI activations in the supplementary eye fields. The neural source of the corollary discharge associated with saccadic suppression, saccadic suppression of stimulus displacement and saccadic adaptation remains to be determined with certainty, but we now have more information as to how these signals arise during the preparation of intentional, goal-directed eye movements.

Pursuit eye movements are slow movements of the eyes required to keep the fovea on an object while it moves in space. A physically stationary background can appear to move in a direction opposite to that of an eye movement. This apparent movement, referred to as the Filehne illusion, usually goes unnoticed, as signals related to the retinal slip are cancelled by a corollary discharge. This corollary discharge or efference copy appears to require processing of motion-sensitive visual area MT/V5, since lesions in this area lead to an enhancement of the Filehne illusion and to vertigo. Pursuit can, under some circumstances, even enhance motion processing, in cases where coherently moving dots, presented among random dots, move in-phase with the pursuit eye movement.

Visual attention and visual working memory are closely related to the control of saccadic eye movements and perceptual stability. Indeed, many parallels suggest that the overt control of eye movements and spatial attention share mechanisms with the covert spatial control of visual attention. Visual working memory is essential to integrate the many percepts that arise during a sequence of saccades. We know that the more robust form of working memory persists across saccades, building up an enriched representation of visual scenes over multiple foveal fixations. However, visual working memory has a limited capacity with a maximal capacity of four items or less. For this reason, it is important that we can shift our gaze quickly and without effort to extract visual information to optimize our present and future behavior.

5.5 Suggested Readings

The interested reader might want to further deepen their understanding of the topic of visual perception during eye movements. For this purpose we strongly recommend to consult the following excellent reviews:

- Findlay and Gilchrist (2003), a textbook on active vision;
- Britten (2008), a good overview of the electrophysiological results on object and self motion;
- Bremmer (2011), a review of the literature on eye movements and visual stability;
- Morrone (2014), a handbook chapter on the changes in visual/spatial perception around the time of saccades.

Several pioneering papers are described above only cursorily, so direct reading of these articles will provide a better insight into the discoveries made there. Many chapters in the present volume also deal with phenomena that are directly related to our topic of visual perception during eye movements and therefore should be consulted.

5.6 Questions Students Should Be Able to Answer

Below are several questions that students should be able to correctly answer after careful reading of the above chapter.

- Which anatomical design aspects of the primate retina are important when discussing interactions between visual perception and eye movements?

- Describe the phenomenon of visual crowding. How does it affect the visual discrimination of peripherally viewed stimuli (e.g., letter recognition)?
- Define the term "corollary discharge". Where in the brain do corollary discharges arise and what role do they play in visual perception?
- Describe the phenomenon of saccadic suppression. How does saccadic suppression affect visual perception? Which neural correlates of saccadic suppression have been described?
- Describe the phenomenon of saccadic suppression of stimulus displacement. When does saccadic suppression of stimulus displacement occur?
- Describe the phenomenon of saccadic mislocalization. How is the spatial representation of briefly flashed stimuli distorted during saccadic eye movements?
- What is meant by the term "saccadic adaptation"? Which experimental paradigms are associated with the phenomenon of saccadic adaptation? Which neural correlates of saccadic adaptation have been discovered?
- How do pursuit eye movements differ from saccadic eye movements? What is the role of motion perception in pursuit eye movements? How is motion perception affected by pursuit eye movements?
- What sort of coordinate transformations are required to locate objects in space during pursuit eye movements? Which neural correlates have been investigated?
- Describe the relationship between visual attention and saccadic eye movements. Which experimental paradigms have been applied to study the role of visual/spatial attention on the planning and execution of saccades?
- How does the limited capacity of visual working memory affect the way we move our eyes? What is meant by the term "transsaccadic memory"? What evidence is there that speaks for or against the existence of transsaccadic memory?

Acknowledgements The authors thank Jale Özyurt, Sebastian M. Frank, John S. Werner and Lothar Spillmann for their helpful comments. Author MWG was supported by a grant from the Deutsche Forschungsgemeinschaft (DFG, GR 988/25-1).

Bibliography

Albright, A. D., & Stoner, G. R. (1995). Visual motion perception. *Proceedings of the National Academy of Sciences of the United States of America, 92,* 2433–2440.

Andersen, R. A., Snyder, L. H., Bradley, D. C., & Xing, J. (1997). Multimodal representation of space in the posterior parietal cortex and its use in planning movements. *Annual Review of Neuroscience, 20*(1), 303–330.

Andersen, R. A., Snyder, L. H., Batista, A. P., Buneo, C. A., & Cohen, Y. E. (1998). Posterior parietal areas specialized for eye movements (LIP) and reach (PRR) using a common coordinate frame. *Novartis Foundation Symposium, 218*, 109–122.

Awater, H., & Lappe, M. (2006). Mislocalization of perceived saccade target position induced by perisaccadic visual stimulation. *Journal of Neuroscience, 26*(1), 12–20.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*(11), 417–423.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Baumann, O., Frank, G., Rutschmann, R. M., & Greenlee, M. W. (2007). Cortical activation during sequences of memory-guided saccades: A functional MRI study. *NeuroReport, 18*(5), 451–455.

Bays, P. M., & Husain, M. (2008). Dynamic shifts of limited working memory resources in human vision. *Science, 321*(5890), 851–854.

Bischof, N., & Kramer, E. (1968). Untersuchungen und Überlegungen zur Richtungswahrnehmung bei willkürlichen sakkadischen Augenbewegungen [Investigations and considerations of directional perception during voluntary saccadic eye movements]. *Psychologische Forschung, 32*(3), 185–218.

Blurton, S. P., Raabe, M., & Greenlee, M. W. (2011) Differential cortical activation during saccadic adaptation. *Journal of Neurophysiology*, *107*, 1738–1747.

Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature, 226,* 177–178.

Brandt, T., Bartenstein, P., Janek, A., & Dieterich, M. (1998). Reciprocal inhibitory visual-vestibular interaction. Visual motion stimulation deactivates the parieto-insular vestibular cortex. *Brain, 121,* 1749–1758.

Bremmer, F. (2011). Multisensory space: From eye-movements to self-motion. *Journal of Physiology, 589*(Pt 4), 815–823.

Bremmer, F., Duhamel, J. R., Ben Hamed, S., & Graf, W. (2002) Heading encoding in the macaque ventral intraparietal area (VIP). *European Journal of Neuroscience, 16*, 1554–1568.

Bremmer, F., Ilg, U. J., Thiele, A., Distler, C., & Hoffmann, K. P. (1997). Eye position effects in monkey cortex. I. Visual and pursuit-related activity in extrastriate areas MT and MST. *Journal of Neurophysiology, 77,* 944–961.

Bremmer, F., Kubischik, M., Hoffmann, K.-P., & Krekelberg, B. (2009). Neural dynamics of saccadic suppression. *Journal of Neuroscience, 29*(40), 12374–12383.

Bridgeman, B., Hendry, D., & Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. *Vision Research, 15*(6), 719–722.

Bridgeman, B., Kirch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics, 29*(4), 336–342.

Britten, K. H. (2008). Mechanisms of self-motion perception. *Annual Review of Neuroscience, 31,* 389–410.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: A comparison of neuronal and psychophysical performance. *Journal of Neuroscience, 12*(12), 4745–4765.

Burr, D. C., Morrone, M. C., & Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. *Nature, 371,* 511–513.

Burr, D. C. (2014). Motion perception: Human psychophysics. In: J. S. Werner & L. Chalupa (Eds.), *The new visual neurosciences* (pp. 763–776). Cambridge, MA, USA: MIT Press.

Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *The Journal of Physiology, 197*(3), 551–566.

Campbell, F. W., & Maffei, L. (1981). The influence of spatial frequency and contrast on the perception of moving patterns. *Vision Research, 21*(5), 713–721.

Chung, S. T. L., Legge, G. E., & Tjan, B. S. (2002). Spatial-frequency characteristics of letter identification in central and peripheral vision. *Vision Research, 42*(18), 2137–2152.

Corbetta, M., Akbudak, E., Conturo, T. E., Snyder, A. Z., Ollinger, J. M., Drury, H. A., et al. (1998). A common network of functional areas for attention and eye movements. *Neuron, 21*(4), 761–773.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Reviews Neuroscience, 3*(3), 215–229.

Cornelissen, F. W., & Greenlee, M. W. (2000). Visual memory for random block patterns defined by luminance and color contrast. *Vision Research, 40*(3), 287–299.

Cowan, N. (2004). *Working memory capacity*. New York: Psychology Press, Taylor & Francis.

Crespi, S., Biagi, L., D'avossa, G., Burr, D. C., Tosetti, M., & Morrone, M. C. (2011). Spatiotopic coding of BOLD signal in human visual cortex depends on spatial attention. *PLoS ONE, 6*(7), e21661.

Curcio, C. A., Sloan, K. R., Packer, O., Hendrickson, A. E., & Kalina, R. E. (1987). Distribution of cones in human and monkey retina: Individual variability and radial asymmetry. *Science, 236,* 579–582.

D'avossa, G., Tosetti, M., Crespi, S., Biagi, L., Burr, D. C., & Morrone, M. C. (2007). Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nature Neuroscience, 10*(2), 249–255.

DeAngelis, G. C. & Angelaki, D. E. (2012). Visual-vestibular integration for self-motion perception. In M. M. Murray & M. T. Wallace (Eds.), *The neural bases of multisensory processes*. Boca Raton, FL: CRC Press.

Desmurget, M., Pélisson, D., Urquizar, C., Prablanc, C., Alexander, G. E., & Grafton, S. T. (1998). Functional anatomy of saccadic adaptation in humans. *Nature Neuroscience, 1*(6), 524–528.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36*(12), 1827–1837.

Deubel, H., Schneider, W. X., & Bridgeman, B. (1996). Postsaccadic target blanking prevents saccadic suppression of image displacement. *Vision Research, 36*(7), 985–996.

Diamond, M. R., Ross, J., & Morrone, M. C. (2000). Extraretinal control of saccadic suppression. *Journal of Neuroscience, 20*(9), 3449–3455.

Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature, 170,* 36–37.

Dodge, R. (1900). Visual perception during eye movement. *Psychological Review, 7,* 454–465.

Dubner, R., & Zeki, S. M. (1971). Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research, 35*(2), 528–532.

Dukelow, S. P., DeSouza, J. F., Culham, J. C., van den Berg, A. V., Menon, R. S., & Vilis, T. (2001). Distinguishing subregions of the human MT+ complex using visual fields and pursuit eye movements. *Journal of Neurophysiology, 86,* 1991–2000.

Eickhoff, S. B., Weiss, P. H., Amunts, K., Fink, G. R., & Zilles, K. (2006). Identifying human parieto-insular vestibular cortex using fMRI and cytoarchitectonic mapping. *Human Brain Mapping, 27*(7), 611–621.

Fechner, G. T. (1889). *Elemente der Psychophysik*. Leipzig: Breitkopf & Härtel Verlag.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford, U.K.: Oxford University Press.

Fischer, B. (1986). The role of attention in the preparation of visually guided eye movements in monkey and man. *Psychological Research, 48*(4), 251–257.

Fischer, B., & Boch, R. (1981). Enhanced activation of neurons in prelunate cortex before visually guided saccades of trained rhesus monkeys. *Experimental Brain Research, 44*(2), 129–137.

Fischer, B., Biscaldi, M., & Gezeck, S. (1997). On the development of voluntary and reflexive components in human saccade generation. *Brain Research, 754,* 285–297.

Frank, S. M., Baumann, O., Mattingley, J. B., & Greenlee, M. W. (2014). Vestibular and visual responses in human posterior insular cortex. *Journal of Neurophysiology, 112*(10), 2481–2491.

Gerardin, P., Miquée, A., Urquizar, C., & Pélisson, D. (2012). Functional activation of the cerebral cortex related to sensorimotor adaptation of reactive and voluntary saccades. *NeuroImage, 61*(4), 1100–1112.

Goldberg, M. E., & Bruce, C. J. (1985). Cerebral cortical activity associated with the orientation of visual attention in the rhesus monkey. *Vision Research, 25*(3), 471–481.

Goldberg, M. E., & Wurtz, R. H. (1972). Activity of superior colliculus in behaving monkey. II. Effect of attention on neuronal responses. *Journal of Neurophysiology, 35*(4), 560–574.

Greenlee, M. W., Schira, M. M., & Kimmig, H. (2002). Coherent motion pops out during smooth pursuit. *NeuroReport, 13*(10), 1313–1316.

Grefkes, C., & Fink, G. R. (2005). The functional organization of the intraparietal sulcus in humans and monkeys. *Journal of Anatomy, 207,* 3–17.

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping, 6,* 316–328.

Guldin, W. O., & Grüsser, O. J. (1998). Is there a vestibular cortex? *Trends in Neurosciences, 21*(6), 254–259.

Guthrie, B. L., Porter, J. D., & Sparks, D. L. (1983). Corollary discharge provides accurate eye position. *Science, 221,* 1193–1195.

Haarmeier, T., Thier, P., Repnow, M., & Petersen, D. (1997). False perception of motion in a patient who cannot compensate for eye movements. *Nature*, *389*, 849–852.

Hamker, F. H., Zirnsak, M., Ziesche, A., & Lappe, M. (2011). Computational models of spatial updating in peri-saccadic perception. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences, 366*(1564), 554–571.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature*, *458*, 632–635.

Henderson, J. M. (1997). Transsaccadic memory and integration during real-world object perception. *Psychological Science, 8,* 51–55.

Helmholtz, H. (1867). *Handbuch der Physiologischen Optik*. Leipzig: Leopold Voss.

Hess, R. F., & Nordby, K. (1986). Spatial and temporal limits of vision in the achromat. *The Journal of Physiology, 371*, 365–385.

Hirsch, J., & Curcio, C. A. (1989). The spatial resolution capacity of human foveal retina. *Vision Research, 29*(9), 1095–1101.

Honda, H. (1989). Perceptual localization of visual stimuli flashed during saccades. *Perception and Psychophysics, 45*(2), 162–174.

Honda, H. (1991). The time courses of visual mislocalization and of extraretinal eye position signals at the time of vertical saccades. *Vision Research, 31*(11), 1915–1921.

Hopp, J. J., & Fuchs, A. F. (2004). The characteristics and neuronal substrate of saccadic eye movement plasticity. *Progress in Neurobiology, 72*(1), 27–53.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology, 160,* 106–154.

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology, 195*(1), 215–243.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience, 22*(16), 7195–7205.

Ilg, U. J., Schumann, S., & Thier, P. (2004). Posterior parietal cortex neurons encode target motion in world-centered coordinates. *Neuron, 43,* 145–151.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 213,* 420–456.

James, W. (1890). *The principles of psychology*. Holt: New York.

Kamitani, Y., & Tong, F. (2005). Decoding the visual and subjective contents of the human brain. *Nature Neuroscience, 8*(5), 679–685.

Khayat, P. S., Spekreijse, H., & Roelfsema, P. R. (2004). Visual information transfer across eye movements in the monkey. *Vision Research, 44*(25), 2901–2917.

Kimmig, H., Ohlendorf, S., Speck, O., Sprenger, A., Rutschmann, R. M., Haller, S., et al. (2008). fMRI evidence for sensorimotor transformations in human cortex during smooth pursuit eye movements. *Neuropsychologia, 46,* 2203–2213.

Kolev, O., Mergner, T., Kimmig, H., & Becker, W. (1996). Detection thresholds for object motion and self-motion during vestibular and visuo-oculomotor stimulation. *Brain Research Bulletin, 40*(5–6), 451–458.

Komatsu, H., & Wurtz, R. H. (1988a). Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *Journal of Neurophysiology, 60*(2), 580–603.

Komatsu, H., & Wurtz, R. H. (1988b). Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *Journal of Neurophysiology, 60*(2), 621–644.

Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience, 20,* 3310–3318.

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

Li, W., & Matin, L. (1990). The influence of saccade length on the saccadic suppression of displacement detection. *Perception and Psychophysics, 48,* 453–458.

Lisberger, S. G., Morris, E. J., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Review of Neuroscience, 10,* 97–129.

Lopez, C., & Blanke, O. (2011). The thalamocortical vestibular system in animals and humans. *Brain Research Reviews, 67*(1–2), 119–146.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*(6657), 279–281.

Mack, A., & Herman, E. (1973). Position constancy during pursuit eye movements: An investigation of the Filehne illusion. *Quarterly Journal of Experimental Psychology,* 25, 71–84.

Matin, L., & Pearce, D. G. (1965). Visual perception of direction for stimuli flashed during voluntary saccadic eye movements. *Science, 148*(3676), 1485–1488.

McLaughlin, S. C. (1967). Parametric adjustment in saccadic eye movements. *Perception and Psychophysics, 2,* 359–362.

Melcher, D., & Kowler, E. (2001). Visual scene memory and the guidance of saccadic eye movements. *Vision Research, 41*(2001), 3597–3611.

Melmoth, D. R., Kukkonen, H. T., Mäkelä, P. K., & Rovamo, J. M. (2000). The effect of contrast and size scaling on face perception in foveal and extrafoveal vision. *Investigative Ophthalmology & Visual Science, 41*(9), 2811–2819.

Merriam, E. P., Genovese, C. R., & Colby, C. L. (2007). Remapping in human visual cortex. *Journal of Neurophysiology, 97,* 1738–1755.

Morgan, M. J., & Ward, R. (1980). Conditions for motion flow in dynamic visual noise. *Vision Research, 20*(5), 431–435.

Morrone, M. C. (2014). Interactions between eye movements and vision: Perception during saccades. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences* (pp. 947–962). Cambridge, MA: MIT Press.

Morrone, M. C., Ross, J., & Burr, D. C. (1997). Apparent position of visual targets during real and simulated saccadic eye movements. *Journal of Neuroscience, 17*(20), 7941–7953.

Müller, R., & Greenlee, M. W. (1994). Effect of contrast and adaptation on the perception of the direction and speed of drifting gratings. *Vision Research, 34*(16), 2071–2092.

Naka, K. I., & Rushton, W. A. (1966). S-potentials from colour units in the retina of fish (Cyprinidae). *Journal of Physiology, 185*(3), 536–555.

Nakayama, K. (1985). Biological image motion processing: A review. *Vision Research, 25,* 625–660.

Newsome, W. T., Britten, K. H., Salzman, C. D., & Movshon, J. A. (1990). Neuronal mechanisms of motion perception. *Cold Spring Harbor Symposia on Quantitative Biology, 55,* 697–705.

Noto, C. T., Watanabe, S., & Fuchs, A. F. (1999). Characteristics of simian adaptation fields produced by behavioral changes in saccade size and direction. *Journal of Neurophysiology, 81*(6), 2798–2813.

Ohlendorf, S., Sprenger, A., Speck, O., Glauche, V., Haller, S., & Kimmig, H. (2010). Visual motion, eye motion, and relative motion: A parametric fMRI study of functional specializations of smooth pursuit eye movement network areas. *Journal of Vision, 10*(14):21, 1–15.

Ohlendorf, S., Sprenger, A., Speck, O., Haller, S., & Kimmig, H. (2008). Optic flow stimuli in and near the visual field centre: A group fMRI study of motion sensitive regions. *PLoS ONE, 3,* e4043.

Ordy, J. M., Massopust, L. C., & Wolin, L. R. (1962). Postnatal development of the retina, electroretinogram, and acuity in the Rhesus monkey. *Experimental Neurology, 5,* 364–382.

Østerberg, G. (1935). Topography of the layer of rods and cones in the human retina. *Acta Ophthalmologica Kbh, 61,* 1–102.

Pasternak, T., & Greenlee, M. W. (2005). Working memory in primate sensory systems. *Nature Reviews Neuroscience, 6*(2), 97–107.

Pelisson, D., Alahyane, N., Panouilleres, M. P., & Tilikete, C. (2010). Sensorimotor adaptation of saccadic eye movements. *Neuroscience and Biobehavioral Reviews, 34*(8), 1103–1120.

Pelli, D. G. (2008). Crowding: A cortical constraint on object recognition. *Current Opinion in Neurobiology, 18*(4), 445–451.

Perry, R. J., & Zeki, S. (2000). The neurology of saccades and covert shifts in spatial attention. *Brain, 123*(11), 2273–2288.

Pierrot-Deseilligny, C., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of memory-guided saccades in man. *Experimental Brain Research, 83,* 607–617.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psychology, 32*(1), 3–25.

Quessy, S., Quinet, J., & Freedman, E. G. (2010). The locus of motor activity in the superior colliculus of the rhesus monkey is unaltered during saccadic adaptation. *Journal of Neuroscience, 30,* 14235–14244.

Raabe, M., Fischer, V., Bernhardt, D., & Greenlee, M. W. (2013). Neural correlates of spatial working memory load in a delayed match-to-sample saccade task. *NeuroImage, 71,* 84–91.

Rashbass, C. (1961). The relationship between saccadic and smooth tracking eye movements. *Journal of Physiology, 159,* 326–338.

Rentschler, I., & Treutwein, B. (1985). Loss of spatial phase relationships in extrafoveal vision. *Nature, 313,* 308–310.

Reppas, J. B., Usrey, W. M., & Reid, R. C. (2002). Saccadic eye movements modulate visual responses in the lateral geniculate nucleus. *Neuron, 35*(5), 961–974.

Reuter-Lorenz, P. A., Hughes, H. C., & Fendrich, R. (1991). The reduction of saccadic latency by prior offset of the fixation point: An analysis of the gap effect. *Perception and Psychophysics, 49*(2), 167–175.

Richards, W. (1969). Saccadic suppression. *Journal of the Optical Society of America, 59,* 617–623.

Riggs, L. A., Armington, J. C., & Ratliff, F. (1954). Motions of the retinal image during fixation. *Journal of the Optical Society of America, 44,* 315–321.

Riggs, L. A., & Schick, A. M. L. (1968). Accuracy of retinal image stabilization achieved with a plane mirror on a tightly fitting contact lens. *Vision Research, 8,* 159–169.

Ross, J., Morrone, M. C., Goldberg, M. E., & Burr, D. C. (2001). Changes in visual perception at the time of saccades. *Trends in Neurosciences, 24*(2), 113–121.

Rovamo, J., Virsu, V., & Nätänen, R. (1978). Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature, 271,* 54–56.

Schlag, J., & Schlag-Rey, M. (1995). Illusory localization of stimuli flashed in the dark before saccades. *Vision Research, 35*(16), 2347–2357.

Schütz, A. C., Braun, D. I., & Gegenfurtner, K. R. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5), 9. https://doi.org/10.1167/11.5.9.

Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science, 20*(2), 207–214.

Smith, A. T., Greenlee, M. W., Singh, K. D., Kraemer, F. M., & Hennig, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *Journal of Neuroscience, 18*(10), 3816–3830.

Smith, A. T., Singh, K. D., & Greenlee, M. W. (2000). Attentional suppression of activity in the human visual cortex. *NeuroReport, 11*(2), 271–277.

Smith, A. T., Williams, A. L., & Singh, K. D. (2004). Negative BOLD in the visual cortex: Evidence against blood stealing. *Human Brain Mapping, 21*(4), 213–220.

Sneve, M. H., Alnaes, D., Endestad, T., Greenlee, M. W., & Magnussen, S. (2012). Visual short-term memory: Activity supporting encoding and maintenance in retinotopic visual cortex. *NeuroImage, 63*(1), 166–178.

Sommer, M. A., & Wurtz, R. H. (2002). A pathway in primate brain for internal monitoring of movements. *Science (New York, NY), 296*(5572), 1480–1482.

Spering, M., Kerzel, D., Braun, D. I., Hawken, M. J., & Gegenfurtner, K. R. (2005). Effects of contrast on smooth pursuit eye movements. *Journal of Vision, 5*(5), 455–465.

Spering, M., Pomplun, M., & Carrasco, M. (2011). Tracking without perceiving: A dissociation between eye movements and motion perception. *Psychological Science, 22*(2), 216–225.

Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology, 43*(6), 482–489.

Stark, L., Kong, R., Schwartz, S., Hendry, D., & Bridgeman, B. (1976) Saccadic suppression of image displacement. *Vision Research*, *16*, 1185–1187.

Steenrod, S. C., Phillips, M. H., & Goldberg, M. E. (2013). The lateral intraparietal area codes the location of saccade targets and not the dimension of the saccades that will be made to acquire them. *Journal of Neurophysiology, 109*(10), 2596–2605.

Steinbach, M. J. (1976). Pursuing the perceptual rather than the retinal stimulus. *Vision Research, 16,* 1371–1376.

Stiles, W. S. & Crawford, B. H. (1933) the luminous efficiency of the eye pupil at different points. *Proceedings of the Royal Society of London. Series B, 112*, 428–450.

Stone, L. S., & Thompson, P. (1992). Human speed perception is contrast dependent. *Vision Research, 32*(8), 1535–1549.

Stone, L. S., Beutter, B. R., & Lorenceau, J. (2000). Visual motion integration for perception and pursuit. *Perception, 29*(7), 771–787.

Strasburger, H., Harvey, L. O., & Rentschler, I. (1991). Contrast thresholds for identification of numeric characters in direct and eccentric view. *Perception and Psychophysics, 49,* 495–508.

Sunaert, S., Van Hecke, P., Marchal, G., & Orban, G. A. (1999). Motion-responsive regions of the human brain. *Experimental Brain Research, 127*(4), 355–370.

Sylvester, R., Haynes, J. D., & Rees, G. (2005). Saccadic eye movements modulate visual responses in the lateral geniculate nucleus. *Current Biology, 15,* 37–41.

Thiele, A., Henning, P., Kubischik, M., & Hoffmann, K. P. (2002). Neural mechanisms of saccadic suppression. *Science, 295*(5564), 2460–2462.

Thompson, P. (1983). Discrimination of moving gratings at and above detection threshold. *Vision Research, 23,* 1533–1538.

Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature, 428,* 751–754.

Tootell, R. B., Reppas, J. B., Kwong, K. K., Malach, R., Born, R. T., Brady, T. J., et al. (1995). Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *Journal of Neuroscience, 15*(4), 3215–3230.

Ungerleider, L. G., Courtney, S. M., & Haxby, J. V. (1998). A neural system for human visual working memory. *Proceedings of the National Academy of Sciences of the United States of America, 95*(3), 883–890.

Vallines, I., & Greenlee, M. W. (2006). Saccadic suppression of retinotopically localized blood oxygen level-dependent responses in human primary visual area V1. *Journal of Neuroscience, 26*(22), 5965–5969.

van Essen, D. C., & Gallant, J. L. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron, 13,* 1–10.

Volkmann, F. C., Schick, A. M., & Riggs, L. A. (1968). Time course of visual inhibition during voluntary saccades. *Journal of the Optical Society of America, 58,* 562–569.

von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip: Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Die Naturwissenschaften, 20*, 464–476.

Watamaniuk, S. N., & Heinen, S. J. (1999). Human smooth pursuit direction discrimination. *Vision Research, 39*(1), 59–70.

Whitney, D., & Levi, D. M. (2011). Visual crowding: A fundamental limit on conscious perception and object recognition. *Trends in Cognitive Sciences, 15*(4), 160–168.

Williams, D. W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research, 24*(1), 55–62.

Wright, M. J., & Gurney, K. N. (1992). Lower threshold of motion for one and two dimensional patterns in central and peripheral vision. *Vision Research, 32,* 121–134.

Zeki, S. M. (1978). Functional specialisation in the visual cortex of the rhesus monkey. *Nature, 274,* 423–428.

Zuber, B. L., & Stark, L. (1966). Saccadic suppression: Elevation of visual threshold associated with saccadic eye movements. *Experimental Neurology, 16,* 65–79.

Chapter 6
Scenes, Saliency Maps and Scanpaths

Tom Foulsham

Contents

T. Foulsham (✉)
Department of Psychology, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, UK
e-mail: foulsham@essex.ac.uk

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_6

Abstract The aim of this chapter is to review some of the key research investigating how people look at pictures. In particular, my goal is to provide theoretical background for those that are new to the field, while also explaining some of the relevant methods and analyses. I begin by introducing eye movements in the context of natural scene perception. As in other complex tasks, eye movements provide a measure of attention and information processing over time, and they tell us about how the foveated visual system determines what to prioritise. I then describe some of the many measures which have been derived to summarize where people look in complex images. These include global measures, analyses based on regions of interest and comparisons based on heat maps. A particularly popular approach for trying to explain fixation locations is the saliency map approach, and the first half of the chapter is mostly devoted to this topic. A large number of papers and models are built on this approach, but it is also worth spending time on this topic because the methods involved have been used across a wide range of applications. The saliency map approach is based on the fact that the visual system has topographic maps of visual features, that contrast within these features seems to be represented and prioritized, and that a central representation can be used to control attention and eye movements. This approach, and the underlying principles, has led to an increase in the number of researchers using complex natural scenes as stimuli. It is therefore important that those new to the field are familiar with saliency maps, their usage, and their pitfalls. I describe the original implementation of this approach (Itti & Koch, 2000), which uses spatial filtering at different levels of coarseness and combines them in an attempt to identify the regions which stand out from their background. Evaluating this model requires comparing fixation locations to model predictions. Several different experimental and comparison methods have been used, but most recent research shows that bottom-up guidance is rather limited in terms of predicting real eye movements. The second part of the chapter is largely concerned with measuring eye movement scanpaths. Scanpaths are the sequential patterns of fixations and saccades made when looking at something for a period of time. They show regularities which may reflect top-down attention, and some have attempted to link these to memory and an individual's mental model of what they are looking at. While not all researchers will be testing hypotheses about scanpaths, an understanding of the underlying methods and theory will be of benefit to all. I describe the theories behind analyzing eye movements in this way, and various methods which have been used to represent and compare them. These methods allow one to quantify the similarity between two viewing patterns, and this similarity is linked to both the image and the observer. The last part of the chapter describes some applications of eye movements in image viewing. The methods discussed can be applied to complex images, and therefore these experiments can tell us about perception in art and marketing, as well as about machine vision.

By the end of the chapter, readers should

- Understand why eye movements are useful for studying natural scene perception.
- Understand some of the measures used for quantifying fixations on images.
- Appreciate the theoretical and neural underpinnings of a saliency map approach.
- Understand the Itti and Koch (2000) model of bottom-up visual saliency.
- Be able to evaluate saliency map models using eye fixation data.
- Appreciate temporal aspects of eye movements in scene PERCEPTION, including fixation duration and order.
- Understand scanpaths and why they have been studied.
- Appreciate comparison methods which look at the scanpath sequence.
- Be familiar with some of the applications of eyetracking experiments in images.

6.1 Introduction

Eye movements are a fundamental part of natural human vision. This is particularly true when we consider complex natural scenes, which comprise more information than we can take in within a single glance. Due to our sampling of the visual field, which is dominated by high acuity at the fovea and decreased resolution everywhere else on the retina, we can only inspect our environment in detail by moving our eyes and bodies to select different regions of interest. This process of actively selecting information gives psychology and neuroscience a uniquely sensitive measure about how people perceive and understand images. However, it also creates difficulties for a visual brain which has to rapidly orient the eyes based on only peripheral information, and then combine the input from multiple fixations so that we can understand a scene and act accordingly.

This chapter describes some of the theory and methods used in the study of eye movements in complex stimuli. I will focus on two particular sets of theoretical questions within this topic, which are related to *saliency maps* and *scanpaths*. However, this chapter could easily be called "Looking at pictures", because the research and methods being discussed are those where we measure people looking at pictures and photographs of scenes. Broadly speaking, the research to be discussed tries to describe *where* people look and *in what order*. Explaining these two things is a complex problem, but it should be easy to see that doing so will involve both the visual appearance of items in the scene (e.g., how bright or colourful something is) and the knowledge or task of the observer. It is worth bearing in mind, from the outset, that these depictions of the world are convenient abstractions for experimenters, but that they may not always reflect the way that we move our eyes in the real environment.

When investigating the viewing of static images, most researchers analyse saccades and fixations (for background on the properties of saccades, see chapters by Pierce et al. and Hutton, this volume). Although scene viewing may elicit other eye movement events, such as fixational eye movements and microsaccades (see Alexander & Martinez-Conde, this volume), saccades are the main way in which we redirect our eye to select particular items. Saccades are easily identifiable from a record of eye position samples, because they have a distinct velocity profile such that the eye rapidly accelerates to a peak velocity of about 500°/s. During saccades, vision is suppressed (see Greenlee & Kimmig, in this volume), and so the processing of visual information takes place largely during fixations, where the eye is relatively still. It is therefore assumed that the location and duration of fixations reflects what is being processed at a given moment in time. This assumption relies on a tight link between overt and covert attention. It also neglects the fact that saccades take at least 100 ms to prepare, which means that at least part of the time during a fixation is devoted to saccade programming. However, in complex stimuli it generally makes sense to talk about attention and fixation as synonymous (consistent with the Active Vision approach: Findlay & Gilchrist, 2003).

Researchers can use the measurement of saccades and fixations in complex scenes to answer many different questions about visual attention and information processing. The top panel of Fig. 6.1 shows the series of fixations and saccades made by a single person viewing a complex scene (despite this sort of diagram, when real saccades are measured with high precision they are not straight and often show curvature). It is clear when we repeat this for many people looking at the same image (see Fig. 6.1, bottom panel) that there are some consistent patterns in *where* people look. One set of questions, therefore, concerns how to represent and predict these patterns. The next section considers a widely used approach to these questions: the saliency map approach.

A related but distinct set of questions concerns the *order* in which people attend to different elements of a picture. Investigating this order requires methods for comparing and manipulating sequential "scanpaths", the topic of both classic and contemporary research which is discussed in the second half of the chapter.

6.2 Historical Annotations

Although saccades and the "path" followed through an image were familiar to early researchers from introspection and observation, it was not until the 20th century that eyetrackers were used to measure this with any precision. Researchers often chose to focus on simpler and more controlled experiments, and it was only later that improvements in technology and computer vision techniques led to more experiments with complex images.

Dodge and Cline (1901) are often credited with developing the first optical eyetracker, which laid the foundation for tracking based on corneal reflections.

Buswell (1935) used an improved eyetracker at the University of Chicago to measure fixations and saccades in two-dimensional images. His observations helped to define the enduring questions of how eye movements were related to image content and viewer cognition.

Fig. 6.1 When someone views a scene, they make a series of eye movements (top panel). These consist of fixations (circles) and saccades (arrows). The fixations have a position, a sequential order and a duration (values in milliseconds). Combining the fixation locations across many observers reveals variability, but also clustering on certain regions (bottom panel)

Yarbus (1967) used a suction cup to record eye movements over a variety of pictures, emphasizing that the active sequence of eye movements changed with the viewer's task.

Noton and Stark (1971) defined the term "scanpath" and incorporated these fixation sequences into a detailed computational model which foreshadowed the later focus on eye movements and embodiment (which links perception to physical and motor states).

Mackworth and Thomas (1962) developed one of the first mobile eyetrackers. Mackworth went on to help conduct several early and important studies on eye movements in scene perception (Loftus & Mackworth, 1978; Mackworth & Morandi, 1967).

Land (1993) conducted several pioneering experiments using mobile eyetracking which placed a new emphasis on the importance of action beyond just "looking at pictures".

Itti, Koch, and Niebur (1998) and Itti and Koch (2000) presented a fully implemented model of visual saliency which could be applied to arbitrary images and tested with human fixation data.

Thanks to advances in eyetracking technology, computing and image processing, it has never been easier to measure fixation locations from an observer viewing a digital image. Modern research investigating how people look at pictures uses a range of devices, measures and models. This chapter reviews two general frameworks for analyzing and explaining the resulting data (saliency maps and scanpaths), and I begin the next section with a review of how we can quantify looking behaviour.

6.3 Spatial Analysis of Fixations: Saliency and Saliency Maps

*6.3.1 Answering the Question "*Where *Do People Fixate?"*

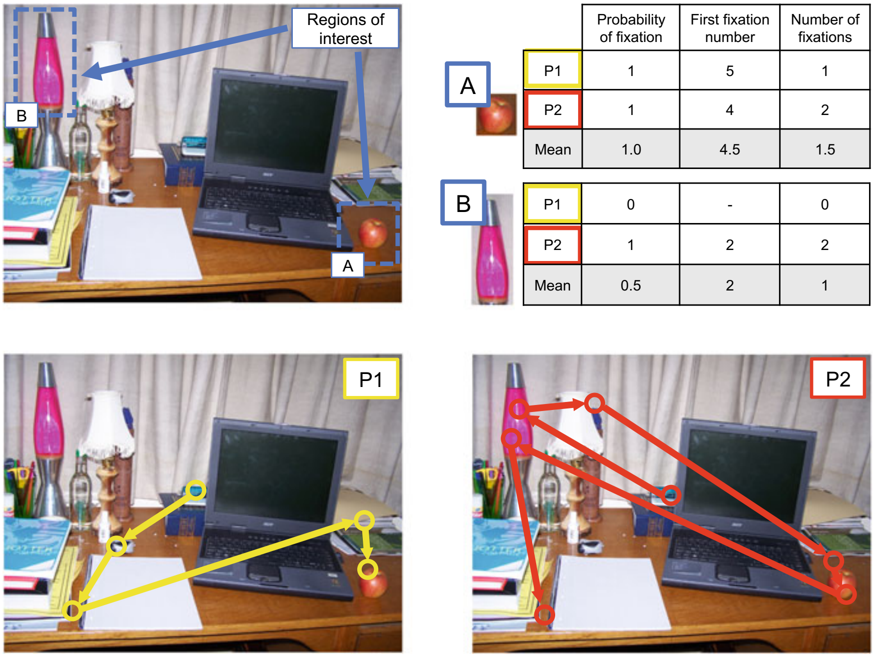
This section considers a particularly influential way of representing and explaining where people look in complex stimuli, with reference to the features in the image. Before discussing this approach in detail, it is worth specifying some of the different measures that researchers typically use to quantify where participants look. Table 6.1 describes some of these measures, their definitions and the way that they are commonly interpreted. These measures are not exhaustive, and different terms are sometimes used for the same measure. This means that it is important for researchers to clearly define their dependent variables when reporting results.

The measures in Table 6.1 can be aggregated to give a summary description of the fixation patterns from many participants and trials. Most modern eyetrackers come with software for automatically calculating these statistics. The process for deriving these measures is straightforward and consists of allocating each fixation to a region of interest, based on the (x, y) screen coordinates of the fixation. The

Table 6.1 Some of the most commonly used measures for quantifying where people look in images

Measure	Definition	Interpretation
Trial level measures		
Average saccade amplitude	The mean (or sometimes median) amplitude of all saccades made, in degrees of visual angle	Greater values = larger shifts between points of interest in a given scene
Fixation dispersion or "spread"	The standard deviation of all *x* and/or *y* fixation coordinates, in degrees of visual angle	Greater values = fixations which are more spread out in space
Region of interest measures		
Probability of fixation	A binary, yes/no variable representing whether or not a region has been fixated at least once in a viewing episode	Regions which are fixated have received more attention than those which are not
First fixation time/number	The time/ordinal fixation number at which the first fixation on a region occurs	Lower values = region which is fixated earlier and prioritised by attention
Number of fixations (on a region)	The total number of fixations landing within a region of interest	Greater values = more interest or attention devoted to this region
First fixation duration	The duration of the first fixation on a region of interest	Greater values = more extensive or elaborate processing of the region on the first look
First gaze duration	The sum duration of all fixations made on a region of interest on the first pass (i.e., before exiting this region)	Greater values = more extensive or elaborate processing of the region on the first look
Total gaze duration/inspection time	The sum duration of all fixations made on a region of interest, including refixations	Greater values = more extensive or elaborate processing of the region over an extended period

region of interest might be defined by a bounding box (most easily, a rectangle), or by a circle with a given centre and radius (thus including all fixations within a certain distance of that point). For example, Fig. 6.2 shows one of the stimuli used in Foulsham and Underwood (2007). In this picture, we identified two regions of interest: a key object (a piece of fruit) and a region containing the brightest, most salient features in the image. Figure 6.2 shows how three measures can be calculated for each person viewing this image (see also Table 6.1). The probability of fixation merely records whether a given area of interest has been fixated or not (1 or 0). When averaged across the two example participants, region A (the apple) has a fixation probability of 1.0 because it was fixated by 100% of the observers. Region B (the



		Probability of fixation	First fixation number	Number of fixations
A	P1	1	5	1
	P2	1	4	2
	Mean	1.0	4.5	1.5
B	P1	0	-	0
	P2	1	2	2
	Mean	0.5	2	1

Fig. 6.2 A region of interest analysis with two regions (*A* and *B*, see top left panel) and example eye movements from two participants, *P1* and *P2* (bottom panels). The top right panel shows some statistics for each region and participant. For example, region B is fixated by P2 but not by P1, giving a (mean) fixation probability of 0.5

lamp) has a fixation probability of 0.5 because only half of the observers looked at this region. The other measures in Fig. 6.2 show that region A is first inspected after 4.5 fixations, and attracted 1.5 fixations, on average. In Foulsham and Underwood (2007), the same regions were present across many pictures in the experiment, and because all participants viewed all these pictures, measures could be averaged across images and then across participants to derive a description of the general tendency for people to look at either of these key areas.

One might well ask why so many measures are necessary to represent where people look. In some cases, several measures will yield the same conclusions. This may be because the measures are formally identical (e.g., the total gaze duration and the sum of all the fixation durations); because they are logically related (e.g., making a greater number of fixations must also mean making a greater number of saccades); or because they are merely correlated (e.g., it is normally, but not always, the case that a greater number of fixations corresponds to a longer total gaze duration).

It is important to remember that whichever measures are used, they are attempting to condense an (often complex) set of spatiotemporal patterns. The actual measures to analyse, which will reflect the hypotheses of the particular study, should be chosen with care. Figure 6.3 depicts some of the questions that researchers should ask

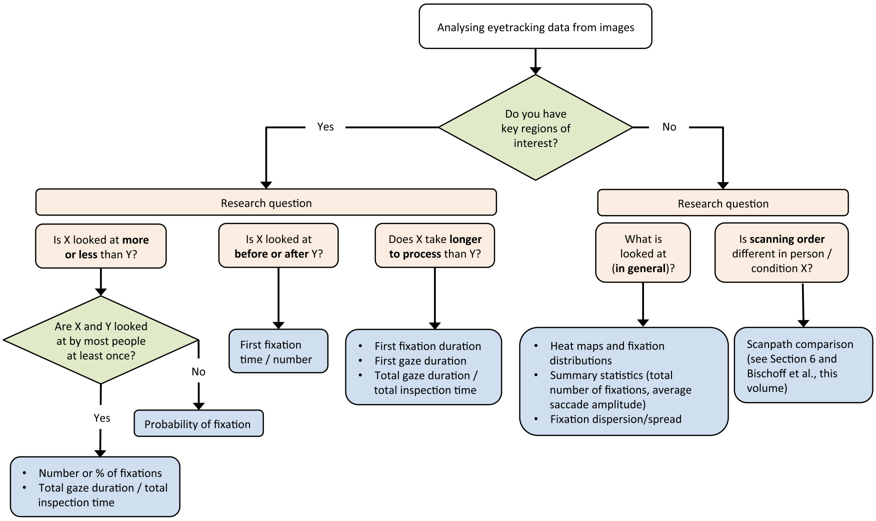


Fig. 6.3 A flowchart depicting some of the ways in which researchers can select eye movement measures that answer their research question

themselves when choosing which measures to use. Referring back to Fig. 6.2, we can see some of the subtleties in using these statistics to compare regions of interest. In one case, there is a region which is fixated by both participants (region A). How early and how frequently this region is inspected can be quantified by calculating the first fixation number and the number of fixations on this item. However, these measures may not make sense if a region is rarely fixated. Calculating the first fixation number is only possible if the region has been inspected at least once (and so there is a missing value for participant P1 looking at region B, see Fig. 6.2). Here it might be more useful to quantify the region fixation probability.

Now that we have discussed some measures of how much a particular region or object is attended to, we can consider how to explain this in terms of a representation of the features in the scene: a saliency map. This approach has been highly influential because it allows researchers to analyse complex natural scenes and relate them to eye movements in a principled way. The underlying image processing techniques are often freely available and relatively easy to use, which means that investigators in a range of fields can produce predictions for their stimuli. For example, users in marketing might wish to determine the saliency of an advertisement before measuring how often observers look at this item. Alternatively, a researcher might want to measure saliency so that they can control for the visual properties of two regions of interest in an experiment (meaning that any differences in how these are inspected must be due to cognitive factors). In the next sections, I discuss the theoretical and methodological background for the saliency map approach. This involves background work in the psychophysics and computational neuroscience of attention. Much of this material is technical, and there remain debates about how best to compute saliency, and whether it actually tells us anything about fixations. However, many of the concepts and methods involved will apply to any spatial model of eye movements in scenes.

6.3.2 The Concept of a Saliency Map Is Founded on Classic Theories of Attention

The term "saliency map" (or "salience map") has its roots in attempts to produce a computational model of visuospatial attention. Based on experiments investigating visual search, and under the framework of feature integration theory, Treisman and Gelade (1980) described the process by which an observer can select a single object amongst an array. When simple features define the target, such as when one has to find a red square among many blue squares, it "pops out" and is found very easily. It is straightforward to imagine a control mechanism which could code for the colour at each location, and filter out only the region where colour = "RED". But how might such a mechanism be implemented in the case of more complex "conjunction" targets, such as a red square amongst red circles and blue squares? The solution in Treisman's model was a "master map" which combined the different basic features (colours, orientations and intensities) which were present at each location. The effortful shifting of attention while looking for the target is then determined by scanning of the master map.

A master map which combines multiple features into an abstract representation of attentional priority has been a fixture in subsequent work on attention and search. In Wolfe's (1994) Guided Search model, an "activation map" carries out the same function, prioritising (i.e., ranking) those locations in a search task that are most likely to contain the target. The activation map is therefore the mechanism by which preattentive, parallel processing of basic features is combined with top-down control in order to guide attention. Koch and Ulman (1985), meanwhile, called the combined representation guiding attention a "saliency map". Koch and Ulman's conceptual paper described the saliency map as a topographical representation which could explain selective attention in a way that was plausible given primate neurophysiology. Importantly, the saliency map was seen as an "early" visual representation which was based mostly on simple visual features. The focus of attention was then determined by a "winner-take-all" process which selects the most salient location.

6.3.3 The Brain Represents Saliency by Integrating Features and Attention

The topographical organisation of neurons in the early visual system was well known to those theorizing about attention in the 1980s. Single cell recording in cortical areas such as V1 showed that neurons were highly spatially selective, coding for particular features, and providing the foundation for a set of basic feature maps. However, it subsequently became clear that elsewhere in the brain there are cells which respond differently to the same stimuli, depending on the current attentional priority. For example, the superior colliculus is crucially involved in the control of eye movements. The activity of cells in this part of the midbrain does not just depend on whether there is a visual stimulus in their receptive field. Instead, cell responses are enhanced when observers are planning to make a saccade to this stimulus (Goldberg & Wurtz, 1972). Thus, deploying attention and eye movements to a certain location increases activity in colliculus neurons coding for this location.

Beyond the general attentional modulation of visual responses, there is mounting evidence that frontal and parietal areas involved in the control of eye movements can be thought of as implementing a saliency map (see Treue, 2003). For example, microstimulation of the frontal eye fields is associated with increased responding in spatially selective regions of V4 (Moore & Armstrong, 2003). Thus, the neural activity integrates relevance—signified by the process of preparing an eye movement—with visual distinctiveness. Modern neuroscience has provided increasing evidence for spatial priority maps in frontal and parietal areas of the brain, as well as increasingly sophisticated discussions about how these are involved in the guidance of behaviour more generally (see Bisley & Goldberg, 2010; Zelinsky & Bisley, 2015, for recent reviews).

It should be clear from this background research that the term saliency map has been used in multiple, overlapping ways: as an abstract, master map for attentional priority; as a neural mechanism for combining visual activity; as a bottom-up predictor of where people will look; and as any heat map type representation of fixations (see Box 1). These terms have not always been applied consistently, and so it is important for researchers to provide a precise definition. However, the main focus of this section will be a particular style of computational model of visual saliency which has been widely used, and which provides an estimate of the bottom-up feature contrast in complex stimuli.

Box 1: Heat Maps and Fixation Distributions

It is often useful to represent the distribution of fixations across an image by presenting a heat map: a spatial density plot showing how frequently each part of the picture has been inspected. The discrete fixations are first transformed into a continuous distribution, often by convolving a binary map with a symmetrical Gaussian function. This can also be thought of as iteratively adding a “blob” at the location of each fixation. The standard deviation (σ) of the Gaussian affects the granularity of the heat map (Fig. 6.4), and should be chosen to reflect error in eye position and the size of the fovea (e.g., 1° of visual angle). The result can be plotted as an “attentional landscape” with “peaks” or “hotspots” showing the places which are fixated most often (Fig. 6.5).

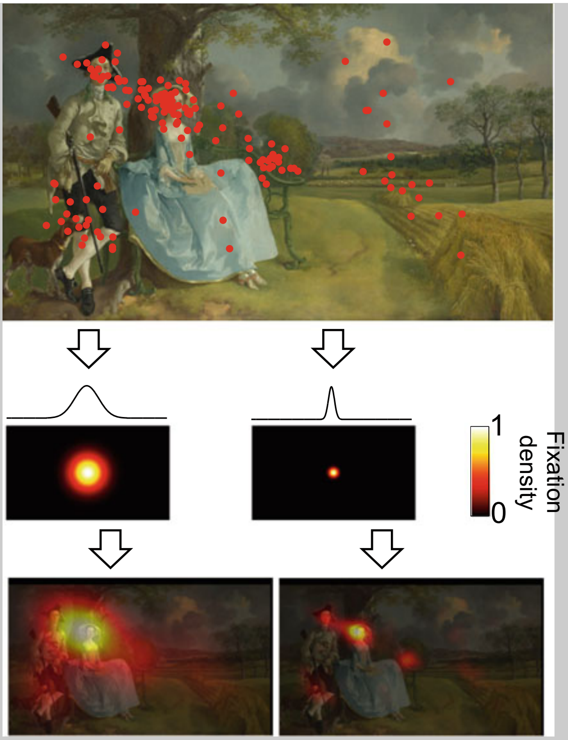


Fig. 6.4 The same fixations, convolved with two differently sized functions. The resulting heatmaps are plotted over the original image

There are a number of somewhat arbitrary factors that can be changed when making such maps. As well as σ, some researchers represent fixation duration (by scaling the height of the Gaussian), and some may produce heat maps for each participant which are then averaged. Wooding (1995) and Le Meur and Baccino (2013) describe heatmaps in more detail.

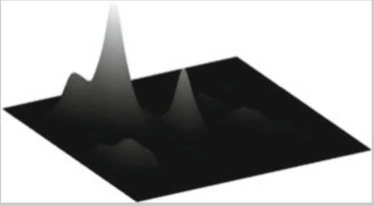


Fig. 6.5 An attentional landscape from the same fixations in Fig. 6.4

6.3.4 The Itti and Koch Saliency Map Provides a Model of Bottom-up Eye Guidance

Why is it that some regions are inspected more than others? One possibility is that there might be a set of visual features, which can be identified before the planning of an eye movement, and which signal those regions that should be prioritised for attention and fixation. A likely candidate feature is *contrast*, in the general sense that something which stands out from its background might be worth looking at, just as a feature search target pops out in simple visual search. This idea underlies the concept of visual saliency that was explicitly modelled by Itti, Koch and Niebur (1998), and then applied to eye movements by Itti and Koch (2000, 2001).

Itti and Koch built on the work of Koch and Ulman (1985), who were inspired by both human psychophysics and primate physiology, and implemented a computational model which combined bottom-up visual features using image processing. Unlike the simple case of searching for a red square amongst blue squares, the problem of extracting and combining features from a natural scene quickly becomes complex. Itti and Koch (2000) proposed a series of steps, implemented and released as a programming toolkit, which could take any arbitrary digital image as input (see Fig. 6.6 for an example, and http://ilab.usc.edu for downloads and software).

The first step extracts basic visual features at a range of spatial scales. Each feature is associated with a map, with the value at each point in the map representing the presence of that feature at that location. In static images, colour, intensity and orientation features are extracted using spatial filters (whereas flicker and motion channels are also available in dynamic stimuli). For example, the intensity map simply represents the amount of light at each point in the image, with black regions having low intensity and very bright regions having high intensity. Colour features are combined to compute colour opponency, while the orientation channel is assembled from edge detectors of different orientations. In order to identify such features at both coarse and fine scales, maps are derived from progressively sub-sampled versions of the image.

The second step is to combine the resulting scaled feature maps in a way which highlights feature contrast. This is accomplished through a centre-surround arrangement which pits fine and coarse feature maps against each other. The result is a "conspicuity map" for each feature, where the locations with the strongest activation are those with features which stand out from the surrounding background. An important consideration for the particular computational implementation at this point is how to combine different features. Should something colourful be given the same priority as something which is brighter than its background? Although several solutions to this are possible, in Itti & Koch's implementation there is normalisation and between-feature competition, such that if there is greater contrast in one feature dimension then it will be emphasised at the expense of features with lower contrast. After normalisation, the conspicuity maps are added together to give an overall saliency map, with "hotspots" showing the most salient regions with the highest feature contrast.

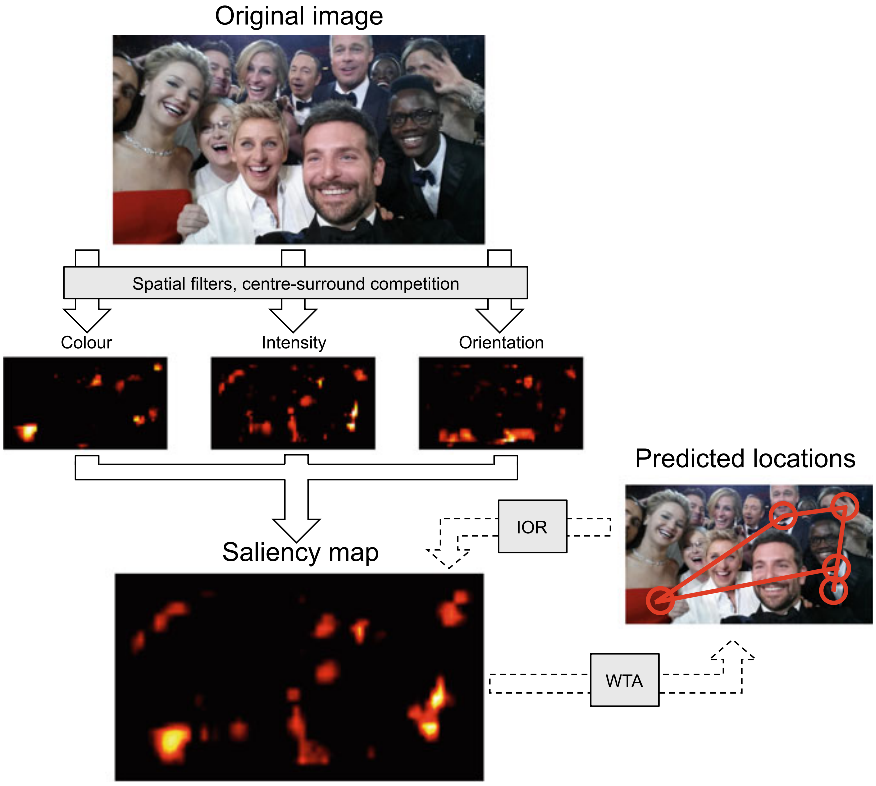


Fig. 6.6 An example of applying the Itti and Koch (2000) saliency map model to a complex image, the "most-tweeted photo ever" from the 2014 Academy Awards. The image is analysed within three different feature channels which are combined in a centre-surround fashion to highlight feature contrast. The resulting saliency map predicts a series of attended locations through a winner-take-all (WTA) process with inhibition of return (IOR). Salient locations include the black-and-white contrast of a tuxedo and a conspicuous red dress

The last step of the model, which makes it particularly suited to applying to eye movement research, is that it uses the saliency map to make explicit predictions about the locations which will be selected by covert and overt attention. Attention moves to the most salient location via a winner-take-all network. The saliency of this location is then suppressed (in a way similar to "inhibition of return", the mechanism proposed by Posner, Rafal, Choate, and Vaughan, (1985), for explaining delayed orienting to previously attended locations). This allows the focus of attention to shift to the next most salient region.

The Itti and Koch saliency model has been very widely cited and applied to many different problems in human vision and computer science (see Borji & Itti, 2013, for a recent review). Its success can be attributed to the fact that it produces real, tractable predictions for any arbitrary visual stimulus, with the algorithms for producing these predictions freely available. I will now describe some of the practical considerations involved in computing a saliency map, before discussing the experiments that have been carried out to test this model with human eye movements.

At this point it is important to note that there are now many different "saliency map models". In the almost 20 years since the original Itti and Koch (2000) model was proposed, it has been regularly revised and improved. Several different implementations have been released, which may differ in both algorithmic detail and actual predictions. Other researchers have proposed different underlying features, or tried to incorporate aspects such as depth or motion. In many cases it is also possible to incorporate some "top-down" modulation, where the model learns or is given information about which features are important (for example by adding a face detector, or training a model to look for a certain object). All of this means that it is important for researchers to be specific about what model and implementation is being used, and with which settings.

Table 6.2 describes some of the related, bottom-up models which have been published, with particular emphasis on those which are available to download (for a much more exhaustive list, see Borji & Itti, 2013; for model evaluations, see Kumerer et al., 2015). Critically, many of the methods I will describe next can be applied to any of these models. Moreover, many of the principled criticisms that have been levied at the saliency map approach are problematic for *any* model which is based solely on bottom-up features.

6.3.5 Guidelines for Eye Movement Researchers Computing a Saliency Map

It is clear from Table 6.2 that there are a bewildering number of models available, and this is not the place for considering them all in detail. One of the major contributions of the modelling community working on saliency maps is that there is a large amount of open data and code available. I recommend that interested readers try out some of the software available. Although the original version requires some knowledge of command line programming and C++, several user-friendly toolboxes have also been developed. For example, Fig. 6.6 was created using Version 3.0 of the Saliency Toolbox (Walther & Koch, 2006). The Graph-Based Visual Saliency implementation by Harel, Koch, and Perona (2006) also provides an implementation of saliency which is relatively easy for beginners to try. Both toolboxes run within MATLAB (Mathworks) and provide a set of functions, documentation and a graphical user interface. In this section I will offer some brief guidelines for eye movement researchers wishing to use such functions with complex images.

The first step in generating saliency map predictions is to load in a digital image for the stimulus. One way of thinking about this simulation is that the model should receive the same visual input as the human observer. Thus the image used should be the same as that seen by participants. The size of the image is one of many things which can change the specific output of the model, as well as affecting the length of

Table 6.2 A selection of "saliency models" which build on the ideas from Itti and Koch (2000)

Model	Noteworthy features	URL
Itti et al. (1998; Itti & Koch, 2000) saliency map model	Original model implemented in C++ as part of the iNVT toolkit	http://ilab.usc.edu/toolkit/
Walther and Koch (2006) saliency toolbox	Implements Itti model but also aims to identify and parse "proto objects"	http://www.saliencytoolbox.net
Harel et al. (2006) graph-based visual saliency	Uses graphical models to identify conspicuous regions. Code also includes Itti model	http://www.vision.caltech.edu/~harel/share/gbvs.php
Itti and Baldi (2005) Bayesian model of surprise	Defines saliency mathematically, according to change in prior beliefs. Tested with eye movements in video	http://ilab.usc.edu/surprise/
Bruce and Tsotsos (2005) AIM: Attention based on Information Maximization	Uses measure from information theory to define salient regions. Applied to visual search (with top-down information)	http://www.cs.umanitoba.ca/~bruce/datacode.html
Le Meur, Le Callet, Barba, and Thoreau (2006)	Using Itti model as a baseline, introduces more detailed stages to mimic human visual system	Code not readily available, but see http://people.irisa.fr/Olivier.Le_Meur/
Zhang, Tong, Marks, Shan, and Cottrell (2008) SUN: Saliency Under Natural statistics	Defines local saliency in Bayesian terms. Combines this with top-down information	http://cseweb.ucsd.edu/~l6zhang/code/imagesaliency.zip
Judd, Ehinger, Durand, and Torralba (2009)	Introduces a widely-used eyetracking dataset and uses machine learning to identify low-level features at fixation (as well as mid- and high-level features such as horizon and face detectors)	http://people.csail.mit.edu/tjudd/WherePeopleLook/index.html
Vig et al. (2014) Ensembles of Deep Networks	Uses "deep learning" to learn the optimal bottom-up features for fixation. Currently one of the best performing models (Kummerer et al., 2015)	http://coxlab.org/saliency/

time taken for the simulation to be complete. Some models may treat grayscale and colour images differently, and it is useful to know how the images are encoded (e.g., as an RGB image).

Next, the image is passed to the saliency functions for analysis. In some cases the processing involved is considerable, meaning that it can take some time. Typically, model output includes a continuous map of the image, where each pixel represents the saliency of the corresponding point in the image. This map can be displayed as a heatmap. However, it is important to be clear how this map is scaled, and the colours that are used in displaying it (for example, how is the minimum or maximum value in the map displayed?).

It is particularly important to understand how the images are scaled so that one can compare the saliency map output to objects in the image or fixated locations (see next section). It is also important to understand that the model essentially pits different parts of an image against each other, calculating the *relative* saliency of each region. Thus, if the research question requires comparing between different objects, both objects need to be present in the same image. It may not be straightforward to compare the saliency maps of two different images, particularly if they show very different variances in terms of the features present and their spatial distribution.

This description was written as a practical guide for people who want to produce saliency-based predictions for a particular image. However, as we shall see in subsequent sections, the saliency map model is not without its critics. Moreover, it is regrettable that there are multiple free parameters which can be changed, many of which are not specified by authors using this software, which may lead to problems replicating the results. It is therefore useful for researchers to experiment with different settings and be clear which differences in saliency are robust and which are highly sensitive to changes in model parameters.

6.3.6 Saliency Map Models Can Be Tested in Multiple Ways, but None Are Perfect

The Itti and Koch (2000) model was billed as a mechanism for shifts of overt and covert attention, making it ideal for producing bottom-up predictions for eye movements in complex stimuli. In this section, I will describe some of the ways in which these predictions have been tested. To begin, we should bear in mind that the model produces both a continuous spatiotopic map, representing the saliency of each location, and a simulated series of attention shifts. Both of these outputs can be used to test the model directly, and these tests fall into three main groups.

The first type of analysis examines the strength of the activation in a feature or saliency map at each fixated location. If the saliency map is a good prediction of the regions which will be inspected, then fixations will select the locations with high values, and avoid those locations where saliency is low. To perform this analysis, we can take the saliency value from each fixated location and average it across multiple fixations and participants. The saliency map is often represented at a lower spatial scale (i.e., it is smaller and represents regions more coarsely than the original image), and so locations will need to be scaled appropriately so that the correspondence between points in the image and on the map is maintained. This is normally just a case of determining the size of the saliency map relative to the image and then scaling position coordinates accordingly. Rather than relying on exact fixation coordinates, one could also take all the values from a small region around fixation, perhaps within 1° of visual angle, compensating for potential errors in eye position measurement, and for the fact that the fovea takes in information from an extended area. These values could be averaged (or summed), or the maximum value could be taken. When evaluating saliency values, it is important to consider the way the map is scaled or normalized (e.g., what is the minimum and maximum value in the map, and how distributed are these values), and differences in the distribution of saliency can make it difficult to compare between different images. If the saliency map is conceptualized as a probability distribution, then all the values will be positive and sum to 1. Other implementations will have a fixed range (e.g., between 0 and 1, or between 1 and 255 which is common in 8-bit digital images). One way to take these into account is to calculate what Peters, Iyer, Itti, and Koch (2005) dub the normalized scanpath salience (NSS; see Fig. 6.7a). To compute this measure, the saliency map is normalized by subtracting the mean saliency across all locations and dividing by the standard deviation of saliency values. This produces a z-score, and thus shows how many standard deviations a particular location is above chance. Using a standard saliency model to predict fixations in outdoor images, Peters et al. reported a mean NSS of 0.69 which was far greater than that expected by chance (an NSS of 0).

The second type of analysis compares the overall distribution of many fixations with the saliency map, in a manner similar to a correlation. If the model is predicting where people look, then there should be a positive relationship between saliency and fixation density. Le Meur and Baccino (2013) give a useful summary of some of the steps and metrics used in this type of analysis, and they also provide some computer code for these analyses. One initial approach is to convert a list of fixation locations (e.g., those from multiple participants viewing over an extended period of time) into a continuous fixation density distribution. Such distributions can be represented as heatmaps, showing the relative frequency with which different regions are inspected (refer back to Box 1). Comparing a saliency map and a fixation distribution can be as simple as calculating a Pearson correlation coefficient between the two. If points with high saliency are also locations with a high density of fixations then the correlation between the two distributions will be positive. However, because the distributions involved may violate parametric assumptions, it is preferable to use a non-parametric

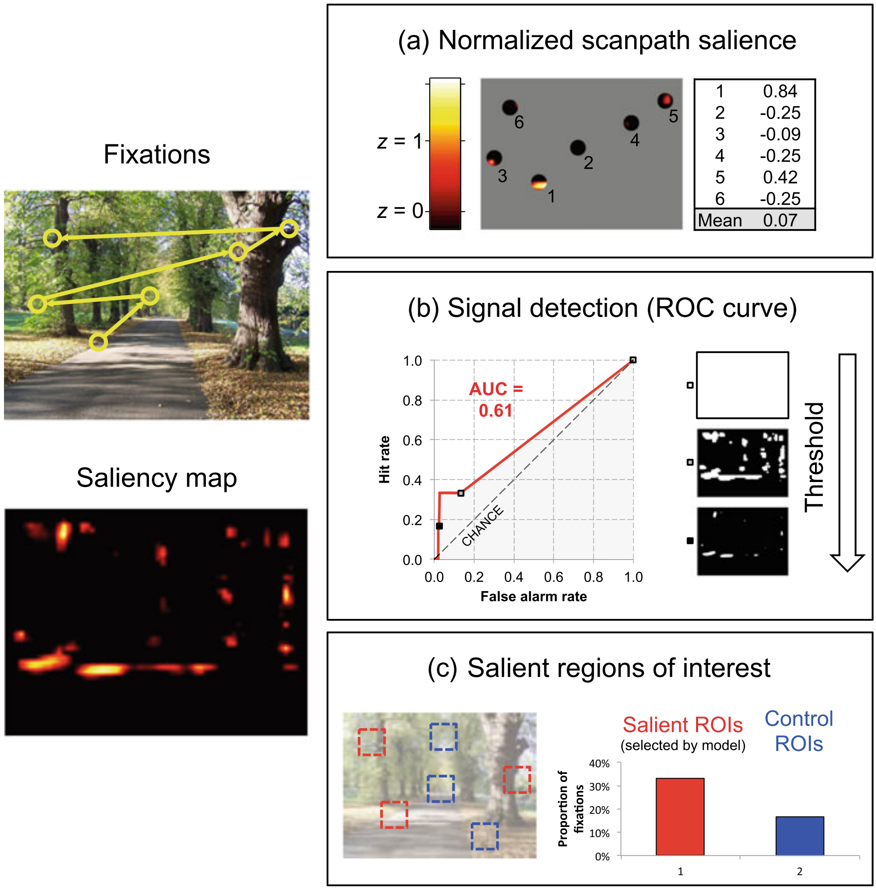


Fig. 6.7 An example of three different ways to compare fixations on a scene to a saliency map. In (**a**), we take the average of the z-transformed map value at each location. In (**b**), an ROC curve is plotted by applying a variable threshold to the map and observing the correctly predicted fixated locations (hit rate) and the false positives. Alternatively, salient regions of interest can be identified from those areas selected by the model, and compared to areas that are not (**c**). The methods are applied to the six example fixations in each case

method for comparing the probability distributions. One such metric is the Kullback-Liebler (K-L) divergence. The K-L divergence is a measure from information theory which quantifies the difference between two probability distributions. The result is a score—the number of bits—indicating how different the two distributions are, with a score of 0 indicating identical distributions. A better match between a saliency map and a fixation distribution would give a lower K-L divergence. This metric is discussed in detail by Tatler, Wade, Kwan, Findlay, and Velichkovsky (2005), who were also among the first to use the signal detection methods which have become standard in those evaluating saliency models.

The signal detection approach uses the saliency map to discriminate between fixated and non-fixated locations, by applying a threshold. Locations with saliency higher than the threshold are classified as fixated points, and the threshold is gradually increased, allowing the hits (correctly identified fixation locations) and false alarms (non-fixated locations classified as fixated) to be tallied. For example, at a low threshold many locations will be selected, leading to many false alarms (as well as some hits). At a high threshold, only the most salient locations will be classified as fixated, and if the saliency map is a very good predictor of fixation these will all be hits. Using the rates of hits and false alarms at each threshold, a receiver operating characteristics (ROC) curve is plotted, and the area under this curve (AUC) quantifies how well the saliency map can discriminate fixated locations (see Fig. 6.7b). This method is robust to differences in the distribution of saliency in fixated and non-fixated locations, and does not rely on parametric assumptions. Critically, because the ROC method is based on the ranks of all the points in the map (i.e., the point with the highest saliency, followed by the next highest, and so on), it is not affected by any monotonic scaling of the actual values. A saliency map which provides no information about fixated locations will lead to an AUC of 0.5. Another advantage of this method, widely used in computer vision, is that it can de-confound effect size and statistical significance. For example, Tatler, Baddeley, and Gilchrist (2005) report an AUC of 0.57 for the predictiveness of a luminance saliency map. This value was statistically very different from chance (according to bootstrapped confidence intervals), but it also leaves much of the difference between fixated and non-fixated locations unexplained (57% is by no means an impressively large effect).

These general analyses are useful for comparing fixations and saliency over an entire image in a theory-neutral way. However, often researchers are interested in certain objects or regions which have a particular significance based on theory or application. The third type of analysis uses the sequence of saliency-predicted shifts of attention to identify key regions in the image, and then determines how often these regions are fixated. For example, in Fig. 6.2 I described an experiment where "target" objects were made more or less salient by placing them in different visual settings. The relative saliency of target objects was determined via the model, according to the number of simulated shifts of attention that were made before they were selected. For example, one might classify a "high saliency" object as one which is selected by the model very early (on the first or second simulated fixation). Such objects could be compared to "low saliency" or "control" objects which were not selected by the model after ten shifts of attention. This approach has a number of advantages. First, it tests one of the key strengths of the saliency map model, as proposed by Itti and Koch (2000), which is that it predicts an actual sequence of fixations and not just a continuous map. Second, it makes it possible to compare between targets which have been matched in other ways (such as according to their semantic meaning by virtue of being the same type of object). As we shall see, the failure of correlational approaches to take semantics into account has led to difficulties in evaluating the saliency map model. In Foulsham and Underwood (2007), we found that in the absence of a strongly constrained task, objects (pieces of fruit) which were more salient according to the model were more likely to be fixated and were fixated earlier. Figure 6.7c gives another example of comparing the fixations on salient and non-salient regions.

When evaluating these analyses, we often have to compare the results to some kind of null hypothesis, indicating the relationship that we would expect if saliency does not predict fixation. For example, the NSS compares saliency at fixation with the average saliency across the whole image, and the AUC compares fixated locations with all other non-fixated locations in the image. However, because neither fixations nor saliency are uniformly distributed, comparing against a "chance" distribution which samples uniformly across locations is problematic. The issue is that such a comparison assumes that all parts of an image are equally likely to be fixated. As we shall see, this may have caused the role of saliency to be overestimated because fixations in the centre of an image are credited to salient features when in fact this is a generic spatial bias which is manifested regardless of the scene. Eye movements across many different images are systematically biased to particular spatial locations in a variety of ways (see Box 2). Whichever metric is used for comparing saliency models and fixations, the best approach is to select a comparison distribution which reflects the general spatial biases inherent in eye movements across images. For example, rather than comparing saliency at fixated locations to the average across the whole scene, they can be compared to values from a "shuffled" dataset which uses positions selected by human fixations in other images. This shuffled dataset thus reproduces the image-general spatial biases, ensuring that only predicting fixation patterns on a specific image is credited to the saliency or feature-based model. Alternatively, one can use a generic, non-uniform comparison distribution (such as that recommended by Clarke and Tatler (2014), which models the general central bias seen in image viewing).

Box 2: Systematic Patterns in Scene Viewing Are Not Random

Tatler and Vincent (2009) showed that we can predict fixation locations just by knowing how the eyes move in general. This is because, regardless of the image, observers show systematic biases towards certain locations and saccades. It is therefore important to investigate how these biases are related to image content and, if they are not, why they arise.

There is a strong central bias in fixations on a screen (Fig. 6.8). This is exacerbated by the practice of cueing participants with a central fixation cross, but may also reflect the eyes' "orbital reserve". Photographers also often place objects and items of interest in the centre.

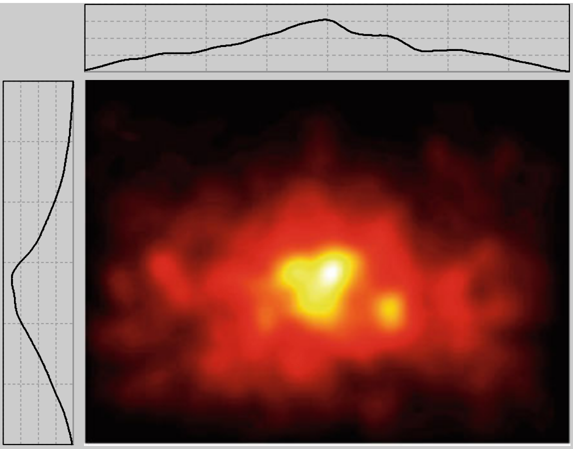


Fig. 6.8 Heat map and histograms showing the relative frequency of fixations across scene space

Saccades are also biased, occurring most often in horizontal directions (Fig. 6.9). This is true even in square images, and follows perception of the layout when the scene is rotated (Foulsham, Kingstone, & Underwood, 2008). There is also a small but consistent bias to make an initial saccade to the left (Fig. 6.10).

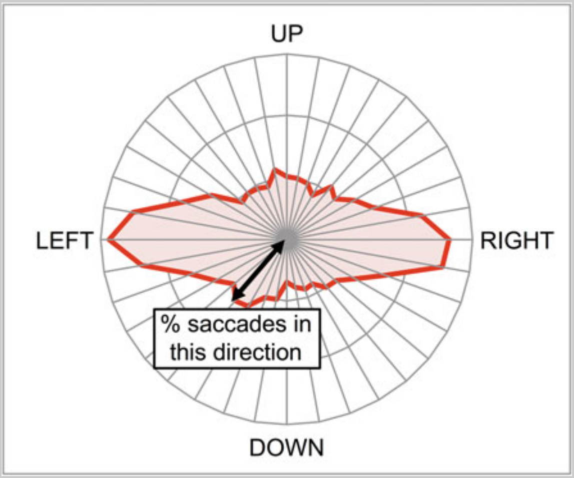


Fig. 6.9 Relative frequency of saccades in each direction (data in Figs 6.8 and 6.9 from Foulsham & Underwood, 2008)

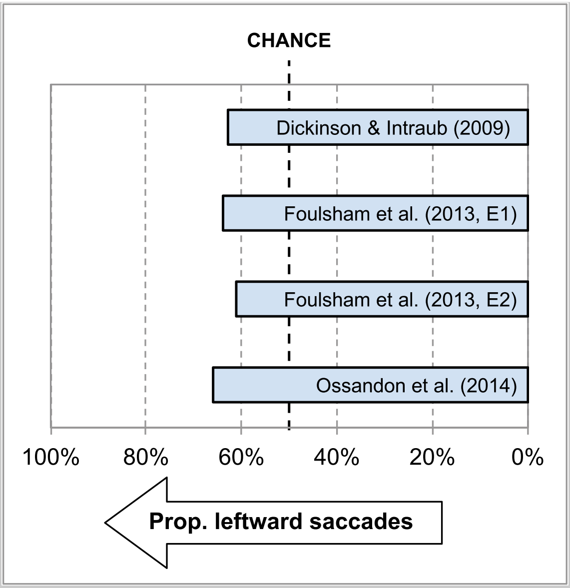


Fig. 6.10 Recent studies reporting a greater than average proportion of leftward initial saccades

6.3.7 *Bottom-up Features Play Only a Minor Role in the Guidance of Eye Movements*

Now that we have discussed methods for testing the predictiveness of the saliency map model, how should we judge its success? When doing this, it is useful to return to the original paper by Itti and Koch (2000), who defined their model as "bottom-up". Although the distinction between bottom-up and top-down is not always clear-cut, their model was meant to simulate early visual processes, and it contained no information about the meaning or relevance of scene regions. A glance back at Fig. 6.6 will convince the reader that the salient items in the "Oscar's selfie" image are not necessarily the most interesting, and neither are they the parts of the image which one intuits will catch viewers' attention, because the model knows nothing about faces or celebrities. Also, because the traditional saliency map approach produces the same predictions regardless of what the viewer is doing, it cannot account for any differences in behaviour based on task. These facts were noted by Itti and Koch

(2000), who acknowledged that it "might be that top-down influences play a significant role in the deployment of attention in natural scenes" (p. 1502), and that this was a key limitation of their model.

There does appear to be a correlation between fixation and saliency, particularly in the contrived case of a "free-viewing" task where participants are merely asked to look at an image with minimal task constraints (e.g., Parkhurst, Law, & Niebur, 2002). We have found that saliency at fixation is higher than chance (as did Peters et al., 2005), and that highly salient objects are looked at more often than less salient objects (Foulsham & Underwood, 2007, 2008). However, it would be incorrect to conclude from this that the literature was supportive of a causal relationship between saliency and eye movements. In fact, the opposite is true, with most recent research arguing that saliency is insufficient for explaining where we look. Some of the limitations of the saliency approach are discussed in detail by Tatler, Hayhoe, Land, and Ballard (2011). The main criticisms are as follows.

First, while there may be a correlation between saliency and fixation, this correlation is really rather weak, even in ideal circumstances. In terms of the area under the ROC curve, values of around 0.6 or 0.7 are most common, meaning that the underlying image statistics provide only a modest amount of information about fixation selection. Considering that "blind" models which do not have any information about image features but only about the systematic way that people move their eyes can lead to classification performance of 0.65 (Tatler & Vincent, 2009) or 0.68 (Ehinger et al. 2009), saliency may not tell us very much extra. Whatever metric is used, it is important to benchmark it against the amount of variability that we could possibly expect to predict. Because the saliency map model is a normative model, which produces the same predictions for everyone, it obviously cannot predict individual differences between observers, and neither can it account for any random component to fixations. A common approach, therefore, is to express the predictiveness of the model as a proportion of the "inter-observer consistency" (i.e., the ease with which fixations from a given observer can be predicted by the fixations of all other observers). This reflects the fact that what we are really trying to predict is the commonalities in looking behavior across people. Peters et al. (2005) report that their baseline saliency model gives an NSS of between 39% and 57% of the inter-observer value. While this is greater than zero, it leaves considerable variance unexplained.

A second problem is that with correlation-type analyses there are numerous other factors which might co-occur with saliency and which cause regions to be fixated. In other words, saliency may not actually cause regions to be fixated in the first place. Visually salient regions are often also semantically informative (Henderson, Brockmole, Castelhano, & Mack 2007) and contain meaningful objects (Elazary & Itti, 2008). Conversely, the regions which are fixated least often—consider the empty patches of sky and road in scenes such as Fig. 6.1—are often both non-salient and not of central focal importance to the meaning of a scene. We cannot, without systematic manipulations, conclude that what makes these regions priorities to fixate is indeed their visual saliency. Although it is a matter of current debate, objects which are manually identified by humans seem to be fixated more often than predicted by saliency, and selected in a way that is more consistent with complex object features than simple edge detection (Einhäuser, Rutishauser, & Koch, 2008; Foulsham & Kingstone, 2013a; but, see Borji, Sihite, & Itti, 2013). There are of course other meaningful objects, such as human faces, which appear to be fixated regardless of their visual saliency (e.g., Birmingham, Bischof, & Kingstone 2009; See Bischoff et al., this volume).

The third major issue for saliency map models is that, when participants view images under realistic task conditions, saliency seems to play very little role and can be immediately overridden. In particular, in visual search, salient regions are completely avoided in preference for regions which look like the target, or areas where it is likely to appear (Ehinger, Hidalgo-Sotelo, Torralba, & Oliva 2009; Foulsham & Underwood, 2007; Henderson, Malcolm, & Schandl, 2009). When the task requires fixating a target in a non-salient region, this task is completed even on the first fixation (Einhauser et al., 2008). Modifying semantically meaningful regions to change their saliency has little or no effect on their likelihood of fixation (Nystrom & Holmqvist, 2008). Beyond picture viewing, it may be best to conceptualise the "relevance" of a to-be-fixated location as depending, not on visual saliency, but on reward in the context of active tasks (Rothkopf, Ballard, & Hayhoe, 2007).

As described in Table 6.2, there has been considerable progress in the development of bottom-up saliency models. Several datasets are freely available for people to test proposed models and evaluate the results, stimulating a competition between models (e.g., the MIT saliency benchmark: http://saliency.mit.edu). Recently, this has involved applying more powerful machine-learning or deep-learning algorithms to learn the optimum features for discriminating fixated and non-fixated regions (e.g., Vig, Dorr, & Cox, 2014). Kümmerer, Wallis, & Bethge (2015) develop a principled way to compare different models and metrics by quantifying the amount of information a model provides, over and above an image-independent baseline. This paper shows that bottom-up models are becoming better at predicting where people look. On the other hand, it also estimates that the very best performing model can still only account for 34% of the information gain which should be explainable based on consistency between participants. Thus the saliency map approach is a long way from being able to explain patterns in where people look.

To summarise, then, although the saliency map model may represent a plausible and tractable way to estimate bottom-up feature contrast, it is not appropriate as a catch-all model for fixations in natural images. It remains an open question whether there are real-world situations where we prioritise saliency at the expense of relevance. Instead, researchers have focused on modeling top-down factors—knowledge of a search target or a scene, and the demands of a task—and combining these with visual saliency (Ehinger et al., 2009; Navalpakkam & Itti, 2005; Zelinsky, 2008).

6.4 Sequential Analysis of Fixations: Scanpaths and Scan Patterns

6.4.1 Temporal and Sequential Analysis of Fixations Can Provide Additional Insights into Eye Movement Control

So far, most of this chapter has dealt with answering questions about *where* people fixate. This is a sensible thing to investigate, because the eyes can only be directed at one location at a time and our eyetrackers can measure this with high precision. However, this has often led to a separation between analyses focused on spatial looking patterns and those based on the *timing* of where people look. This is surprising given that in other domains the question of when the eyes are moved is especially important (e.g., in reading; see Hyönä & Kaakinen, this volume). In this part of the chapter, I will describe one way of combining spatial and temporal information: by analyzing scanpaths. First, we should consider fixation duration and other temporal measures that have been investigated in scene viewing.

Turning back to Fig. 6.1 at the beginning of the chapter, we can see that each fixation on an image has a duration as well as an order in the scanning sequence. Typically, fixations on complex images have an average duration of around 300 ms, but they can vary considerably between different observers and different images. Very short fixations (less than 80 ms or so) are rare and, because it must take longer than this for the visual system to program the next saccade, these outliers are sometimes excluded. Understanding the causes of the variability in fixation duration (what makes some fixations longer than others) is difficult in natural scenes because we are rarely in control of precisely where someone is looking or the information available at that point. However, with the support of evidence from reading and picture viewing it is generally assumed that longer fixations reflect more difficult, more extensive, or more effortful processing of the details at that location. For example, objects which are out-of-place are associated with longer fixation durations (Underwood & Foulsham, 2006), and fixations are longer on average when trying to remember the details of a scene than when searching around for a specific object (Mills, Hollingworth, Van der Stigchel, Hoffman, & Dodd, 2011). Conversely, we would not expect fixation duration to be associated with superficial visual properties of an object which do not change its meaning (and hence in Foulsham & Underwood, 2007, we found that salient objects were fixated *earlier* but not for a longer duration).

There are a number of difficulties with analyzing fixation durations for particular areas of interest in complex scenes. One issue is that, although it is normally a safe assumption that fixation duration reflects processing at the current location, in some cases prolonged fixation duration may indicate covert scanning or changes in the periphery. Another is that objects are often fixated more than once, and so it is really the cumulative time spent on an object which is a better measure of processing (see

Fig. 6.3). Multiple, consecutive fixations on an object are normally described as a "gaze" (but also as a "dwell" or a "run"), and thus the first gaze duration describes the sum duration of all fixations before moving the eyes away from a particular region of interest. The pattern of fixations and refixations may be complex. For example, participants might make a short initial fixation near an object, before making a small "corrective" saccade to a better position for a longer fixation. The interaction of fixation position on an object and fixation duration has been of recent interest for several researchers (Foulsham & Kingstone, 2013a; Nuthmann & Henderson, 2010).

A good way to introduce some control into experiments in scene perception is by using gaze-contingent displays, where parts of the scene are changed in real time and in response to eye movements. Henderson and Pierce (2008) masked the scene at a critical point during a saccade and observed the results on the following fixation (the "scene onset delay" paradigm). Surprisingly, they observed that not all fixations were affected by this change. Thus, although some fixations were under direct control in response to what the observer could currently see, others were "pre-programmed". Such evidence has been used by Nuthmann, Smith, Engbert, and Henderson (2010) to propose a model of fixation durations in scenes which is driven by the underlying saccade programming.

Participants in scene viewing experiments often view a picture for several seconds at a time. As we might expect if observers are learning about the image and changing their priorities, eye movements may change over this period. For this reason, some researchers choose to look specifically at the first few fixations, arguing that this is the point when participants are drawn to certain details for the first time. Fixations in the first 1–2 s of the viewing period tend to have a shorter average duration (and these are associated with larger average saccades; Unema, Pannasch, Joos, & Velichkovsky 2005). It has been proposed that such changes reflect a switch between "ambient" or "global" exploration and "focal" or "local" scanning. However, it is not known how ubiquitous these patterns are or how they are related to scene content.

It is clear from the research reviewed in this section that it is important to think about *when* fixations are made, as well as *where*. Rather than aggregating across a number of fixations made at different times, we can also examine the presence of particular sequential patterns in where observers look. This is discussed in the following sections.

6.4.2 Classic Research Emphasized Cognitive Constraints on Scanning Patterns

The focus of recent research on visual saliency and other image-based accounts of fixation is curious, because classic research emphasised the opposite: that where we look and the order in which we look there is determined top-down. Buswell

(1935) and Yarbus (1967) studied participants' eye movements while viewing paintings and images. They found that people tended to look at semantically meaningful regions: faces, objects and details important for understanding the scene. As we discussed in the previous section, it is unlikely that a purely saliency-based account of where people look can explain the concentration of fixations on such items. Moreover, Yarbus is often credited with showing that the pattern of eye movements made depended on the task that the participant was performing (although Buswell had in fact already observed this). In particular, Yarbus showed observers the same painting (Ilya Repin's *An Unexpected Visitor*) with a number of different questions in mind. In a widely-reproduced figure, he depicted the series of fixations and saccades made in these different conditions, confirming that the viewing pattern in each case was very different. For example, when asked to give the ages of the people depicted in the scene, most fixations were on the faces of these characters. In contrast, when asked to "estimate the material circumstances of the family", many more fixations were made on the paintings, furniture and decoration in the room. Therefore the places where people looked were selected according to the information required by the task.

Yarbus' illustrations place a key emphasis on both where the observers were looking and the order in which they look there. Indeed, as discussed by Tatler, Wade, Kwan, Findlay, and Velichkovsky (2010) in their review of Yarbus' impact, he was particularly interested in the way that eye movement scanning might be cyclical or idiosyncratic. When dividing up the viewing of an image into different time slices, he claimed that observers repeated a pattern of inspection. For example, when looking at a face, participants would iterate through the key features (eyes, mouth, nose) before starting again. Although different participants tended to look at similar locations, when the same person looked at an image on multiple occasions the resulting scanning patterns were even more alike.

In this section of the chapter, I will review methods and theory which examine these "scanpaths" or "scan patterns" in more detail. As we shall see, this goes beyond merely describing what is looked at, to quantifying the sequence of eye movements as a whole.

6.4.3 Scanpath Theory Links Eye Movements to a Cognitive Model for Visual Recognition

As we have seen, the viewing patterns of a number of people looking at an image may cluster on certain regions of detail. What determines the order or temporal structure of their fixations? Noton and Stark (1971) advanced the theory that the execution of a sequence of eye movements is intimately bound up in the processes of encoding and recognizing an image. Noton and Stark were influenced by Yarbus' observation of cyclical and repeating eye movements, and they used recordings from those inspecting simple line drawings. In particular, in each observer's pattern of inspection they claimed to see a "scanpath": a "fixed path, followed intermittently but repeatedly by a subject's eye while he views a pattern" (p. 933). This observation was based on judging the similarity within and between viewers, and in particular on recordings from the same person looking at the same image on two occasions. Noton and Stark proposed that some elements of an individual's scanpath was repeated on subsequent viewings.

This qualitative observation was used to support a theory of human pattern recognition. According to "scanpath theory", when a visual pattern is encountered for the first time, its local features are stored in memory. Crucially, these sensory memory traces are combined with a representation of the eye movements made to view the pattern—a motor memory trace. The result is a sensorimotor network: a series of connections between nodes representing features and those representing the eye movements. When the pattern is encountered again at a later date, this network is reactivated, so that successive eye movements in the sequence are repeated, and the visual features are verified against the memory trace.

On the one hand, scanpath theory was an overly simplistic and impractical account of how people see and remember. It is not clear how features are integrated, or how such networks could scale up to all the different sorts of patterns that we can recognize, and all the situations in which we view them. On the other hand, it predated much more modern accounts of embodied perception, such as Barsalou (1999), which see motor simulation as a crucial part of perception. For those interested in eye movements, the notion of scanpaths as generated top-down, based on individual memories and representations is a radically different approach to the research on visual saliency discussed in the first part of this chapter.

6.4.4 Methods for Scanpath Comparison Can Preserve Spatial and Temporal Information About Viewing Behaviour

I began this chapter by discussing some methods for quantifying the eye movements from multiple observers looking at an image. In general, these methods were based on the spatial position of fixations. For example, heatmaps tell us something about the spatial spread and consistency of fixations, and how they align with items in the scene. If we are interested in the *order* in which certain regions are fixated, one option is to evaluate measures of fixation timing (e.g., how early objects get fixated in different scenes or conditions, see Table 6.1 and Fig. 6.2). However, this requires clear regions of interest and analyses only a small number of fixations (i.e., those which land on the regions). They are also generally aggregated across many trials and observers. If scanpaths really do feature predictable sequences of fixations, then these methods will probably not be able to detect them.

An alternative way of analyzing scanpaths is to compare them more holistically, over space and time. The central problem here is how to take two scanpaths (A and B) and compute their pairwise similarity. This problem has been addressed in detail by several authors (Cristino, Mathot, Theeuwes, & Gilchrist, 2010; Dewhurst et al., 2012; Foulsham & Underwood, 2008; Privatera & Stark, 2000; See Bischoff et al., this volume).

Perhaps the most straightforward approach would be to measure the linear distance between the fixations in each scanpath. Similar scanpaths will contain fixations which are close to each other in space. This is the approach that was taken by Mannan, Ruddock, and Wooding (1995) when they studied fixations in pictures with various levels of image manipulation. But to which of the fixations in scanpath B should a given fixation in scanpath A be compared? There are a number of options (the closest, the average of all others, that which occurred at the same time), but all of these result in un-intuitive values when there are big differences in the distributions of the fixations in A and B. Moreover, measuring linear distance cannot easily take into account the sequence of the fixations, meaning that two scanpaths which visit the same locations in a completely different order will seem highly similar.

Instead, a range of authors have proposed scanpath comparison methods based on aligning fixation sequences (see Box 3, and Bischoff et al., this volume). These have the advantage that they represent both spatial and sequential aspects of the scanpaths. However, they can be complex to calculate, and there is currently no clear consensus on which method to use. The preferred analysis will depend on whether there are clear regions of interest and the aim of the comparison. Whichever method is chosen, it is important to evaluate similarity values with care. Often, the aim of the analysis is to investigate whether scanpaths are more (or less) similar than some kind of chance or baseline expectation. Because of the global biases discussed earlier in the chapter, scanpaths will probably be similar, to a certain degree, even when they originate from different scenes and observers. A good approach, therefore, is to also compute some control comparisons, against which the experimental values can be evaluated.

Box 3: Sequential Scanpath Comparison

Several of the scanpath comparison algorithms that have been developed are based on the "string edit" or Levenshtein distance, which was applied to eye movements by researchers such as Brandt and Stark (1997). First, scanpaths are coded as sequences of characters based on regions of interest (Fig. 6.11). Then, the similarity between these sequences is quantified by calculating the number of steps or edits required to transform one into the other.

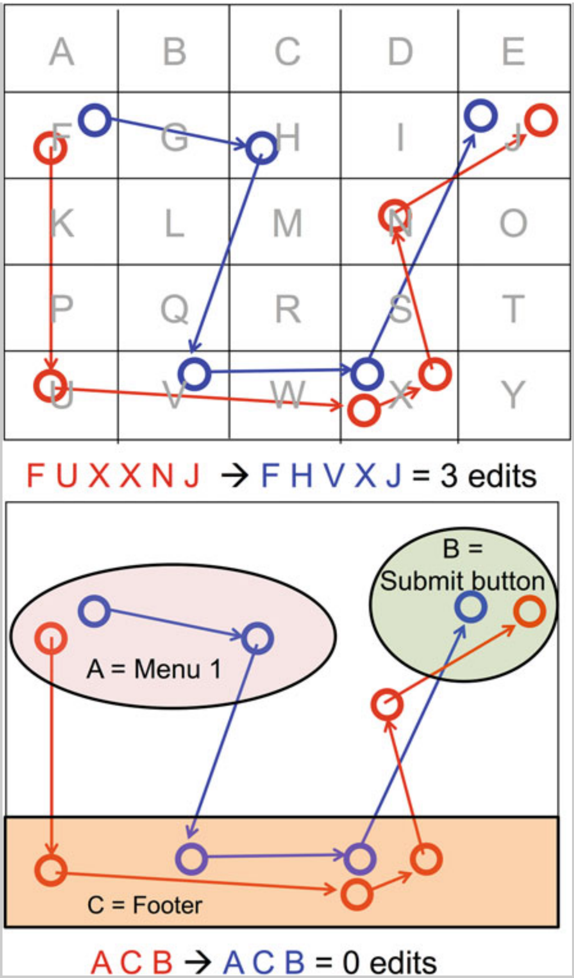
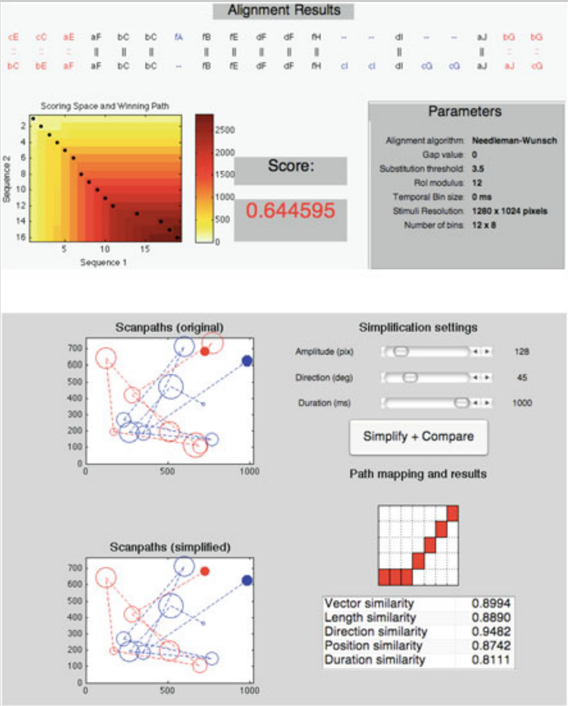


Fig. 6.11 Two scanpaths are compared as strings, using either a grid or predetermined regions of interest

Cristino et al. (2010) proposed a more sophisticated approach (ScanMatch) which takes into account fixation durations and spatial and non-spatial relationships between ROIs.

Dewhurst et al. (2012), described an alternative (MultiMatch), which aligns and compares scanpaths as simplified vectors. This allows scanpaths to be compared across multiple dimensions. Both ScanMatch and MultiMatch are available as MATLAB toolboxes (Fig. 6.12).



Score: 0.644595

Parameters

Alignment algorithm: Needleman-Wunsch
Gap value: 0
Substitution threshold: 3.5
RoI modulus: 12
Temporal Bin size: 0 ms
Stimuli Resolution: 1280 x 1024 pixels
Number of bins: 12 x 8

Amplitude (pix): 128
Direction (deg): 45
Duration (ms): 1000

Vector similarity	0.8994
Length similarity	0.8890
Direction similarity	0.9482
Position similarity	0.8742
Duration similarity	0.8111

Fig. 6.12 ScanMatch and MultiMatch toolboxes give similarity scores for a pair of scanpaths

6.4.5 Tests of Scanpath Theory Confirm that Participants Are Idiosyncratic in Where They Look

Now that we have considered how scanpaths can be represented and compared, we can return to the central predictions of scanpath theory: (i) that participants will repeat eye movement sequences when recognizing a previously seen image; and (ii) that this recapitulation will trigger successful recognition. Noton and Stark (1971) addressed (i) by making qualitative judgements about the similarity of scanpaths at encoding and test. Subsequent work used transition matrices or the string edit distance to quantify similarity, confirming that there was some similarity between viewings (Choi, Mosely, & Stark, 1995; Stark & Ellis, 1981). However, these experiments used rather simple stimuli and few observers, and it was not clear how the similarity measurements should be interpreted.

In Foulsham and Underwood (2008) and Foulsham et al. (2012), we returned to the question of scanpath similarity, using 45 complex natural scenes which were viewed by 21 observers as part of a memory test. Figure 6.13 shows an example of the resulting scanpaths. Using a variety of different comparison methods, we confirmed

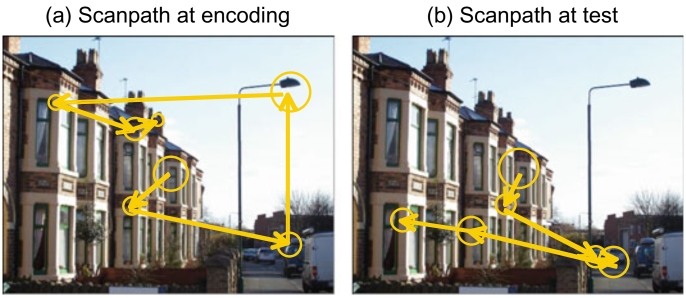


Fig. 6.13 An example of two scanpaths made by the same person viewing the same scene, once while learning the image (**a**), and once when recognizing it later during a memory test (**b**). Figure redrawn from Foulsham et al. (2012). Note that in this figure fixation duration is represented by the size of the circles

earlier reports that participants did show some similarity between viewings. This similarity was certainly not 100%, and because the image is the same, any differences must be due to memory and the demands of the task. However, scanpaths from the same person viewing the same image were more similar than two different people viewing the same image. In this sense, the eye movements of a particular individual seem to be idiosyncratic.

It is clear that the scanpaths made by the same person viewing an image on multiple occasions are not identical, and indeed, it is possible for people to recognize some images without making any eye movements. Therefore the strong version of scanpath theory can be rejected. Moreover, the modest similarity that does occur could be caused by a variety of factors: the re-presentation of salient items; a consistent reaction to scene elements; or idiosyncratic, systematic tendencies in eye movements which persist over time. So what of prediction (ii) above, that eye movement recapitulation enhances recognition memory? The evidence that images with closely repeating scanpaths are recognized more accurately is inconclusive (Foulsham et al., 2012). It is also not clear whether repeating eye movements could be said to be causing memory, or the other way around. A much better test of the role of eye movements in image memory, therefore, is to manipulate scanpaths and observe the results.

6.4.6 Manipulating Scanpaths Can Affect Memory

Scanpaths in picture viewing arise when participants freely view an image, and thus they provide a measure of unconstrained, natural behavior. However, often a researcher may need to introduce constraints, so that information in the image is acquired in a certain order or central and peripheral material is controlled. Manipulating fixations in natural images can therefore be a useful tool. One possibility is to use a gaze-contingent procedure, which changes presentation based on voluntary eye movements. For example, Zelinsky and Loschky (2005) investigated memory in scenes by ending the trial when a certain number of objects had been fixated. Foulsham, Gray, Nasiopoulos, and Kingstone (2013) used “moving windows” of different shapes to encourage a different pattern of scanning, with vertical "slits" leading to more vertical eye movements.

Holm and Mantyla (2007) coerced participants viewing landscape paintings to follow a certain scanpath. Observers were required to fixate a sequence of dots, superimposed on the image. Later on in the experiment, looking back to the same locations was associated with explicit recognition. In Foulsham and Kingstone (2013b), we used a similar technique to probe the causal role of eye movements, providing a direct test of scanpath theory. In our first experiment, we manipulated the scanpath during the learning phase of the experiment, as had been done by Holm and Mantyla. Rather than freely viewing each image, participants saw a series of square patches, one at a time, which simulated a sequence of fixations. After viewing 48 scenes in this fashion, there was a memory test where the task was to indicate whether the displayed image had appeared in the first half of the experiment. During this "test" block, the full scenes were displayed and the question was whether unconstrained eye movements would follow the sequence from that image in the learning block. The results showed that indeed the imposed scanpath affected viewing at test, but only when the image was subsequently recognized correctly.

In subsequent experiments, we reversed the procedure, allowing free viewing during the learning phase but constraining the scanpath at test (Foulsham & Kingstone, 2013b, Experiments 2–5). This meant that the sequence of fixations at test could be manipulated to be either the same as those made by that observer when first seeing the image, or drawn from different locations. These experiments were technically challenging and required that the fixations be written to a data file during the learning phase, and then retrieved to constrain the test phase. If repeating a scanpath really activates a stored memory of the pattern, then we would expect recognition of previously seen images to be better when we forced participants to look at the image in the same sequence as they had on the first encounter. This was the case when recognition was compared with a condition presenting random patches. However, when the control condition presented patches from the eye movements made by another individual, or the same fixation locations but in a scrambled order, there was no memory advantage. In other words, there is not anything especially useful about repeating your own scanpath, in its set order. Fixations select memorable locations, but in the case of complex scenes replaying a scanpath does not seem to have a clear causal effect on memory in the way proposed in scanpath theory.

6.4.7 Eye Movements Provide Information About Recognition and Imagery

Scanpath theory seems to be unable to account for the particular relationship between eye movements and memory observed in Foulsham and Kingstone (2013b), because it places special importance on one's own scanning pattern, which we found does not necessarily enhance memory. Eye movements vary in multiple ways, and although some of this variability appears to be systematic within an observer, supporting a link with memory processes has proven difficult. However, scanpaths clearly are useful measures, both for characterising viewing patterns in a holistic fashion and for investigating memory. At the most basic level, eye movements and attention are gatekeepers to our sensory input, and so they must constrain what we end up seeing and storing in memory. Eye movements and memory continue to be explored in a variety of stimuli, particularly in face recognition (Althoff & Cohen, 1999; Schwedes & Wentura, 2012).

Scanpath theory and methods have also been applied to visual imagery, where there continues to be a debate about the extent to which eye movements are functional. In a typical experiment, participants view an image or other spatial pattern, and they then visualise or imagine this image later while looking at a blank screen. In these conditions, observers seem to spontaneously make eye movements, which may partially reflect items in the remembered image. Laeng and Teodorescu (2002) and Johansson, Holsanova, Dewhurst, and Holmqvist (2012) argued that preventing eye movements during retrieval led to poorer imagery, although it is not always clear whether the precise pattern of eye movements made is important for functional benefits.

6.5 Applications and Implications

This chapter has described a range of methods for investigating eye movements in scenes through region-of-interest analyses, comparisons with saliency maps, and quantifying scanpaths. Because the focus has been on relatively complex stimuli which vary across many dimensions, these techniques are well suited to applying to a range of different contexts. Eyetracking has become a popular tool for researchers far beyond the traditional realm of cognitive psychology and vision, and here I will mention a few particularly active interdisciplinary topics.

6.5.1 Perception of Art

Like Yarbus and Buswell, some of my examples have been drawn from experiments where participants viewed works of art (e.g., Box 1, which shows Gainsborough's painting *Mr and Mrs Andrews*). As with more theoretical work in scene perception, eyetracking experiments with art have addressed the consistency between observers, as well as how this may be affected by the techniques of the artist and viewer expertise.

Moving beyond the truism that people look at areas of detail in a painting, DiPaola, Riebe, and Enns (2010) investigated the subtle techniques that Rembrandt and other artists use to guide the eyes. DiPaola et al. selectively modified the rendering of portraits to test the idea that artists enhance the detail of one side of the face in order to induce a particular viewing pattern. Sure enough, participants spent more time on a textured eye region than on the other side of the face, and this affected judgements of the quality of the art. Thus artists may have evolved techniques which affect eye movements, and this may even be true in the colour balance of abstract artworks (Nodine, Locher, & Krupinski, 1993).

Eyetracking has also been used to evaluate the viewer's experience of looking at a piece of art. Nodine, Locher, and Krupinski (1993) were among the first to investigate how artistic training might affect this experience as well as viewing patterns.

The results suggested that untrained viewers focus more on individual objects, while experts were more sensitive to composition. This was reflected in more "specific" scanpaths targeted at relationships between objects. More recent work has shown that artistically-trained individuals spend more time on structural features (Vogt & Magnussen, 2007). Naïve observers may also change their gaze patterns when speaking about their interpretation of a painting, becoming more systematic (Klein et al., 2014). Thus eye movements continue to show promise in this interdisciplinary field. Indeed, eyetracking has even been incorporated into a large exhibition in an art gallery, resulting in data from thousands of visitors (Wooding, 2002).

6.5.2 Marketing and Websites

There is now a large industry which seeks to evaluate and improve marketing materials by using techniques from visual attention research and eyetracking (see section D of this volume). In advertisements, experimental research has often looked at the way in which observers scan combined text and images. For example, Rayner, Rotello, Stewart, Keir, and Duffy (2001) found that people normally first read the text associated with print advertisements, before looking at the product image. Wedel and Pieters (2008) provide a useful review of how eyetracking has been applied to marketing, as a now dominant measure of the attention paid to branding. Although it is clear that there are things that marketers can do to increase the attention paid to their advertisements (e.g., by making the brand and advertisement larger), the impact on actual purchasing is less persuasive. Current research using eyetracking to investigate marketing has also investigated sequential scanpaths (Pieters, Rosbergen, & Wedel, 1999) and the role of bottom-up saliency (Van der Lans, Pieters, & Wedel, 2008). Moreover, the central fixation bias appears to have an effect on the products which are noticed and ultimately chosen (Atalay, Bodur, & Rasolofoarison, 2012).

The visual exploration of websites has also provided a measure for marketers. For example, a large number of studies and findings are discussed by Nielsen and Pernice (2010). Among the most popular claims from this research are that participants show an "F-shaped" pattern when reading web content (distinguished by a large horizontal exploration which tapers off further down the page) and that observers frequently ignore advertisements ("banner blindness"). More relevant for the present chapter, both saliency and scanpaths have been investigated in the context of websites (Holmberg, Sandberg, & Holmqvist, 2014; Josephson & Holmes, 2002).

6.5.3 Computer Vision

Computational models of visual saliency, benchmarked against human eye movement data, continue to be extremely popular in a range of computer vision applications. Borji and Itti (2013) review some of the many models that have been developed, along with their applications. Because it is assumed that fixation (and attention) is the first stage in cognitive processing, visual saliency may help build computers which can recognize objects (Walther, Rutishauser, Koch, & Perona, 2005). Saliency can also be the first step in processes of image and video segmentation.

Interestingly, the focus on eye movements has both influenced and been influenced by developments in robotics. In particular, the acknowledgement that artificial visual systems benefitted from an active sensor (which moves like the eye), contributed to the complementary movements known as Active Vision (Ballard, 1991; Findlay & Gilchrist, 2003). Saliency has been implemented as a way for robots to cut down on their visual input and learn to move and localize themselves in space (e.g., Siagan & Itti, 2009).

6.6 Conclusions and Limitations

In this chapter, I have discussed eyetracking results from people looking at pictures, and I have done so largely within the framework of saliency and scanpaths. Visual saliency provides an explicit and implemented model of going from simple visual features to predictions for complex scenes and images. However, it is clearly also a very limited approach for describing actual eye movement behavior. The weak empirical effects of saliency on fixations do, however, allow us to focus on ways of representing and manipulating task knowledge so that more explicit predictions can be made about where people will look in a given situation.

Eye movement data can be challenging to analyse. The research I have discussed shows that we should not forget the many ways that eye movements over pictures may vary, in terms of fixation position, saccade directions and so on. One often comes back to the observation that eye movements are very far from being uniformly distributed across a display, and this systematicity comes from both the stimulus and the observer. In particular, researchers should be mindful that some “ways of looking” may emerge more often, based purely on biases in the oculomotor system (the origins of which may be learned or biological). Although there does not appear to be a straightforward relationship between memory and scanpaths, there is a regular and idiosyncratic component to the way that we move our eyes.

There are of course limitations to the scope of what I have discussed, and to the conclusions which can be made from observing eye movements in scenes. There are many other descriptions of individual differences, emotional states and task instructions making a difference to where people look in images, and these fall under the general umbrella of top-down attention. It remains to be seen whether or not these effects interact with visual saliency. Although most research on scene perception has been carried out with healthy, typical observers, complex images have also been used to investigate neuropsychological and developmental disorders (e.g., Freeth, Foulsham, & Chapman, 2011; Foulsham, Barton, Kingstone, Dewhurst, & Underwood, 2009; Tseng et al., 2013).

Recording where people look in complex images can tell us a lot about attention and cognitive processing, but it cannot tell us everything. Fixation and attention are not synonymous. Covert attention can certainly be allocated separately from fixation, but there are few studies examining this in complex stimuli, and thus its role in naturalistic viewing is unknown. Importantly, we should also be cautious at using results from picture-viewing to make conclusions about attention in the real world. The vast majority of the research that I have described uses small, static scenes, presented without context in a highly constrained setting. In other words, it is far removed from the real world, which contains cues such as sound and motion and in which observers are immersed and free to move, where concepts such as a "central bias" may be meaningless. Nevertheless, those continuing to investigate looking at pictures will find much to discover.

6.7 Suggested Readings

Henderson (2003) gives a good summary of the basic methods and findings in scene perception.

Foulsham (2015) reviews the state of the art regarding eye movements in scene perception, covering some of the same material as this chapter but with more detail and scope.

Itti and Koch (2001) provide a comprehensive introduction to computational modeling of attention, including a shorter description of their saliency map model.

Tatler et al. (2011) provide a detailed critique of saliency models and the picture viewing paradigm.

Le Meur and Baccino (2013) summarise some of the commonly used methods for comparing fixations and model predictions.

Holmqvist et al. (2011) is a detailed textbook on eye movement methodology, with particularly comprehensive coverage of scanpath comparison and other measures.

6.8 Questions to Students

Why are there so many measures based on eye movements in scenes? Are they all necessary?

What is the best way of defining "saliency" in terms of eye movement control?

How useful is the Itti and Koch saliency map model for predicting human eye movements?

Is analysis of scanpaths necessary for investigating cognition and scene perception?

Are scanpaths random? If not, why?

What can experiments investigating eye movements in scene perception tell researchers working in marketing and other applied domains?

References

Althoff, R. R., & Cohen, N. J. (1999). Eye-movement-based memory effect: A reprocessing effect in face perception. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 25*(4), 997–1010.

Atalay, A. S., Bodur, H. O., & Rasolofoarison, D. (2012). Shining in the center: Central gaze cascade effect on product choice. *Journal of Consumer Research, 39*(4), 848–866.

Ballard, D. H. (1991). Animate vision. *Artificial Intelligence, 48*(1), 57–86.

Barsalou, L. W. (1999). Perceptions of perceptual symbols. *Behavioral and Brain Sciences, 22*(04), 637–660.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49*(24), 2992–3000.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience, 33,* 1.

Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*(1), 185–207.

Borji, A., Sihite, D. N., & Itti, L. (2013). Objects do not predict fixations better than early saliency: A re-analysis of Einhäuser et al.'s data. *Journal of Vision*, *13*(10), 18.

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9*(1), 27–38.

Bruce, N., & Tsotsos, J. (2005). Saliency based on information maximization. In *Advances in Neural Information Processing Systems* (pp. 155–162).

Buswell, G. T. (1935). *How people look at pictures*. University of Chicago Press.

Choi, Y. S., Mosley, A. D., & Stark, L. W. (1995). String editing analysis of human visual search. *Optometry and Vision Science, 72*(7), 439–451.

Clarke, A. D., & Tatler, B. W. (2014). Deriving an appropriate baseline for describing fixation behaviour. *Vision Research, 102,* 41–51.

Cristino, F., Mathot, S., Theeuwes, J., & Gilchrist, I. (2010). ScanMatch: A novel method for comparing saccade sequences. *Behavior Research Methods, 42*(3), 692–700.

Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with MultiMatch, a vector-based approach. *Behavior Research Methods, 44*(4), 1079–1100.

Dickinson, C. A., & Intraub, H. (2009). Spatial asymmetries in viewing and remembering scenes: Consequences of an attentional bias? *Attention, Perception, & Psychophysics, 71*(6), 1251–1262.

DiPaola, S., Riebe, C., & Enns, J. (2010). Rembrandt's textural agency: A shared perspective in visual art and science. *Leonardo, 43*(2), 145–151.

Dodge, R., & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review, 8*(2), 145.

Ehinger, K. A., Hidalgo-Sotelo, B., Torralba, A., & Oliva, A. (2009). Modelling search for people in 900 scenes: A combined source model of eye guidance. *Visual cognition, 17*(6–7), 945–978.

Einhäuser, W., Rutishauser, U., & Koch, C. (2008a). Task-demands can immediately reverse the effects of sensory-driven saliency in complex visual stimuli. *Journal of Vision, 8*(2), 2.

Einhäuser, W., Spain, M., & Perona, P. (2008b). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14), 18.

Elazary, L., & Itti, L. (2008). Interesting objects are visually salient. *Journal of Vision, 8*(3), 3.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford University Press.

Foulsham, T. (2015). Scene perception. To appear in Fawcett, Risko & Kingstone (Eds.). *The Handbook of Attention*. MIT Press.

Foulsham, T., Barton, J. J., Kingstone, A., Dewhurst, R., & Underwood, G. (2009). Fixation and saliency during search of natural scenes: the case of visual agnosia. *Neuropsychologia, 47*(8), 1994–2003.

Foulsham, T., Dewhurst, R., Nyström, M., Jarodzka, H., Johansson, R., Underwood, G., et al. (2012). Comparing scanpaths during scene encoding and recognition: A multi-dimensional approach. *Journal of Eye Movement Research, 5*(4), 3.

Foulsham, T., Gray, A., Nasiopoulos, E., & Kingstone, A. (2013). Leftward biases in picture scanning and line bisection: A gaze-contingent window study. *Vision Research, 78,* 14–25.

Foulsham, T., & Kingstone, A. (2013a). Optimal and preferred eye landing positions in objects and scenes. *The Quarterly Journal of Experimental Psychology, 66*(9), 1707–1728.

Foulsham, T., & Kingstone, A. (2013b). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology: General, 142*(1), 41.

Foulsham, T., Kingstone, A., & Underwood, G. (2008). Turning the world around: Patterns in saccade direction vary with picture orientation. *Vision Research, 48*(17), 1777–1790.

Foulsham, T., & Underwood, G. (2007). How does the purpose of inspection influence the potency of visual salience in scene perception? *Perception, 36,* 1123–1138.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*(2), 6.

Freeth, M., Foulsham, T., & Chapman, P. (2011). The influence of visual saliency on fixation patterns in individuals with Autism Spectrum Disorders. *Neuropsychologia, 49*(1), 156–160.

Harel, J., Koch, C., & Perona, P. (2006). Graph-based visual saliency. In *Advances in Neural Information Processing Systems* (pp. 545–552).

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in cognitive sciences, 7*(11), 498–504.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. L. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. W. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Amsterdam: Elsevier.

Henderson, J. M., Malcolm, G. L., & Schandl, C. (2009). Searching in the dark: Cognitive relevance drives attention in real-world scenes. *Psychonomic Bulletin & Review, 16*(5), 850–856.

Henderson, J. M., & Pierce, G. L. (2008). Eye movements during scene viewing: Evidence for mixed control of fixation durations. *Psychonomic Bulletin & Review, 15*(3), 566–573.

Holm, L., & Mäntylä, T. (2007). Memory for scenes: Refixations reflect retrieval. *Memory & Cognition, 35*(7), 1664–1674.

Holmberg, N., Sandberg, H., & Holmqvist, K. (2014). Advert saliency distracts children's visual attention during task-oriented internet use. *Frontiers in Psychology*, *5*.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press.

Itti, L., & Baldi, P. F. (2005). Bayesian surprise attracts human attention. In *Advances in Neural Information Processing Systems* (pp. 547–554).

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10), 1489–1506.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20*(11), 1254–1259.

Johansson, R., Holsanova, J., Dewhurst, R., & Holmqvist, K. (2012). Eye movements during scene recollection have a functional role, but they are not reinstatements of those produced during encoding. *Journal of Experimental Psychology: Human Perception and Performance, 38*(5), 1289.

Josephson, S. & Holmes, M. E. (2002). Visual attention to repeated internet images: Testing the scanpath theory on the world wide web. In *ETRA '02: Proceedings of the 2002 Symposium on Eye Tracking Research and Applications*, (pp. 43–49). New York: ACM.

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. In *Proceedings of the 12th International Conference on Computer Vision*, (pp. 2106–2113). IEEE.

Klein, C., Betz, J., Hirschbuehl, M., Fuchs, C., Schmiedtová, B., et al. (2014). Describing art—An interdisciplinary approach to the effects of speaking on gaze movements during the beholding of paintings. *PLoS ONE, 9*(12), e102439.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology, 4*(4), 219–227.

Kümmerer, M., Wallis, T. S., & Bethge, M. (2015). Information-theoretic model comparison unifies saliency metrics. *Proceedings of the National Academy of Sciences, 112*(52), 16054–16059.

Laeng, B., & Teodorescu, D.-S. (2002). Eye scanpaths during visual imagery reenact those of perception of the same visual scene. *Cognitive Science, 26*(2), 207–231.

Land, M. F. (1993). Eye-head coordination during driving. In *Systems, man and cybernetics,* (pp. 490–494). IEEE.

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: Strengths and weaknesses. *Behavior Research Methods, 45*(1), 251–266.

Le Meur, O., Le Callet, P., Barba, D., & Thoreau, D. (2006). A coherent computational approach to model the bottom-up visual attention. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28,* 802–817.

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. *Journal of Experimental Psychology: Human Perception and Performance, 4*(4), 565.

Mackworth, N. H., & Morandi, A. J. (1967). The gaze selects informative details within pictures. *Perception and Psychophysics, 2*(11), 547–552.

Mackworth, N. H., & Thomas, E. L. (1962). Head-mounted eye-marker camera. *Journal of the Optical Society of America (1917–1983)*, *52*, 713.

Mannan, S., Ruddock, K., & Wooding, D. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision, 9*(3), 363–386.

Mills, M., Hollingworth, A., Van der Stigchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*(8), 17.

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature, 421*(6921), 370–373.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research, 45*(2), 205–231.

Nielsen, J., & Pernice, K. (2010). *Eyetracking web usability*. New Riders.

Nodine, C. F., Locher, P. J., & Krupinski, E. A. (1993). The role of formal art training on perception and aesthetic judgment of art compositions. *Leonardo*, 219–227.

Noton, D., & Stark, L. (1971). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research, 11*(9), 929.

Nuthmann, A., & Henderson, J. M. (2010). Object-based attentional selection in scene viewing. *Journal of Vision, 10*(8), 20.

Nuthmann, A., Smith, T. J., Engbert, R., & Henderson, J. M. (2010). CRISP: A computational model of fixation durations in scene viewing. *Psychological Review, 117*(2), 382.

Nyström, M., & Holmqvist, K. (2008). Semantic override of low-level features in image viewing-both initially and overall. *Journal of Eye-Movement Research, 2*(2), 1–2.

Ossandon, J. P., Onat, S., & Koenig, P. (2014). Spatial biases in viewing behavior. *Journal of Vision*, *14*(2), 20–20

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42*(1), 107–123.

Peters, R. J., Iyer, A., Itti, L., & Koch, C. (2005). Components of bottom-up gaze allocation in natural images. *Vision Research, 45*(18), 2397–2416.

Pieters, R., Rosbergen, E., & Wedel, M. (1999). Visual attention to repeated print advertising: A test of scanpath theory. *Journal of Marketing Research*, 424–438.

Posner, M. I., Rafal, R. D., Choate, L. S., & Vaughan, J. (1985). Inhibition of return: Neural basis and function. *Cognitive Neuropsychology, 2*(3), 211–228.

Privitera, C. M., & Stark, L. W. (2000). Algorithms for defining visual regions-of-interest: Comparison with eye fixations. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*(9), 970–982.

Rayner, K., Rotello, C. M., Stewart, A. J., Keir, J., & Duffy, S. A. (2001). Integrating text and pictorial information: Eye movements when looking at print advertisements. *Journal of Experimental Psychology: Applied, 7*(3), 219.

Rothkopf, C. A., Ballard, D. H., & Hayhoe, M. M. (2007). Task and context determine where you look. *Journal of Vision, 7*(14), 16.

Schwedes, C., & Wentura, D. (2012). The revealing glance: Eye gaze behavior to concealed information. *Memory & Cognition, 40*(4), 642–651.

Siagian, C., & Itti, L. (2009). Biologically inspired mobile robot vision localization. *IEEE Transactions on Robotics, 25*(4), 861–873.

Stark, L., & Ellis, S. R. (1981). Scanpath revisited: Cognitive models of direct active looking. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 193–226). Hillsdale, NJ: Erlbaum.

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005). Visual correlates of fixation selection: Effects of scale and time. *Vision Research, 45*(5), 643–659.

Tatler, B. W., Hayhoe, M. M., Land, M. F., & Ballard, D. H. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of vision, 11*(5), 5.

Tatler, B. W., & Vincent, B. T. (2009). The prominence of behavioural biases in eye guidance. *Visual Cognition, 17*(6–7), 1029–1054.

Tatler, B. W., Wade, N. J., Kwan, H., Findlay, J. M., & Velichkovsky, B. M. (2010). Yarbus, eye movements, and vision. *i-Perception*, *1*(1), 7.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136.

Treue, S. (2003). Visual attention: The where, what, how and why of saliency. *Current Opinion in Neurobiology, 13*(4), 428–432.

Tseng, P. H., Cameron, I. G., Pari, G., Reynolds, J. N., Munoz, D. P., & Itti, L. (2013). High-throughput classification of clinical populations from natural viewing eye movements. *Journal of Neurology, 260*(1), 275–284.

Underwood, G., & Foulsham, T. (2006). Visual saliency and semantic incongruency influence eye movements when inspecting pictures. *Quarterly Journal of Experimental Psychology, 59*(11), 1931–1949.

Unema, P. J., Pannasch, S., Joos, M., & Velichkovsky, B. M. (2005). Time course of information processing during scene perception: The relationship between saccade amplitude and fixation duration. *Visual Cognition, 12*(3), 473–494.

Van der Lans, R., Pieters, R., & Wedel, M. (2008). Competitive brand salience. *Marketing Science, 27*(5), 922–931.

Vig, E., Dorr, M., & Cox, D. (2014). Large-scale optimization of hierarchical features for saliency prediction in natural images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2798–2805).

Vogt, S., & Magnussen, S. (2007). Expertise in pictorial perception: Eye-movement patterns and visual memory in artists and laymen. *Perception, 36*(1), 91–100.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19*(9), 1395–1407.

Walther, D., Rutishauser, U., Koch, C., & Perona, P. (2005). Selective visual attention enables learning and recognition of multiple objects in cluttered scenes. *Computer Vision and Image Understanding, 100*(1), 41–63.

Wedel, M., & Pieters, R. (2008). A review of eye-tracking research in marketing. *Review of Marketing Research, 4*(2008), 123–147.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic bulletin & review*, *1*(2), 202–238.

Wooding, D. S. (2002). Eye movements of large populations: II. Deriving regions of interest, coverage, and similarity using fixation maps. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 518–528.

Wurtz, R. H., & Goldberg, M. E. (1972). Activity of superior colliculus in behaving monkey. III. Cells discharging before eye movements. *Journal of Neurophysiology*, *35*(4), 575–586.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum press.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*(4), 787.

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences, 1339*(1), 154–164.

Zelinsky, G. J., & Loschky, L. C. (2005). Eye movements serialize memory for objects in scenes. *Perception and Psychophysics, 67*(4), 676–690.

Zhang, L., Tong, M. H., Marks, T. K., Shan, H., & Cottrell, G. W. (2008). SUN: A Bayesian framework for saliency using natural statistics. *Journal of Vision, 8*(7), 32.

Chapter 7
Eye Movements During Reading

Jukka Hyönä and Johanna K. Kaakinen

Contents

Abstract The present chapter is a brief introduction to research on eye movements during reading. In the chapter, we will describe what the eye movement methodology has revealed about word recognition, syntactic parsing of sentences, and comprehension of longer segments of text. The typical eye movement pattern in reading is one where the reader makes a sequence of left to right eye movements from one word to the next so that most words are fixated at least once. Words that typically do not receive a fixation are short, high-frequency function words, such as articles and prepositions. Fixation time spent on individual words and larger text units can then be used to tap into the cognitive processes that make reading possible. A wealth of data has been accumulated on using fixation time measures to investigate word recognition processes during reading. These studies have shown that fixation times

---

J. Hyönä (✉) · J. K. Kaakinen
Department of Psychology, University of Turku, 20014 Turku, Finland
e-mail: hyona@utu.fi

J. K. Kaakinen
Turku Institute for Advanced Studies, University of Turku, Turku, Finland

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_7

reliably reflect the underlying mental processes ongoing during fixations on words. It has been shown, among other things, that orthographic, phonological and semantic features all influence the time spent fixating a word. This also applies to developing readers, who differ from mature readers in making more and longer fixations, shorter saccades and more regressions. The differences in eye movement patterns between developing and mature readers are primarily explained by lexical processing efficiency, not by low-level oculomotor factors. Eye movements have also been used to study processing of non-foveal text information. This is referred to as parafoveal word processing. Here the interesting questions are how far to the periphery useful information can be extracted and what kind of information can be gleaned from words around the fixated word. It has been found that the readers' perceptual span is strongly biased toward the reading direction, that is, to the right when reading from left to right. Developing readers have smaller perceptual span than proficient readers. Adult readers extract parafoveally visual, orthographic and phonological information from the word to the right of the fixated word. Recent evidence suggests that they are also capable of extracting parafoveally semantic information concerning word meanings. Developmental studies of parafoveal word processing are just beginning to accrue. In the study of sentence parsing, that is, assigning a syntactic structure to a sentence, eye-tracking has become the gold standard. Having reliable and sensitive measures of the timing of effects is pertinent in testing predictions derived from different competing theories of sentence parsing. That sentences (and clauses) are regarded as processing units is demonstrated by the sentence wrap-up effect. Readers pause for longer time at sentence-final words before proceeding to the next sentence, presumably in order to integrate the words into a coherent sentence meaning. Eye movement studies of sentence parsing have demonstrated that parsing takes place incrementally as readers proceed through the sentence. Syntactic ambiguity, syntactic complexity and syntactic violations, among other things, have been shown to influence the eye movement patterns during sentence comprehension. Reading texts longer than single sentences has been the least researched area in the study of eye movements during reading. However, with the help of eye fixation measures developed to study global text processing, eye movement studies on text comprehension have recently been gaining increased popularity. To date, eye fixation patterns have been used to study how readers construct and update a mental representation of the text contents. For example, studies have been conducted on how readers solve inconsistencies in text and comprehend ironic statements. Moreover, research has shown that the reading goal has pervasive effects on eye movements during reading. Among other things, readers make more and longer fixations when reading text segments relevant to their reading goal. Also developing readers are capable of using a reading goal to adjust their intake of text information as they proceed through the text.

Keywords Eye movements · Reading · Word recognition · Reading development · Text comprehension · Task effects · Individual differences

7.1 Introduction and Learning Objectives

We as readers experience reading as if we smoothly slid our eyes from one text line to the next. However, our intuition does not hold true. Stable periods where the eyes stay relatively still are frequently interspersed with abrupt, fast movements of the eyes. Stable periods are called fixations and movements are called saccades, which are the fastest motor movements humans can make. A typical eye movement pattern of an adult reading a sentence in Finnish is depicted in Fig. 7.1.

During reading we also experience large parts of the written text to be readily available for us to take in. Also this intuition is incorrect. The foveal area of the eye where the visual acuity is best is very limited in scope, only about 2° of visual angle. Hence, we can typically identify only one or two words at a time. In order to bring the foveal area to an optimal location for word identification, we need to make frequent saccades from one word to the next. Even though we make frequent fixations in the text, the individual fixations are short, typically lasting for 200–250 ms among competent readers. Thus, adult readers make 4–5 fixations in a second. A typical saccade extends about 7–10 letters in reading an alphabetic language. However, saccade length depends on the orthographic properties of the language; for example, they are much shorter when reading a logographic script like Chinese. Box 1 provides a summary of the basic aspects of eye movements during reading (Rayner, 1998, 2009).

Box 1: Basic Features of Eye Movements During Reading

Fixation = the period of time when the eyes remain relatively stationary in one place in text. Intake of visual input takes place during fixations. Average fixation duration in reading is about 200–250 ms.

Saccade = a fast, ballistic eye movement that take the eyes from one word to another. Due to the limitations of the foveal area (which extends only about 2° of visual angle around the point of fixation), readers need to make a series saccades to visually sample the written text. Saccades typically last for about 20–40 ms in reading; its average length in reading alphabetic script is about 7–10 letters (somewhat depending on the script).

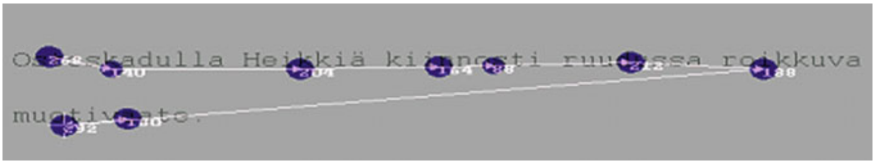


Fig. 7.1 A typical eye movement pattern of an adult reading a sentence in Finnish. Dark circles depict fixations, the number attached to it its duration, and the white arrows depict saccades and their direction

Return sweep = a long-range saccade that take the eyes from the end of one text line to the beginning of next line. It is typically followed by a short corrective saccade to the left (when reading from left to right).
Saccadic suppression = during saccades, no visual input is acquired, so in that sense the reader is functionally blind during the saccades. It is utilized in the boundary paradigm (see Box 2). If a change is made in the text during a saccade, the reader does not notice the actual change taking place.
Regression = a saccade launched to the opposite direction from the normal reading direction. Short, corrective regressions often appear after a return sweep or after launching a saccade that lands too far to the right of the word's center (i.e., the optimal viewing position). Longer regressions (for more details, see the Section on *Eye movements in text reading*) are made in the service of comprehension monitoring: when the reader (a) has misunderstood something, (b) has forgotten something, (c) needs to resolve an inconsistency between two text elements, (d) wants to refresh his/her mental representation of the text by taking a second sample of a text region, or (e) is not ready to move on to a new text region (e.g., from one sentence to the next).

As the readers fixate on nearly every word (on longer words even more than once), an eye movement record provides a researcher with a real-time protocol of how the reader proceeds with his/her eyes through the text. Thus, readers' eye fixation patterns have successfully been used to investigate various mental processes ongoing during reading. This is made possible by a close link existing between where the eyes are gazing at and what the mind is engaged with. This phenomenon is referred to as the eye-mind hypothesis (Just & Carpenter, 1980). Difficulty in processing may be reflected in longer and/or more fixations on the text region requiring extra effort to be comprehended. The reader may also go back in text when realizing that (s)he has misunderstood something in the previous text or (s)he would like to double-check his/her understanding of previous text. All this is faithfully reflected in readers' eye movements that can then be used to tap into the mental processes ongoing during reading. To date, readers' eye movement recordings have been very productively put to use for studying word recognition, the size and nature of the effective visual field, and syntactic parsing of sentence and clause structure among competent adult readers. Fewer eye movement studies have examined comprehension processes ongoing when reading longer texts.

Recently, an increased number of studies have been devoted to using eye movements to investigate how the reading skill evolves during the initial stages of skill development (for a review, see Blythe, 2014; Blythe & Joseph, 2011). The interest here has been to study (a) how reading development is reflected in the eye movement record and (b) what added value may be gained by using eye-tracking to study early reading development. Hence, in this chapter we also review recent studies focusing on reading development among normally developing children (for atypical development, see Klein et al., this volume).

In what follows, we review key findings of eye movement studies on word recognition, sentence parsing and text comprehension (for more comprehensive reviews, see Rayner, 1998, 2009) among competent readers and to some extent also among developing readers. Our focus is on reading alphabetic scripts. Readers interested in Chinese reading may consult, for example, the special issue edited by Liversedge, Hyönä, and Rayner (2013) and the review of Zang, Liversedge, Bai, and Yang (2011).

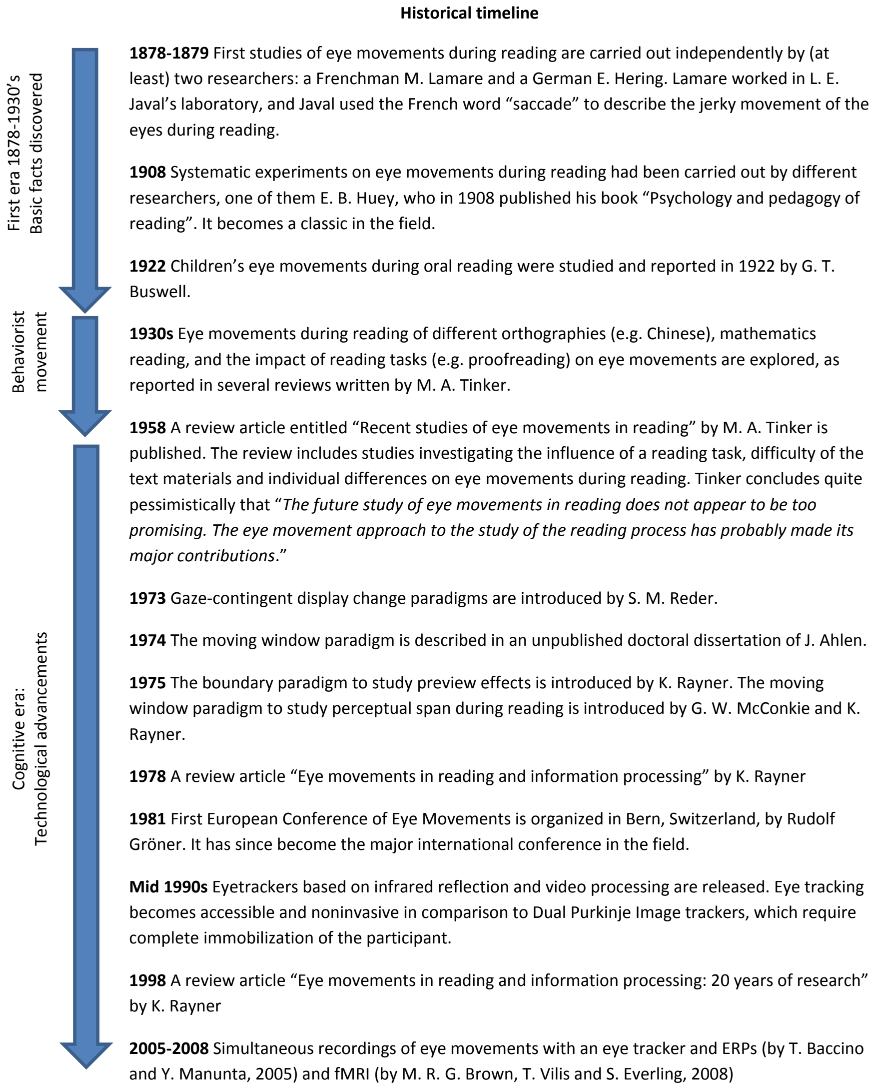
After reading this chapter, the learner knows:

- What the major determinants of fixation time are on individual words.
- What type of information is extracted from the parafovea during reading.
- What kind of eye movement paradigms have been developed to study reading, and what is the logic behind them.
- How visual attention and eye movements are related to each other.
- What serial and parallel models of eye movement guidance are and how they differ from each other.
- How eye movements of younger, less skilled readers differ from those of more mature adult readers.
- How sentence comprehension processes are reflected in readers' eye movements.
- How good comprehension (in comparison to poor comprehension) is reflected in eye movements during text reading.
- How the reading task influences readers' eye movements.

7.2 Historical Annotations

The history of eye movement research can be roughly divided into three eras: (1) the early days (from the end of 1870s to 1930s) when the basic facts about the nature of eye movements during reading were discovered, (2) the time of the behaviorist movement (1930s–1956) that was characterized by fewer, yet very interesting, studies of eye movements and reading, and (3) the cognitive era (from 1956 onwards), which is characterized by remarkable methodological and technological advancements. A brief timeline of the most important events is presented in Fig. 7.2 (see also Rayner, 1978, 1998; Wade & Tatler, 2009).

The history of observing eye movements during reading date back to the end of 1870s, when the "jerky" nature of readers' eye behavior was documented independently by two researchers, a Frenchman M. Lamarre, and a German researcher E. Hering (Wade & Tatler, 2009). The first eye tracking devices were fairly crude mechanical systems that were based on detecting eye movements via a mirror and concurrently detecting sounds produced by saccadic movements via a rubber tube placed on the eyelid (see Wade & Tatler, 2009). Despite their simplicity, the systems were accurate enough to provide basic facts about eye movements during reading. These initial findings are reported in a book by Huey (1908), which can be regarded as one of classics in the field.



Historical timeline

1878-1879 First studies of eye movements during reading are carried out independently by (at least) two researchers: a Frenchman M. Lamare and a German E. Hering. Lamare worked in L. E. Javal's laboratory, and Javal used the French word "saccade" to describe the jerky movement of the eyes during reading.

1908 Systematic experiments on eye movements during reading had been carried out by different researchers, one of them E. B. Huey, who in 1908 published his book "Psychology and pedagogy of reading". It becomes a classic in the field.

1922 Children's eye movements during oral reading were studied and reported in 1922 by G. T. Buswell.

1930s Eye movements during reading of different orthographies (e.g. Chinese), mathematics reading, and the impact of reading tasks (e.g. proofreading) on eye movements are explored, as reported in several reviews written by M. A. Tinker.

1958 A review article entitled "Recent studies of eye movements in reading" by M. A. Tinker is published. The review includes studies investigating the influence of a reading task, difficulty of the text materials and individual differences on eye movements during reading. Tinker concludes quite pessimistically that "*The future study of eye movements in reading does not appear to be too promising. The eye movement approach to the study of the reading process has probably made its major contributions*."

1973 Gaze-contingent display change paradigms are introduced by S. M. Reder.

1974 The moving window paradigm is described in an unpublished doctoral dissertation of J. Ahlen.

1975 The boundary paradigm to study preview effects is introduced by K. Rayner. The moving window paradigm to study perceptual span during reading is introduced by G. W. McConkie and K. Rayner.

1978 A review article "Eye movements in reading and information processing" by K. Rayner

1981 First European Conference of Eye Movements is organized in Bern, Switzerland, by Rudolf Gröner. It has since become the major international conference in the field.

Mid 1990s Eyetrackers based on infrared reflection and video processing are released. Eye tracking becomes accessible and noninvasive in comparison to Dual Purkinje Image trackers, which require complete immobilization of the participant.

1998 A review article "Eye movements in reading and information processing: 20 years of research" by K. Rayner

2005-2008 Simultaneous recordings of eye movements with an eye tracker and ERPs (by T. Baccino and Y. Manunta, 2005) and fMRI (by M. R. G. Brown, T. Vilis and S. Everling, 2008)

Fig. 7.2 A time line of the most important events in the history of eye movement research

After the initial stages, there was a slight decrease in the use of eye movement recordings to study reading, which has been connected with the rise of the behaviorist movement. However, the research topics that were covered at that time are still timely. For example, Tinker (1958) reviews studies that examined the influence of text difficulty on eye fixations, compared eye movement patterns during reading in 14 different languages, observed differences in fixation times induced by various reading tasks, and investigated developmental trends in eye movements during reading.

As the cognitive revolution started in the mid-1950s, several methodological and technological advancements helped to spark new research. The introduction of gaze-contingent display paradigms (see Box 2) provided new tools for studying the control of eye movements during reading. Since then, eye tracking devices have become easier to use without compromises to the data quality, and trackers have now even been combined with brain imaging techniques, such as EEG (e.g. Baccino & Manunta, 2005) and fMRI (e.g., Brown, Vilis, & Everling, 2008).

7.3 Theories of Eye Guidance in Reading

Over the years, eye movement research in reading has gathered a wealth of data that has allowed scientists to propose detailed, mathematically formulated models of eye guidance during reading. The proposed theories model local eye movement patterns pertaining to word processing. They are built to model (a) foveal word processing (i.e., eye movement patterns on words when they are fixated) as well as (b) parafoveal word processing (i.e., the extent to which processing extends to words adjacent to the fixated word). Next, key aspects of the two most prominent models are described. This is followed by a summary of what is currently known about foveal and parafoveal word processing, as revealed by readers' eye movements.

The most influential model of eye guidance during reading is the E-Z Reader model put forth by Reichle et al. (1998, 2003; Pollatsek et al., 2006). It models fixation durations on words and saccade programming between words. It is a serial model in that attention is allocated to only one word at a time. Another key tenet of the model is that shifting of attention and saccade programming are decoupled. When a word is fixated, attention is allocated to that word in order to access its identity. Once this lexical access is imminent, and processing has reached a state where the word looks familiar and is about to be identified, a saccade is programmed to the next word. However, attention stays at the fixated word until the lexical access has completed. The familiarity check is reached sooner when the word is highly familiar to the reader or is predictable from the prior text context. This way the model simulates the well-established word frequency and predictability effects discussed in more detail below.

Due to saccadic latency, some time elapses before a saccade is executed after it is programmed. This is called the labile stage: a saccade is initiated but not yet executed, and it is still possible to cancel and reprogram it. If lexical access is completed for the currently fixated word, which we will call the Word N, and a saccade to the next word (Word N + 1) has not yet been executed, attention shifts to Word N + 1 before the eyes. This way processing of Word N + 1 is initiated prior to its foveal inspection. This is the time period when Word N + 1 is parafoveally processed. During this time, the eyes and attention are decoupled. In the relatively rare cases when there is sufficient time for the familiarity check to complete also for Word N + 1, the following saccade to Word N + 1 may be canceled and reprogrammed to Word N + 2, in which case Word N + 1 is skipped. This will take place if the familiarity check for Word N + 1 occurs during the labile stage when the saccade may still be reprogrammed. If it completes after that, during the so-called non-labile stage, the saccade can no longer be reprogrammed, but a saccade is executed to Word N + 1, followed by a short fixation on it prior to saccading to Word N + 2. The E-Z Reader model also takes into account visual acuity in modeling fixation durations on words. The further the individual letters of the word are from the fixation center, the longer the familiarity check takes.

As described above, the E-Z Reader model assumes words to be attended serially one at a time. Attention does not shift to the next word before the currently fixated word is identified. Its prime competitor, the SWIFT (Autonomous **S**accade-Generation **W**ith **I**nhibition by **F**oveal **T**argets) model (Engbert et al., 2005) challenges this assumption by postulating that multiple words (up to four) may be simultaneously attended. This is achieved on the basis of a dynamically determined processing gradient. Not only the fixated word is activated, but also activation related to Word N − 1 may linger on, if it is not completely identified. Moreover, Word N + 1, and possibly Word N + 2 may receive activation, depending on the scope of the attentional gradient at any one time. Due to the dynamic nature of the attentional gradient, processing may also be limited to a single word, comparable to serial processing assumed by E-Z Reader.

Another key difference between the two models is that SWIFT assumes saccade initiation to be a stochastic process, rather than determined by word processing difficulty, as assumed by E-Z Reader. Yet, word processing can impact on saccade initiation by inhibiting a saccade when lexical access is incomplete. This way SWIFT is able to simulate effects of word frequency and predictability known to affect fixation time on words. However, it is done differently from E-Z Reader. While E-Z Reader assumes lexical processing to be the driving force of eye guidance during reading, in SWIFT lexical processing influences saccade programming by delaying stochastic saccade programming in response to processing difficulty.

A third key difference between the models concerns the mechanism responsible for determining saccade targets. As described above, according to E-Z Reader the target for the next saccade from Word N is Word N + 1, unless the familiarity check has sufficiently early reached completion also for Word N + 1, in which case Word N + 2 is targeted. According to SWIFT, on the other hand, saccade targets are selected on the basis of the dynamic lexical activation pattern within the attentional gradient. Target selection occurs as a competition between words that possess a variable degree of lexical activation. "Activation is built up in a preprocessing stage and decreases during a later lexical completion process. The relative amount of activation will determine the probability that a word is selected as a saccade target" (Engbert et al., 2005, p. 782). In other words, the word with the highest lexical activation at a given time will be selected as the target for the next saccade. For example, if the fixated word is difficult to process, lexical access is not completed and its activation remains high. As a consequence, it may win the competition for the next saccade, resulting in a refixation of that word. On the other hand, when a word is easy to process, the lexical access is achieved rapidly resulting in decay in activation. In such a case, a word in the parafovea will have more activation than the fixated word; thus, a saccade is targeted to it. In such a way, using the same underlying mechanism SWIFT is in a position to model all types of saccades: saccades to the following word, refixations of the same word, word skipping, and regressions back to a previous word.

It is important to note that both E-Z Reader and SWIFT model can successfully simulate key findings related to eye movements in reading (discussed below). They do that by postulating different mechanisms governing fixation durations and saccade programming. They are both specified in a series of mathematical equations that make possible to estimate their predictions to the observed data. They both also make new predictions for which no data are available. Thus, they have the potential to move research forward in a theory-driven way. Finally, it should be noted that E-Z Reader and SWIFT are not the only models of eye guidance in reading. An interested reader may consult Reichle (in press), who provides a comprehensive review of all major theories of reading, including also theories not making recourse to readers' eye movements.

7.4 Eye Movement Paradigms Used to Study Reading

Eye movement recordings can be utilized in various ways to study reading-related cognitive processes. In a typical reading study, participants are asked to read single sentences silently for comprehension. Fixation durations on words and saccade lengths are then computed to examine how different word characteristics, which can be experimentally varied, influence the eye movement patterns (Rayner, 1998). Comprehension is periodically tested by either asking participants to paraphrase the sentence, or by asking readers to respond to statements regarding the sentence contents. Reading of texts consisting of more than one sentence can be studied by presenting multiple lines of text on one page, and the text may also extend to more than one page. Depending on the accuracy of the eye tracker, the line spacing is adjusted to be wide enough (i.e., 2.5 line spacing) in order to reliably differentiate which line of text the reader is currently fixating on. In addition to computing fixation durations on words, it may be necessary to compute fixation durations on phrases or sentences (Hyönä, Lorch, & Rinck, 2003). After reading, a measure of comprehension (e.g. free recall) is collected to check that the reader was engaged with the reading task; it is also fruitful to examine the relationship between comprehension and the way that the text was inspected.

In addition to simply recording eye movements during reading of sentences or texts, the different applications of the gaze-contingent display change paradigm (see Box 2) can be used to examine in more detail the cognitive processes underlying reading. In these paradigms, eye gaze is constantly tracked with an eye tracker, and changes to the text display are made depending on the location and direction of the movement of the eye gaze. Studies utilizing these paradigms reveal interesting facts about the interplay between vision and cognition during reading: how much information can be extracted on one fixation, how much time is needed for different types of information to be extracted from a word, what kind of information can be extracted from the parafovea, and whether regressive eye movements are needed for comprehension.

Box 2: Gaze-Contingent Eye Movement Paradigms

Examples of gaze-contingent eye movement paradigms are presented in Table 7.1. In the examples, a circle indicates fixation location.

In the **disappearing text paradigm** (Liversedge et al., 2004), the fixated word either disappears or is masked after the fixation onset, e.g. after a 40 ms delay. By varying the time of disappearance, it is used to estimate the minimum exposure time needed for reading to continue with normal speed.

In the **fast priming paradigm** (Sereno & Rayner, 1992), the target word is initially replaced with a prime stimulus. After a short delay (e.g. 30 ms) after fixation onset on the word, the prime is replaced with the target. The paradigm is used to examine what kinds of primes facilitate reading.

In the **moving window paradigm** (McConkie & Rayner, 1975), only a specified area (e.g., 11 characters to the right and left) around the point of fixation is displayed normally while other parts are masked. The window moves in synchrony with the eyes. By varying the size of the window and the type of mask (e.g., X's, visually similar or dissimilar characters) and comparing the reading times in the window and normal reading conditions, it is possible to define the size of the area from which a reader can efficiently extract and utilize information.

The **boundary paradigm** (Rayner, 1975) makes use of the **saccadic suppression.** Saccadic suppression means that during a saccade the intake of visual information is suspended and the reader is practically blind. If a change in the visual environment is made during a saccade or very soon after the eyes have landed (< 6 ms after the end of a saccade, McConkie & Loschky, 2002), the reader does not become consciously aware of it. The target word ("*sentence*" in the example of Table 7.1) is initially masked with a character string ("*somkasoc*"), and when the reader's eyes cross an invisible boundary in the text, the mask is replaced with the actual target word. If the reader has extracted information from the target word preview prior to its change to the correct form, one should observe increased fixation time on the target word, even though the reader is not consciously aware of this. The size of the slow-down in eye fixation time, i.e. the difference between normal condition in which no change was made and a change condition is called the **preview effect**.

In the **trailing mask paradigm** (Schotter, Tran, & Rayner, 2014), previously fixated words are replaced with a mask as soon as the reader moves away from the word so that if the reader would return back to already read words, no useful information is available. Words to the right of the fixation point are presented normally.

Table 7.1 Examples of the gaze-contingent display change paradigms. The gray circle marks the fixation location

Paradigm	Example
Normal reading	`This is an example sentence to demonstrate the paradigm.`

Disappearing text	Initially:
	`This is an example sentence to demonstrate the paradigm.`
	After a delay:
	`This is an example ○ to demonstrate the paradigm.`

Fast priming	Initially:
	`This is an example s n○ nc to demonstrate the paradigm.`
	After a delay:
	`This is an example sentence to demonstrate the paradigm.`

Moving window	`Xxxx is an example sentencx xx xxxxxxxxxxx xxx xxxxxxxx.` `Xxxx xx xx xxample sentence to demoxxxxxxx xxx xxxxxxxx.`
Invisible boundary	`This is an example somkasoc to demonstrate the paradigm.` `This is an example sentence to demonstrate the paradigm.`
Trailing mask	`Xxxx xx xx example sentence to demonstrate the paradigm.` `Xxxx xx xx xxxxxxx sentence to demonstrate the paradigm.`

Two paradigms have proven useful in examining foveal word processing: the disappearing text paradigm and the fast priming paradigm. In the disappearing text paradigm (see Box 2), a target word is initially presented normally. However, after a fixation lands on the word, the word disappears from screen. By varying the length of the delay of the disappearance, it is possible to infer how long visual exposure time is required for the word identification to proceed normally. In the fast priming paradigm, a prime stimulus is initially presented in place of the target word. After a short delay (e.g. 30 ms) from fixation onset on the word, the prime is replaced with the target word. In the example presented in Table 7.1, the prime comprises all the consonants of the word to examine whether consonants play a privileged role in early stages of word processing. If so, presenting the vowels first followed shortly by the consonants would delay the word recognition compared to the situation depicted in Table 7.1.

Two main paradigms used to study parafoveal information processing during reading are the *text window paradigm* and the *boundary paradigm* described in Box 2. In the text window paradigm, the reader sees only a certain amount of useful information around the point of fixation. The window moves in synchrony with the eyes, and by varying the size of the window, it is possible to estimate the size of the effective field of vision, or as it is often called, the size of a reader's perceptual span.

In the boundary paradigm (see Box 2), the target word is initially masked with a letter string. However, when the eye gaze crosses an invisible boundary placed in the end of the preceding word, the target word appears normally. By manipulating the type of preview it is possible to infer what type of information is extracted from the target word before it is actually fixated. For example, visual similarity is manipulated in the boundary paradigm by replacing letters in the parafovea with visually similar (e.g., k with h) or dissimilar letters (e.g., g with t). The main index is the parafoveal *preview benefit*, which assesses the amount of facilitation in processing gained by different kinds of parafoveal previews. The preview benefit is simply computed as the difference in fixation time between a full preview condition, in which the target word was presented normally, and the preview condition. Another measure to assess parafoveal processing is the so-called *parafoveal-on-foveal effect* (Kennedy, 2000). It measures the extent to which parafoveally available information affects fixation time on the previous word (Drieghe, 2011).

7.5 Eye Movements During Foveal and Parafoveal Processing of Words

In this section, we summarize what is known about foveal and parafoveal processing of single words, as revealed by readers' eye movement patterns. Foveal word processing refers to cognitive processes carried out for the currently fixated word that falls onto the foveal vision. Parafoveal word processing, in turn, refers to processing done in the parafoveal region extending up to 5° of visual angle to the right and left of the current fixation. In what follows, we first discuss foveal word processing among adult readers, followed by a section focusing on young, developing readers.

7.5.1 Foveal Word Processing Among Competent Adult Readers

When readers fixate a word, fixation time spent on the word faithfully reflect cognitive processes needed to identify the word. The most frequently used eye fixation measure to tap into the foveal word processing is gaze duration, which sums up the durations of fixations made on the word when it is first encountered and before a saccade is launched away from the word (typically to the subsequent word). A robust and consistent finding has been the *word frequency effect*. Written words that appear infrequently in the language are read with longer gaze durations than words whose frequency is high (e.g., Inhoff & Rayner, 1986; Rayner & Duffy, 1986). The effect materializes either as longer durations of individual fixations made on the infrequent word or as an increased probability of making a refixation on the word, or both. An intriguing observation has been made by Liversedge et al. (2004) using the disappearing text paradigm (see Box 2; see also Rayner, Inhoff, Morrison, Slowiaczek, & Bertera, 1981). In their experiment, high- and low-frequency words were presented for foveal inspection during sentence reading for 60 ms, after which the word disappeared and the readers fixated on an empty space between two parafoveal words. Gaze durations on the empty space were longer when the empty space replaced a low-frequency word than a high-frequency word. This is compelling evidence for the view that fixation times reflect the mental processes ongoing during reading. The process of accessing a mental representation for a low-frequency word takes more time to complete than that for a high-frequency word. This mental process is reflected in the fixation time on the empty space even when the to-be-identified word is no longer visually present. The study also demonstrates that only a relatively short exposure time (60 ms or so) is enough to acquire sufficient visual input in the system for the word recognition process to proceed normally.

Another key finding is the word predictability effect: words that are highly predictable from the prior sentence or discourse context receive shorter fixation times than contextually unpredictable words (e.g., Balota, Pollatsek, & Rayner, 1985; Calvo & Meseguer, 2002; Hyönä, 1993; Rayner & Well, 1996). Also local, lexically-based predictability influences fixation time on words (Vainio, Hyönä, & Pajunen, 2009). When a verb strongly constrains the identity of the upcoming word(s) (e.g. "he *hunched* his back"), gaze duration is shorter on the highly constrained phrase ("his back") than on the same phrase preceded by a less constraining verb (e.g., "he *hurt* his back").

Alphabetic scripts are based on principles of converting spoken language codes to written counterparts. In orthographically completely transparent scripts there is a direct mapping between letters (i.e., graphemes) and sounds (i.e., phonemes). In other words, each grapheme corresponds to only one phoneme, and each phoneme is always represented by the same grapheme. A prime example of such script is Finnish where the phoneme-grapheme correspondence is practically 100%. Examples of less straightforward or more opaque alphabetic languages are English and Danish. An example of a phonologically opaque word in English is "choir", for which rule-based grapheme to phoneme mapping would yield an output significantly different from the correct one.

When readers of alphabetic scripts process words in sentences, perhaps unsurprisingly the identification process entails a phonological recoding phase, which is reflected in fixation times in words. For example, Inhoff and Topolski (1994) found longer fixation times on words that had an irregular (e.g., "weird") than regular (e.g., "mood") spelling. The effect was short-lived, as it was obtained for the initial fixation made on the word but not for gaze duration (i.e., the summed duration of fixations made before fixating away from the word). This finding is evidence for the early activation of phonological codes during word recognition. Further evidence supporting early activation of phonological representations comes from studies where homophones were inserted in target sentences. Homophones are words that sound the same but are written differently (e.g., "bear" and "bare"). Rayner, Pollatsek, and Binder (1998) observed no difference in first fixation duration between the correct word form and its homophonic counterpart, despite the fact that the homophonic word was different in meaning. Signs of meaning activation were observed in gaze duration, which was longer for the homophonic than the correct word. Analogous results were observed for French by Sparrow and Miellet (2002) who found no difference in first fixation duration between correctly spelled words and homophonic non-words. Finally, Sereno and Rayner (2000) demonstrated that phonological recoding is more robust for infrequent than frequent words. They observed a phonological regularity effect for infrequent words but not for frequent words.

Written words are identified via the individual letters they contain. Thus, it is not surprising that the characteristics of letters and letter clusters are also reflected in fixation times on words. Word-external letters appear to be more relevant for successful word identification than word-internal letters. This became evident from the study of White, Johnson, Liversedge, and Rayner (2008). White et al. jumbled up letters both word-internally (e.g., "problem" vs. "probelm") and word-externally ("problme" and "*rpoblem*"). The transposed-letter conditions produced longer fixation times than the correct condition. The transposed letter effect in fixation times was greater for the word-external than word-internal transpositions, indicating that letters in the beginning and end of a word are more crucial for word identification than letters in the middle of the word.

Another finding is that consonants play a more significant role in word recognition than vowels. This was demonstrated by Lee, Rayner, and Pollatsek (2001, 2002) using the so-called *fast priming paradigm* (Sereno & Rayner, 1992). In their version of the paradigm, when a word was fixated the presentation of one of the word's letters, either a consonant or a vowel, was delayed for 30 or 60 ms. They found that delaying the presentation of a consonant for 30 ms resulted in significantly longer gaze durations on the word than delaying the presentation of a vowel. Such effect was not present in the 60-ms presentation condition. The pattern of results was interpreted to suggest that in the early stages of foveal word processing, consonants play a more significant role than vowels.

Where in the word the reader initially lands with his/her eyes also has a significant impact on eye behavior on words and on the word processing efficiency. If the eyes land in the word center, the reader is less likely to make a refixation on the word, and the gaze duration is much shorter than is the case when the eyes initially land in the word beginning or end. This was first demonstrated by O'Regan, Lévy-Schoen, Pynte, and Brugaillere (1984) for isolated words, but it was subsequently extended to reading words in sentences (e.g., Hyönä & Bertram, 2011; Nuthmann, Engbert, & Kliegl, 2005; Vitu, O'Regan, & Mittau, 1990; Vitu, McConkie, Kerr, & O'Regan, 2001). Thus , word center is the optimal viewing position (OVP) for smooth word processing. Presumably, this is due to all or most letters of the word falling on the fovea, at least when the word is relatively short. Fortunately, OVP is also close to the preferred landing position in reading (Rayner, 1979); that is, readers are likely to launch the first saccade into the word so that it lands close to the word center. The preferred landing position typically departs somewhat from the optimal viewing location by being a bit closer to the word beginning. A saccade to a word is launched on the basis of the length information extracted from the parafovea. Thus, when a word is short, the amplitude of the saccade into the word is shorter than is the case when programming a saccade to a long word (McConkie, Kerr, Reddix, & Zola, 1988).

The research has also identified an intriguing phenomenon that runs counter to the OVP effect. When readers make only one fixation on the word, this single fixation is longest when positioned in the word center and shortest when located toward word beginning or end. In other words, if only fixation is made on a word, word center is *not* the most optimal viewing position. This finding has been coined the inverted optimal viewing position (IOVP). Its exact nature is not yet known. It is at least partly explained by mislocated fixations that land toward the word beginning or end. The idea here is that the fixation at the word beginning is intended to land on the previous word and the fixation at the word end is intended to land on the subsequent word. Because a corrective saccade is quickly programmed, the fixation duration on these mislocated fixations is short. The relatively long duration of a single fixation that lands in the middle of the word may be also related to the amount of perceptual information that needs to processed: there is more visual information to be gleaned in the middle of the word than at the word edges (Vitu, Lancelin, & Marrier d'Unienville, 2007).

7.5.2 Foveal Word Processing Among Developing Readers

A significant part of the seminal work on eye movements in reading dealt with differences in eye movement patterns as a function of age and reading development (Buswell, 1922; Taylor, 1965). The seminal work demonstrated that less mature readers make more and longer fixations, shorter saccades and more regressive fixations that take the eyes leftward in text (see Fig. 7.3). More recent work (for a review, see Blythe & Joseph, 2011) has confirmed these global effects. These global differences between developing and mature readers also feature in word-level reading: younger and less experienced readers make more refixations on words and skip over words less frequently than older and more experienced readers. These developmental differences do not reflect differences in maturation of the oculomotor system, but are instead a reflection of differences in lexical processing efficiency. This becomes apparent, for example, in model simulations carried out by Reichle et al. (2013) and Mancheva et al. (2015). They applied the E-Z Reader model of eye movement control in reading (see Sect. 7.4) to account for developmental differences in readers' eye movement patterns. Reichle et al. (2013) showed that the main difference between children's and adults' eye behavior in reading can be explained by overall lexical processing speed. In other words, the crucial difference between children and adults is that adults are faster in word identification. This conclusion was supported by Mancheva et al. (2015), who observed that the model parameter indexing lexical processing efficiency correlated strongly with children's lexical skills as measured by offline tests, particularly with orthographic processing ability. Also other studies have shown that tasks that tap into linguistic processing efficiency are better predic-

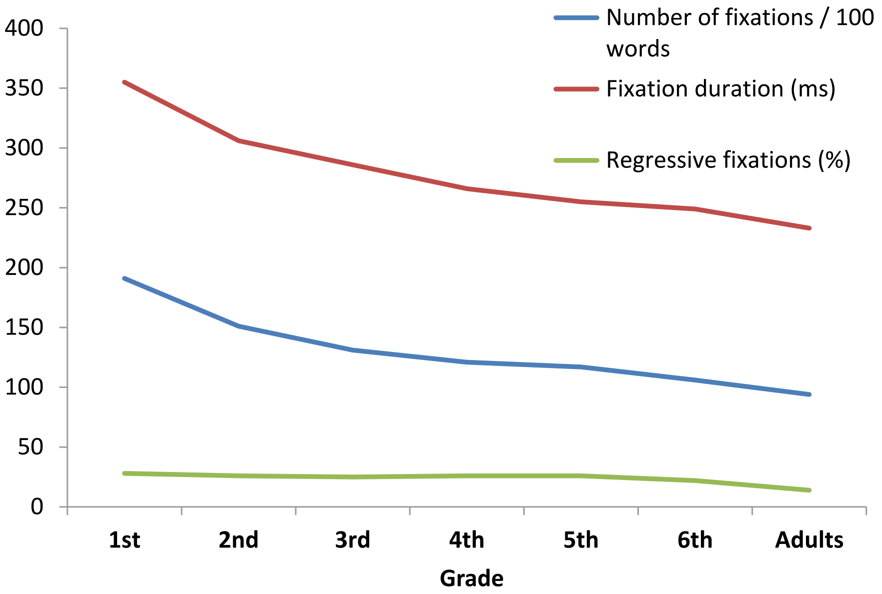


Fig. 7.3 The developmental pattern of some eye movement characteristics during reading from the first to sixth grade and in comparison to skilled adult readers (as reported in Rayner, Ardoin, & Binder, 2013)

tors of eye fixation times on words than tasks that measure oculomotor efficiency (Huestegge, Radach, Corbic, & Huestegge, 2009).

Further proof for the claim that the oculomotor system is well developed among normally developing children when they begin to read comes from the study of McConkie et al. (1991). They showed that already first-grade elementary school children demonstrate the preferred landing position effect obtained among skilled readers (see also Joseph, Liversedge, Blythe, White, & Rayner, 2009). As described above, the phenomenon refers to the finding that the initial fixation in the word lands close to the word center, which is the optimal location for word recognition. Huestegge et al. (2009) found the initial fixation location to be nearer the word beginning among 2nd graders than 4th graders. They ascribed the difference in the 2nd graders' tendency to read words with more than one fixation, whereas 4th graders read words more often with a single fixation. By applying such a refixation strategy, it makes sense for the 2nd graders to launch the initial fixation closer to the word beginning.

Given the findings demonstrating that children's eye fixation patterns reflect their lexical processing efficiency, it is understandable that also children's eye fixation patterns show effects of word frequency and word length. Hyönä and Olson (1995) found that these two factors interacted so that gaze durations on words were particularly long when the words were long and infrequent. Huestegge et al. (2009) in turn found the word frequency and word length effects to be smaller for 4th graders than

2nd graders. Similarly, Joseph et al. (2009) observed stronger word length effects in gaze duration and refixation probability for children than adults.

Studies using the disappearing text paradigm (Liversedge et al., 2004; Box 2) suggest that already very young readers are able to efficiently extract visual information. Blythe, Liversedge, Joseph, White, and Rayner (2009) found that as short as 40–75 ms presentation times were sufficient for reading to proceed normally even among 7-year-old children. A typical finding in studies employing the disappearing text paradigm is that readers stop making refixations on "words" (i.e., the empty space where the word was briefly presented). This is understandable, as there is no visual input, there is nothing to direct a second fixation to. Blythe, Häikiö, Bertram, Hyönä, and Liversedge (2010) used the disappearing text paradigm to study reading of short (4 letters) and longer (8 letters) words among 8–9 year-olds, 10–11 year-olds and adults. The study demonstrated that the youngest (8–9 year-old) children regressed back to the longer words in order to get another sample of the word (in the paradigm, the word reappears when a saccade is launched away from it). This effect was reduced for older children (10–11 year-olds) and absent for adults. The need for obtaining a second visual sample of longer words among young readers is interpreted to reflect younger readers' smaller perceptual span (Häikiö, Bertram, Hyönä, & Niemi, 2009; Rayner, 1986).

7.5.3 Parafoveal Word Processing Among Competent Adult Readers

In addition to extracting text information from the foveated words, readers also glean useful information from the parafovea that extends 5° of visual angle to the right and left around the fixation point.

A key finding that has emerged from the text window (see Box 2) studies is that readers' perceptual span is heavily biased to the right (when reading from left to right; McConkie & Rayner, 1976). The perceptual span extends up to 15 letters to the right of fixation, whereas the major determiner of the left boundary appears to be the beginning of the currently fixated word (Rayner, Well, & Pollatsek, 1980; Rayner, Well, Pollatsek, & Bertera, 1982). Recent work has slightly modified the latter conclusion by showing that the leftward span can extend to the previous word when the reader's attention is not fully disengaged from the word to the left (i.e., processing of Word $N - 1$ is not complete) of the currently fixated word (Apel, Henderson, & Ferreira, 2012; Binder, Pollatsek, & Rayner, 1999; McGowan, 2015). This rightward asymmetry in perceptual span is an attentional effect and not related to the brain's hemispheric specialization. This has become apparent from studies conducted in Hebrew and Arabic, which are read from right to left. It has been found that the perceptual span of Hebrew and Arabic readers is asymmetric (greater) to the left, that is, toward the reading direction (Jordan et al., 2014; Pollatsek, Bolozky, Well, & Rayner, 1981).

A number of studies have examined what type of information readers glean from the parafovea to the right of fixation (when reading from left to right). The primary method in these studies has been the boundary paradigm (see Box 2). A meta-analysis of studies using the boundary paradigm is reported by Vasilev and Angele (2017). These studies have shown that word length information is extracted up to 15 letters to the right of fixation (McConkie & Rayner, 1975). When correct length information is provided of the words in the parafovea, reading is sped up in comparison to the situation when the parafoveal word length information is incorrect (Inhoff, Radach, Eiter, & Juhasz, 2003; White, Rayner, & Liversedge, 2005). Length information is also utilized in saccadic programming. The incoming saccade to the following word is longer when the word is also longer (e.g., McConkie et al., 1988). As discussed above, parafoveal word length information is utilized to position the eyes toward the word center to optimize word recognition.

Orthographic information is also picked up from the parafovea. The overall visual shape is perceived of parafoveal words (e.g., McConkie & Rayner, 1975). This becomes apparent in the fixation times being shorter when the parafoveal previews are visually similar than dissimilar to the correct word. Also letter identity information is processed parafoveally. Using the moving window technique (see Box 2), Häikiö, Bertram, Hyönä, and Niemi (2009) demonstrated that the letter identity span extends up to 9 letters to the right of fixation. The conclusion is based on a comparison between the normal condition (no window) and the condition where outside the window around the fixation letters were replaced with visually similar letters preserving their overall visual shape but changing their identity. By applying the boundary paradigm (see Box 2), Johnson et al. (2007) used parafoveal previews in which the adjacent letters were transposed either word-externally (e.g., *leement* or *elemetn* instead of *element*) or word-internally (*elemnet*). Using 7-letter words as their target stimuli, they showed that transposition led to increased gaze durations on the target word if they were preceded by transposed-letter previews affecting the word-external letters, but not the word-internal letters. This led Johnson et al. conclude that "readers are able to extract information from the first five letters of the word to the right of fixation plus the word-final letter" (p. 222). The conclusion that more parafoveal orthographic information is extracted from word-external than word-internal letters is also supported by the study of Briihl and Inhoff (1995). This is presumably due to word-internal letters suffering more from visual crowding than word-external letters.

Parafoveally available orthographic information can also affect saccadic programming. Hyönä (1995a) found the initial saccade into the word to land closer to the word beginning if there was a highly infrequent letter cluster in the word beginning, compared to the word beginning hosting a frequent letter combination (see also Radach, Inhoff, & Heller, 2004; White & Liversedge, 2006). Similarly, White and Liversedge (2004) found the initial fixation to land closer to the word beginning if the word beginning contained a misspelling. Moreover, when a word contained a misspelling in the beginning, a regression was frequently launched toward the word beginning after an initial fixation on the word.

There is ample evidence that phonological information is extracted from the word appearing in the parafovea. Evidence comes from boundary paradigm (see Box 2) studies with different kinds of phonological preview manipulations. Pollatsek, Lesch, Morris and Rayner (1992) and Miellet and Sparrow (2004) presented homophonic previews of target words (e.g., target word *site* was parafoveally previewed as *cite*), which were found to facilitate subsequent foveal processing of the target words. Chace, Rayner, and Well (2005) obtained a parafoveal homophone effect for skilled readers but not for less skilled readers. Henderson, Dixon, Petersen, Twilley, and Ferreira (1995) manipulated phonological regularity in the word beginning and observed preview words with phonologically regular initial trigrams to benefit subsequent foveal processing more than preview words with irregular initial trigrams. Finally, Ashby, Treiman, Kessler, and Rayner (2006) manipulated phonological vowel concordance (*cherg* -> *chirp* vs. *chord* -> *chirp*) between the preview string (non-word) and the target word. Concordant previews led to shorter gaze durations on the target words than discordant previews, which provides further converging evidence for parafoveal phonological processing.

There has been a long-standing debate about whether readers can parafoveally extract lexical-semantic information. One reason for the sometimes heated debate is that its presence or absence has important implications to the competing theoretical models of eye guidance during reading. As explained in Sect. 4, a key difference between these models is the extent to which word identification is assumed to be serial versus parallel (Engbert & Kliegl, 2011; Reichle, 2011). Evidence for parafoveal lexical-semantic effects speaks for parallel identification of more than one word. Thus, it may be used as evidence for parallel models and against serial models.

Evidence for parafoveal lexical-semantic effects has been sought via parafoveal-on-foveal effects by manipulating the frequency or semantic plausibility of Word N $+$ 1 and measuring their effects on fixation time on Word N. The evidence has been mixed (Drieghe, 2011; Hyönä, 2011). Also the boundary paradigm (see Box 2) has been applied by manipulating the semantic relatedness of parafoveal previews to the intended word. Earlier evidence primarily spoke for an absence of parafoveal semantic effects (Rayner, White, Kambe, Miller, & Liversedge, 2003). More recently, however, Hohenstein and Kliegl (2014) observed a parafoveal semantic effect using the standard boundary paradigm (i.e., semantically related and unrelated previews were parafoveally available for the entire time Word N $-$ 1 was fixated). Hohenstein, Laubrock, and Kliegl (2010) observed fast parafoveal semantic priming using a modified boundary paradigm where the target word was initially replaced with random string of consonants. When Word N $-$ 1 was fixated, the random letter string was first replaced with a semantically related or unrelated word for a variable amount of time (20–125 ms, depending on the experiment), followed by the target word. They found gaze duration on the target word to be shorter in the semantically related than unrelated condition, when the parafoveal preview was present for the first 125 ms of the fixation on Word N $-$ 1 (or for only 80 ms when the parafoveal preview appeared in boldface). The results were taken as evidence in favor of parallel word processing in reading. Finally, Schotter (2013) observed a parafoveal semantic preview effect for synonyms (e.g., *curlers* was replaced with *rollers*) but not for semantic associates (e.g., *curlers* was replaced with *styling*). The study suggests that similarity in meaning between the target word and its parafoveal preview influences the degree to which parafoveal semantic effects may be observed. Moreover, Schotter speculates that it may be easier to find parafoveal semantic effects in languages with regular phoneme-grapheme correspondence rules, such as German (Hoehenstein & Kliegl, 2014; Hohenstein et al., 2010), than in less regular languages, such as English. The idea is that as foveal word processing is made relatively easy in regular languages, more attentional resources may be devoted to parafoveal word processing.

7.5.4 Parafoveal Word Processing Among Developing Readers

Recently, there has been an increased interest in studying parafoveal processing among developing readers. The seminal study of Rayner (1986) demonstrated that the readers' perceptual span develops as a function of reading ability. The perceptual span of 6th grade children was observed to be analogous to that of adults, while the perceptual span of 2nd and 4th grade children was smaller. Häikiö et al. (2009) replicated these developmental trends. Rayner demonstrated that the perceptual span for word length information extends to 11 letters to the right of fixation for 2nd and 4th graders, and it grows up to 14–15 letters among 6th graders and adults. The perceptual span for global letter feature information is somewhat narrower, extending to 7 letters for 2nd graders and to 11–12 letters from the 4th grade onwards. Häikiö et al. studied the perceptual span for letter identity information and found it to grow from 5 letters to the right for 2nd graders to 7 letters among 4th graders and 9 letters among 6th graders and adults. In a moving window study, Sperlich, Schad, and Laubrock (2015) demonstrated that the growth of perceptual span (for letter feature information) in the first stages of the development of reading skill during Grades 1–3 takes place between the second and third school year, whereas little growth is visible between Grades 1 and 2.

Regarding parafoveal word processing, there appears to be little differences between developing and mature readers. The study of Häikiö, Bertram, and Hyönä (2010) was one of the first where the boundary paradigm (see Box 2) was applied to the study of developing readers. Studies with adult readers had demonstrated that readers extend their attentional span more strongly across a spatially unified letter cluster (i.e., an unspaced compound word such as *basketball*) than across a linguistic unit that comprises two words (i.e., spaced compound words such as *tennis ball*) (Hyönä, Bertram, & Pollatsek, 2004; Juhasz, Pollatsek, Hyönä, Drieghe, & Rayner, 2009). Häikiö et al. (2010) extended these results also to developing readers: Surprisingly, despite the fact that 2nd grade readers' perceptual span is significantly smaller than that of adult readers, even they displayed the same effect.

The results of Tiffin-Richards and Schroeder (2015) suggest a developmental shift in parafoveal processing when reading an orthographically regular language

(German). As reading skill develops, readers shift from using parafoveal phonological information to using parafoveal orthographic information. By applying the boundary technique (see Box 2), Tiffin-Richards and Schroeder compared children's and adults' parafoveal processing of phonological and orthographic information in German. Their main finding was that children but not adults showed parafoveal phonological effects. In contrast, adults demonstrated effects indexing orthographic parafoveal processing, while children showed these effects only under specific conditions.

7.6 Eye Movements During Sentence Comprehension

Identification of individual words is not enough for successful reading comprehension, but the successive words need to be integrated to understand the meaning of a whole clause and sentence. This process takes place incrementally, as readers form meanings of successive words as they move forward in text. Yet, eye-tracking studies have shown that readers pause for longer time at sentence (and also clause) boundaries presumably to integrate the sentence meaning (Just & Carpenter, 1980; Rayner, Kambe, & Duffy, 2000; White, Warren, & Reichle, 2011). This phenomenon is called the *sentence wrap-up effect*. Readers do not proceed to the following sentence until they have secured a sufficient understanding of the currently read sentence (or clause). Integrative processing at sentence boundaries may also be reflected in regressive fixations launched to earlier parts of the sentence (Hyönä, 1995b; Kaakinen & Hyönä, 2007) or sometimes to an earlier sentence. For example, when readers process a text with unfamiliar text contents, they are more likely to initiate a regression to an earlier part of the sentence, particularly when the sentence contains information pertinent to their reading goal (Kaakinen & Hyönä, 2007). By regressing back in text, readers provide themselves with another opportunity to visually sample a text region they might find difficult to understand and/or important to form a good grasp of its meaning.

Successful sentence comprehension also requires that syntactic relations between words are sorted out in order to achieve a correct interpretation of the sentence. To do so, the reader needs to parse the syntactic structure of the sentence. Syntactic parsing entails that the reader identifies the actor of the action depicted by the main verb, whom the action is directed to, where the action takes place, etc. Syntactic information is conveyed in sentences, for example, by word order, morphological case marking, verb argument structure (whether a verb is transitive or intransitive), and animacy of the depicted entities (whether or not they refer to animate entities capable of initiating the action depicted by the verb). These processes are reflected in readers' eye movements (Clifton, Staub, & Rayner, 2007; Clifton & Staub, 2011).

That syntactic parsing is typically incremental in nature is nicely demonstrated by the so-called *garden-path effect* (Frazier & Rayner, 1982). The effect reflects a misparse of a sentence that is syntactically locally ambiguous. One such sentence in English is "*Since Jay always jogs a mile seems like a long distance to him*". When reading a sentence like this, the readers are "led down the garden path", as they typically attach "*a mile*" to the first clause as the sentence object ("*Since Jay always jogs a mile*"). However, that is not the intended meaning; instead "*a mile*" should be considered the syntactic subject of the second clause. Readers realize this when they fixate the word "*seems*", which is fixated for a long time and is frequently followed by a regression to an earlier part of the sentence (Mitchell, Shen, Green, & Hogdson, 2008) and a series of rereading fixations made in the service of correcting the initial misparse. This pattern of results is taken as evidence for the so-called *late closure principle* of the garden path theory (Frazier & Rayner, 1982). Apart from syntactic ambiguity, previous research has examined effects of syntactic complexity, syntactic prediction and syntactic violations on readers' eye movement patterns (Clifton & Staub, 2011).

As the eye movement record provides a real-time protocol of processing as it evolves through time, the eye movement data has been used to tease apart the time course of sentence parsing. The relative degree of delay in the observed effects has consequences to the theories of sentence parsing. However, there is no uniform pattern in the timing of effects, with some researchers finding an early syntactic effect obtained in the duration of first fixation made on the critical text region, while other researchers observe only delayed effects, for example in the probability of regression or in the fixation time on the region following the critical region (for an extensive summary of these studies, see Clifton et al., 2007).

7.7 Eye Movements During Text Reading

Understanding a single sentence is not sufficient, but readers also need to integrate the meaning of successive sentences to construct the meaning of whole text paragraphs and even larger text segments. Reading a text consisting of a full page (or several pages) of text means that a reader has to navigate through several lines of text. This is quite different from reading a single sentence that typically extends through only one or two lines; thus, certain eye movement patterns are typical for reading a text page. Moreover, the increased cognitive demands of understanding a text instead of single sentences are reflected in eye movements. In the following, we will outline the typical characteristics of eye movement patterns related to text reading.

As the reader navigates through the lines of text, the eyes move from the beginning of each line of text to the end of it, and then to the beginning of new line. A *return sweep* refers to an eye movement initiated from the end of a line towards the beginning of a new line. This is such a long saccade that it often is inaccurate, typically undershooting the target in the beginning of the new line. Thus, the reader makes a corrective eye movement, and typically there is an "extra", short fixation close to the beginning of the new line of text.

Research suggests that readers do not extract meaningful information, such as word meanings, from the lines below the currently fixated line of text (Pollatsek, Raney, Lagasse, & Rayner, 1993). Occasionally words can be identified one or two lines below the currently fixated line, but three lines down is already too far for word identification to occur. However, even though word identification from more than two lines of text down is not likely, readers do extract information about the layout of the text page. Evidence for this comes from studies conducted by Cauchard, Eyrolle, Cellier, and Hyönä (2010a, b). Cauchard et al. (2010a) examined whether reading is affected by a text window (see Box 2) that restricts the visibility of a text page to the fixated line and two lines above and below the fixated line. The text itself contained organizational signals, such as subheadings and paragraph breaks, typical for expository texts; such signals are used to cue the content structure of the text. It was found that comprehension was poorer in the text window condition in comparison to the normal reading condition. Readers also tended to make more regressions to headings in the normal than in the window condition. These results suggest that readers make use of the page layout information available in the peripheral vision during reading in order to guide long-distance look-backs, which in turn may be crucial for comprehension. In another study using a similar window paradigm (Cauchard et al., 2010b), participants were asked to look for answers to specific questions in text. Readers displayed longer search times, more and longer eye fixations and shorter saccades if they were denied a preview of organizational signals such as headings and paragraph marks in the periphery, indicating that readers do make use of these cues when they navigate through text.

As noted above, it is not enough to identify the individual words and parse the sentence structure, but the information conveyed by the sentences should also be integrated into a coherent memory representation (see, e.g., Kintsch, 1998). These more global integrative processes influence the eye movement patterns already at the word level. Gaze durations on words are shorter during reading of passages than single unrelated sentences (Radach, Huestegge, & Reilly, 2008). Total fixation time on the words, on the other hand, is longer during passage reading than sentence reading. These results indicate that presenting words in a text facilitates the initial encoding of words but increases the need to reread words in text. Increased rereading rate presumably reflects integrative processing at the text or paragraph level. Moreover, saccade amplitudes are greater and saccades land further into the words during passage than sentence reading. In sum, it seems that both the temporal and spatial aspects of eye movements differ between sentence and text reading.

As mentioned earlier, at sentence boundaries readers integrate the information presented in the sentence before they move on to the next sentence, producing increased fixation times on the last word of a sentence (Rayner, Kambe, & Duffy, 2000). If there is a greater need to obtain a well-integrated memory representation of the text, wrap-up times are increased. For example, when a reader encounters text information that is highly pertinent to the reader's reading goal, wrap-up times are longer than when the sentence contains information that is not relevant to the reader (Kaakinen & Hyönä, 2007).

Problems in integrating sentence information to the evolving memory representation of the text are also reflected in sentence wrap-up times. For example, when reading an ambiguous text passage without a title that would indicate what the passage is all about, wrap-up times are increased in comparison to a condition where the title is given before reading (Wiley & Rayner, 2000). When reading a passage describing, for instance, a space trip to the moon, comprehending what it is all about is more difficult if the reader does not know the topic of the text in advance. This is reflected in the eye movement patterns as increased gaze durations on words, more regressive eye movements, and longer sentence wrap-up times.

During reading of longer texts, readers may make regressive eye movements, also called look-backs, to previously read parts of text in order to integrate text information to memory. Instead of being a signature of inefficient reading, as is often believed, regressions to earlier parts of text seem to be fundamental for successful comprehension. Using the trailing mask paradigm (see Box 2), Schotter, Tran, and Rayner (2014) demonstrated that denying readers the opportunity to resample words was harmful for sentence comprehension. Moreover, previous studies suggest that readers who make look-backs to informative parts of the text gain better memory of the information presented in the text than readers who do not make look-backs. This was demonstrated in the study of Hyönä, Lorch, and Kaakinen (2002), where adult readers read expository texts that followed a typical expository text structure and contained subheadings, which marked the topic of the following paragraph. Readers who tended to look back to headings, typically from the end of the paragraph, gained good memory of the information presented in the text. Selective look-backs to informative parts of the text are most likely strategic in nature, meaning that they reflect a conscious decision to reread parts of text that the reader believes will help in constructing a good memory representation of the text (Hyönä & Nurminen, 2006). On the other hand, readers who unselectively reread parts of text showed poorer memory for text information (Hyönä et al., 2002).

Look-backs may also be triggered by problems in integrating text information to the previous text context. An ironic statement is an example of a situation in which an utterance does not literally and directly fit into the context in which it is presented. Consider a phrase "*What a great weather for a picnic!*". If this sentence is presented in a passage describing a rainy and windy day, hence carrying an ironic meaning, readers make more look-backs to it than if it is presented in a context describing a beautiful sunny day (Kaakinen, Olkoniemi, Kinnari, & Hyönä, 2014). Look-backs in this case are assumed to reflect attempts to resolve the incongruence between the literal meaning of the phrase and the context in which it is presented.

7.7.1 Task Effects in Text Reading

Readers often have a specific goal in mind when reading longer texts, such as reading in order to learn new information on a topic or looking for certain type of information. Previous research shows that adult readers adjust their intake of text information to meet the demands of the reading task (e.g., Heller, 1982; Laycock, 1955; Radach et al., 2008; Rothkoph & Billington, 1979; Shebilske & Fisher, 1983; Wotschack & Kliegl, 2013). For example, when expecting difficult rather than easy questions after reading, readers make shorter saccades and fixate more function words (prepositions, articles, etc.) in text (Wotschack & Kliegl, 2013). Moreover, the reading task may influence the local processing of the text information such that readers show different eye movement patterns in different regions of the same text. Rothkoph and Billington (1979) showed that when looking for answers to specific questions, readers made more and longer eye fixations in sentences that contained question-relevant than question-irrelevant information. In other words, the reading task may not only influence the global processing of text information by inducing a more careful reading strategy (e.g., Wotschask & Kliegl, 2013) but also the local processing of text information can be affected by increasing the amount of time spent on specific parts of the text (e.g., Rothkopf & Billington, 1979).

Readers are sensitive to the goal relevance of text information and tend to selectively attend to information that is pertinent to their goal (Kaakinen & Hyönä, 2007, 2008, 2014; Kaakinen, Hyönä, & Keenan, 2003). Kaakinen et al. (2003) asked adult participants to read two expository texts describing various diseases so that they could explain critical facts about one of the diseases described in each text to somebody else. The instructions thus made information related to one disease highly relevant to the readers, whereas other information presented in the text could be considered irrelevant. The results of the study showed that the *relevance effect*, that is, the difference in fixation time between reading the sentences as task-relevant or task-irrelevant, emerged early in processing, as revealed by the progressive fixation time during first-pass reading of the sentence (for eye movement measures used in text comprehension studies, see Hyönä, Lorch, & Rinck, 2004). This means that readers reacted to task relevance as soon as they could identify text information as relevant (or irrelevant). The relevance effect was also observed in later look-back time, indicating that readers also later reread relevant sentences more than irrelevant sentences.

In another study examining relevance effects at the level of individual words (Kaakinen & Hyönä, 2007), it was observed that these effects appear already during the initial processing of words. Gaze durations for words (i.e., the time spent fixating on a word before proceeding to the next word) were longer in task-relevant than in task-irrelevant sentences, even for words at the beginning of the target sentences. Moreover, readers tended to skip over more words within irrelevant than relevant sentences.

Moreover, it was found that task relevance influences the magnitude of the parafoveal preview effect, suggesting that readers' attentional span is zoomed into the currently fixated word when reading relevant information (Kaakinen & Hyönä, 2014). This result suggests that the attentional requirements of encoding relevant information to memory restricts the amount of information the reader can process during one fixation.

However, there are individual differences in how well readers can adjust their eye movements to meet the task demands. In a seminal study, Laycock (1955) investigated individual differences in the ability to "read as fast as possible without missing important points". The results showed that while all readers were able to speed up reading without showing detrimental effects on comprehension, a group of more flexible readers differed from less flexible readers particularly in the rate and number of fixations during speeded reading: they were faster and made fewer fixations. Laycock concluded that some readers are better able to control their eye movements and possibly to increase their attentional span according to the task demands than others.

Previous research suggests that individual differences in working memory capacity (WMC) also play a role in how well readers adjust their text intake to the task at hand (Kaakinen, Hyönä, & Keenan, 2002, 2003). It seems that high-WMC readers are more effective than low-WMC readers in guiding their attention selectively to task-relevant text information and away from task-irrelevant information. In these studies, for high-WMC readers the relevance effect emerged already during the first-pass reading of sentences, while for the low-WMC readers the effect was only observed later in the look-back fixation time (i.e., the time spent rereading the target sentences after first fixating on a subsequent sentence).

Moreover, WMC is related to the ease with which task-relevant information is encoded to memory when readers have relevant prior knowledge about the text contents. Kaakinen et al. (2003) asked participants to read a text describing familiar diseases and another text describing unfamiliar diseases with the instructions to pay special attention to one of the diseases described in the text. In addition to tracking the readers' eye movements during reading, a free recall was collected after reading. The results showed that when reading the text describing diseases readers had ample prior knowledge of, high-WMC readers did not show longer eye fixation time on relevant than on irrelevant text segments, even though they showed better recall of relevant than irrelevant text information. Low-WMC readers, on the other hand, demonstrated a relevance effect in the eye fixation times as well as in the text recall. This pattern of results suggests that high-WMC readers can efficiently encode task-relevant information to memory when the encoded information is familiar to them.

7.7.2 Developmental Trends in Text Reading

Relatively little is known about developmental trends in eye movements during reading of longer passages of text. However, the existing studies show that eye tracking has great potential for revealing developmental trends in reading comprehension skill.

The study by van der Schoot, Vasbinder, Horsley and van Lieshout (2008) suggests that 10–12 year-old children are sensitive to the relevance of text information. In the study, participants were asked to read a story either from a science or gossip journalist's perspective. The stories used in the experiment contained both science-related information as well as descriptions of the social relationships between story characters. When told to pretend to be science journalists children showed longer eye fixation times on science-related than on gossip-related words in the story, whereas when reading from the gossip journalist's perspective, children showed longer eye fixation times on gossip-relevant information. The relevance effect was observed already in the first-pass fixation times on words. Moreover, children's comprehension ability (as measured by an independent test) was positively correlated with increased time spent on task-relevant text information. After controlling for the effects of word decoding skill and vocabulary, the time spent regressing back to the task-relevant words was positively correlated with comprehension ability.

In the study of Kaakinen, Lehtola, and Paattilammi (2015), groups of 2nd (8–9 years), 4th (10–11 years) and 6th graders (12–13 years) read age-appropriate science textbook materials either in order to answer a "why" question presented before reading, or for general comprehension. The results showed that already 2nd graders adjusted their text scanning patterns to the task at hand by showing slower first-pass fixation times when preparing to answer a why-question than when reading for comprehension. In older age groups the effect of the reading task was seen as increased look-back times within the text, i.e., readers did more look-backs when reading to answer why-questions than when reading for comprehension. It is surprising to find that already 2nd graders adjusted their reading behavior in response to task instructions. Perhaps this is made possible by Finnish readers being relatively skillful word decoders already at young age due to Finnish having completely regular letter-sounds correspondence rules (Seymour et al., 2003). It would be interesting to study these effects among children reading less regular orthographies.

What seems to be differentiating between developing good and poor comprehenders is the ability to strategically look back to task-relevant information. Van den Broek, White, Kendeou, and Carlson (2009) used eye-tracking to examine reading strategies of successful and struggling young readers (4th, 7th and 9th graders). They found that the groups differed with respect to where in text look-backs were directed: readers with good comprehension skills reread specific, informative parts of text, whereas struggling readers reread text unselectively. A study by van der Schoot, Reijntjes, and van Lieshout (2012) suggests that young readers (10–12 year-olds) with good comprehension skill are more likely than poor comprehenders to look back to inconsistencies in text. In their study, participants read stories describing a character (e.g., "*Mary is a vegetarian*" or "*Mary is a fast-food addict*"). The story character's actions (e.g., "*Mary ordered cheeseburger and fries*") described in the passage were either consistent or inconsistent with the character description, and the distance between the action and the character description was manipulated. Both good and poor comprehenders were sensitive to inconsistencies when the inconsistent character action directly followed the character description. However, when the description and action were separated by several sentences, only readers with good comprehension skills tended make regressions to the contradictory information. This reflects their more integrated representation of the entire text.

7.8 Outlook

As it has become apparent from the present chapter, the application of the eye-tracking technique to the study of reading has been a success story. We also see a bright future ahead of us. We see five research avenues that are likely to gain momentum in the future. First, developmental eye movement studies are likely to significantly increase our understanding of the acquisition and development of reading skills (see Schröder et al., 2015). Second, a lot can be learned from cross-linguistic studies where reading processes are directly compared across different languages and orthographies (see Liversedge et al., 2015). Third, eye-tracking has not been fully exploited in investigating text comprehension processes. Thus, we expect more eye movement studies to appear on reading longer texts. Fourth, eye-tracking studies on reading texts presented via electronic media (Hyönä, 2010) are highly likely to gain popularity simply for the mere reason that a lot of reading is nowadays done via electronically available texts. Moreover, linear reading of printed books is increasingly complemented by non-linear reading, for example, of hypertexts. This is likely to change some key aspects of reading related to higher-order comprehension processes. Finally, the combination of eye-tracking data with other data reflecting moment-to-moment processing (e.g., EEG, fMRI, motion capture, psychophysiological measures) will further advance our understanding of the reading process. All in all, we feel confident that a lot can still be learned from applying the eye-tracking method to study reading as it takes place in different orthographies and reading environments.

7.9 Suggested Readings

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C. Jr. (2012). *Psychology of reading* (second edition). Hove, UK: Psychology Press.

- *This monograph provides a comprehensive coverage of the cognitive processes involved in reading, including chapters on eye movements in reading.*

Liversedge, S.P., Gilchrist, I.D., & Everling, S. (2011). *The Oxford handbook of eye movements*. Oxford, UK: Oxford University Press.

- *This edited volume contains two sections relevant to eye movements in reading (Part 6: Eye movement control during reading; Part 7: Language processing and eye movements). Relevant chapters of this volume are referred to in the text.*

Reichle, E. D. (in press). *Computational models of reading: A handbook*. Oxford, UK: Oxford University Press.

- *This monograph provides a comprehensive coverage of all major theories of reading.*

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*, 372–422.

- *This review article provides a comprehensive summary of the eye movement research done in reading up to late 1990s.*

Rayner, K., Ardoin, S. P., & Binder, K. S. (2013). Children's eye movements in reading: A commentary. *School Psychology Review, 42*, 223–233.

– *This is an introduction to a special issue on children's eye movements during reading.*

Schröder, S., Hyönä, J., & Liversedge, S.P. (2015). Developmental eye-tracking research in reading: Introduction to the Special Issue. *Journal of Cognitive Psychology*, *27*, 500–510.

– *This Special Issue contains original research articles dealing with developmental aspects of eye movements in reading. The emphasis is on young, developing readers.*

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science, 23,* 201–207.

– *This article reviews the literature on developmental changes in eye movements during reading.*

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford, UK: Oxford University Press.

– *This chapter summarizes eye movement studies conducted on syntactic parsing during reading.*

Hyönä, J., Lorch, R.F.Jr., Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam: Elsevier.

– *The chapter introduces eye movement measures to study the processing of long texts presented as multiple-line screens.*

7.10 Questions Students Should Be Able to Answer

(1) Why is eye-tracking a useful tool to study reading? Mention at least three reasons.
(2) Is speedreading feasible based on what you have read in this chapter? Why do you think that way?
(3) How do eye movement patterns differ during reading of single sentences and longer text paragraphs? Why?
(4) What have eye movement studies revealed about the development of reading skill?

Bibliography

Ahlen, J. (1974). *Spatial and temporal aspects of visual information processing during reading*. Unpublished doctoral dissertation, University of Illinois Medical Center.

Apel, J. K., Henderson, J. M., & Ferreira, F. (2012). Targeting regressions: Do readers pay attention to the left? *Psychonomic Bulletin & Review, 19,* 1108–1113.

Ashby, J., Treiman, R., Kessler, B., & Rayner, K. (2006). Vowel processing during silent reading: Evidence from eye movements. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 32,* 416–424.

Baccino, T., & Manunta, Y. (2005). Eye-fixation-related potentials: Insight into parafoveal processing. *Journal of Psychophysiology, 19,* 204–215.

Balota, D. A., Pollatsek, A., & Rayner, K. (1985). The interaction of contextual constraints and parafoveal visual information in reading. *Cognitive Psychology, 17,* 364–390.

Binder, K. S., Pollatsek, A., & Rayner, K. (1999). Extraction of information to the left of the fixated word in reading. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 1162–1172.

Blythe, H. I. (2014). Developmental changes in eye movements and visual information encoding associated with learning to read. *Current Directions in Psychological Science, 23,* 201–207.

Blythe, H. I., Häikiö, T., Bertram, R., Liversedge, S. P., & Hyönä, J. (2010). Reading disappearing text: Why do children refixate words? *Vision Research, 51,* 84–92.

Blythe, H. I., & Joseph, H. S. S. I. (2011). Children's eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 643–662). Oxford, UK: Oxford University Press.

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., & Rayner, K. (2009). Visual information capture during fixations in reading for children and adults. *Vision Research, 49,* 1583–1591.

Briihl, D., & Inhoff, A. W. (1995). Integrating information across fixations during reading: The use of orthographic bodies and exterior letters. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21,* 55–67.

Brown, M. R. G., Vilis, T., & Everling, S. (2008). Isolation of saccade inhibition processes: Rapid event-related fMRI of saccades and nogo trials. *NeuroImage, 39,* 793–804.

Buswell, G. T. (1922). *Fundamental reading habits: A study of their development*. Chicago: Chicago University Press.

Calvo, M. G., & Meseguer, E. (2002). Eye movements and processing stages in reading: Relative contribution of visual, lexical and contextual factors. *Spanish Journal of Psychology, 5,* 66–77.

Cauchard, F., Eyrolle, H., Cellier, J.-M., & Hyönä, J. (2010a). Vertical perceptual span and the processing of visual signals in reading. *International Journal of Psychology, 45,* 40–47.

Cauchard, F., Eyrolle, H., Cellier, J. M., & Hyönä, J. (2010b). Visual signals vertically extend the perceptual span in searching a text: A gaze-contingent window study. *Discourse Processes, 47,* 617–640.

Chace, K. H., Rayner, K., & Well, A. D. (2005). Eye movements and phonological parafoveal preview: Effects of reading skill. *Canadian Journal of Experimental Psychology, 59,* 209–217.

Clifton, C., & Staub, A. (2011). Syntactic influences on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 895–909). Oxford, UK: Oxford University Press.

Clifton, C., Jr., Staub, A., & Rayner, K. (2007). Eye movements in reading words and sentences. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 341–371). Amsterdam: Elsevier.

Drieghe, D. (2011). Parafoveal-on-foveal effects on eye movements during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 839–855). Oxford, UK: Oxford University Press.

Engbert, R., & Kliegl, R. (2011). Parallel graded attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 787–800). Oxford, UK: Oxford University Press.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology, 14,* 178–210.

Häikiö, T., Bertram, R., & Hyönä, J. (2010). Development of parafoveal processing within and across words in reading: Evidence from the boundary paradigm. *Quarterly Journal of Experimental Psychology, 63,* 1982–1998.

Häikiö, T., Bertram, R., Hyönä, J., & Niemi, P. (2009). Development of the letter identity span in reading: Evidence from the eye movement moving window paradigm. *Journal of Experimental Child Psychology, 102,* 167–181.

Heller, D. (1982). Eye movements in reading. In R. Groner & P. Fraisse (Eds.), *Cognition and eye movements* (pp. 139–154). Amsterdam, The Netherlands: North Holland.

Henderson, J. M., Dixon, P., Petersen, A., Twilley, L. C., & Ferreira, F. (1995). Evidence for the use of phonological representations during transsaccadic word recognition. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 82–97.

Hering, E. (1879a). Über Muskelgeräusche des Auges. Sitzberichte der kaiserlichen Akademie der Wissenschaften in Wien. *Mathematischnaturwissenschaftliche Klasse, 79*, 137–154.

Hering, E. (1879b). Der Raumsinn und die Bewegungen des Auges. In L. Hermann (Ed.), *Handbuch der Physiologie* (Vol. 3, pp. 343–601)., Part 1 Physiologie des Gesichtssinnes Leipzig: Vogel.

Hohenstein, S., & Kliegl, R. (2014). Semantic preview benefit during reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 40,* 166–190.

Hohenstein, S., Laubrock, J., & Kliegl, R. (2010). Semantic preview benefit in eye movements during reading: A parafoveal fast-priming study. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 36,* 1150–1170.

Huestegge, L., Radach, R., Corbic, D., & Huestegge, S. M. (2009). Oculomotor and linguistic determinants of reading development: A longitudinal study. *Vision Research, 49,* 2948–2959.

Huey, E. B. (1908). *The psychology and pedagogy of reading*. New York: Macmillan.

Hyönä, J. (1993). Effects of thematic and lexical priming on readers' eye movements. *Scandinavian Journal of Psychology, 34,* 293–304.

Hyönä, J. (1995a). Do irregular letter combinations attract readers' attention? evidence from fixation locations in words. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 68–81.

Hyönä, J. (1995b). An eye movement analysis of topic-shift effect during repeated reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21,* 1365–1373.

Hyönä, J. (2011). Foveal and parafoveal processing during reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 819–838). Oxford, UK: Oxford University Press.

Hyönä, J., & Bertram, R. (2011). Optimal viewing position effects in reading Finnish. *Vision Research, 51,* 1279–1287.

Hyönä, J., Bertram, R., & Pollatsek, A. (2004). Are long compound words identified serially via their constituents? evidence from an eye-movement-contingent display change study. *Memory & Cognition, 32,* 523–532.

Hyönä, J., Lorch, R. F., Jr., & Kaakinen, J. K. (2002). Individual differences in reading to summarize expository text: Evidence from eye fixation patterns. *Journal of Educational Psychology, 94,* 44–55.

Hyönä, J., Lorch, R. F., Jr., & Rinck, M. (2003). Eye movement measures to study global text processing. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 313–334). Amsterdam: Elsevier Science.

Hyönä, J., & Nurminen, A. M. (2006). Do adult readers know how they read? Evidence from eye movement patterns and verbal reports. *British Journal of Psychology, 97,* 31–50.

Hyönä, J., & Olson, R. K. (1995). Eye fixation patterns among dyslexic and normal readers: Effects of word length and word frequency. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 21,* 1430–1440.

Inhoff, A. W., Radach, R., Eiter, B. M., & Juhasz, B. (2003). Distinct subsystems for the parafoveal processing of spatial and linguistic information during eye fixations in reading. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 56A,* 803–827.

Inhoff, A. W., & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception and Psychophysics, 40,* 431–439.

Inhoff, A. W., & Topolski, R. (1994). Use of phonological codes during eye fixations in reading and in on-line and delayed naming tasks. *Journal of Memory and Language, 33,* 689–713.

Javal, L. É. (1878a). Essai sur la physiologie de la lecture. *Annales d'Oculistique, 79,* 97–117.

Javal, L. É. (1878b). Essai sur la physiologie de la lecture. *Annales d'Oculistique, 80,* 135–147.

Javal, L. É. (1878c). Essai sur la physiologie de la lecture. *Annales d'Oculistique, 80,* 240–274.

Jordan, T. R., Almabruk, A. A. A., Gadalla, E. A., McGowan, V. A., White, S. J., Abedipour, L., et al. (2014). Reading direction and the central perceptual span: Evidence from Arabic and English. *Psychonomic Bulletin & Review, 21,* 505–511.

Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research, 49,* 2078–2086.

Juhasz, B. J., Pollatsek, A., Hyönä, J., Drieghe, D., & Rayner, K. (2009). Parafoveal processing within and between words. *The Quarterly Journal of Experimental Psychology, 62*(7), 1356–1376.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87,* 329–354.

Kaakinen, J. K., & Hyönä, J. (2007). Perspective effects in repeated reading: An eye movement study. *Memory & Cognition, 35,* 1323–1336.

Kaakinen, J. K., & Hyönä, J. (2008). Perspective-driven text comprehension. *Applied Cognitive Psychology, 22,* 319–334.

Kaakinen, J. K., & Hyönä, J. (2014). Task relevance induces momentary changes in the functional visual field during reading. *Psychological Science, 25,* 626–632.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2002). Perspective effects on on-line text processing. *Discourse Processes, 33,* 159–173.

Kaakinen, J. K., Hyönä, J., & Keenan, J. M. (2003). How prior knowledge, working memory capacity, and relevance of information affect eye-fixations in expository text. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 29,* 447–457.

Kaakinen, J. K., Lehtola, A., & Paattilammi, S. (2015). The influence of a reading task on children's eye movements during reading. *Journal of Cognitive Psychology, 27,* 640–656.

Kaakinen, J. K., Olkoniemi, H., Kinnari, T., & Hyönä, J. (2014). Processing of written irony: An eye movement study. *Discourse Processes, 51,* 287–311.

Kennedy, A. (2000). Attention allocation in reading: Sequential or parallel? In A. Kennedy, R. Radach, D. Heller, & J. Pynte (Eds.), *Reading as a perceptual process* (pp. 193–220). Oxford, UK: Elsevier.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Laycock, F. (1955). Significant characteristics of college students with varying flexibility in reading rate: I. Eye movements in reading prose. *Journal of Experimental Education, 23,* 311–319.

Lee, H. W., Rayner, K., & Pollatsek, A. (2001). The relative contribution of consonants and vowels to word identification during silent reading. *Journal of Memory and Language, 44,* 189–205.

Lee, H. W., Rayner, K., & Pollatsek, A. (2002). The processing of consonants and vowels in reading: Evidence from the fast-priming paradigm. *Psychonomic Bulletin & Review, 9,* 766–772.

Liversedge, S. P., Gilchrist, I. D., & Everling, S. (2011). *The Oxford handbook of eye movements*. Oxford, UK: Oxford University Press.

Liversedge, S. P., Hyönä, J., & Rayner, K. (2013). Eye movements during Chinese reading. *Journal of Research in Reading, 36,* S1–S3.

Liversedge, S. P., Rayner, K., White, S. J., Vergilino-Perez, D., Findlay, J. M., & Kentridge, R. W. (2004). Eye movements when reading disappearing text: Is there a gap effect in reading? *Vision Research, 44*(1013), 1024.

Mancheva, L., Reichle, E. D., Lemaire, B., Valdois, S., Ecalle, J., & Guérin-Dugué, A. (2015). An analysis of reading skill development using E-Z Reader. *Journal of Cognitive Psychology, 27,* 657–676.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations in words. *Vision Research, 28,* 1107–1118.

McConkie, G. W., & Loschky, L. C. (2002). Perception onset time during fixations in free viewing. *Behavioral Research Methods, Instruments, and Computers, 34,* 481–490.

McConkie, G. W., & Rayner, K. (1975). The span of effective stimulus during a fixation in reading. *Perception and Psychophysics, 17,* 578–586.

McConkie, G. W., & Rayner, K. (1976). Asymmetry in the perceptual span in reading. *Bulletin of the Psychonomic Society, 8,* 365–368.

McConkie, G. W., Zola, D., Grimes, J., Kerr, P. W., Bryant, N. R., & Wolff, P. M. (1991). Children's eye movements during reading. In J. E. Stein (Ed.), *Vision and visual dyslexia* (pp. 251–262). London: Macmillan Press.

McGowan, V. A. (2015). The processing of spatial information during reading: Processing of previously read text and effects of adult age. Doctoral thesis. UK: University of Leicester.

Miellet, S., & Sparrow, L. (2004). Phonological codes are assembled before word fixation: Evidence from boundary paradigm in sentence reading. *Brain and Language, 90,* 299–310.

Mitchell, D. C., Shen, X., Green, M. J., & Hodgson, T. L. (2008). Accounting for regressive eye-movements in models of sentence processing: A reappraisal of the selective reanalysis hypothesis. *Journal of Memory and Language, 59,* 266–293.

Nuthmann, A., Engbert, R., & Kliegl, R. (2005). Mislocated fixations during reading and the inverted optimal viewing position effect. *Vision Research, 45,* 2201–2217.

O'Regan, J. K., Lévy-Schoen, A., Pynte, J., & Brugaillere, B. (1984). Convenient fixation location within isolated words of different lengths and structure. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 250–257.

Pollatsek, A., Bolozky, S., Well, A. D., & Rayner, K. (1981). Asymmetries in the perceptual span for Israeli readers. *Brain and Language, 14,* 174–180.

Pollatsek, A., Lesch, M., Morris, R. K., & Rayner, K. (1992). Phonological codes are used in integrating information across saccades in word identification and reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 148–162.

Pollatsek, A., Reichle, E. D., & Rayner, K. (2006). Tests of the E-Z Reader model: Exploring the interface between cognition and eye-movement control. *Cognitive Psychology*, *52*, 1–56.

Pollatsek, A., Raney, G. E., Lagasse, L., & Rayner, K. (1993). The use of information below fixation in reading and in visual search. *Canadian Journal of Experimental Psychology, 47,* 179–200.

Radach, R., Huestegge, L., & Reilly, R. (2008). The role of global top-down factors in local eye-movement control in reading. *Psychological Research, 72,* 675–688.

Radach, R., Inhoff, A., & Heller, D. (2004). Orthographic regularity gradually modulates saccade amplitudes in reading. *European Journal of Cognitive Psychology, 16,* 27–51.

Rayner, K. (1975). The perceptual span and peripheral cues in reading. *Cognitive Psychology, 7,* 65–81.

Rayner, K. (1978). Eye movements in reading and information processing. *Psychological Bulletin, 85,* 618–660.

Rayner, K. (1979). Eye guidance in reading: Fixation locations within words. *Perception, 8,* 21–30.

Rayner, K. (1986). Eye movements and the perceptual span in beginning and skilled readers. *Journal of Experimental Child Psychology, 41,* 211–236.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124,* 372–422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Rayner, K., Ardoin, S. P., & Binder, K. S. (2013). Children's eye movements in reading: A commentary. *School Psychology Review, 42,* 223–233.

Rayner, K., & Duffy, S. A. (1986). Lexical complexity and fixation times in reading: Effects of word frequency, verb complexity, and lexical ambiguity. *Memory & Cognition, 14,* 191–201.

Rayner, K., Inhoff, A. W., Morrison, R., Slowiaczek, M. L., & Bertera, J. H. (1981). Masking of foveal and parafoveal vision during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 167–179.

Rayner, K., Kambe, G., & Duffy, S. A. (2000). The effect of clause wrap-up on eye movements during reading. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 53A,* 1061–1080.

Rayner, K., Pollatsek, A., Ashby, J., & Clifton, C., Jr. (2012). *Psychology of reading* (2nd ed.). Hove, UK: Psychology Press.

Rayner, K., Pollatsek, A., & Binder, K. S. (1998). Phonological codes and eye movements in reading. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 24,* 476–497.

Rayner, K., & Well, A. D. (1996). Effects of contextual constraint on eye movements in reading: A further examination. *Psychonomic Bulletin & Review, 3,* 504–509.

Rayner, K., Well, A. D., & Pollatsek, A. (1980). Asymmetry of the effective visual field in reading. *Perception and Psychophysics, 27,* 537–544.

Rayner, K., Well, A. D., Pollatsek, A., & Bertera, J. H. (1982). The availability of useful information to the right of fixation in reading. *Perception and Psychophysics, 31,* 537–550.

Rayner, K., White, S. J., Kambe, G., Miller, B., & Liversedge, S. P. (2003b). On the processing of meaning from parafoveal vision during eye fixations in reading. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 213–234). Amsterdam: Elsevier.

Reder, S. M. (1973). On-line monitoring of eye position signals in contingent and noncontingent paradigms. *Behavior Research Methods and Instrumentation, 5,* 218–228.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*, 125–157.

Reichle, E. D., Rayner, K., & Pollatsek, A. (2003). The E-Z Reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences*, *26*, 445–476.

Reichle, E. D. (2011). Serial-attention models of reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 767–786). Oxford, UK: Oxford University Press.

Reichle, E. D. (in press). *Computational models of reading: A handbook*. Oxford, UK: Oxford University Press.

Reichle, E. D., Liversedge, S. P., Drieghe, D., Blythe, H. I., Joseph, H. S. S. L., White, S. J., et al. (2013). Using E-Z Reader to examine the concurrent development of eye-movement control and reading skill. *Developmental Review, 33,* 110–149.

Rothkopf, E. Z., & Billington, M. J. (1979). Goal-guided learning from text: Inferring a descriptive processing model from inspection times and eye movements. *Journal of Educational Psychology, 71,* 310–327.

Schotter, E. R. (2013). Synonyms provide semantic preview benefit in English. *Journal of Memory and Language, 69,* 619–633.

Schotter, E. R., Angele, B., & Rayner, K. (2012). Parafoveal processing in reading. *Attention, Perception, & Psychophysiology, 74,* 5–35.

Schotter, E. R., Tran, R., & Rayner, K. (2014). Don't believe what you read (only once): Comprehension is supported by regressions during reading. *Psychological Science, 25,* 1218–1226.

Schröder, S., Hyönä, J., & Liversedge, S. P. (2015). Developmental eye-tracking research in reading: Introduction to the special issue. *Journal of Cognitive Psychology*, *27*, 500–510.

Sereno, S. C., & Rayner, K. (1992). Fast priming during eye fixations in reading. *Journal of Experimental Psychology: Human Perception and Performance, 18,* 173–184.

Sereno, S. C., & Rayner, K. (2000). Spelling-sound regularity effects on eye fixations in reading. *Perception and Psychophysics, 62,* 402–409.

Seymour, P. H. K., Aro, M., & Erskine, J. M. (2003). Foundation literacy acquisition in European orthographies. *British Journal of Psychology, 94,* 143–174.

Shebilske, W. L., & Fisher, D. F. (1983). Eye movements and context effects during reading of extended discourse. In K. Rayner (Ed.), *Eye movements in reading: Perceptual and language processes* (pp. 157–179). New York, NY: Academic Press.

Sparrow, L., & Miellet, S. (2002). Activation of phonological codes during reading: Evidence from errors detection and eye movements. *Brain and Language, 81,* 509–516.

Sperlich, A., Schad, D. J., & Laubrock, J. (2015). When preview information starts to matter: Development of the perceptual span in German beginning readers. *Journal of Cognitive Psychology,* in press.

Staub, A., & Rayner, K. (2007). Eye movements and on-line comprehension processes. In G. Gaskell (Ed.), *The Oxford handbook of psycholinguistics* (pp. 327–342). Oxford, UK: Oxford University Press.

Taylor, S. E. (1965). Eye movements while reading: Facts and fallacies. *American Educational Research Journal, 2,* 187–202.

Tiffin-Richards, S. P., & Schroeder, S. (2015). Children's and adults' parafoveal processing in German: Phonological and orthographic effects. *Journal of Cognitive Psychology*, in press.

Tinker, M. A. (1958). Recent studies of eye movements in reading. *Psychological Bulletin, 55,* 215–231.

Vainio, S., Hyönä, J., & Pajunen, A. (2009). Lexical predictability exerts robust effects on fixation duration, but not on initial landing position during reading. *Experimental Psychology, 56,* 66–74.

van den Broek, P., White, M. J., Kendeou, P., & Carlson, S. (2009). Reading between the lines. Developmental and individual differences in cognitive processes in reading comprehension. In K. Wagner, C. Schatschneider, & C. Plythian-Sence (Eds.), *Beyond decoding: The behavioral and biological foundations of reading comprehension* (pp. 107–123). New York, NY: Guilford Press.

van der Schoot, M., Reijnjes, A., & van Lieshout, E. C. D. M. (2012). How do children deal with inconsistencies in text? An eye fixation and self-paced reading study in good and poor comprehenders. *Reading and Writing, 25,* 1665–1690.

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N + 1 and N + 2 during reading: A critical review and Bayesian meta-analysis. *Psychonomic Bulletin & Review, 24,* 666–689.

Vitu, F., Lancelin, D., & Marrier d'Unienville, V. (2007). A perceptual-economy account for the inverted-optimal viewing position effect. *Journal of Experimental Psychology: Human Perception and Performance, 33,* 1220–1249.

Vitu, F., McConkie, G. W., Kerr, P., & O'Regan, J. K. (2001). Fixation location effects on fixation durations during reading: An inverted optimal viewing position effect. *Vision Research, 41,* 3513–3533.

Vitu, F., O'Regan, J. K., & Mittau, M. (1990). Optimal landing position in reading isolated words and continuous text. *Perception and Psychophysics, 47,* 583–600.

Wade, N. J., & Tatler, B. W. (2009). Did Javal measure eye movements during reading? *Journal of Eye Movement Research*, *2(5):5*, 1–7.

White, S. J., Johnson, R. L., Liversedge, S. P., & Rayner, K. (2008). Eye movements when reading transposed text: The importance of word beginning letters. *Journal of Experimental Psychology: Human Perception and Performance, 34,* 1261–1276.

White, S. J., & Liversedge, S. P. (2004). Orthographic familiarity influences initial eye fixation positions in reading. *European Journal of Cognitive Psychology, 16,* 52–78.

White, S. J., & Liversedge, S. (2006). Foveal processing difficulty does not modulate non-foveal orthographic influences on fixation positions. *Vision Research, 46,* 426–437.

White, S. J., Rayner, K., & Liversedge, S. P. (2005). The influence of parafoveal word length and contextual constraints on fixation durations and word skipping in reading. *Psychonomic Bulletin & Review, 12,* 466–471.

White, S. J., Warren, T., & Reichle, E. D. (2011). Parafoveal preview during reading: Effects of sentence position. *Journal of Experimental Psychology: Human Perception and Performance, 37,* 1221–1238.

Wiley, J., & Rayner, K. (2000). Effects of titles on the processing of text and lexically ambiguous words: evidence from eye movements. *Memory & Cognition, 28,* 1011–1021.

Wotschack, C., & Kliegl, R. (2013). Reading strategy modulates parafoveal-on-foveal effects in sentence reading. *Quarterly Journal of Experimental Psychology, 66,* 548–562.

Zhang, C., Liversedge, S. P., Bai, X., & Yan, G. (2011). Eye movements during Chinese reading. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 961–978). Oxford, UK: Oxford University Press.

Part II
Methods

Chapter 8
Eye Tracking Methodology

S. B. Hutton

Contents

Abstract This chapter has two main aims. The first is to introduce readers to the range of eye tracking technologies currently available, and describe the basic principles on which they operate. The second is to provide readers with an understanding of the main determinants of eye tracking data quality and the ways in which this can be quantified. A greater understanding of how eye tracking technology works, and the key determinants of data quality has two important benefits. Firstly, it will improve the likelihood of researchers being able to maximise the quality of the eye tracking data they generate themselves, using eye tracking technology that is appropriate for their research goals. Secondly it will increase their ability to critically evaluate eye tracking research produced by other researchers. Holmqvist et al. (2011) identify several distinct categories of eye tracker users, including usability and media consultants as well as those interested in human-computer interaction and gaze controlled interfaces. This chapter assumes that the majority of readers are academic researchers, probably working in the fields of psychology or cognitive neuroscience and related

---

S. B. Hutton (✉)
Sussex University, Brighton, UK
e-mail: s.hutton@sussex.ac.uk

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_8

disciplines, and as such are most likely interested in using eye tracking technology to establish point of gaze and oculomotor dynamics—and also be concerned with issues of accuracy, precision, sampling rate and timing. Section 8.2 of this chapter concerns eye tracking technology, and starts with a brief historical overview of early eye tracking techniques, followed by a description of some of the less common technologies that can still be found in research published today—albeit often in relatively niche areas. The vast majority of commercial eye trackers that are currently available are video based—and as such this approach is covered in most detail. Video-based eye tracking methodologies can be divided (broadly) into two categories: stationary, screen-based systems and mobile head-mounted systems. Clearly these two types of equipment are generally used in very different research scenarios—and differences in the technology concerned make comparisons across these categories difficult if not meaningless. As such they are treated separately in the chapter. Section 8.2 may provide some useful information for readers who are considering purchasing an eye tracker—but it is important to note that this is not its primary purpose—nor is it meant as an exhaustive description of the pros and cons of all currently available eye tracking techniques and commercial models for all possible research scenarios. Not only would such an undertaking become rapidly outdated, it would also involve comparing apples with oranges. Indeed, care has been taken to avoid mentioning specific manufacturers or models where possible. Hopefully any readers interested in purchasing an eye tracker will, after reading this chapter, be equipped with sufficient knowledge to make informed decisions as to which type of eye tracker would be most appropriate given their research goals—and be able to critically evaluate performance claims made by manufacturers and ask the right questions of sales people. Those already in possession of an eye tracker may gain a better understanding of how it works, and its capabilities and limitations, and be more confident that they are using it to its full potential. Section 8.3 considers eye tracking software—not only is software a central component of most commercially available eye trackers, and an important determinant of data quality, it is also one of the main factors determining the ease with which the technology can be used, and the range of uses to which it can be put. Again, this entire section is intentionally generic, and is not intended as an exhaustive evaluation of all currently available software. Section 8.4 addresses the other key topic of the chapter—data quality. It starts with an attempt to define important key terms such as "accuracy" and "precision". The section then considers how such concepts might be quantified. The final part of this section is intended to offer practical advice for maximising data quality—including some basic information on the importance of setting up participants and getting a good calibration. The precise setup and calibration details will differ depending on the type of eye tracker used—so the advice contained in this section again intentionally generic and aims to outline basic principles and good practices that apply to all or most eye tracking scenarios.

8.1 Introduction and Learning Objectives

The number of papers published that use eye tracking technology is increasing every year. Researchers who are interested in using gaze information to test hypotheses no longer have to build their own equipment from scratch, and have a wide range of techniques and commercial systems available to them. Many eye trackers are far simpler to operate and less uncomfortable for participants than the systems available even one or two decades ago, and some are capable of tracking eye movements with exceptionally high levels of accuracy and precision. Finally, advances in software mean that the analysis of oculomotor data is a far more efficient and rapid process than before. Current video-based eye tracking systems vary greatly in cost, and differ enormously with respect to their capabilities. Researchers need to make sure that any system they use (and the software that accompanies it) is capable of delivering data that will allow them to draw meaningful conclusions given their research goals.

The chapter has two key learning objectives: (1) To understand how various eye tracking techniques work (in particular the principles underlying video-based eye tracking). (2) To appreciate the various determinants of data quality, and understand how data quality can be measured and improved.

8.2 Eye Tracking Techniques

8.2.1 Historical Approaches

Wade and Tatler (2005) provide a fascinating history of the origins of eye tracking methodology in their book "The Moving Tablet of the Eye". Anyone interested in finding out more about early eye tracking pioneers and the techniques they developed should read this book. Techniques for tracking eye movements were first developed in the late 19th century. Pioneers such as Huey (1898) and Delbarre (1898) attached a crude "contact lens" to the anaesthetised eye of their participants. In Huey's apparatus, a lightweight aluminium rod attached to the contact lens connected to a pivot wheel, to which a second rod was attached. The movements of this second rod marked out the eye's movements on a rotating drum containing "smoke paper". Ingeniously, the second rod was electrified, creating sparks from its tip to the drum, allowing precise measurement of the eyes' speed. In Huey's words:

> *"In order to measure the speed [of the eye] an electric current from an induction coil was passed through the pointer to the drum. This current was interrupted at very regular short intervals by the vibrations of an electrically driven tuning-fork, the snap of the spark from the pointer's tip thus displacing a dot of soot on the paper record at each interruption. As the pointer flitted over the drum during the reading, a tracing was thus produced like that shown in Fig.* 8.1*".*

Fixations appear as "blobs" on the resulting trace—where many sparks have been emitted at roughly the same location, and saccades as a sequence of more widely

Fig. 8.1 The equipment used by Huey (left) and resulting recording (right)

spaced dots. By counting the dots between the blobs it is possible to measure saccade duration exceptionally accurately (a spark, and hence dot on the paper, was made every 10 ms by the "electronic tuning fork").

In the 20th century several alternative eye tracking techniques were developed—some involving more sophisticated contact lens arrangements in which rods were replaced with tiny mirrors, such as were used by Yarbus (1967). The reflection of a light source on these mirrors could be recorded on film and used to recover point of gaze. These approaches would still no doubt have been very uncomfortable for the participant—and the comparatively heavy contact lenses will certainly have altered the eye's dynamics to some extent.

A less invasive technique was pioneered by Dodge (see Dodge & Cline, 1901; Diefendorf & Dodge, 1908). His "Photochronograph" used a slowly falling photographic plate to record the reflection of a vertical strip of light on the cornea. In many ways this device was the precursor of modern video based oculography—the concept of measuring a reflection on the eye is employed by most of the video based eye trackers which currently dominate the commercial market.

Several other technical approaches were developed, mainly during the latter half of the 20th century. As many of these techniques (such as EOG) are still in use today, albeit with updated technology, they are described briefly in the following section.

8.2.2 Current Techniques

Currently available commercial eye tracking systems make use of a wide range of technologies. As mentioned previously, by far the most common are video-based eye trackers and consequently the majority of this chapter will focus on this technique. There are, however, several alternative approaches still actively used in research. These are outlined in the sections below.

8.2.2.1 Electro-Oculography (EOG)

As illustrated in Fig. 8.2, the eye acts as a dipole—a rod magnet—being slightly more negatively charged at the retina compared to the cornea (Mowrer, Ruch, & Miller, 1936). The magnitude of this corneoretinal potential is in the range 0.4–1.0 mV, and is generally attributed to the greater metabolic activity in the retina. Electro-oculograms are recorded via pairs of electrodes which are typically placed above and below the eye or to its left and right. When the eye rotates (horizontally or vertically) a potential difference occurs between the electrode it rotates towards and the one it rotates away from. This potential difference is measured and can be plotted and interpreted as gaze position on the horizontal or vertical axis (Fig. 8.3).

EOG has the advantage of being relatively cheap, and the technology is well established and readily available. The temporal resolution is high (the analogue signal can be sampled as often as desired). With some minor adaptations such as special electrodes the technique can also be used in neuroimaging environments such as MEG/MRI. Finally, EOG is capable of measuring eye movements whilst the eye is closed—essential if studying the eye movements that occur during sleep.

The EOG signal is measured with respect to the head—meaning that the technique is appropriate in head-fixed scenarios or for researchers for whom eye rotation rather than point of gaze is of primary importance. It remains a reasonably common technique in hospital settings, and is sometimes used by EEG researchers, who place

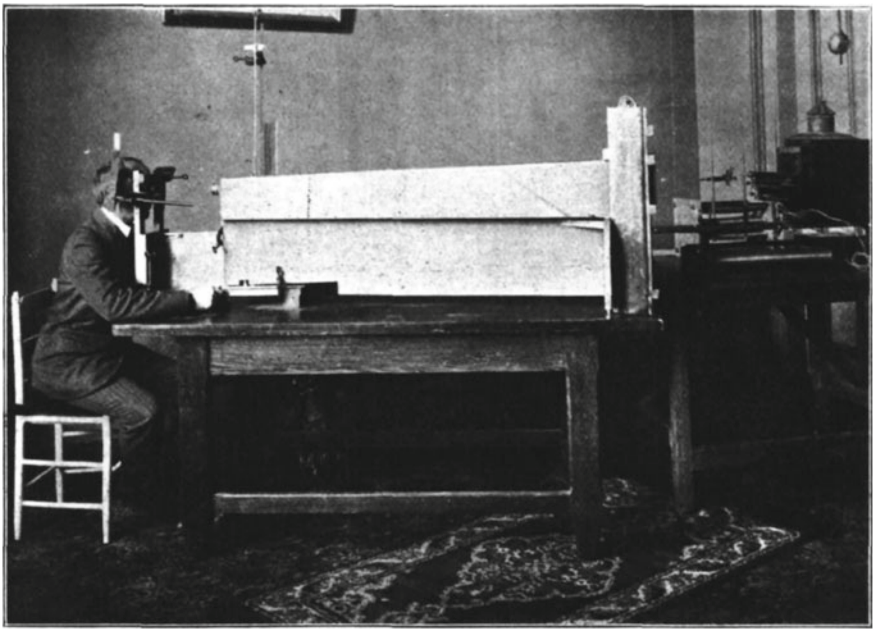


Fig. 8.2 Diefendorf and Dodge's Photochronograph

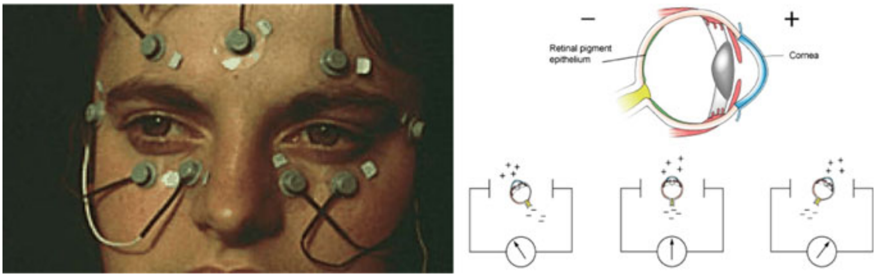


Fig. 8.3 Example of EOG electrode placement (left) and illustration of basic principles (right)

additional electrodes close to the eyes in order to pick up blinks or saccades which may confound their EEG data.

The main disadvantage of EOG is spatial accuracy because EOG is very prone to drift artefacts over time, typically due to impedance changes at one or more of the electrodes. These changes often result from sweating and mean that recalibration is frequently required. In addition, contraction of facial/neck muscles can also influence the signal, adding noise to the data. Spatial accuracy in EOG may also be compromised by the fact that the polarity of the eye is to some extent dependent on luminance changes, with increases in luminance increasing the level of signal. Finally, the signal is only truly linear for relatively small (less than about 10°) eye movements.

8.2.2.2 Limbus Reflection

Until video-based eye tracking techniques became dominant, infrared limbus trackers were fairly common. The basic principle is straightforward, and the components are relatively cheap. In essence a source of (typically infrared) light is shone at the eye, and a sensor (or array of sensors) measures the amount reflected back (see Fig. 8.4). The technique takes advantage of the limbus—the border of the coloured iris and predominantly white sclera of the eye. In essence the white sclera reflects more light than the coloured iris, and as the eye rotates more or less sclera is exposed to the IR emitters, and more or less IR is reflected onto the sensors.

Limbus trackers have a number of positive features. They are comparatively cheap and provide a direct measure of eye rotation within the head, which is important for researchers looking at basic oculomotor dynamics. As with EOG they are essentially analogue devices, and the analogue signal can be sampled at high frequency, allowing saccade dynamics to be observed in detail. However, as with DPI trackers, they require effective head restraint if point of gaze is required. If gaze shifts from a central target to a target 8° to the right, but the participants head simultaneously rotates 1° in the direction of the target, then Limbus trackers will only report a movement of 7°. One limbus based eye tracker has an ingenious solution to this issue. The system is head-mounted and the saccade targets are presented by what

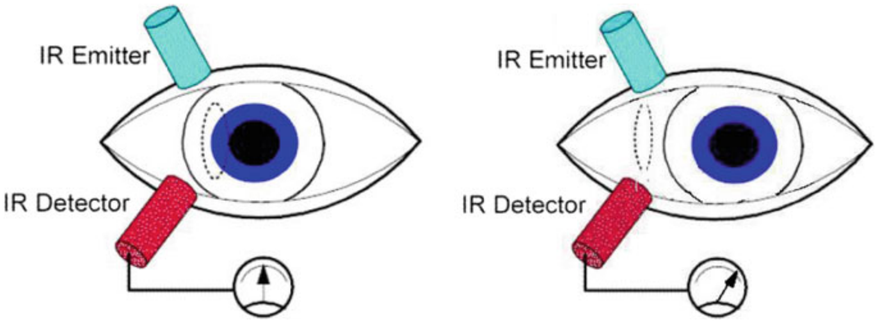


Fig. 8.4 The basic principle of limbus eye trackers. More IR is reflected in the image on the right because the limbus has moved away from the sensor as the eye rotates to its left

amounts to a laser pointer attached to the device. Participants can stand in front of any reasonably uniform surface such as an interior wall, and the targets are projected onto the wall. As such any head movement will result in the target also shifting, so the eye will always have to rotate the same amount as the target displacement in order for the target to be fixated.

Another limitation of limbus tracking devices is that due to the relationship between the emitters and sensors it is generally simplest to record either horizontal or vertical eye movements at any one time (although recording vertical eye movements in one eye and horizontal eye movements in the other is possible).

8.2.2.3 Dual Purkinje Trackers

Dual Purkinje Image (DPI) eye trackers provide high spatial accuracy. They are electro-mechanical devices and take advantage of the fact that a point of light shone at the eye results in four separate reflections—from the front and back of the cornea (the first and second Purkinje images) and from the front and back of the lens (the third and fourth Purkinje images). A detailed description of the mechanics of DPI systems is far beyond the scope of this chapter. The key principle to grasp is that rotational movements of the eye result in significant displacements of the locations of the first and fourth Purkinje images with respect to each other—whereas translational movements (such as those that arise from small head movements) result in no such displacement. The displacements between the first and fourth image occur during eye rotation because the two images are at different distances from the centre of rotation of the eye.

The two Purkinje images are separated by the optical components of the DPI eye tracker, and projected to photoreceptors via a complex series of lenses and mirrors, some of which are attached to servo-motors so that their angle can be rapidly adjusted. The signal from the photo-receptors is used to drive the motors so that the angle of the mirrors keeps the position of each Purkinje image “fixed” on its sensor. The

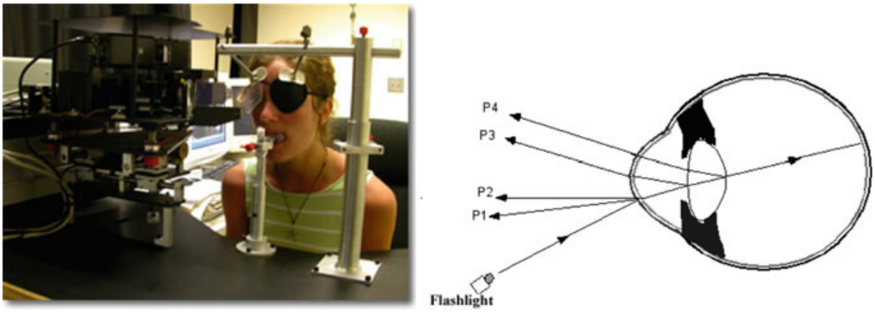


Fig. 8.5 A DPI eye tracker (left). Illustration of the 4 Purkinje Images (right)

"output" of the system is the difference in electrical signal sent to the various servo-motors in order to keep the two Purkinje images centred on their sensors. If both Purkinje images move the same amount (e.g. as would occur with a translational head movement) then the difference in the electrical signal sent to the two motors is zero—they both had to rotate their mirrors the same distance. When the eye rotates however, the two Purkinje images move a different distance—and hence the two motors have to move the two mirrors a different amount in order to stabilise their respective images. After a calibration procedure, the output of the system can be used to provide an exceptionally accurate estimate of eye position (Fig. 8.5).

DPI eye trackers are accurate up to 1 min of arc (a 60th of a degree of visual angle). As such the systems are perfectly capable of resolving very small micro-saccades. They were often used by psycholinguists, for whom spatial accuracy is of paramount importance (so that they can unambiguously determine which word, or even which letter within a word is being fixated).

On the downside, the nature of the optics means that whilst small head are tolerated, larger head movements (e.g. more than about 5 mm) can result in the eye becoming non-trackable. As a result, participants are often required to use bite-bars, which can quickly become uncomfortable.

Whilst DPI trackers provide a very accurate measure of eye position during fixations, they provide a slightly less accurate indication of eye position at the end of fast saccadic movements. This is because the fourth Purkinje image is reflected from the lens, which is suspended via muscles behind the cornea. At the end of a saccade, which is a very rapid movement that terminates very suddenly (see chapter by Pierce and colleagues in this volume), the lens "wobbles" slightly, causing small oscillatory changes in the position of the 4th Purkinje image, which are reported as a change in eye position by the tracker (see e.g. Deubel & Bridgeman, 1995). These "wobbles" are now referred to as "Post Saccadic Oscillations" (Hooge, Nyström, Cornelissen, & Holmqvist, 2015) and can cause ambiguity as to the precise timing of the onset of the subsequent fixation. As will be seen in Sect. 8.2.3, a related issue occurs with video based eye trackers which track the centre of the pupil.

Whilst exceptionally accurate (at least during fixations), DPI eye trackers are comparatively expensive—no doubt due to the complex optical and electromechanical components they involve. They also have a relatively limited operational range (in terms of how far the eye can rotate before one or other of the Purkinje images is lost), and this range can be very dependent on individual differences in pupil size as the fourth Purkinje image is seen through the pupil. Improvements in video based eye tracking systems—which offered similar levels of accuracy combined with a lower cost and greater ease of use and flexibility—resulted in many researchers moving away from DPI tracking, although there are still plenty of systems in active use.

8.2.2.4 Search Coils

Scleral search coils are essentially large contact lenses that contain a very thin filament wire wound round them (see Fig. 8.6). In humans the lenses are attached after the application of a topical anaesthetic. In non-human primates they may be surgically attached—removing any possibility of the lens slipping. The participant is placed within a magnetic field (created by larger wire coils, often housed within a "Necker Cube" type frame). When the coil in the contact lens moves, the magnetic field induces voltage change within the coil which can be directly measured. A figure of 8 winding in the lens allows torsional (rotational) eye movements to be measured, whilst a separate circular winding allows horizontal and vertical movements to be recorded.

The resulting voltage changes provide a very accurate measure of eye rotation within the head, and along with DPI trackers, scleral search coils systems are considered another "gold standard" in terms of eye tracking accuracy. There are, however, some disadvantages, not least the relatively invasive nature of the technique. Corneal scarring (albeit mild) can occur during lens insertion. The equipment is also relatively expensive. In addition, the lenses have been shown to slightly alter the kinematics of the eye (Frens & van der Geest, 2002), and even the smallest amount of slip can lead to inaccurate data.

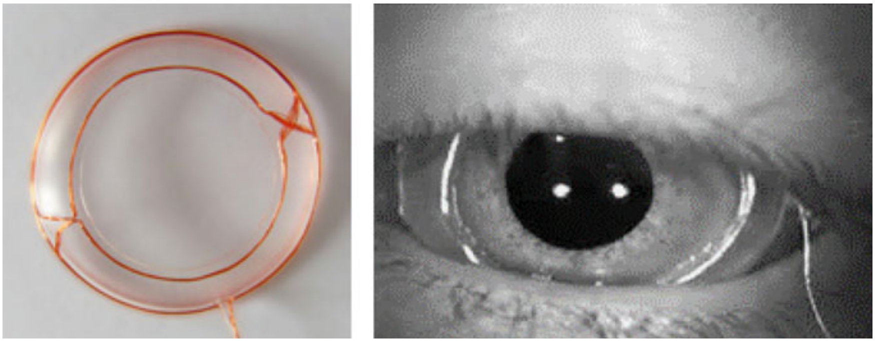


Fig. 8.6 Figure of 8 winding suitable for measuring torsional eye movements (left). Search coil in use (right)—note the trailing wires

An interesting variant of the scleral search coil technique is the dual induction system developed by Bour, van Gisbergen, Bruijns, and Ottes (1984). Rather than silicon lenses containing coiled wire, a gold annulus is placed on the eye. The distortions in the magnetic field induced by rotations of the eye are picked up by a secondary sensor placed in front of the eye. This technique avoids the wires trailing out from the lens—but is still comparatively invasive, and is not able to track torsional eye movements.

As with DPI trackers, the comparative ease of use and non-invasive nature of video based eye tracking has meant that scleral search coil techniques are used less often in human populations than they used to be—although it remains a relatively common technique for tracking the eyes of non-human primates. A recent publication by Kimmel, Mammo, and Newsome (2012) described a simultaneous comparison of search coil and video based eye tracking. They found very close agreement between the positional data from both systems and concluded that the video based system was appropriate for many if not most applications for which search coils have been used.

8.2.3 Video Based Eye Trackers

8.2.3.1 Basic Principles

Video based eye trackers are in some respects the modern descendants of Dodge's photochronograph in that they rely on (video) photography. They may also take advantage of a principle employed by DPI eye trackers to compensate for small head movements, that of two landmarks on the eye moving differently with respect to each other during rotation, but similarly during translation.

Rather than track two Purkinje images, most video based eye trackers track the first Purkinje image and/or the centre of the pupil (see Fig. 8.7).

The sampling rate of video based eye trackers is a function of the camera speed, that is the number of images it can capture per second. Commercial systems operate at anything from 30 to 2000 Hz. As is discussed in Sect. 8.4.1.3, sampling rate is a potentially important determinant of data quality. Only high speed systems are capable of delivering the high levels of spatial and temporal resolution previously associated with DPI and search coil techniques.

Video based eye tracking relies on image processing software that is able to rapidly identify, in each image, the precise location of the necessary landmarks (e.g. pupil and corneal reflection). When identifying the centre of the pupil two approaches can be taken. The most common is known as "dark pupil" tracking, and takes advantage of the fact that under infrared illumination, the pupil appears as a comparatively dark circle compared to the rest of the eye (e.g. the iris and sclera). The alternative approach is "bright pupil" tracking—which takes advantage of the well-known "red eye" effect—when light is shone directly through the pupil it bounces off the back of the retina. Most video based eye trackers rely on an infrared light source to illuminate

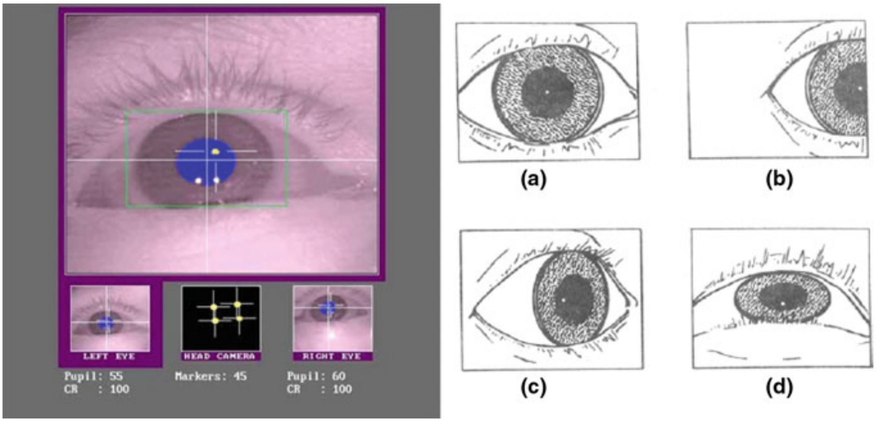


Fig. 8.7 Video based eye tracker using the centre of the pupil and first corneal reflection to track eye position (left). Illustration of the change in position of pupil and corneal reflection as seen by camera pointed straight at the eye (**a**) for rotational (**c**, **d**) and translational (**b**) movements. (Right, From Young & Sheena, 1975)

the eye and provide the corneal reflection. Systems that use infrared illumination and cameras are relatively untroubled by variations in natural light levels.

Complex image processing algorithms identify the centre of landmarks such as the pupil and corneal reflection, usually in real time, and after a calibration procedure has been performed the relative locations of the landmarks can be used to determine point of gaze. The precise details of these algorithms are generally proprietary, and a detailed discussion of the various approach is beyond the scope of this chapter.

As with DPI trackers, small head movements (which are translational) are differentiated from actual eye rotations by the different effect these movements have on the relative location of the pupil and CR centre on the camera's sensor—head movements result in both the CR and pupil shifting to the same extent, whereas eye rotations result in the pupil and CR moving a differing amount.

8.2.3.2 Monocular Versus Binocular Systems

Depending on the make and model, video based eye trackers may either track one eye or both eyes (or give the user the choice of tracking monocularly or binocularly). Most head mounted eye trackers are binocular. One advantage of binocular eye tracking is that vergence movements can be quantified and therefore some estimate of depth of focus can be established. This is particularly important for head mounted systems as participants are free to focus on objects at different distances, whereas for desktop systems the monitor is kept at a fixed distance. In general our eye movements are conjugate. In other words, both eyes move in the same direction at the same time, and the majority of eye tracking studies in the fields of Psychology and Cognitive Neuroscience report monocular data. However, when we change our focal depth eye movements can be disconjugate—looking at something straight ahead that is very close requires both eyes to rotate nasally compared to when looking at an object in the distance. Researchers interested in depth perception therefore require a binocular eye tracker. If presenting stimuli using a 3D monitor, or stereoscope, it is important to check that the eye tracker is still able to operate. For example, some 3D shutter glass systems require a strobing infrared light to synchronise the glasses with the monitor—and this could interfere with some video based eye trackers.

Another situation in which binocular data is useful is for detecting micro-saccades (see chapter by Alexander and Martinez-Conde in this volume). These saccades can be very close to the minimum spatial resolution of eye trackers, and therefore ambiguity as to whether a detected event is or is not a micro-saccade is reduced if the same movement occurred, in the same direction, in both eyes simultaneously. One of the most widely used algorithms for detecting micro-saccades (Engbert & Kliegl, 2003) can make use of binocular data.

8.2.3.3 Stationary Versus Mobile Systems

As mentioned in the introduction, video based eye trackers can be broadly categorised into stationary vs mobile systems. Stationary systems generally track eye movements whilst participants view stimuli presented (typically on a computer monitor) in front of them. These systems vary in the extent that head movements can be tolerated, requiring either some method to fix the head's position with a chin/forehead rest for instance, or allowing head movements within a limited "head box". Mobile systems allow participants to move around freely and look at objects other than a computer monitor, with the eye tracking components worn as glasses or a lightweight head-set. Stationary systems are typically used in research scenarios in which stimuli such as saccade targets, images or videos are presented to participants on a computer monitor. Mobile systems are generally more suitable for more "real world" type tasks in which participants are free to interact with people and objects around them as they go about a specific task (e.g. playing a sport/shopping/making a cup of tea).

Due to their need for small and lightweight components, most mobile eye trackers tend to have lower sampling rates and poorer spatial resolution than stationary systems (see next two sections).

8.2.3.4 Stationary Video Eye Trackers

There are a number of commercially available stationary systems, with large variations in cost, functionality and performance, but all operate on the same basic principles outlined above. Most are used in standard "laboratory" settings, in which a participant sits at a table in front of a monitor, but some stationary systems can also be used in neuroimaging environments. Their distance from the head means that they do not interfere with EEG electrodes for example. Some systems can operate in fMRI settings, with an MR compatible camera and illuminator being placed in a position where they can see the eyes either directly, or via a mirror above the head-coil.

Stationary systems can be divided further into two subtypes—remote eye trackers (that typically sit on a desktop) and tower mounted systems (which may be free standing or clamped to the desktop). Tower mounted systems combine the eye tracker with a head-restraint solution, and allow the camera to be placed closer to the participant's eye. In many cases the camera is mounted above the eye, looking down to a reflection of the eyes in a "hot mirror"—a mirror which is transparent to visible light (so that the participant can look through it) but which reflects infrared light. By placing the camera above the eye, out of the field of view, tower mounted systems have the advantage that the participant can interact with the stimuli being presented (e.g. by pressing on a touch screen) without interfering with the view from the camera or the IR light source. Tower mounted systems also often have a much larger trackable range.

Some systems require the participants head to remain relatively stationary, with some kind of chin rest or other head-restraint. Other "head free" systems allow greater freedom of movement, and some systems employ distinct algorithms for head fixed versus head free eye tracking. In general, greater accuracy can be obtained during head-fixed recordings. Research goals should determine which approach is most appropriate for any given project. Some populations (for example young children, infants or some patient groups) may not be able to use head-supports. Similarly, if the research goal is to explore eye movements during a relatively natural interaction with a computer (as in certain usability or HCI settings) then head-restraint is again undesirable. If head free eye tracking is required, it is important to consider the size of the "head box", that is the area in which the head is free to move whilst eye tracking is maintained.

One potential limitation of high speed video-based eye trackers that rely on the pupil centre to compute the point of gaze, is that the pupil is continually changing size. There are two ways in which this can be problematic. The first concerns dilations and constrictions—these are typically driven by luminance changes (including local luminance changes in the image being viewed), but can also be elicited by factors such as cognitive work load, arousal or fatigue (see chapter on pupillometry by Alnaes and Lang in this volume). It might be tempting to assume that as the centre of a large circle is in the same location as the centre of a small circle, pupil size should not influence the reported location of gaze. Unfortunately, pupils do not dilate or constrict symmetrically around their centre (see, e.g., Wildenman & Schaeffel, 2013). This means that large luminance changes both within and between trials can potentially cause slight spatial inaccuracy in the reported gaze location. The second potential problem that tracking the centre of the pupil can cause is post-saccadic oscillations. As mentioned previously, these are an issue in DPI trackers because the fourth Purkinje image is reflected from the back of the lens, which can wobble at the end of saccades. The wobbling of the lens may possibly influence the shape of the pupil, and thus its computed center. Video eye trackers with low sampling rates do not suffer from this issue as they lack the temporal (and often spatial) resolution necessary to resolve these tiny changes in the location of the pupil's centre. Nyström, Andersson, Magnusson, Pansell, and Hooge (2015a), Nyström, Hooge, and Holmqvist (2015b) provide a comprehensive discussion of the issue.

8.2.3.5 Mobile Video Eye Trackers

Some mobile eye trackers are worn as glasses—with the critical eye tracking components housed within the frame. Other mobile systems have the components attached, one way or another, to some kind of head-set—this may be as simple as a modified spectacle frame on which is mounted one or more miniature cameras and possibly a mirror, or a more complex headset carrying more or larger components. As with stationary video based eye trackers, the key components of mobile systems are cameras, capable of viewing the eyes, and one or more infrared light sources capable of illuminating the eyes and providing a corneal reflection. Due to the obvious need for key components (e.g., cameras) to be lightweight, head mounted eye trackers tend to have lower specifications than desktop systems.

A key feature of most mobile eye trackers is that in addition to the camera or cameras that are recording the eye, they contain a "scene camera", that is a camera that is capable of recording the scene the participant is looking at, capturing the participants view of the world, and over which a gaze cursor can be plotted both in real time and any subsequent analysis stage. This feature is necessary for any use in which the person being eye tracked is interacting with the real world, rather than simply observing stimuli presented on a computer screen.

Data quality is more likely to be an issue for mobile eye trackers than stationary eye trackers. This may be due in part to the relatively low specifications of the cameras and other components used (in order that the systems remain wearable), but perhaps more important are the challenging scenarios in which mobile trackers tend to be used, for example driving or sport. Such scenarios clearly limit the control the experimenter has over potentially important variables. For example, bright sunlight can be particularly problematic for mobile eye trackers. Not only can it "bleach out" the infrared illumination systems, it can also induce a dramatic reduction in pupil size, and even squinting behaviour—making the eye difficult or impossible to track.

8.3 Software Used in Eye Tracking Research

A critical component of modern eye tracking methodology is software, which can be divided into three broad categories: Firstly, there is the software used by the eye tracking system itself. At a minimum this software provides a user interface and controls the actual recording of the eye movement data. Secondly, there is software for stimulus presentation. This is a requirement for most, but by no means all eye tracking research. Finally, there is software that facilitates the analysis of eye tracking data.

The aim of this brief section is not to provide an exhaustive list of the capabilities of what are generally rapidly evolving pieces of software. Nor is it meant to be a comprehensive comparison between various approaches and solutions. Its central purpose is to make readers aware that when evaluating published eye tracking research, or when considering which type or model of eye tracker would be most appropriate for their needs, they should not limit their evaluation to hardware considerations. People looking to buy an eye tracking system should look in detail at the capabilities of the software itself: Does it provide the functionality you need? How it is licensed? Will future upgrades incur a cost? People evaluating research should consider the extent to which critical factors (such as how fixations and saccades are defined) are controlled by the eye tracking software.

An informed understanding the eye tracking software—what it is doing and how it is doing it—can also lead to improvements in participant set up and ultimately data quality. By understanding what your equipment's software is trying to do, you are better placed to know when (and why) it might be struggling—and what you can do to help.

8.3.1 Software Controlling the Eye Tracker

One of the key functions of eye tracking software is to provide the user with an interface through which participant set up and calibration can be controlled. Important points to consider are its ease of use (particularly if working with populations such as infants or patients) and its flexibility (can different calibration targets be used for example?). The software will also save the eye movement data to a file.

At the heart of all video based eye tracking software are image processing algorithms. Essentially the software receives a series of images of the eye(s) from the eye tracking camera(s), and needs to identify (at a minimum) the location within that image of the centre of the pupil and the corneal reflection(s). Manufacturers differ in the extent to which the various algorithms (filters, parsing rules etc.) used by the eye tracking software are revealed to customers, and in the extent to which these can be modified, but these low level algorithms are often proprietary, and may not made available to researchers. Their robustness is one of the key determinants of eye tracking data quality, and as is argued in Sect. 8.4, setting up participants optimally involves ensuring that the algorithms are able to do their job.

Eye tracking software also often handles communication with the presentation software. Such communication is important as it allows stimulus events (such as the onset of a target or image, or the offset of a fixation cross) that are controlled by the presentation software to be logged in the eye tracking data. Without such communication calculating saccade latencies, and when fixations occur with respect to the stimulus onset can be difficult. Another important aspect of this communication is whether it is bidirectional. In other words, can the eye tracking software make the results of its gaze calculations available to the presentation software? Such communication is critical for gaze contingent tasks (in which for example the location of a saccade target is shifted during the saccade itself). If such bi-directional communication is required, then it is important to establish the end to end delay (that is, the delay between the eye changing position and that change being calculated by the eye tracking software and sent back to the stimulus presentation software; see Sect. 8.4).

8.3.2 Stimulus Presentation Software

There are two broad categories of stimulus presentation software—software supplied by the eye tracking manufacturer, and generic stimulus presentation software (which can be further divided into commercial software such as E-Prime, Matlab, NBS-Presentation, and free solutions such as OpenSesame or PsychoPy). When considering stimulus presentation software, it is important to evaluate their timing accuracy (their ability to present stimuli at a known point in time) and synchronisation abilities (the extent to which the time that a stimulus event occurs is accurately flagged in the eye tracking data stream).

Software provided by manufacturers ranges from basic products that essentially implement some kind of "slideshow" of text, images or possibly videos, to powerful and sophisticated software capable of implementing a very wide range of experimental scenarios. The main advantage of software provided by the eye tracking manufacturer is that integration between the stimulus presentation and recording software is generally handled very straightforwardly, for example if calibration routines and gaze contingent triggers are easily set up. Manufacturer supplied display software may also offer straightforward integration with any analysis software that they provide, for example so that interest areas defined in the experiment are available automatically at the analysis stage. Another advantage is that the manufacturers may be able to offer more support to customers using their software for stimulus presentation.

Commercial stimulus presentation software is often highly featured and well supported. Most modern eye trackers will provide libraries or some means by which commercial stimulus presentation packages can communicate and interface with the eye tracker—for example to start or stop recording and signal key stimulus events such as fixation onset or target onset etc. Commercial programs differ with respect to their timing capabilities, as well as their ability to accurately display stimuli of different types (e.g. audio or video). Users should make sure that they choose presentation software that is capable of implementing the eye tracking tasks they plan to run.

A final distinction that can be made is between presentation software that is essentially a programming environment (Matlab with Psychtoolbox or Presentation) and software that is uses a graphical interface (e.g. E-Prime or PsychoPy). Whilst programming environments are often very versatile, they can take a relatively long time to learn. Conversely, while software that uses a graphical approach can be very easy to work with, and allow functional experiments to be produced quickly, they may lack flexibility required for more complex experimental designs.

8.3.3 Data Analysis Software

Much like EEG and fMRI, eye tracking can very quickly result in very large amounts of raw data. A relatively modest experiment containing 24 trials in which eye movements were recorded for 8 s per trial, would result in 192,000 samples from a high speed eye tracker recording at 1000 Hz. Each sample would contain, at a minimum, a timestamp and the X and Y location of gaze (and often pupil area) at that point in time. Multiply the number of data points by the number of participants and the disadvantages of analysing the data without some kind of software solution are clear.

For researchers happy to deal in sample level data, programs capable of dealing with large quantities of data such as Matlab and R can be used. Sample level position data can readily be converted to velocity, allowing the data to be parsed into saccades and fixations using whatever criteria the researcher feels are appropriate. Such software also allows for flexibility in terms of applying filters should these be desired or be felt necessary due to data quality issues.

There are some commercial software solutions designed specifically for eye movement data analysis, that can work with data from a number of different eye trackers (either in its raw form, or converted to ASCII format). There are some free eye movement data analysis packages—but they are typically aimed at relatively discreet types of analysis. For example iMAP (Caldara & Meillet, 2011) is a powerful package of Matlab functions that allow sophisticated fixation maps (heatmaps) to be created and compared statistically. Most analysis software, however, is supplied by manufacturers, and will work only with data collected on their eye tracking systems.[1]

As with stimulus presentation software, analysis solutions provided by manufacturers may vary enormously in their sophistication. At the most basic level are programs that allow a single data file to be loaded, perform some basic parsing to allow fixations to be identified, and plot these over the stimuli that were presented. Such software may perform some other basic visualisations—scan paths for example (see chapter by Foulsham in this volume)—and calculate basic metrics such as fixation count and average fixation duration. More sophisticated software will allow data to be loaded from more than one participant, and will likely involve some integration with the stimulus presentation software. Often such software will also allow the creation of interest areas (also known as regions of interest) and be capable of outputting reports that provide a range of summary data at the interest area level (such as dwell time). The most sophisticated analysis software will be tightly integrated with both the recording and stimulus presentation software, and be able to present the data with a range of visualisations, including, for instance, the creation of heat maps etc. Other functions may include the ability to group data by trial variables, "clean" data, limit analyses to specific interest periods, create dynamic interest areas, and will be capable of outputting hundreds of useful dependent variables for analysis.

---

[1] A reasonably up to date list of software for analysis can be found here: http://www.eyemovementresearch.com/software/.

8.4 Data Quality

No matter what eye tracking technology is used, or what population is being tested, all researchers should be aiming to record the best quality data that they can. Good quality data makes subsequent analysis far simpler and any effects easier to describe and prove statistically. It is, however, important to bear in mind that different aspects of data quality may be more or less important to researchers depending on their research goals. For example, psycholinguists interested in which part of which word was fixated during reading, and exactly when and for how long it was fixated, will require exceptional spatial and temporal accuracy and precision. Researchers using a dot-probe like task, or a preferential looking task with infants, may only need to know whether the first eye movement after target presentation was to the left or to the right, in which case spatial accuracy and precision will be less important. Due to space limitations this following discussion should be considered introductory. Readers interested in a more detailed discussion on this topic can find much useful information in Reingold (2014) and Holquvist and Nystrom (2010).

As will become clear, some determinants of data quality are a function of the eye tracker itself such as its sampling rate, noise level, the filtering algorithms it employs. However, other determinants are, to some extent at least, within the control of the experimenter and the participant. The aim of this section is to ensure that researchers working with eye trackers have sufficient understanding of those aspects of data quality that they have some control over to be able to take positive steps to ensure they are getting the best data they can, given their system's potential.

8.4.1 Key Terms

Any attempt to discuss eye tracking data quality involves stepping into a terminological minefield. Terms such as "accuracy" and "precision" are often used by different researchers to mean rather different things. This section starts with an attempt to define various key terms (highlighted in bold), as they are used in this chapter—which, it should be born in mind, is not necessarily how they are always used in published research or manufacturer's literature.

As a starting point, I follow Reingold (2014) in defining **data quality** as "the fidelity with which the continuous variation in the eye movement signal is reflected in the values measured and reported by the eye tracker". In other words, we have something real that changes over time (the position of the eye) and we have our measurement of that position. In crude terms data quality can be considered as the extent to which the measurement agrees with reality, with high levels of agreement equating to high data quality. Any attempt to measure data quality, defined as such, obviously faces an immediate problem: When recording from an actual eye we have no independent estimate of its "real" eye position other than our measurement of it via the eye tracker. This point (and possible solutions) is returned to in Sect. 8.4.2

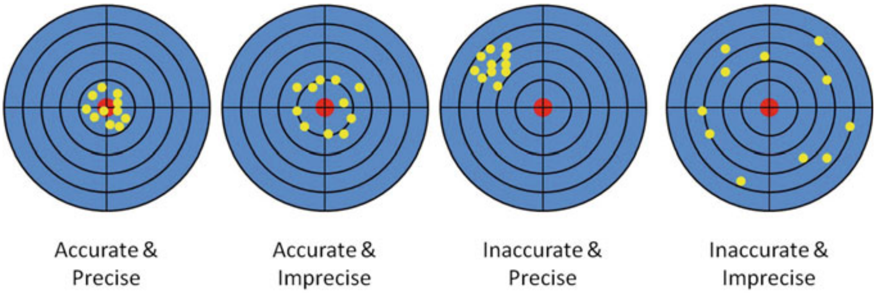


Fig. 8.8 illustration of the difference between, and independence of, accuracy and precision

on measuring data quality. The remainder of this section will attempt to provide definitions of several other important concepts.

8.4.1.1 Accuracy and Precision

Perhaps the two most important (and widely used) concepts in any discussion of eye tracking data quality are "accuracy" and "precision". These two concepts are, as Reingold (2014) points out, often treated (erroneously) by researchers as equivalent. If a single actual (true) value is repeatedly measured, accuracy can be defined as the mean difference between the measured and the true value. In contrast, precision (or reproducibility or repeatability), can be considered as the extent to which repeated measurements of a set of true values produces the same or similar set of measured values—regardless of the accuracy of these values. In other words the measured values may not be accurate, but if they are consistent then precision is high. The concept is illustrated in Fig. 8.8:

8.4.1.2 Noise

Another concept, which can be related to both accuracy and precision, is often referred to as "noise". This can be considered as variations (unwanted or of unknown source) in the gaze position signal, a signal which, at least during fixations, is often implicitly assumed by researchers to be stable. It is important to understand that there are several potential sources of noise in eye tracking data. One major source of noise is the human oculomotor system itself. The eye is never truly still; even during fixations movements such as drift, tremor and micro-saccades are readily observed. Such oculomotor "noise" is real and can be accurately measured, but it can be considered as noise in some research contexts because it adds variation to what many researchers assume is a steady signal (for instance, the eye during fixation). The second source of noise can be termed environmental, caused for example by vibrations disturbing the eye tracking camera, or electromagnetic interference. The final source of noise is the eye tracking equipment itself. When an artificial eye is tracked, and environmental factors are adequately controlled, the contribution to overall noise levels made by the eye tracker itself can be isolated and the “spatial resolution” (the smallest change in position that can be detected) of the system can be calculated. High end video based eye trackers have spatial resolutions of around 0.01–0.05° of visual angle.

8.4.1.3 Temporal Resolution, Latency and Delays

Our discussion of data quality has thus far focussed on the spatial dimension, but timing issues are often equally important. Perhaps the simplest concept to grasp is that of “temporal resolution”, being essentially the sampling rate of the eye tracker, as typically expressed in Hz and corresponding to the number of times the eye tracker is capable of sampling the eye’s position each second. As mentioned previously, video based eye trackers have sampling rates ranging from around 30 to 2000 Hz. Sampling rates are important for a variety of reasons, and high sampling rates in general confer benefits rather than disadvantages.

If fixation duration, or saccade onset latency are critical variables, then having an eye tracker that can establish these durations or timings to the nearest 1 or 2 ms (e.g. a 1000 or 500 Hz system), as opposed to the nearest 16.67 ms (as is the case for a 60 Hz eye tracker) is preferable. Holmqvist et al. (2011) point out that the effects of low sampling frequency can be compensated for by collecting sufficient data. But collecting sufficient data to overcome the limitation of a slow eye tracker can be a high price to pay if the expected difference in values (e.g. fixation durations) between conditions is small, and in many experimental settings is simply not an option.

High sampling frequencies are also needed if you are interested in detecting microsaccades, or accurately measuring saccade velocity or amplitude. Another advantage of relatively fast sampling rates is that it can reduce “recovery time”. During eye tracking recordings it is not unusual for the eye tracker to periodically be unable to track the eye, with blinks being by far the most common reason. When the eyelid descends it obscures the pupil and corneal reflection and eye tracking is not possible. Head rotations that obscure the camera’s view of the eye, or head movements that take the eye out of the camera’s field of view can also result in “tracking loss”. Eye trackers with high sampling frequencies are much quicker to “reacquire” the eye than trackers with low sampling frequencies.

Another important term is “end to end delay” which can be defined as the time between an actual movement of the eye taking place, and the eye tracking system signalling that the movement has occurred. Having a low latency eye tracker is particularly important for users who want or need to run gaze-contingent tasks in which some aspect of the stimulus display is dependent on the participants gaze position. Examples of gaze contingent tasks include the “moving window” type tasks popular in psycholinguistics, and saccade adaptation tasks, in which the target location is shifted during the saccade itself.

A related concept is “temporal precision”, which can be defined as the standard deviation of eye tracking latencies. Poor temporal precision can occur if the computer running the eye tracking software allows the operating system to “hog” the processor, taking priority over the gaze calculations. Poor temporal precision can create huge difficulties for researchers, particularly when trying to synchronise their eye movement data with stimulus events. It can also contribute ambiguity to the onset and offset of fixations and saccades.

8.4.1.4 The Interaction Between Noise and Sampling Rate: Velocity Noise

The velocity of an object can be defined as its change in position per unit time. Eye velocity is typically measured in degrees of visual angle per second. Velocity can be calculated from position data, using a variety of models. A detailed treatment of these models is beyond the scope of this chapter, but in essence velocity can be thought of as the speed of the eye, and is therefore related to the difference in eye position between samples.

Velocity is often calculated, either during recording, or after data collection, and can be used for defining the onset and offset of saccades (and hence the onset and offset of fixations). Because velocity is based on the differences between positions, any positional noise will contribute to what is known as velocity noise. Reingold (2014) illustrates the critical relationship between sampling speed and velocity noise with the following diagram. It is clear that a moderate amount of spatial “noise” can have a dramatic impact on the number, duration and timings of fixation and saccade events. Holmqvist et al. (2012) provide another important demonstration and discussion of this issue.

In the top panel of Fig. 8.9, moderate noise has been added to two samples of positional data recorded at 200 Hz (indicated by asterisks). When the positional data is converted to velocity, the velocity noise causes both the false detection of a saccade, and errors in the duration and onset of two fixations.

8.4.1.5 Filtering

Modern eye trackers may employ some kind of filtering to attenuate variations in the eye tracker signal that are not related to the eye movement itself (e.g., noise). Filtering can be applied by the tracker software during the recording itself, and/or by the analysis software after the data has been recorded. Different eye trackers use different filtering techniques, typically either “heuristic” or rule of thumb filters, or some mathematical filtering process (e.g., a moving average filter). It is important to note, however, that certain filters may also distort the signal, for instance, by changing the latencies of certain events (e.g., saccades) or filtering out others (e.g., micro-saccades). It is therefore important to understand the impact of any filtering on your data.

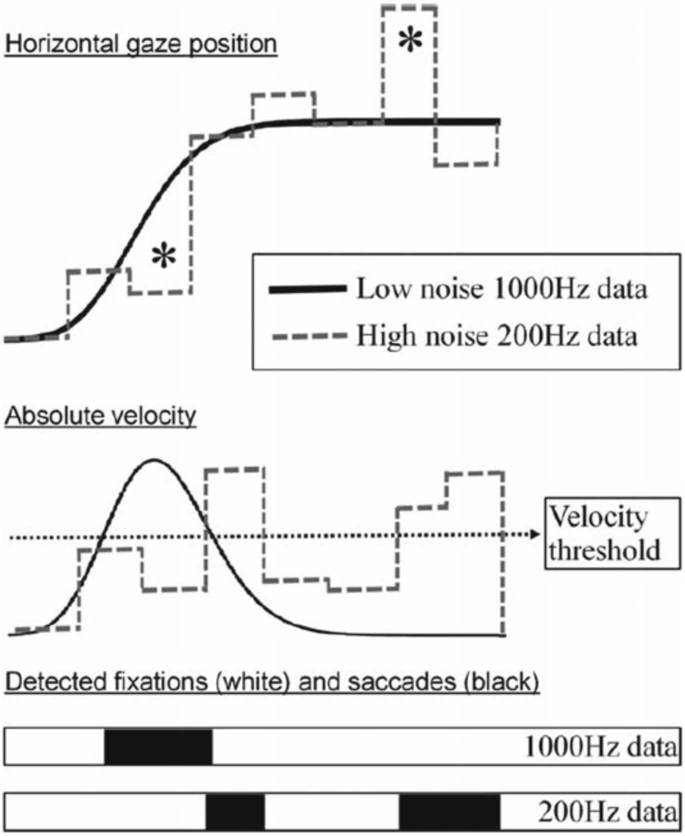


Fig. 8.9 Illustration of the effect of velocity noise on parsing. From Reingold (2014)

8.4.1.6 Parsing

Another issue when considering data quality concerns what exactly we are referring to by "data". Eye trackers output "sample level" information, in other words a single estimate of the eye's location (or point of gaze) for each measurement taken (e.g., 60 times per second for an eye tracker with a 60 Hz sampling rate or 1000 times per second for a 1000 Hz system). Whilst low level sample data is of enormous importance to many researchers, particularly those concerned with the basics of oculomotor control, saccade dynamics and curvature, smooth pursuit etc., it is of less importance to many others, at least in the sense that they are unlikely to want to deal with their data at the sample level. This is particularly the case for those researchers whose primary measures are based on fixations, including their location and duration. For example, an infant researcher may want to know which of two simultaneously presented faces was fixated first, and for longest. A psycholinguist may want to know which words were fixated, in which sequence and for how long.

Such analyses generally do not require looking at data at the sample level. They do, however, require that the sample data be parsed, which means that a set of rules is required to allocate any given sample to some event-state, typically fixation, saccade or blink. The properties of these events (e.g., the location of a fixation, its onset time, its duration, the amplitude of a saccade) are the key metrics of interest. If dealing with parsed data it is important to understand the parsing process, in order to maximise the potential of it delivering meaningful decisions, as well as understand its limitations.

Some video based eye trackers parse the data in real time. Others leave the parsing until after data collection—at which point the parsing is typically performed by the analysis software. There are two broad approaches to parsing. One is a fixation detection approach, with fixations typically being defined by properties such as duration and dispersion. The other is known as a saccade picking approach, with the onset/offset of saccades being defined by velocity/acceleration thresholds. High speed systems tend to use the saccade picking approach as, with sufficiently high temporal resolution and low velocity noise, the onset of a saccade is unambiguous. Typically, saccade onset will be determined by a relatively simple set of rules (e.g. eye velocity increases above X degrees per second). When a saccade picking approach is used, fixations are defined simply as being those samples that are not in a saccade or blink. In other words, the onset of a fixation is defined by the end of the previous saccade, and the end of a fixation is defined by the onset of the subsequent saccade.

8.4.2 Measuring Data Quality

As discussed in Sect. 8.4.1, if data quality is defined as the correspondence between the measurement and reality then we are faced with a problem, namely finding some “true” measure of eye position. In the relatively sparse literature that has directly addressed eye tracking data quality two different approaches have been employed. The first is to use some kind of exceptionally accurate and precise eye tracking technique to act as a proxy for the “real” eye position, and measure eye movements simultaneously with both the “gold standard” technique and the eye tracker whose data quality is of interest. As discussed in Sect. 2, whilst both DPI and scleral search coil techniques have been claimed as “gold standards”, there is no such thing as a perfect eye tracker. As mentioned in Sect. 8.2 of this chapter, search coils are known to influence the kinematics of eye movements and lens slippage can also be problematic. DPI trackers report tiny movements of the lens at the end of saccades as changes in eye position. There are also physical limitations with respect to which eye tracking techniques can be performed simultaneously. For example, the infrared emitters in limbus tracking systems could potentially interfere with the infrared illuminators often used in video-based systems. Similarly, it would be very difficult to simultaneously record using a DPI tracker and a video based system given the need for both systems to have an uninterrupted view of the eye being tracked.

The alternative approach is to use some kind of artificial eye whose absolute position at any point in time can be known (typically because it is invariant). This approach (see Reingold, 2014) also neatly avoids another problem inherent in assessing data quality, namely the fact that any spatial inaccuracy in eye movement data is a product of both the eye tracker related error, and participant related error. Eye tracker related error can only be truly estimated if the precise position of the eye being tracked is known—as with an artificial eye. Some manufacturers will report values for spatial resolution that are based on tracking an artificial eye. It should be pointed out that artificial eyes may be easier or harder to track than actual eyes, and as such, and evaluation of eye tracking accuracy and precision should attempt to utilise data from both.

One interesting variant of the artificial eye is one that uses two LEDs to mimic the corneal reflection. If these LEDs are in different locations, rapidly switching from one to the other simulates a saccade. Such devices are particularly useful for measuring latency and end-to-end delay. Whilst a variety of static artificial eyes are commonly used, recent advances now allow "robotic" artificial eyes to be used instead of human eyes (Reingold, 2014).

8.4.2.1 Measuring Accuracy

Accuracy can be calculated as the average difference in the reported location of gaze to the actual location of the target. In practical terms it is often measured at the fixation level. In other words, it is the reported location of the fixation that is compared to the actual location of the target. As is explained in Alexander and Martinez-Conde's chapter in this volume, the eye is never truly still, and there are a large number of different approaches to defining what exactly constitutes a fixation. As such, the algorithm involved in computing the fixation onset or offset and location is a critical determinant of accuracy. Assuming that fixation duration has been established, one relatively straightforward solution to defining a fixation's location is to simply take the average X and Y location of all samples that are included within its duration.

In practical terms, accuracy can be measured by asking participants to repeatedly fixate a target, or targets with known locations. Accuracy can be expressed as the difference in position of the fixation and the target, in either the X or Y dimensions (or both if Pythagoras' theorem is used to calculate the angular distance between the fixation and the target locations).

8.4.2.2 Measuring Precision

Precision is generally reported at the sample level and can be expressed either as the standard deviation of the X, Y locations of samples, or in terms of the distances between samples—in which case it is often expressed as Root Mean Square (Sample to Sample), or RMS-S2S.

8.4.3 Maximising Data Quality—Participant Set up and Calibration

Without doubt, one of the most important determinants of data quality (particularly spatial accuracy) is the initial set up and calibration procedure, but it is one that often gets rushed or is done sub-optimally for a variety of reasons.

Obviously, the precise details for optimal participant set up will differ as a function of the eye tracker being used—for example setting up a head-mounted eye tracker is very different to setting up a desktop based system. However, various basic principles are common to all video-based systems, desktop or head mounted. The camera needs a good view of the eye and the eye itself needs to be well illuminated with both pupil and corneal reflection(s) visible. Similarly, a number of factors such as glasses, eye makeup and droopy eyelids can be potentially problematic for all video based eye trackers.

Ideally set-up and calibration would involve a knowledgeable and practiced experimenter, and an alert and motivated participant. Techniques for setting up and calibrating infants and other potentially non-motivated populations are provided in Sect. 8.4.3.2.

8.4.3.1 Set up

Many data quality problems can be avoided if the participant and eye tracker are set up optimally. There are several factors to consider, the relative importance of each depending to some extent on the make and model of eye tracker being used. The following advice is meant to be as generic as possible.

One key factor is the height of participant with respect to the monitor on which the stimuli are to be presented. Many researchers assume that the eye should be aligned with the centre of the screen. In fact, high-end desktop based eye trackers may actually function best if the level of the eye is closer to the top of the screen. Having a height adjustable chair or adjustable table is critical.

Another important factor is distance of the eye from the camera. This is important for two reasons. Firstly, if the eye is too far from the camera, the pupil and corneal reflection will appear smaller, and occlude fewer sensor pixels. The image processing algorithms will then have less data to work with and will produce less stable estimates of their centres. In other words, noise in the data will increase and data quality will decrease. Secondly, if the eye is too close to the camera it is possible that eye movements to targets placed towards the edge of the screen will exceed the trackable range of the system. All eye trackers have a trackable range, often described in degrees of visual angle, which correspond to the maximum rotation the eye can make before tracking becomes inaccurate or impossible (as the corneal reflection will "fall off" the cornea and onto the sclera). Whilst head mounted systems will also have a trackable range, it often less of an issue as the participants are free to turn their head towards objects of interest, thus minimising the rotation of the eye itself. For desktop systems, however, if the monitor is placed too close to the participant, it can be very easy to engineer a situation in which the corners or edges of the monitor fall outside of the trackable range of the system. A good rule of thumb is that at about 60 cm distance 1 cm on the screen equates to 1° of visual angle. Hence a 40 cm wide monitor 60 cm from the participant's eyes would subtend ±20° of visual angle horizontally.

Before starting the calibration, it is always a good idea to check the stability of the set-up by asking the participants to look at the four corners of the calibration space. If the eye tracking software allows it, make sure that the pupil and corneal reflections are visible at each corner, and their shapes are not in any way distorted. One very common issue is for the corneal reflection to become distorted or disappear altogether in one of the corners. This is often a sign that something is sub-optimal with respect to the geometrical arrangement between the participant, the eye tracking camera and the monitor. The simplest solution is often to simply move the monitor slightly further away, or alternatively adjust the height of the monitor or participant such that they are not having to rotate their eyes up or down so much in order to view the problematic area. So for example of the problem is with the top corners, lowering the monitor or raising the participant would probably provide a solution. If the bottom corners then become problematic, the monitor should be moved further away.

Factors that can complicate participant set up or increase "noise" and some possible solutions:

(1) Make up: Eye make-up, particularly mascara, can occasionally cause problems as it tends to be the same colour as the pupil (e.g. black) and can therefore potentially interfere with the image processing algorithms that are attempting to locate the pupil's edge. The extent to which mascara can be an issue depends very much on the eye tracker. Particularly large or even false eye lashes, which can curve such that they overlap with the pupil from the camera's view point can be problematic for even the most robust video based system.

(2) Glasses: Glasses with particularly thick or dirty lenses can obscure the view of the eye from the camera. In addition, glasses may diffuse some of the infrared light typically used to illuminate the eye and provide the corneal reflection. This can be particularly problematic if the lenses have been coated to reflect infrared light. The reflection of the infrared light source from the glasses can sometimes obscure the eye. In many cases simple repositioning of the participant's head by asking him or her to thrust their chin forward towards the camera will move the reflection out of the way. Other tips include bringing the camera closer and tilting it up at a steeper angle. If tracking monocularly, often the simplest thing to do is try the other eye. Often the reflection will turn out not to be problematic for that eye.

(3) Contact lenses: Soft contact lenses tend to remain fixed to the cornea, and are generally not a problem for most video based eye trackers. Hard lenses, on the other hand, are generally smaller, and tend to "float" over the eye—particularly during saccades. As a result, the edge of a lens can get close or even cut across

the pupil, distorting its shape as it appears to the camera, and confusing the algorithms that are attempting to calculate the pupils centre. The position of the lens can also impact on the position of the corneal reflection. In general, the best solution is to ask the participant if they have any glasses they can wear as an alternative.

(4) Monitor distance: It is surprising how often users set equipment up in ways that mean it they eye is forced to rotate beyond the trackable range of the camera in order to view the corners of the monitor. The camera should be placed at the distance from the eye recommended by the manufacturer, and the monitor positioned at a distance that ensures it falls within the trackable range.

(5) Eye (iris) colour: Video based eye trackers rely on being able to distinguish the boundary of the pupil and iris. Certain eye colours (particularly brighter or paler irises which provide less contrast with the pupil under infrared illumination) may be problematic for some models of eye tracker. In very young infants (3 to 6 month olds) the iris lacks pigment, again making it difficult for some eye trackers to discriminate between the pupil and iris under infra-red illumination (a similar issue occurs in people with albinism). Some eye trackers can operate with different frequencies of infrared illumination which improve the pupil/iris contrast in very young infants.

(6) Pupil size: As discussed in Sect. 8.2, a somewhat unfortunate property of pupils is that they do not dilate symmetrically around their centre. As a result, it is important to try to avoid drastic changes in luminance across trials or during the course of a single trial (e.g. black fixation cross on white background followed by face on a black background).

(7) Ocular dominance: Approximately 80–90% of participants are right eye dominant. Perhaps the simplest test of ocular dominance, which works well with both adults and children (who can find winking difficult) is to ask them to extent their arms and form a small opening in their hands through which they look at a distant object (such as the experimenter's nose). Then ask the participant to bring their hands up to their face, whilst continuing to keep the experimenter's nose in view through the opening in their hands. They will bring their hands to their dominant eye. The "extent" of ocular dominance varies greatly between individuals—but tracking from a non-dominant eye can be a cause of accuracy problems. Symptoms of tracking from a non-dominant eye include significant (readily observable) drift during fixations, and poor/inaccurate calibration models/difficulty with validation (because the non-dominant eye tends to go to a slightly different place each time).

A useful approach if recording monocularly is to start calibrating the right eye. If calibration models are poor/attempts at validation give bad results, then one possibility is that you are in fact tracking from the non-dominant eye, so try calibrating the left eye and see if things improve.

(8) Ptosis (droopy eye lids): This can be a particular problem in elderly participants (although it can also occur with tired or drowsy younger participants). The top eyelid can droop sufficiently that it obscures the top of the pupil – potentially causing problems for the image processing algorithms that are try-

ing to determine its centre. A related problem can occur in low light levels if the pupil dilates sufficiently that its edge is partly obscured, even in the absence of a drooping eyelid. Some video based eye trackers allow the user to switch between "centroid-based" and model based estimates of the pupil centre. Model based solutions are often better when the pupil is partially obscured as the model "fills in" the missing edge. Genuine ptosis (as opposed to drowsiness) cannot be corrected other than by taping up the eyelids, a procedure that can be both embarrassing and uncomfortable for elderly participants, and which should be avoided if at all possible. Occasionally repositioning the camera, so that it is looking up into the eye at a steeper angle can help.

8.4.3.2 Calibration

The calibration procedure is the most important determinant of subsequent spatial accuracy—but as with participant set up, it is often rushed, or unwittingly done sub-optimally—with seemingly little or no awareness of the consequences this can have. The aim of the calibration procedure is to provide a mapping function that allows raw eye tracking data (the location of the pupil and CR centres on the camera sensor for example) to be converted to screen (typically pixel) co-ordinates. It usually involves presenting targets at a sequence of known locations (anywhere between 3 and 25 depending on the system and needs of the researcher—although some head mounted eye trackers allow single point calibrations). Calibrations can be one-dimensional (e.g. horizontal only) or two-dimensional.

Once the values of the raw eye tracking data at each location are known various mathematical/modelling approaches can be applied to allow the point of regard to be established for any intermediate raw eye movement values. The simplest approach would be to use a linear regression. Limiting our discussion to the horizontal dimension for simplicity's sake, assume that the participant fixated the following three target positions on the X axis: -200, 0, 200 (assume that the units are pixels and 0 represents the centre of the screen). If the raw X signals from the eye tracker when the eye fixated those 3 locations (averaged over say 100 ms of a fixation) were 100,500 and 900 respectively, the following regression equation ($y = mx + c$) describes the relationship:

$$\text{ScreenPixelX} = 0.5 * \text{ EyeTrackerX} - 250$$

Thus a raw eye tracker value of 700 would equate to a screen pixel position of $(0.5 * 700) - 250 = 100$.

In reality the calibration procedure may involve considerably more complicated non-linear regression or modelling techniques, but the above example illustrates the central concept: the calibration serves to provide some mapping function that allows raw tracker values to be converted to more meaningful units related to the calibrated space (typically a monitor, and thus often expressed in screen pixels). Therefore, it is one of the most important determinants of the spatial accuracy of any eye movement recording.

One important consequence of any calibration that involves regression is that any inaccuracy within the model may become amplified towards the extreme edges of the calibrated space. Taking the simplified example above, assume that when the participant was supposed to be looking at the target 200 pixels to the right of the screen, the experimenter pressed the "Accept" button before they eye made a secondary saccade to get on target. As a result, the participant was actually fixating pixel 195, and the raw tracker value for that calibration target was 880.

The (inaccurate) mapping function is now

$$\text{ScreenPixelX} = 0.51 * \text{EyeTrackerX} - 253.$$

When the participant looks at the centre of the screen (500 in EyeTrackerX; 0 in ScreenPixelX) the equation will now return a value for ScreenPixelX of 3.84, approximately 4 pixels out from the "true" gaze position. However, when the participant looks at the target on the right (900 in EyeTrackerX, 200 in ScreenPixelX) the equation returns a value of ~208, in other words an 8 pixel discrepancy between the true eye position and the reported eye position, that is twice as much error as when the participant looked at the centre.

The purpose of this somewhat laboured explanation is to try to highlight the importance of having an accurate calibration model for getting spatially accurate data. Note that a faulty calibration model will not have any impact on precision, in the case above every time the participant looks at the target at +200 the eye tracker will say they are looking at +208.

Reingold (2014) points out that at the heart of the calibration procedure lies an assumption that is almost certainly false. The assumption is that at the point at which the raw eye tracking data is sampled and assigned to the target's location, the eye was looking exactly at the centre of the target. This assumption is clearly false as we know that the eye is never truly still (see chapter by Alexander and Martinez-Conde in this volume). In addition, there is the simple fact that the participant's actual point of gaze may not be on the target's centre. Foveal vision is capable resolving in high acuity an area of the visual scene 1–2° in diameter—typically larger than the size of a standard adult calibration target. In other words, in order to "see" the calibration target, the participant does not necessarily have to fixate its exact centre. As Reingold (2014) points out, we have limited control in accurately directing our gaze and limited awareness of our actual gaze position. So whilst the calibration procedure will necessarily assume that the participant is looking at the centre of the target, this may not necessarily be the case. Another potential source of error during calibration, as in the worked example above, is that the experimenter presses the accept button (or the tracking software automatically accepts the fixation) before the participant had an opportunity to make a secondary saccade that would have taken their eye closer to the target. In other words, even if the participant was actually capable of looking at the exact centre of the target and keeping gaze fixed in exactly that location, the calibration still requires the operator (or eye tracking software) to make some decision as to whether this is in fact the case (in the absence of a calibration model that provides any information on accuracy), and this decision itself can be erroneous.

So, armed with this knowledge, what can be done to improve the chances of getting an accurate calibration model and maximising the accuracy of your data? Perhaps the most important, and often neglected factor under the experimenter's control is participant motivation. It is necessary to ensure that the participant is alert and trying their best to fixate the centre of each target. The experimenter can assist them in their endeavours by providing them with calibration points with clearly defined—but most importantly, it will be required to tell them what to do. In other words, it is important to explain to participants the importance of the calibration procedure, and that they are supposed to look closely at the centre of each calibration point in turn. Obviously such motivation cannot be expected of children or infants—so other strategies are required. These are covered in the following section. Some eye tracking systems allow the experimenter to choose between "manual calibration" (in which the experimenter accepts each point in turn) and "automatic calibration" (in which the eye tracking software determines when to accept the fixation). It is usually preferable to use the manual approach, particularly with un-practiced participants.

Most modern eye tracking systems provide the user with some kind of feedback as to the robustness of the calibration model. Users need to understand what this feedback is telling them, and use it to ensure that their calibration model is as good as it can be—given the limitations of both the human oculomotor system and their tracker. Corners are often weak points in calibrations—often because of poor set-up. If one or more corners are consistently "out" then consider moving the monitor further back, or adjusting the monitor or participant height.

Calibrating gaze positions of infants and young children

The key to successful infant eye tracking is to ensure that the screen you are using to present your stimuli is by far the most interesting thing in the infant or child's immediate environment. Visual stimuli should be bright/high contrast, and if possible accompanied by a sound. Children and particularly infants will rapidly tire of the same stimulus, so a good strategy is to ensure that you have a sufficient supply of attractive and noisy "attention grabbers" which can be employed whenever their attention wanes. It can help to have a cartoon running on the screen during camera set up if your system allows this.

Calibration targets should be chosen with particular care. Whilst large and ideally animated targets are best at attracting infant's attention, if they are too large, or contain no obvious focal point, the calibration process can be compromised as the infant may look at the bottom left of the target at one position and the top right of the target at another position. One useful approach is to use "looming/shrinking" targets—these are animations which start off small and then expand to attract the infants' attention, and then rapidly shrink back to the calibration location—infants' eyes will tend to follow the object as its shrinks, and the trick is to accept the calibration at the point when the target is at its smallest.

When eye tracking infants it may be preferable to limit the number of calibration points—a 3 or 5 point calibration may suffice. This may have a slight cost in terms of spatial accuracy over some parts of the screen—but many researchers feel that some data is better than no data.

In summary, ascertaining the best possible data quality is an obvious requirement in the recording of eye movements if researchers are to fully exploit the great potential of eye tracking technology in the study of psychological processes. Section 8.4 was meant to provide some practical guidelines in achieving this important goal.

8.5 Suggested Readings

Wade, N. J., & Tatler, B. W. (2005). The moving tablet of the eye: the origins of modern eye movement research. Oxford: Oxford University Press.

– *A fascinating history of eye tracking research.*

Reingold EM. (2014) Eye Tracking Research and Technology: Towards Objective Measurement of Data Quality. Vis Cogn. 22(3):635-652.

– *A very interesting and thorough discussion of the issues surrounding data quality.*

8.6 Questions Students Should Be Able to Answer

What techniques are currently used to track eye movements?
What are the main advantages of video-based eye tracking?
Why do video-based eye trackers often track the corneal reflection as well as the pupil?
What is the “trackable range” of an eye tracker?
What does “end to end delay” refer to?
How do you decide which of the available eye tracking systems is suitable for your planned study?
How can data quality be measured?
How can an eye tracker be precise but not accurate?
What steps can the experimenter take to ensure the best data quality?

References

Bour, L. J., van Gisbergen, J. A., Bruijns, J., & Ottes, F. P. (1984). The double magnetic induction method for measuring eye movement–results in monkey and man. *IEEE Transactions on Biomedical Engineering, 31*(5), 419–427.

Caldara, R., & Miellet, S. (2011). iMap: a novel method for statistical fixation mapping of eye movement data. *Behavior Research Methods, 43*(3), 864–878.

Delabarre, E. B. (1898). A method of recording eye-movements. *American Journal of Psychology, 9*(4), 572–574.

Deubel, H., & Bridgeman, B. (1995). Fourth Purkinje image signals reveal eye-lens deviations and retinal image distortions during saccades. *Vision Research, 35*(4), 529–538.

Diefendorf, A. R., & Dodge, R. (1908). An experimental study of the ocular reactions of the insane from photographic records. *Brain, 31,* 451–489.

Dodge and Cline. (1901). The angle velocity of eye movements. *Psychological Review, 8,* 145–157.

Engbert, R., & Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. *Vision Research, 43*(9), 1035–1045.

Frens, M. A., & van der Geest, J. N. (2002). Scleral search coils influence saccade dynamics. Journal of Neurophysiol. *2*, 692–698.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: University Press.

Hooge, I., Nyström, M., Cornelissen, T., & Holmqvist, K. (2015). The art of braking: Post saccadic oscillations in the eye tracker signal decrease with increasing saccade size. *Vision Research, 112,* 55–67.

Huey, E. B. (1898). Preliminary experiments in the physiology and psychology of reading. *American Journal of Psychology, 9*(4), 575–886.

Kimmel, D.L., Mammo, D., Newsome, W.T. (2012). Tracking the eye non-invasively: Simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. Frontiers in behavioral neuroscience *14;6*, 49.

Mowrer, O. R., Ruch, R. C., & Miller, N. E. (1936). The corneoretinal potential difference as the basis of thegalvanometric method of recording eye movements. *American Journal of Physiology, 114,* 423.

Nyström, M., Andersson, R., Magnusson, M., Pansell, T., & Hooge, I. (2015a). The influence of crystalline lens accommodation on post-saccadic oscillations in pupil-based eye trackers. *Vision Research, 107,* 1–14.

Nyström, M., Hooge, I., & Holmqvist, K. (2015b). Post-saccadic oscillations in eye movement data recorded with pupil-based eye trackers reflect motion of the pupil inside the iris. *Vision Research, 92,* 59–66.

Reingold, E. M. (2014). Eye tracking research and technology: Towards objective measurement of data quality. *Visual cognition, 22*(3), 635–652.

Wade, N. J., & Tatler, B. W. (2005). *The moving tablet of the eye: The origins of modern eye movement research*. Oxford: Oxford University Press.

Wildenmann, U., & Schaeffel, F. (2013). Variations of pupil centration and their effects on video eye tracking. *Ophthalmic and Physiological Optics, 33*(6), 634–641.

Yarbus, A. L. (1967). *Eye movements and vision*. Springer.

Chapter 9
Advanced Statistical Methods for Eye Movement Analysis and Modelling: A Gentle Introduction

Giuseppe Boccignone

Contents

Abstract In this Chapter we consider eye movements and, in particular, the resulting sequence of gaze shifts to be the observable outcome of a stochastic process. Crucially, we show that, under such assumption, a wide variety of tools become available for analyses and modelling beyond conventional statistical methods. Such tools encompass random walk analyses and more complex techniques borrowed from the Pattern Recognition and Machine Learning fields. After a brief, though critical, probabilistic tour of current computational models of eye movements and

---

G. Boccignone (✉)
Department of Computer Science, Universitá di Milano, Milano, Italy
e-mail: giuseppe.boccignone@unimi.it
URL: http://boccignone.di.unimi.it

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_9

visual attention, we lay down the basis for gaze shift pattern analysis. To this end, the concepts of Markov Processes, the Wiener process and related random walks within the Gaussian framework of the Central Limit Theorem will be introduced. Then, we will deliberately violate fundamental assumptions of the Central Limit Theorem to elicit a larger perspective, rooted in statistical physics, for analysing and modelling eye movements in terms of anomalous, non-Gaussian, random walks and modern foraging theory. Eventually, by resorting to Statistical Machine Learning techniques, we discuss how the analyses of movement patterns can develop into the inference of hidden patterns of the mind: inferring the observer's task, assessing cognitive impairments, classifying expertise.

9.1 Introduction

Consider Fig. 9.1: it shows typical *scan paths* (in this case a succession of saccades and fixations) produced by two human observers on a natural image: circular spots and lines joining spots graphically represent fixations and gaze shifts between subsequent fixations, respectively.

When looking at scan paths, the first question arising is: How can we characterise the shape and the statistical properties of such trajectories? Answering this question entails a *data analysis* issue. The second question is: What factors determine the shape and the statistical properties? and it relates to the *modelling* issue.

From a mere research practice standpoint these two issues need not be related (yet, from a more general theoretical standpoint such attitude is at least debatable). A great deal of research can be conducted by performing an eye tracking experiment based on a specific paradigm, and then analysing data by running standard statistical tools (e.g., ANOVA) on scan path "features" such as fixation frequency, mean fixation time, mean saccadic amplitudes, scan path length, etc. The "data-driven" attitude can be preserved even in the case where standard tools are abandoned in favour of more complex techniques borrowed from the Pattern Recognition and Machine Learning fields; for instance, in the endeavour of inferring or classifying the observer's mental task or the expertise behind his gaze shifts (e.g., Boccignone, Ferraro, Crespi, Robino, & de'Sperati, 2014; Henderson, Shinkareva, Wang, Luke, & Olejarczyk, 2013).

In the same vein, it is possible to set up a gaze shift model and successively assess its performance against eye tracking data in terms of classic statistical analyses. For instance, one might set up a probabilistic dynamic model of gaze shifting; then "synthetic" shifts can be generated from the model-based simulation. The distribution of their features can so be compared against the feature distribution of human gaze shifts - on the same stimuli - by exploiting a suitable goodness-of-fit test (e.g., Boccignone & Ferraro, 2014; Liberati et al., 2017).

Clearly, the program of following the data lies at the heart of scientific methodology. When trying to understand a complex process in nature, the empirical evidence is essential. Hypotheses must be compared with the actual data, but the empirical evidence itself may have limitations; that is, it may not be sufficiently large or accu-

Fig. 9.1 Different scan paths on a pair of images eye-tracked from different human observers. Left, free viewing of a natural scene; right, natural scene embedding a face. The area of yellow disks marking fixations between saccades is proportional to fixation time (images and eye tracking data from the Fixations in FAces dataset)

rate either to confirm or rule out hypotheses, models, explanations, or assumptions, even when the most sophisticated analytical tools are used.

For eye movement patterns, this issue may be in some cases particularly delicate. Such patterns are, in some sense, a summary of all the motor and perceptual activities in which the observer has been involved during data collection. As sketched in Fig. 9.2, from a functional standpoint, there are several interacting action/perception loops that drive eye movements. These factors act on different levels of representation and processing: salience, for instance, is a typical bottom-up process, while plans are typical top-down processes (Schütz, Braun, & Gegenfurtner, 2011).

In principle, all such activities should be taken into account when analysing and modelling actual eye movements in visual attention behaviour. Clearly, this is a mind-blowing endeavour.

This raises the question of what is a computational model and how it can support more advanced analyses of experimental data. In this Chapter we discuss a minimal phenomenological model.

At the most general level, the aim of a computational model of visual attention is to answer the question *Where to Look Next?* by providing:

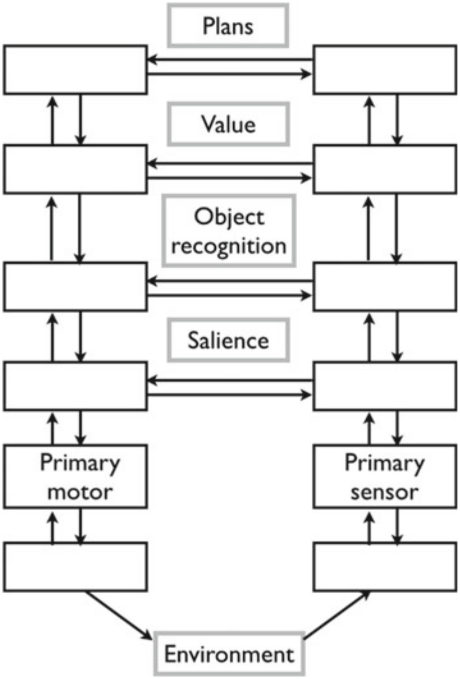


Fig. 9.2 Framework for the control of eye movements. There are several interacting layers of control that influence target selection: the scheme highlights, top to bottom, the contributions of plans, value, object recognition and salience to target selection. The left hand route summarises the motor components, the right hand one, the perceptual components. Figure modified after (Schütz et al., 2011)

1. at the *computational theory level* (in the sense of Marr, 1982; defining the input/output computation at time t), an account of the mapping from visual data of a complex natural scene, say $\mathcal{D}$ (raw image data, or more usefully, features), to a sequence of gaze locations $\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots$, under a given task $\mathbf{T}$, namely

$$\mathcal{D} \underset{\mathbf{T}}{\mapsto} \{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}, \tag{9.1}$$

   where the sequence $\{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}$ can be used to define a scan path (as illustrated in Fig. 9.3);
2. at the *algorithmic level*, Marr (1982), a procedure that simulates such mapping (we will not specifically address here the third level of neural realisation Marr, 1982).

Under this conceptualisation, when considering for instance the input $\mathcal{D}$ in the form of a static scene (a picture), either the raw time series $\{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}$ or fixation duration and saccade (length and direction) are the only two observable behaviours of the underlying control mechanism. When, $\mathcal{D}$ is a dynamic or time varying scene (e.g. a video), then pursuit needs to be taken also into account. Thus, it is convenient to adopt the generic terms of gaze shifts (either pursuit or saccades) and gaze shift amplitudes. Fixation duration and shift amplitude vary greatly during visual scanning of the scene. As previously discussed, such variation reflects

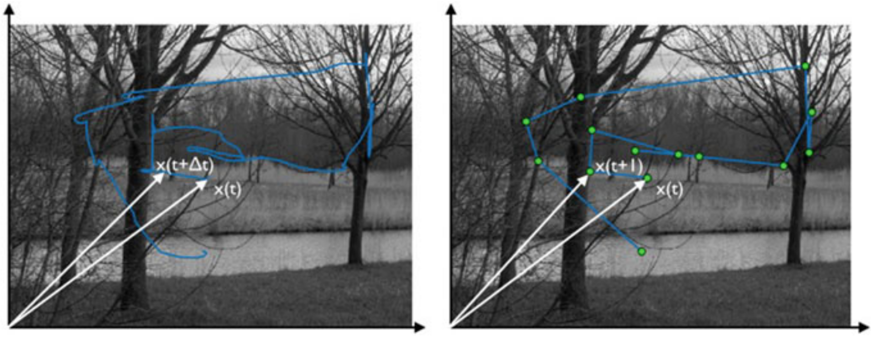


Fig. 9.3 Scan path of an eye-tracked human observer rendered as a temporal sequence of gaze position represented by time-varying location vectors $\mathbf{x}(t)$. The left image shows the continuous raw-data trajectory; the right image, the discretized sequence of fixations and saccades. Images and data are from the Doves dataset (Van Der Linde, Rajashekar, Bovik, & Cormack, 2009), which is freely available on the Web

moment-to-moment changes in the visual input, processes occurring at different levels of representation, the state of the oculomotor system and stochastic variability in neuromotor force pulses.

We can summarize this state of affairs by stating that fixation duration and the time series $\{\mathbf{x}_F(1), \mathbf{x}_F(2), \dots\}$ (or equivalently, gaze shift lengths and directions) are random variables (RVs) that are generated by an underlying random process. In other terms, the sequence $\{\mathbf{x}_F(1), \mathbf{x}_F(2), \dots\}$ is the realisation of a stochastic process, and the goal of a computational theory is to develop a mathematical model that describes statistical properties of eye movements as closely as possible.

Is this minimalist approach to computational modelling of gaze shifts a reasonable one? The answer can be positive if "systematic tendencies" between fixation durations, gaze shift amplitudes and directions of successive eye movements exist and such sequential dependencies can be captured by the stochastic process model. Systematic tendencies in oculomotor behaviour can be thought of as regularities that are common across all instances of, and manipulations to, behavioural tasks. In that case useful information about how the observers will move their eyes can be found.

Indeed, such systematic tendencies or "biases" in the manner in which we explore scenes with our eyes are well known in the literature. One example is provided in Fig. 9.4 showing the amplitude distribution of saccades and microsaccades that typically exhibit a positively skewed, long-tailed shape. Other paradigmatic examples of systematic tendencies in scene viewing are (Tatler & Vincent, 2008, 2009): initiating saccades in the horizontal and vertical directions more frequently than in oblique directions; small amplitude saccades tending to be followed by long amplitude ones and vice versa.

Such biases may arise from a number of sources. Tatler and Vincent (2009) have suggested the following: biomechanical factors, saccade flight time and landing accuracy, uncertainty, distribution of objects of interest in the environment, task parameters.

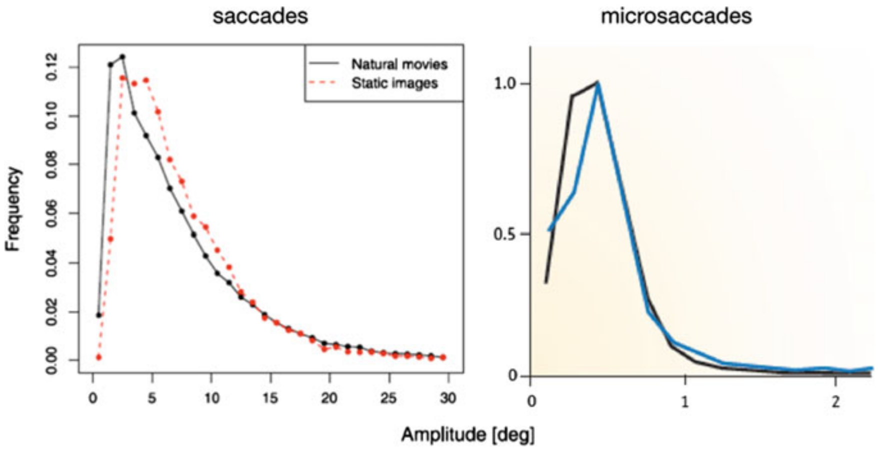


Fig. 9.4 Amplitudes distribution of saccades (on natural movies and static images, left) and microsaccades (right, recordings via video - black - and search coil - blue). In both cases amplitudes follow a positively skewed, long-tailed distribution. Figure modified after (Dorr, Martinetz, Gegenfurtner, & Barth, 2010; Martinez-Conde, Otero-Millan, & Macknik, 2013)

Understanding biases in how we move the eyes can provide powerful new insights into the decision about where to look in complex scenes. In a remarkable study (Tatler & Vincent, 2009), Tatler and Vincent have shown that a model based solely on these biases and therefore blind to current visual information can outperform salience-based approaches (in particular, they compared against the well known model proposed by Itti, Koch, & Niebur (1998), Walther and Koch (2006) - see Tom Foulsham's Chapter in this book for an introduction, and the following Sect. 9.3 for a probabilistic framing of saliency models).

Summing up, the adoption of an approach based on stochastic processes bring about significant advantages. First, the analysis and modelling of eye movements can benefit of all the "tools" that have been developed in the field of stochastic processes and time series. For example, the approach opens the possibility of treating visual exploration strategies in terms of *random walks*, e.g., (Carpenter & Williams, 1995; Engbert, 2006; Engbert, Mergenthaler, Sinn, & Pikovsky, 2011). Indeed, this kind of conceptual shift happened to the modern developments of econophysics (Mantegna et al., 2000) and finance (Paul & Baschnagel, 2013). Further, by following this path, visual exploration can be reframed in terms of *foraging* strategies an intriguing perspective that has recently gained currency (Boccignone & Ferraro, 2004, 2014; Cain, Vul, Clark, & Mitroff, 2012; Clavelli, Karatzas, Lladós, Ferraro, & Boccignone, 2014; Wolfe, 2013). Eventually, by embracing the stochastic perspective leads to the possibility of exploiting all the results so far achieved in the "hot" field of Statistical Machine Learning.

Thus, in this Chapter, we pursue the following learning objectives

1. Casting eye movement analysis and modelling in probabilistic terms (Sect. 9.3);
2. Understanding the essential concepts of stochastic process, such as Markov processes, and microscopic/macroscopic levels of description (Sects. 9.4 and 9.5);

3. Setting the basics of random walk analyses and modelling of eye movements either within the scope of the Central Limit Theorem or beyond, towards anomalous walks and diffusions (Sect. 9.6);
4. Moving from the analyses of scan path patterns to the inference of mental patterns by introducing the basic tools of modern probabilistic Machine learning (Sect. 9.7).

As to the eye movements concepts exploited in the modelling review of Sect. 9.3, it is worth referring to the related Chapters of this book.

For all the topics covered hereafter we assume a basic calculus level or at least a familiarity with the concepts of differentiation and integration. Box 1 provides a brief introductory note. However, find an A-level text book with some diagrams if you have not seen this before. Similarly, we surmise reader's conversance with elementary notions of probability and statistics.

Box 1: Interlude: differential and integral calculus with no pain

Differential calculus deals with the concept of **rate of change**. The rate of change of a function $f(x)$ is defined as the ratio of the change in f to the change in x. Consider Fig. 9.5 showing a plot of f as a function of x. There are intervals during which f increases and other intervals where f decreases. We can quantify the ups and downs of the changes in the values of f by estimating the slope, i.e., the change in the variable f over a given interval Δx, say between x_1 and x_2. Denote the interval or average slope by

$$\frac{\Delta f}{\Delta x} = \frac{f(x_2) - f(x_2)}{x_2 - x_1} = \frac{f(x + \Delta x) - f(x)}{\Delta x} = \frac{\text{rise}}{\text{run}},$$

with $\Delta x = x_2 - x_1$. What happens as the interval Δx becomes smaller and smaller and approaches zero, formally, $\Delta x \to 0$?

In that case the interval or average rate of change shrinks to the *instantaneous rate of change*. This is exactly what is computed by the **derivative** of f with respect to x:

$$\frac{df}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.$$

If you prefer thinking in a geometric way, the derivative at a point x provides the slope of the tangent of the curve at x.

As an example, we calculate the derivative of the function $f(x) = x^2$. First, write the term $f(x + \Delta x)$:

$$f(x + \Delta x) = (x + \Delta x)^2 = x^2 + 2x\Delta x + \Delta x^2$$

Then, subtract $f(x)$ and divide by Δx:

$$\frac{f(x+\Delta x)-f(x)}{\Delta x}=\frac{x^2+2x\Delta x+\Delta x^2-x^2}{\Delta x}=2x+\Delta x$$

Now in the limit $\Delta x \rightarrow 0$ we shrink Δx to zero, i.e.,

$$\lim_{\Delta x\rightarrow 0} 2x+\Delta x=2x.$$

Eventually,

$$\frac{d(x^2)}{dx}=2x.$$

If differential calculus has to do with rates of change, **integral calculus** deals with sums of many tiny incremental quantities. For instance, consider a continuous function f such as the one plotted in Fig. 9.6 and the following sum

$$\sum_{i=1}^{n} f(x_i)\Delta x = f(x_1)\Delta x + f(x_2)\Delta x + \cdots + f(x_n)\Delta x.$$

Here the uppercase greek letter $\sum$ indicates a sum of successive values defined by i and where $\Delta x=\frac{b-a}{n}$ and $x_i=a+i\Delta x$. Note that the term

$$f(x_i)\Delta x = \text{height}\times\text{width}=\delta A_i$$

computes the area δA_i of the i-th rectangle (see Fig. 9.6). Thus, the (Riemann) sum written above approximates the area defined by the continuous function f within the left and right limits a and b, as a the sum of tiny rectangles covering the area under f. The sum transforms into the **integral**

$$\int_a^b f(x)dx=\lim_{\Delta x\rightarrow 0}\sum_{i=1} f(x_i)\Delta x$$

when Δx shrinks to 0 (i.e. in the limit $\Delta x \rightarrow 0$) and the number n of intervals grows very large ($n \rightarrow \infty$).

There is a deep connection between integration and differentiation, which is stated by the **fundamental theorem of calculus**: the processes of integration and differentiation are reciprocal, namely, the derivative of an integral is the original integrand.

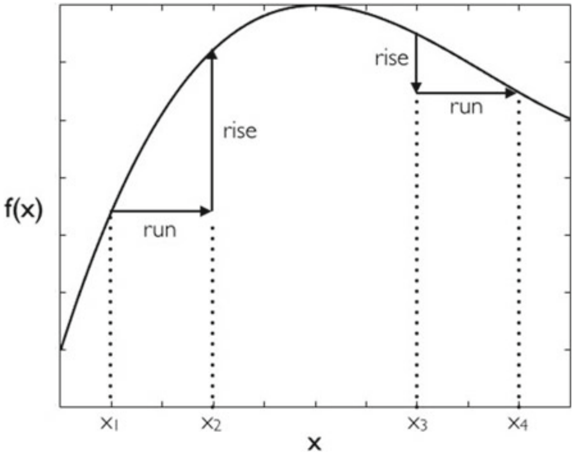


Fig. 9.5 A plot of the values of a function f as a function of x, showing a region of negative rate of change or slope (between x_1 and x_2) and a region of positive change (between x_3 and x_4)

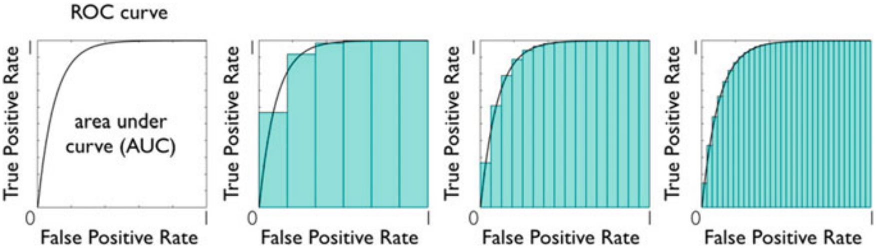


Fig. 9.6 An illustration of the integral concept by using Receiver Operating Characteristic (ROC) analysis, one of the methods to evaluate salience map algorithms. Continuous saliency maps are processed as a binary classifier applied on every pixel: the image pixels of the ground truth, as well as those of the prediction, are classified as fixated (or salient) or as not fixated (not salient). A simple threshold operation is used for this purpose. The ROC curve is exploited to display the classification result for a varying threshold used, each threshold value originating a number of False Positives and True Positives. An ROC curve is shown in the leftmost graph, and it has been obtained by plotting the False Positive Rate (FPR) as a function of the True Positive Rate (TPR). The ROC area, or the **area under curve** (AUC), provides a measure indicating the overall performance of the classification. The second graph shows the approximate calculus of the AUC as the (Riemann) sum of approximating rectangles as discussed in Box 1. Third and fourth graphs demonstrate how the computed AUC becomes more and more precise for increasing number of rectangles ($n \rightarrow \infty$) and diminishing rectangle widths ($\Delta x \rightarrow 0$). In such limit the sum $\sum$ becomes the integral $\int$

9.2 Historical Annotations

Stochastic modelling has a long and wide history encompassing different fields. The notion of stochastic trajectories possibly goes back to the scientific poem "*De Rerum Natura*" ("On the Nature of Things", circa 58 BC) by Titus Lucretius Carus:

> All things keep on in everlasting motion, / Out of the infinite come the particles, / Speeding above, below, in endless dance.

Yet, it is towards the end of the nineteenth century that a major breakthrough occurred. As Gardiner put it (Gardiner, 2009):

> Theoretical science up to the end of the nineteenth century can be viewed as the study of solutions of differential equations and the modelling of natural phenomena by deterministic solutions of these differential equations. It was at that time commonly thought that if all initial data could only be collected, one would be able to predict the future with certainty.

Quantum theory, on the one hand, and the concept of chaos (a simple differential equation, due to any error in the initial conditions that is rapidly magnified, can give rise to essentially unpredictable behaviour) on the other, have undermined such a Laplacian conception. However, even without dealing with quantum and chaotic phenomena, there are limits to deterministic predictability. Indeed, the rationale behind this Chapter is that of "limited predictability" (Gardiner, 2009) mostly arising when *fluctuating* phenomena are taken into account. As a matter of fact, stochastic processes are much closer to observations than deterministic descriptions in modern science and everyday life. Indeed, it is the existence of fluctuations that calls out for a statistical account. Statistics had already been used by Maxwell and Boltzmann in their gas theories. But it is Einstein's explanation (Einstein, 1905) of the nature of Brownian motion (after the Scottish botanist Robert Brown who observed under microscope, in 1827, the random highly erratic motion of small pollen grains suspended in water), which can be regarded as the beginning of stochastic modelling of natural phenomena.[1] Indeed, Einstein's elegant paper is worth a look, even by the non specialist, since containing all the basic concepts which will make up the subject matter of this Chapter: the Markov assumption, the Chapman-Kolmogorov equation, the random or stochastic differential equation for a particle path, the diffusion equation describing the behaviour of an ensemble of particles, and so forth. Since then, research in the field has quickly progressed. For an historically and technically detailed account the reader might refer to Nelson's "*Dynamical Theories of Brownian Motion*",[2] (Nelson, 1967).

To make a long story short, Einstein's seminal paper has provided inspiration for subsequent works, in particular that by Langevin (1908) who, relying upon the analysis of a single particle random trajectory, achieved a different derivation of Einstein's results. Langevin's equation was the first example of the stochastic differential equation, namely a differential equation with a random term and whose solution is, in some sense, a random function. Langevin initiated a train of thought that, in 1930, culminated in the work by Uhlenbeck and Ornstein (1930), representing a truly dynamical theory of Brownian motion. Although the approach of Langevin was improved and expanded upon by Ornstein and Uhlenbeck, some more fundamental problems remained, markedly related to the differentiability and integrability

---

[1] Actually, the first who noted the Brownian motion was the Dutch physician, Jan Ingen-Housz in 1794, in the Austrian court of Empress Maria Theresa. He observed that finely powdered charcoal floating on an alcohol surface executed a highly random motion.

[2] Freely available at https://web.math.princeton.edu/~nelson/books/bmotion.pdf.

of a stochastic process. The major contribution to the mathematical theory of Brownian motion has been brought by Wiener (1930), who proved that the trajectories of a Brownian process are continuous almost everywhere but are not differentiable anywhere. These problems were addressed by Doob (who came to probability from complex analysis) in his famous paper of 1942 (Doob, 1942). Doob aimed at applying the methods and results of modern probability theory to the analysis of the Ornstein-Uhlenbeck distribution. His efforts, together with those of Itô, Markov, Kac, Feller, Bernstein, Lévy, Kolmogorov, Stratonovich and others lead the *theory of random processes* to become an important branch of mathematics. A nice historical account of stochastic processes from 1950 to the present is provided by Meyer (2009).

Along with theoretical achievements, many more cases of random phenomena materialised in science and engineering. The major developments came in the 1950s and 1960s through the analysis of electrical circuits and radio wave propagation. A great deal of highly irregular electrical signals were given the collective name of "noise": uncontrollable fluctuations in electric circuits (e.g, thermal noise, namely the distribution of voltages and currents in a network due to thermal electron agitation); scattering of electromagnetic waves caused by inhomogeneities in the refractive index of the atmosphere (fading). The fundamental theorem of Nyquist is based on the principle of thermal equilibrium, the same used by Einstein and Langevin (2002). Beyond those early days the theory of random processes has become a central topic in the basic training of engineers, and lays the foundation for spectral representation and estimation of signals in noise, filtering and prediction, entropy and information theory (Papoulis & Pillai, 2002). Clearly, electrical noise, albeit very important, is far from a unique case. As other examples, one might consider the pressure, temperature and velocity vector of a fluid particle in a turbulent flow. A substantial overlap between the topics of neuroscience and stochastic systems has been acknowledged (Laing & Lord, 2010).

Interestingly, beyond the realm of the natural sciences and engineering, analyses of the random character of stock market prices started to gain currency in the 1950s. Osborne "rediscovered" the Brownian motion of stock markets in 1959 (Osborne, 1959). Computer simulations of "microscopic" interacting-agent models of financial markets have been performed as early as 1964 (Stigler, 1964). Brownian motion became an important model for the financial market: Paul Samuelson for his contributions on such topic received the 1970 Nobel Prize in Economics; in 1973, Merton and Scholes, in collaboration with the late Fischer Black, have used the geometric Brownian motion to construct a theory for determining the price of stock options; their achievements were also honoured by the Nobel Prize in 1997. The theory represents a milestone in the development of mathematical finance and today's daily capital market practice. Interestingly enough, such body of work builds on the early dissertation of a PhD student of Henri Poincaré, named Louis Bachelier. In 1900 Bachelier defended his thesis entitled "*Théorie de la Spéculation*" at the Sorbonne University of Paris (Bachelier, 1900). He had developed, five years before Einstein, the theory of the random walk as a suitable probabilistic description for price fluctuations on the financial market. Unfortunately, such humble application was not acknowledged by the scientific community at that time; hence, Bachelier's work fell into complete oblivion until the early 1940s (when Itô used it as a motivation to introduce his calculus). Osborne himself made no mention of it (Osborne, 1959). For an historical account of the role played by stochastic processes in the development of mathematical finance theory, see Jarrow and Protter (2004).

A relevant step, which is of major importance for this Chapter, was taken by Richardson's work (1926). Twenty years later Einstein and Langevin works, he presented empirical data related to the "superdiffusion" of an admixture cloud in a turbulent atmosphere being in contradiction with the normal diffusion. Such anomalous diffusion can be explained as a deviation of the real statistics of fluctuations from the Gaussian law. Subsequently, anomalous diffusion in the form of Lévy flights has been discovered in many other physical, chemical, biological, and financial systems (Lévy flights, as we will see, are stochastic processes characterised by the occurrence of extremely long jumps, so that their trajectories are not continuous anymore). The first studies on the subject were those of Kolmogorov (1941) on the scale invariance of turbulence in the 1940s. This topic was later on addressed by many physicists and mathematicians, particularly by Mandelbrot (the father of fractal mathematics). In the 1960s he applied it not only to the phenomenon of turbulence but also to the behaviour of financial markets (Mandelbrot, 1963). As Mandelbrot lucidly summarised (Mandelbrot, 1963):

> Despite the fundamental importance of Bachelier's process, which has come to be called "Brownian motion," it is now obvious that it does not account for the abundant data accumulated since 1900 by empirical economists, simply because the empirical distributions of price changes are usually too "picked" to be relative to samples from Gaussian populations.

An historical but rather technical perspective on anomalous diffusion and Lévy flights is detailed by Dubkov, Spagnolo, and Uchaikin (2008); a more affordable presentation is outlined by Schinckus (2013). Today, these kinds of processes are important to characterise a multitude of systems (e.g., microfluidics, nanoscale devices, genetic circuits that underlie cellular behaviour). Lévy flights are recognised to underlie many aspects of human dynamics and behaviour (Baronchelli & Radicchi, 2013). Eye movement processes make no exception, as we will see.

Nowadays, the effective application of the theory of random processes and, more generally, of probabilistic models in the real world is gaining pace. The advent of cheap computing power and the developments in Markov chain Monte Carlo simulation produced a revolution within the field of Bayesian statistics around the beginning of the 1990s. This allowed a true "model liberation". Computational tools and the latest developments in approximate inference, including both deterministic and stochastic approximations, facilitate coping with complex stochastic process based models that previously we could only dream of dealing with (Insua, Ruggeri, & Wiper, 2012). We are witnessing an impressive cross-fertilisation between random process theory and the more recently established areas of Statistical Machine Learning and Pattern Recognition, where the commonalities in models and techniques emerge, with Probabilistic Graphical Models playing an important role in guiding (Barber, Cemgil, and Chiappa, 2011).

9.3 A Probabilistic Tour of Current Computational Models of Eye Movements and Visual Attention (with Some Criticism)

Many models in psychology and in the computational vision literature have investigated limited aspects of the problem of eye movements in visual attention behaviour (see Box 2, for a quick review). And, up to now, no model has really succeeded in predicting the sequence of fixations of a human observer looking at an arbitrary scene (Frintrop et al., 2010).

Box 2: Visual attention models: a brief critical review

In the field of psychology, there exists a wide variety of theories and models on visual attention [see, e.g., the review by Heinke and Humphreys (2005)]. Among the most influential for computational attention systems: the well known Treisman's **Feature Integration Theory** (FIT) (Treisman, 1998; Treisman & Gelade, 1980); Wolfe's **Guided Search Model** (Wolfe, 1994), aiming at explaining and predicting the results of visual search experiments; Desimone and Duncan's **Biased Competition Model** (BCM, Desimone & Duncan, 1995), Rensink's triadic architecture (2000), the **Koch and Ullman's model** (Koch & Ullman, 1985), and Tsotsos' **Selective Tuning** (ST) model [?].

Other psychophysical models have addressed attention modelling in a more formal framework. One notable example is Bundensen's **Theory of Visual Attention** (TVA, Bundesen, 1998), further developed by Logan into the **CODE theory** of visual attention (CTVA, Logan, 1996). Also, theoretical approaches to visual search have been devised by exploiting Signal Detection Theory (Palmer, Verghese, & Pavel, 2000).

At a different level of explanation, other proposals have been conceived in terms of connectionist models, such as **MORSEL** (Multiple Object Recognition and attentional SELection, Mozer, 1987), **SLAM** (SeLective Attention Model) (Phaf, Van der Heijden, & Hudson, 1990), **SERR** (SEarch via Recursive Rejection) (Humphreys & Muller, 1993), and **SAIM** [Selective Attention for Identification Model by Heinke and Humphreys (2003)] subsequently refined in the Visual Search SAIM (VS-SAIM) (Heinke & Backhaus, 2011).

To a large extent, the psychological literature was conceived and fed on simple stimuli, nevertheless the key role that the above models continue to play in understanding attentive behaviour should not be overlooked. For example, many current computational approaches, by and large, build upon the bottom-up salience based model by Itti et al. (1998), which in turn is the computational counterpart of Koch and Ullman and Treisman's FIT models. The seminal work of Torralba (2003), draws on an important component of Rensink's triadic architecture (2000), in that it considers contextual information such as gist - the abstract meaning of a scene, e.g., a city scene, etc. - and layout - the spatial arrangement of the objects in a scene. More recently, Wischnewski, Belardinelli, Schneider, and Steil (2010) have presented a computational model that integrates Bundensen's TVA (Bundesen, 1998).

However, in the last three decades, psychological models have been adapted and extended in many respects, within the **computational vision** field where the goal is to deal with attention models and systems that are able to cope with natural complex scenes rather than simple stimuli and synthetical images [e.g., see Frintrop, Rome, & Christensen, 2010 and the most recent review by Borji and Itti (2013)]. The adoption of complex stimuli has sustained a new brand of computational theories, though this theoretical development is still at an early stage: up to this date, nobody has really succeeded in predicting the sequence of fixations of a human observer looking at an arbitrary scene (Frintrop et al., 2010). This is not surprising given the complexity of the problem. One might think that issues of generalisation from simple to complex contexts are nothing more than a minor theoretical inconvenience; but, indeed, the generalisation from simple to complex patterns might not be straightforward. As it has been noted in the case of attentive search, a model that exploits handpicked features may fail utterly when dealing with realistic objects or scenes (Zelinsky, 2008).

Current approaches within this field suffer from a number of limitations: they mostly rely on a low-level salience based representation of the visual input, they seldom take into account the task's role, and eventually they overlook the eye guidance problem, in particular the actual generation of gaze-shifts (but see Tatler, Hayhoe, Land, and Ballard (2011) for a lucid critical review of current methods). We will discuss such limitations in some detail in Sect. 9.3.

The issue of devising a computational model of eye guidance as related to visual attention - i.e. answering the question *Where to Look Next?* in a formal way - can be set in a probabilistic Bayesian framework (see Box 3 for a brief introduction). Tatler and Vincent (2009) have re-phrased this question in terms of Bayes' rule:

$$\overbrace{P(\mathbf{x} \mid \mathcal{D})}^{\text{posterior prob. of gaze shift}} = \frac{\overbrace{P(\mathcal{D} \mid \mathbf{x})}^{\text{data likelihood under the shift}}}{P(\mathcal{D})} \overbrace{P(\mathbf{x})}^{\text{gaze shift prior}}, \qquad (9.2)$$

Box 3: Dangerous relationships: A rendezvous with Bayesian Probabilities

We assume the readers to be already familiar with the elementary notions (say, undergrad level) of probability and **random variables** (RVs). Thus, a warning. Sometimes we talk about probabilities of events that are "out there" in the world. The face of a flipped coin is one such event. But sometimes we talk about probabilities of events that are just possible **beliefs** "inside the head." Our belief about the fairness of a coin is an example of such an event. Clearly, it might be bizarre to say that we randomly sample from our beliefs, like we sample from a sack of coins. To cope with such embarrassing situation, we shall use probabilities to express our information and beliefs about unknown quantities. $P(A)$ denotes the probability that the event A is true. But event A could stand for logical expressions such as “there is a red car in the bottom of the scene” or “an elephant will enter the pub”. In this perspective, probability is used to quantify our uncertainty about something; hence, it is fundamentally related to information rather than repeated trials. Stated more clearly: we are adopting the Bayesian interpretation of probability in this Chapter.

Fortunately, the basic rules of probability theory are the same, no matter which interpretation is adopted (but not that smooth, if we truly addressed inferential statistics). For what follows, we just need to refresh a few.

Let X and Y be RVs, that is numbers associate to events. For example, the quantitative outcome of a survey, experiment or study is a RV; the amplitudes of saccades or the fixation duration times recorded in a trial are RVs. In Bayesian inference a RV (either discrete or continuous) is defined as an unknown numerical quantity about which we make probability statements. Call $P(X, Y)$ their joint probability. The **conditional probability** of X given Y is:

$$P(X \mid Y) \equiv \frac{P(X, Y)}{P(Y)} \text{ if } P(Y) \neq 0. \tag{9.3}$$

In Bayesian probability we always deal with conditional probabilities: at least we condition on the assumptions or set of **hypotheses** $\mathcal{H}$ on which the probabilities are based. In data modelling and Machine Learning, the following holds (Mackay, 2002):

You cannot do inference without making assumptions

Then, the rules below will be useful:

Product rule (or chain rule)

$$P(X, Y \mid \mathcal{H}) = P(X \mid Y, \mathcal{H})P(Y \mid \mathcal{H}) \tag{9.4}$$

Sum rule (marginalisation)

$$P(Y \mid \mathcal{H}) = \sum_X P(X, Y \mid \mathcal{H}) \text{ (discrete RVs)} \tag{9.5}$$

$$P(Y \mid \mathcal{H}) = \int_X P(X, Y \mid \mathcal{H}) dX \text{ (continuous RVs)} \tag{9.6}$$

Bayes’ rule (see Fig. 9.7 for a simple example)

$$P(X \mid Y, \mathcal{H}) = \frac{P(Y \mid X, \mathcal{H})P(X \mid \mathcal{H})}{P(Y \mid \mathcal{H})} \leftrightarrow \text{posterior} = \frac{\text{likelihood} \times \text{prior}}{\text{evidence}} \tag{9.7}$$

To avoid burying the reader under notations, we have used $P(\cdot)$ to denote both the probability of a discrete outcome (**probability mass function**, PMF) and the probability of a continuous outcome (**probability density function**, pdf). We let context make things clear. Also, we may adopt the form $X = x$ for a specific choice of value (or outcome) of the RV X. Briefer notation will sometimes be used: for example, $P(X = x)$ may be written as $P(x)$. A bold $\mathbf{X}$ might denote a set of RVs or a random vector/matrix.

The "bible" of the Bayesian approach is the treatise of Jaynes (2003). A succinct introduction with an eye to **inference and learning** problems can be found in Chap. 2 of the beautiful book by MacKay (2002), which is also available for free online, http://www.inference.phy.cam.ac.uk/mackay/itila/.

where $\mathbf{x} = \mathbf{x}_F(t) - \mathbf{x}_F(t-1)$ is the random vector representing the gaze shift (in Tatler & Vincent, 2009, saccades), and $\mathcal{D}$ generically stands for the input data. As Tatler and Vincent put it, "The beauty of this approach is that the data could come from a variety of data sources such as simple feature cues, derivations such as Itti's definition of salience, object-or other high-level sources".

In Eq. 9.2, the first term on the right hand side accounts for the likelihood of particular visual data (e.g., features, such as edges or colors) occurring at a gaze shift target location normalized by $P(\mathcal{D})$ the pdf of these visual data occurring in the environment. As we will see in brief, this first term bears a close resemblance to approaches previously employed to evaluate the possible involvement of visual features in eye guidance.

Most interesting, and related to issues raised in the introductory Section, is the Bayesian prior $P(\mathbf{x})$, i.e., the probability of shifting the gaze to a location *irrespective of the visual information* at that location. Indeed, this term will encapsulate any systematic tendencies in the manner in which we explore scenes with our eyes. The striking result obtained by Tatler and Vincent (2009) is that if we learn $P(\mathbf{x})$ from actual observer's behaviour, then we can **sample** gaze shifts (cfr. Box 4), i.e.,

$$\mathbf{x}(t) \sim P(\mathbf{x}), \quad t = 1, 2, \ldots \tag{9.8}$$

so to obtain scan paths that, blind to visual information, out-perform feature-based accounts of eye guidance (Tatler & Vincent, 2009): 0.648 area under the receiver operator curve (AUC, which has been illustrated in Fig. 9.6) as opposed to 0.593 for edge information (namely, an orientation map computed from edge maps constructed over a range of spatial scales, by convolving the image with four oriented odd-phase Gabor filters) and 0.565 for salience information as derived through the Itti et al. model (1998).[3]

---

[3] More precisely, they used the latest version of Itti's salience algorithm, available at http://www.saliencytoolbox.net (Walther & Koch, 2006), with defaults parameters setting. One may argue that since then the methods of saliency computation have developed and improved significantly so far. However, if one compares the predictive power results obtained by salience maps obtained within the very complex computational framework of deep networks, e.g., via the PDP system (with

Box 4: When God plays dice: the art (and magic) of sampling

Eye movements can be considered a natural form of sampling. Another example of actual physical sampling is tossing a coin, as in the example illustrated in Fig. 9.8, or throwing dice. Nevertheless, we can (and need to) **simulate** sampling that occurs in nature (and thus the underlying process). Indeed, for both computational modelling and analysis we assume of being capable of the fundamental operation of generating a sample $\mathbf{X} = \mathbf{x}$ from a probability distribution $P(\mathbf{X})$. We denote the sampling action via the $\sim$ symbol:

$$\mathbf{x} \sim P(\mathbf{X}). \tag{9.9}$$

For instance, tossing a coin like we did in the example of Fig. 9.7 can be simulated by sampling $\mathbf{x}$ from a **Bernoulli distribution**, $\mathbf{x} \sim Bern(\mathbf{X}; \theta)$, where θ is the parameter standing for the coin bias ($\theta = \frac{1}{2} = 0.5$ denotes a fair coin).

Surprisingly, to simulate nature, we need a minimal capability: that of generating realisations of RVs uniformly distributed on the interval [0, 1]. In practical terms, we just need a programming language or a toolbox in which a `rand()` function is available implementing the $\mathbf{u} \sim Uniform(0, 1)$ operation. Indeed, given the RVs $\mathbf{u}$, we can generate the realisations of any other RV with appropriate “transformations” of $\mathbf{u}$.

There is a wide variety of “transformations” for generating samples, from simple ones (e.g. inverse transform sampling and rejection sampling) to more sophisticated, like those relying on **Markov Chain Monte Carlo** methods (e.g., **Gibbs sampling** and **Metropolis sampling**). Again, MacKay’s book (2002) provides a very clear introduction to the art of random sampling.

You can qualitatively assess the results of your computational sampling procedure using sample **histograms**. Recall from your basic statistic courses that an histogram is an empirical estimate of the probability distribution of a continuous variable. It is obtained by “binning” the range of values – that is, by dividing the entire range of values into a series of small intervals –, and then counting how many values fall into each interval. Intuitively, if we look at the empirical distribution of the set of samples $\{\mathbf{x}(t)\}_{t=1}^{T}$ obtained for a large number T of sampling trials $\mathbf{x}(t) \sim P(\mathbf{X})$, $t = 1, 2, \ldots, T$, we expect the shape of the histogram to approximate the originating theoretical density. Examples are provided in Fig. 9.8 where 1000 samples have been generated experimenting with the **Uniform distribution**, the **Gaussian distribution** and the **Cauchy distribution**, respectively.

---

fine tuning) (Jetley, Murray, & Vig, 2016), against a simple central bias map (saliency inversely proportional to distance from centre, blind to image information), one can read an AUC performance of 0.875 against 0.780 on the large VOCA dataset (Jetley et al., 2016) (on the same dataset, the Itti et al. model achieves 0.533 AUC). Note that a central bias map can be computed in a few Matlab lines (Mathe & Sminchisescu, 2013).

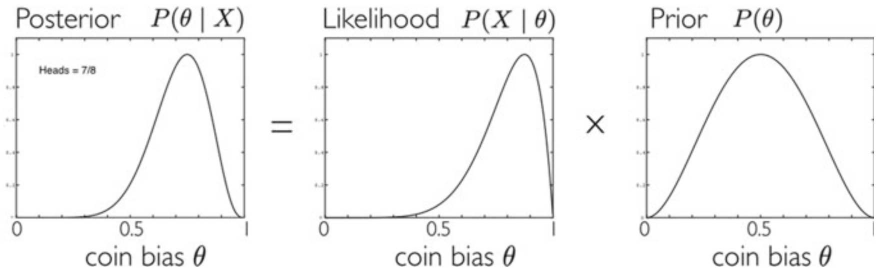


Fig. 9.7 An illustration of the use of the Bayes' rule for inferring the bias of a coin on the basis of coin tossing results. The prior probability $P(\theta)$ for the coin bias θ captures the assumption that the coin is likely to be a fair one (the pdf is "peaked" on $\theta = 0.5$). However, 7 heads occur after 8 tosses. Such experimental result is captured by the shape of the likelihood $P(X \mid \theta)$ strongly biased to the right. Bayes' rule computes the posterior pdf $P(\theta \mid X)$ by "updating" the initial prior through the "observed" likelihood (the evidence term is not shown in the figure and it has been treated as a normalisation factor to constrain probabilities between 0 and 1)

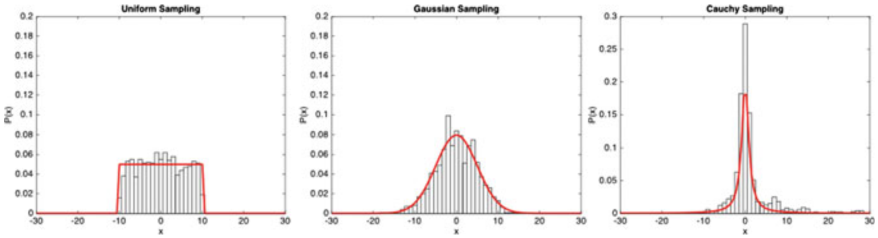


Fig. 9.8 From left to right, the empirical distributions (histograms) for $n = 1000$ samples drawn from Uniform, Gaussian and and Cauchy pdfs. Uniform and Gaussian sampling have been performed via the Matlab functions `rand()` and `randn()`, respectively; samples from the Cauchy pdf have been generated resorting to Metropolis sampling. Each histogram is overlaid with the generating theoretical density depicted as a continuous red curve

Learning is basically obtained by empirically collecting through eye tracking the observer's behaviour on an image data set (formally, the joint pdf $P(\mathbf{x}, \mathcal{D})$ and then factoring out the informative content of the specific images, briefly, via marginalisation,i.e., $P(\mathbf{x}) = \sum_{\mathcal{D}} P(\mathbf{x}, \mathcal{D})$.

Note that the apparent simplicity of the prior term $P(\mathbf{x})$ hides a number of subtleties. For instance, Tatler and Vincent expand the random vector $\mathbf{x}$ in terms of its components, amplitude l and direction θ. Thus, $P(\mathbf{x}) = P(l, \theta)$. This simple statement paves the way to different options. First easy option: such RVs are marginally independent, thus, $P(l, \theta) = P(l)P(\theta)$. In this case, gaze guidance, solely relying on biases, could be simulated by expanding Eq. (9.8) via independent sampling of both components, i.e. at each time t, $l(t) \sim P(l(t)), \theta(t) \sim P(\theta(t))$. Alternative option: conjecture some kind of dependency, e.g. amplitude on direction, so that $P(l, \theta) = P(l \mid \theta)P(\theta)$. In this case, the gaze shift sampling procedure would turn into the sequence $\widehat{\theta}(t) \sim P(\theta(t)), l(t) \sim P(l(t) \mid \widehat{\theta}(t))$. Further: assume that there is some persistence in the direction of the shift. This gives rise to a stochastic process in which subsequent directions are correlated, i.e., $\theta(t) \sim P(\theta(t) \mid \theta(t-1))$, and so on.

To summarise, by simply taking into account the prior $P(\mathbf{x})$, a richness of possible behaviours and analyses are brought into the game. To further explore this perspective, we recommend the thorough and up-to-date review by Le Meur and Coutrot (2016).

Unfortunately, most computational accounts of eye movements and visual attention have overlooked this issue. We noticed before, by inspecting Eq. (9.2) that the term $\frac{P(\mathcal{D}|\mathbf{x})}{P(\mathcal{D})}$ bears a close resemblance to many approaches proposed in the literature. This is an optimistic view. Most of the approaches actually discard the dynamics of gaze shifts, say $\mathbf{x}_F(t) \rightarrow \mathbf{x}_F(t+1)$, implicitly captured through the shift vector $\mathbf{x}(t)$. In practice, most models are more likely to be described by a simplified version of Eq. (9.2):

$$\overbrace{P(\mathbf{x}_F \mid \mathcal{D})}^{\text{posterior prob. of gazing at}} = \frac{\overbrace{P(\mathcal{D} \mid \mathbf{x}_F)}^{\text{data likelihood under gaze at}}}{P(\mathcal{D})} \overbrace{P(\mathbf{x}_F)}^{\text{prior prob. of gazing at}} , \tag{9.10}$$

By careful inspection, it can be noted that the posterior $P(\mathbf{x}_F \mid \mathcal{D})$ answers the query "What is the probability of *fixating* location $\mathbf{x}_F$ given visual data $\mathcal{D}$?". Further, the prior $P(\mathbf{x}_F)$ accounts for the probability of *fixating* location $\mathbf{x}_F$ irrespective of the visual information at that location. The difference between Eqs. 9.2 and 9.10 is subtle. But, as a matter of fact, Eq. 9.10 bears no dynamics. In probabilistic terms we may re-phrase this result as the outcome of an assumption of independence:

$$P(\mathbf{x}) = P(\mathbf{x}_F(t) - \mathbf{x}_F(t-1)) \simeq P(\mathbf{x}_F(t) \mid \mathbf{x}_F(t-1)) = P(\mathbf{x}_F(t)).$$

To make things even clearer, let us explicitly substitute $\mathbf{x}_F$ with a RV $\mathbf{L}$ denoting locations in the scene, and $\mathcal{D}$ with RV $\mathbf{F}$ denoting features (whatever they may be); then Eq. 9.10 boils down to the following

$$\overbrace{P(\mathbf{L} \mid \mathbf{F})}^{\text{posterior prob. of selecting location L}} = \frac{\overbrace{P(\mathbf{F} \mid \mathbf{L})}^{\text{feature likelihood under location L}}}{P(\mathbf{F})} \overbrace{P(\mathbf{L})}^{\text{prior prob. of location L}} \tag{9.11}$$

The feature-based **Probabilistic Graphical Model** underlying this query (see Box 5 for a brief PGM overview) is a very simple one and is represented on the left of Fig. 9.10. As it can be seen, it is a subgraph of the object-based model PGM (Fig. 9.10, centre), which is the one previously discussed in Box 5.

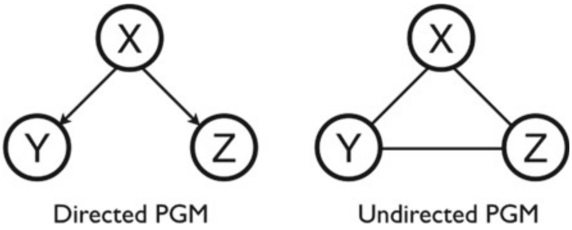


Fig. 9.9 Probabilistic Graphical Models: a directed PGM (left, a.k.a. Bayesian Network) and an undirected PGM (right, a.k.a. Markov Random Field). Nodes represents RVs and arcs express probabilistic relationships between RVs

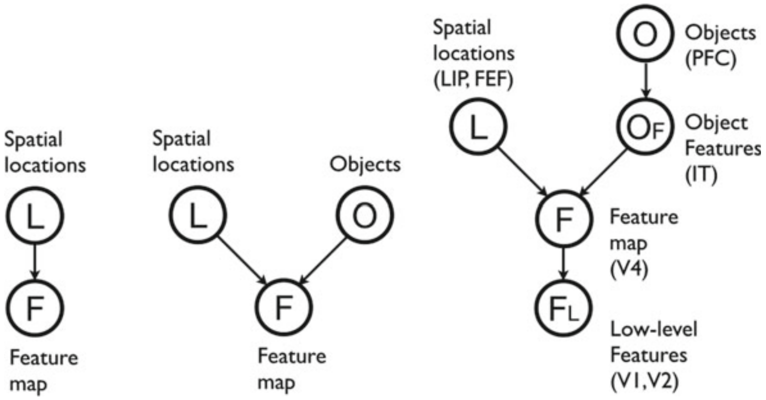


Fig. 9.10 PGMs of increasing level of representational complexity (left to right) that can account for most models proposed in the computational vision field. Left: feature-based model. Centre: object-based model. Right: the Bayesian model by Chikkerur, Serre, Tan, and Poggio (2010), which extends the object-based model in the centre and maps the resulting PGM structure to brain areas underpinning visual attention: early visual areas V1 and V2, V4, lateral intraparietal (LIP), frontal eye fields (FEF), inferotemporal (IT), prefrontal cortex (PFC)

Box 5: Probabilistic Graphical Models (PGM)

A PGM (Koller & Friedman, 2009) is a graph-based representation (see Fig. 9.9) where **nodes** (also called vertices) are connected by **arcs** (or **edges**). In a PGM, each node represents a RV (or group of RVs), and the arcs express probabilistic relationships between these variables. Graphs where arcs are arrows are **directed** PGM, a generalisation of **Bayesian Networks** (BN), well known in the Artificial Intelligence community. The other major class of PGMs are **undirected** PGM (Fig. 9.9, right), in which the links have no directional significance, but are suitable to express soft constraints between RVs. The latter are also known as **Markov Random Fields** (MRF), largely exploited in Computer Vision.

We shall focus on directed PGM representations where arrows represent conditional dependencies (Fig. 9.9, left). For instance the arrow $X \rightarrow Y$ encodes the probabilistic dependency of RV Y on X quantified through the conditional probability $P(Y \mid X)$. Note that arrows do not generally represent causal relations, though in some circumstances it could be the case. We will mainly exploit PGMs as a descriptive tool: indeed, (1) they provide a simple way to visualise the structure of a probabilistic model and can be used to design and motivate new models; (2) they offer insights into the properties of the model, including conditional independence properties, which can be obtained by inspection of the graph.

PGMs capture the way in which the joint distribution over all of the RVs can be decomposed into a product of factors each depending only on a subset of the variables. Assume that we want to describe a simple object-based attention model (namely, the one presented at the centre of Fig. 9.10), so to deal with: (i) objects (e.g., red triangles vs. blue squares), (ii) their possible locations, and (iii) the visual features sensed from the observed scene. Such "world' can be described by the joint pdf $P(Objects, Location, Features)$ which we denote, more formally, through the RVs O, L, F: $P(Objects, Location, Features) \equiv P(O, L, F)$. Recall that - via the product rule - the joint pdf could be factorised in a combinatorial variety of ways, all equivalent and admissible:

$$\begin{aligned} P(O, L, F) &= P(O \mid L, F)P(L \mid F)P(F) \\ &= P(L \mid O, F)P(O \mid F)P(F) \\ &= P(F \mid O, L)P(O \mid L)P(L) \\ &= \cdots \end{aligned} \tag{9.12}$$

The third factorisation is, actually, the meaningful one: the likelihood of observing certain features (e.g, color) in the visual scene depends on *what* kind of objects are present and on *where* they are located; thus, the factor $P(F \mid O, L)$ makes sense. $P(L)$ represents the prior probability of choosing certain locations within the scene (e.g., it could code the *center bias* effect Tatler, 2007). Eventually, the $P(O \mid L)$ factor might code the prior probability of certain kinds of objects (e.g., we may live in a world where red triangles are more frequent than blue squares). As to $P(O \mid L)$ we can assume that the object location and object identity are independent, formally, $P(O \mid L) = P(O)$, finally leading to

$$P(O, L, F) = P(F \mid O, L)P(O)P(L). \tag{9.13}$$

This factorisation is exactly that captured by the structure of the directed PGM presented at the centre of Fig. 9.10. Indeed, the graph renders the most suitable factorisation of the unconstrained joint pdf, under the assumptions and the constraints we are adopting to build our model. We can "query" the PGM for making any kind of **probabilistic inference**. For instance, we could ask what is the posterior probability $P(O, L \mid F)$ of observing certain objects at certain locations given the observed features. By using the definition of conditional probability and Eq. 9.13:

$$P(O, L \mid F) = \frac{P(O, L, F)}{P(F)} = \frac{P(F \mid O, L)P(O)P(L)}{P(F)} \tag{9.14}$$

Complex computations for **inference and learning** in sophisticated probabilistic models can be expressed in terms of graph-based algorithms. PGMs are a formidable tool to such end, and nowadays are widely adopted in modern probabilistic **Machine Learning** and **Pattern Recognition**. An affordable introduction can be found in Bishop (2006). The PGM "bible" is the textbook by Koller and Friedman (2009).

Surprisingly enough, this simple model is sufficiently powerful to account for a large number of visual attention models that have been proposed in computational vision. This can be easily appreciated by setting $P(\mathbf{F} \mid \mathbf{L}) = const.$, $P(\mathbf{L}) = const.$ so that Eq. 9.11 reduces to

$$\overbrace{P(\mathbf{L} \mid \mathbf{F})}^{\text{posterior prob. of selecting location L}} \quad \propto \quad \overbrace{\frac{1}{P(\mathbf{F})}}^{\text{salience at location L}} \quad . \tag{9.15}$$

Equation 9.15 tells that the probability of fixating a spatial location $\mathbf{L} = (x, y)$ is higher when "unlikely" features ($\frac{1}{P(\mathbf{F})}$) occur at that location. In a natural scene, it is typically the case of high contrast regions (with respect to either luminance, color, texture or motion) and clearly relates to entropy and information theory concepts (Boccignone, Ferraro, & Caelli, 2001). This is nothing but the most prominent salience-based model in the literature proposed by Itti et al. (1998), which Eq. 9.15 only re-phrases in probabilistic terms.

A thorough reading of the recent review by Borji and Itti (2013) is sufficient to gain the understanding that a great deal computational models so far proposed are more or less variations of this leitmotif (experimenting with different features, different weights for combining them, etc.). The weakness of such a pure bottom-up approach has been largely discussed (see, e.g. Einhäuser, Spain, & Perona, 2008; Foulsham & Underwood, 2008; Tatler et al., 2011). Indeed, the effect of early saliency on attention is likely to be a correlational effect rather than an actual causal one (Foulsham & Underwood, 2008; Schütz et al., 2011), though salience may be still more predictive than chance while preparing for a memory test as discussed by Foulsham and Underwood (Foulsham & Underwood, 2008).

Thus, recent efforts have tried to go beyond this simple stage with the aim of climbing the representational hierarchy shown in Fig. 9.2. This entails a first shift from Eq. 9.15 (based on an oversimplified representation) back to Eq. 9.11. Torralba (2003) have shown that using prior knowledge on the typical spatial location of the search target, as well as contextual information (the "gist" of a scene) to modulate early saliency improves its fixation prediction.

Next shift is exploiting object knowledge for top-down "tuning" early salience; thus, moving to the PGM representation at the centre of Fig. 9.10. Indeed, objects and their semantic value have been deemed as fundamental for visual attention and eye guidance (e.g., Bundesen, 1998; Heinke & Backhaus, 2011; Mozer, 1987; Rensink, 2000, but see Scholl, 2001 for a review). For instance, when dealing with faces within the scene, a face detection step can provide a reliable cue to complement early conspicuity maps, as it has been shown by Cerf, Harel, Einhäuser, and Koch (2008), deCroon, Postma, and van den Herik (2011), Marat, Rahman, Pellerin, Guyader, and Houzet (2013), or a useful prior for Bayesian integration with low level cues (Boccignone et al., 2008). This is indeed an important issue since faces may drive attention in a direct fashion (Cerf, Frady, & Koch, 2009). The same holds for text regions (Cerf et al., 2008; Clavelli et al. 2014) Other notable exceptions are those provided by Rao, Zelinsky, and Hayhoe (2002), Sun, Fisher, Wang, and Gomes (2008), the Bayesian models discussed by Borji, Sihite, and Itti (2012) and Chikkerur et al. (2010). In particular the model by Chikkerur et al., which is shown at right of Fig. 9.10 is the most complete to the best of our knowledge (though it does not consider contextual scene information (Torralba, 2003), but the latter could be easily incorporated). Interestingly enough, the authors have the merit of making the effort of providing links between the structure of the PGM and the brain areas that could support computations.

Further, again in the effort of climbing the representational hierarchy (Fig. 9.2), attempts have been made for incorporating task and value information (see Clavelli et al., 2014; Schütz et al., 2011 for a brief review, and Tatler et al., 2011 for a discussion).

Now, a simple question arises: where have the eye movements gone?

To answer such question is useful to summarise the brief overview above. The common practice of computational approaches is to conceive the mapping (9.1), as a two step procedure:

1. obtain a suitable representation $\mathcal{R}$, i.e., $\mathcal{D} \underset{\mathbf{T}}{\mapsto} \mathcal{R}$;
2. use $\mathcal{R}$ to generate the scanpath, $\mathcal{R} \underset{\mathbf{T}}{\mapsto} \{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}$.

Computational modelling has been mainly concerned with the first step: deriving a representation $\mathcal{R}$ (either probabilistic or not). The second step, that is $\mathcal{R} \mapsto \{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}$, which actually brings in the question of *how* we look rather than *where*, is seldom taken into account.

In spite of the fact that the most cited work in the field, that by Itti et al. (1998), clearly addressed the *how* issue (gaze shifts as the result of a Winner-Take-All, WTA, sequential selection of most salient locations), most models simply overlook the eye movement problem. The computed representation $\mathcal{R}$ is usually evaluated in terms of its capacity for predicting the image regions that will be explored by covert and overt attentional shifts according to some evaluation measure (Borji & Itti, 2013). In other cases, if needed for practical purposes, e.g. for robotic applications, the problem of oculomotor action selection is solved by adopting some deterministic choice procedure. These usually rely on selecting the gaze position $\mathbf{x}$ as the argument that maximises a measure on the given representation $\mathcal{R}$ (in brief, see Walther & Koch, 2006 for using the $\arg\max_{\mathbf{x}} \mathcal{R}$ operation[4] and Boccignone and Ferraro (2014; Tatler et al., 2011, for an in-depth discussion).

Yet, another issue arises: the variability of visual scan paths. When looking at natural movies under a free-viewing or a general-purpose task, the relocation of gaze can be different among observers even though the same locations are taken

---

[4] $\arg\max_x f(x)$ is the mathematical shorthand for “find the value of the argument x that maximizes $f(\cdot)$”.

into account. In practice, there is a small probability that two observers will fixate exactly the same location at exactly the same time. Such variations in individual scan paths (as regards chosen fixations, spatial scanning order, and fixation duration) still hold when the scene contains semantically rich "objects" (e.g., faces, see Fig. 9.1). Variability is even exhibited by the same subject along different trials on equal stimuli. Further, the consistency in fixation locations between observers decreases with prolonged viewing Dorr et al. (2010). This effect is remarkable when free-viewing static images: consistency in fixation locations selected by observers decreases over the course of the first few fixations after stimulus onset (Tatler et al., 2011) and can become idiosyncratic.

Note that, the WTA scheme (Itti et al., 1998; Walther & Koch, 2006, or the selection of the proto-object with the highest attentional weight (Wischnewski et al., 2010) are deterministic procedures. Even when probabilistic frameworks are used to infer where to look next, the final decision is often taken via the maximum a posteriori (MAP) criterion,[5] which again is an arg max operation (e.g., Boccignone et al., 2008; Chernyak & Stark, 2001; Elazary & Itti, 2010; Najemnik & Geisler, 2005), or variants such as the robust mean (arithmetic mean with maximum value) over candidate positions (Begum, Karray, Mann, & Gosine, 2010). As a result, for a chosen visual data input $\mathcal{D}$ the mapping $\mathcal{R} \underset{\mathbf{T}}{\mapsto} \{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}$ will always generate the same scanpath across different trials.

There are few notable exceptions to this current state of affairs (see Boccignone & Ferraro, 2014 for a discussion). In Kimura, Pang, Takeuchi, Yamato, and Kashino (2008) simple eye-movement patterns, in the vein of Tatler and Vincent (2009), are straightforwardly incorporated as a prior of a dynamic Bayesian network to guide the sequence of eye focusing positions on videos. The model presented in Ho Phuoc, Guérin-Dugué, and Guyader (2009) embeds at least one parameter suitable to be tuned to obtain different saccade length distributions on static images, although statistics obtained by varying such parameter are still far from those of human data. The model by Keech and Resca (2010) mimics phenomenologically the observed eye movement trajectories and where randomness is captured through a Monte Carlo selection of a particular eye movement based on its probability; probabilistic modelling of eye movement data has been also discussed in Rutishauser and Koch (2007). However, both models address the specific task of conjunctive visual search and are limited to static scenes. Other exceptions are given, but in the very peculiar field of eye-movements in reading (Feng, 2006). Other works have addressed the variability issue in the framework of foraging random walks (Boccignone & Ferraro, 2004, 2013a, b, 2014; Clavelli et al. 2014; Napoletano, Boccignone, & Tisato, 2015).

What we need at least is to bring stochasticity back into the game. As Canosa put it (Canosa, 2009):

> Where we choose to look next at any given moment in time is not completely deterministic, yet neither is it completely random.

---

[5] Given a posterior distribution $P(X \mid Y)$ the MAP rule is just about choosing the argument $X = x$ for which $P(X \mid Y)$ reaches its maximum value (the arg max) ; thus, if $P(X \mid Y)$ is a Gaussian distribution, then the arg max corresponds to the mode, which for the Gaussian is also the mean value.

9.4 Stochastic Processes and Eye Movements

When we randomly sample a sequence $\{\mathbf{x}(t = 1), \mathbf{x}(t = 2), \mathbf{x}(t = 3), \ldots\}$ of gaze shifts from the pdf $P(\mathbf{x})$ (cfr., Eq. 9.8), we set up a stochastic process. For example, the *ensemble* of different scan paths on the same viewed image can be conceived as the record of a stochastic process (Fig. 9.11).

Stochastic processes are systems that evolve probabilistically in time or more precisely, systems in which a certain time-dependent random variable $\mathbf{X}(t)$ exists (as to notation, we may sometimes write $\mathbf{X}_t$ instead of $\mathbf{X}(t)$) The variable t usually denotes time and it can be integer or real valued: in the first case, $\mathbf{X}(t)$ is a discrete time stochastic process; in the second case, it is a continuous time stochastic process. We can observe realisations of the process, that is we can measure values

$$\mathbf{X}(t_1) = \mathbf{x}_1, \mathbf{X}(t_2) = \mathbf{X}_2, \mathbf{X}(t_3) = \mathbf{x}_3, \ldots,$$

at times $t_1, t_2, t_3, \ldots$. The set $\mathcal{S}$ whose elements are the values of the process is called **state space**.

Thus, we can conceive the stochastic process $\mathbf{X}(t)$ as an ensemble of paths as shown in Fig. 9.3 or, more simply, as illustrated in Fig. 9.12: here, for concreteness, we show four series of the raw x coordinates of different eye-tracked subjects gazing at picture shown in Fig. 9.3. Note that if we fix the time, e.g., $t = t_1$, then $\mathbf{X}(t_1)$ boils

Fig. 9.11 An ensemble of scan paths recorded from different observers while viewing the same image. For visualisation purposes, only five trajectories are shown, different colors coding individual trajectories. If such ensemble is considered to represent the outcome of a stochastic process, the fundamental question that should be answered is: What is the probability $P(\mathbf{x}, t)$ of gazing at location $\mathbf{x}$ at time t? Images and data are from the Doves dataset (Van Der Linde et al., 2009)

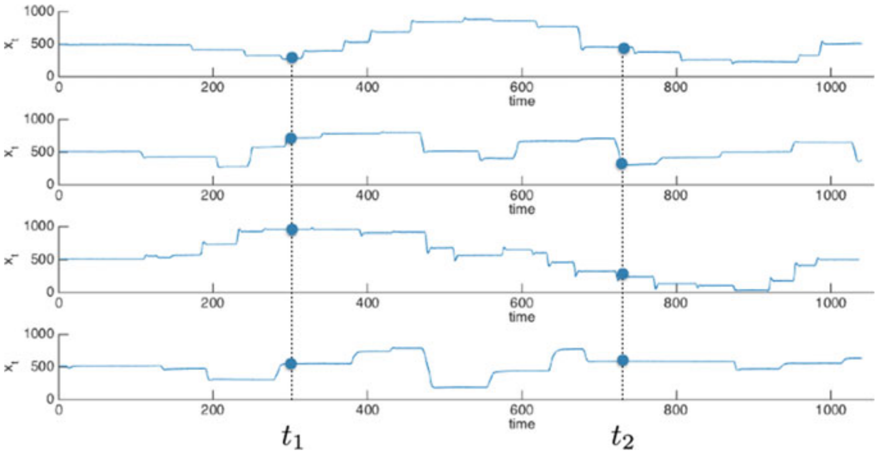


Fig. 9.12 An ensemble of paths representing a stochastic process. Each path represents the sequence in time of raw x coordinates from different scan paths recorded on the same picture (cfr. Fig. 9.11). We can conceive the trajectories (namely, time series) of such ensemble as realisations of a stochastic process

down to a RV (vertical values); the same holds if we choose one path $\mathbf{x}$ and we (horizontally) consider the set of values $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots,$ at times $t_1, t_2, t_3, \ldots$.

To sum up, a stochastic process can be regarded as either a family of realisations of a random variable in time, or as a family of random variables at fixed time. Interestingly enough, referring back to Sect. 9.2, notice that Einstein's point of view was to treat Brownian motion as a distribution of a random variable describing position, while Langevin took the point of view that Newton's law's of motion apply to an individual realisation.

In order to be more compact with notation, we will use Huang's abbreviation (Huang, 2001)

$$k \leftrightarrow \{\mathbf{x}_k, t_k\},$$

where, e.g., $P(1)$ succintly stands for $P(\mathbf{x}_1, t_1)$.

To describe the process completely we need to know the correlations in time, that is the hierarchy of pdfs (but see Box 6, for a discussion of correlation):

$$\begin{aligned} P(1) &: \text{ the 1 point pdf} \\ P(1,2) &: \text{ the 2 points pdf} \\ P(1,2,3) &: \text{ the 3 points pdf} \\ &\qquad\qquad\qquad \ldots \end{aligned} \tag{9.16}$$

up to the n point joint pdf. The n point joint pdf must imply all the lower k point pdfs, $k < n$:

$$P(1, \dots, k) = \int P(1, \dots, n) d\mathbf{x}_{k+1} d\mathbf{x}_{k+2} \dots d\mathbf{x}_n \tag{9.17}$$

where $P(1, \dots, n) d\mathbf{x}_{k+1} d\mathbf{x}_{k+2} \dots d\mathbf{x}_n$ stands for the joint probability of finding that $\mathbf{x}$ has a certain value

$$\begin{aligned} \mathbf{x}_{k+1} &< \mathbf{x} \le \mathbf{x}_{k+1} + d\mathbf{x}_{k+1} \text{ at time } t_{k+1} \\ \mathbf{x}_{k+2} &< \mathbf{x} \le \mathbf{x}_{k+2} + d\mathbf{x}_{k+2} \text{ at time } t_{k+2} \\ &\dots \end{aligned}$$

For instance, referring to Fig. 9.12, we can calculate the joint probability $P(1, 2)$ $d\mathbf{x}_1 d\mathbf{x}_2$ by following the vertical line at t_1 and t_2 and find the fraction of paths for which $\mathbf{x}(t_1) = \mathbf{x}_1$ within tolerance $d\mathbf{x}_1$ and $\mathbf{x}(t_2) = \mathbf{x}_2$ within tolerance $d\mathbf{x}_2$, respectively[6]

Summing up, the joint probability density function, written in full notation as

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \dots; \mathbf{x}_n, t_n),$$

is all we need to fully characterise the statistical properties of a stochastic process and to calculate the quantities of interest characterising the process (see Box 6).

Box 6: How to observe a stochastic process

Consider a series of time signals. The signal fluctuates up and down in a seemingly erratic way. The measurements that are in practice available at one time of a measurable quantity $x(t)$ are the mean and the variance. However, the latter do not tell a great deal about the underlying dynamics of what is happening. A fundamental question in time series analysis is: to what extent the value of a RV variable measured at one time can be predicted from knowledge of its value measured at some earlier time? Does the signal at t_0 influence what is measured at a later time $t_0 + t$? We are not interested in any specific time instant t_0 but rather in the typical (i.e., the statistical) properties of the fluctuating signal. The amount of dependence, or history in the signal can be characterised by the **autocorrelation** function

$$C_{xx}(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_0^T x(t) x(t + \tau) dt. \tag{9.18}$$

This is the time average of a two-time product over an arbitrary large time T, which is then allowed to become infinite. Put simply, is the integral of the product of the time series with the series simply displaced with respect to itself by an amount τ. An autocorrelated time series is predictable, probabilistically,

---

[6] This gives an intuitive insight into the notion of $P(1, 2)$ as a *density*.

because future values depend on current and past values. In practice, collected time series are of finite length, say N. Thus, the estimated autocorrelation function is best described as the **sample autocorrelation**

$$c_{xx}(\Delta) = \frac{1}{N} \sum_{n=0}^{N-|\Delta|-1} x(n)x(n+\Delta) \tag{9.19}$$

Measurements of $C_{xx}(\tau)$ are used to estimate the time-dependence of the changes in the joint probability distribution, where the **lag** is $\tau = t - t_0$. If there is no statistical correlation $C_{xx}(\tau) = 0$. The rate at which $C_{xx}(\tau)$ approaches 0 as τ approaches ∞ is a measure of the **memory** for the stochastic process, which can also be defined in terms of **correlation time**:

$$t_{corr} = \frac{1}{C_{xx}(0)} \int_0^{+\infty} C_{xx}(\tau)d\tau. \tag{9.20}$$

The autocorrelation function has been defined so far as a **time average** of a signal, but we may also consider the **ensemble average**, in which we repeat the same measurement many times, and compute averages, denoted by symbol $\langle\rangle$. Namely, the correlation function between $x(t)$ at two different times t_1 and t_2 is given by

$$\langle x(t_1), x(t_2)\rangle = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x_1 x_2 P(x_1, t_1; x_2, t_2) dx_1 dx_2. \tag{9.21}$$

For many systems the ensemble average is equal to the time average, $\langle x\rangle = \int_{-\infty}^{+\infty} x_1 P(x_1, t)dx_1 \approx \lim_{T\to\infty} \frac{1}{T} \int_0^T x(t)dt = \overline{x(t)}$. Such systems are termed **ergodic**. Ergodic ensembles for which the probability distributions are invariant under time translation and only depend on the relative times $t_2 - t_1$ are **stationary** processes. If we have a stationary process, it is reasonable to expect that average measurements could be constructed by taking values of the variable x at successive times, and averaging various functions of these. Correlation and memory properties of a stochastic process are typically investigated by analysing the autocorrelation function or the spectral density (**power spectrum**) $S(\omega)$, which describes how the power of a time series is distributed over the different frequencies. These two statistical properties are equivalent for stationary stochastic processes. In this case the *Wiener-Kintchine theorem* holds

$$S(\omega) = \frac{1}{2\pi} \int_{-\infty}^{-\infty} \exp(-i\omega\tau) C_{xx}(\tau)d\tau \tag{9.22}$$

$$C_{xx}(\tau) = \int_{-\infty}^{-\infty} \exp(i\omega\tau) S(\omega)d\omega \tag{9.23}$$

It means that one may either directly measure the autocorrelation function of a signal, or the spectrum, and convert back and forth, which by means of the Fast Fourier Transform (FFT) is relatively straightforward. The sample power spectral density function is computed via the FFT of c_{xx}, i.e. $s(\omega) = FFT(c_{xx}(\Delta))$, or viceversa by the inverse transform, $c_{xx}(\Delta = IFFT(s(\omega))$.

The **dynamics**, or *evolution* of a stochastic process can be represented through the specification of **transition probabilities**:

$$\begin{gathered} P(2 \mid 1) : \text{probability of finding } 2, \text{ when } 1 \text{ is given;} \\ P(3 \mid 1, 2) : \text{probability of finding } 3, \text{ when } 1 \text{ and } 2 \text{ are given;} \\ P(4 \mid 1, 2, 3) : \text{probability of finding } 4, \text{ when } 1, 2 \text{ and } 3 \text{ are given;} \\ \dots \end{gathered}$$

Transition probabilities for a stochastic process are nothing but the conditional probabilities suitable to predict the future values of $\mathbf{X}(t)$ (i.e., $\mathbf{x}_{k+1}, \mathbf{x}_{k+2}, \dots \mathbf{x}_{k+l}$, at $t_{k+1}, t_{k+2}, \dots t_{k+l}$), given the knowledge of the past ($\mathbf{x}_1, \mathbf{x}_2, \dots \mathbf{x}_k$, at $t_1, t_2, \dots t_k$). The conditional pdf explicitly defined in terms of the joint pdf can be written:

$$P(\overbrace{\mathbf{x}_{k+1}, t_{k+1}; \dots; \mathbf{x}_{k+l}, t_{k+l}}^{\text{future}} \mid \underbrace{\mathbf{x}_1, t_1; \dots; \mathbf{x}_k, t_k}_{\text{past}}) = \frac{P(\mathbf{x}_1, t_1; \dots; \mathbf{x}_{k+l}, t_{k+l})}{P(\mathbf{x}_1, t_1; \dots; \mathbf{x}_k, t_k)}. \tag{9.24}$$

assuming the time ordering $t_1 < t_2 < \dots < t_k < t_{k+1} < \dots < t_{k+l}$.

By using transition probabilities and the product rule, the following update equations can be written:

$$\begin{aligned} P(1, 2) &= P(2 \mid 1)P(1) \qquad (9.25) \\ P(1, 2, 3) &= P(3 \mid 1, 2)P(1, 2) \\ P(1, 2, 3, 4) &= P(4 \mid 1, 2, 3)P(1, 2, 3) \\ &\dots \end{aligned}$$

The transition probabilities must satisfy the normalisation condition $\int P(2 \mid 1)d\mathbf{x}_2 = 1$. Since $P(2) = \int P(1, 2)d\mathbf{x}_1$ and by using the update Eq. (9.25), the following evolution (integral) equation holds

$$P(2) = \int \overbrace{P(2 \mid 1)}^{\text{propagator}} P(1)d\mathbf{x}_1 \tag{9.26}$$

where $P(2 \mid 1)$ serves as the *evolution kernel* or *propagator* from state 1 to state 2, i.e., in full notation, from state $(\mathbf{x}_1, t_1)$ to state $(\mathbf{x}_2, t_2)$.

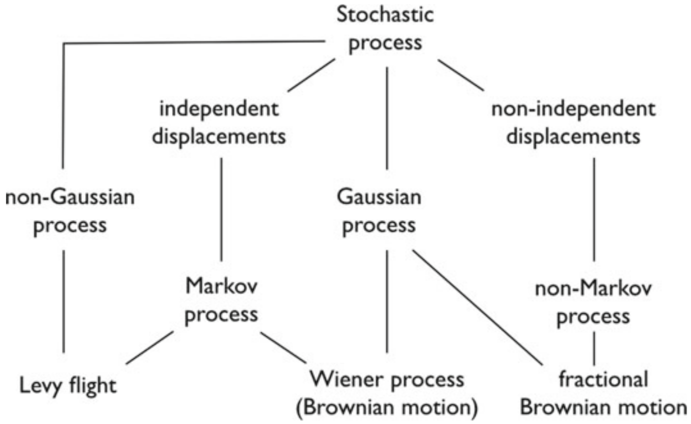


Fig. 9.13 A conceptual map of stochastic processes that are likely to play a role in eye movement modelling and analyses

A stochastic process whose joint pdf does not change when shifted in time is called a (strict sense) **stationary process**:

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \ldots; \mathbf{x}_n, t_n) = P(\mathbf{x}_1, t_1 + \tau; \mathbf{x}_2, t_2 + \tau; \ldots; \mathbf{x}_n, t_n + \tau) \qquad (9.27)$$

$\tau > 0$ being a time shift. Analysis of a stationary process is frequently much simpler than for a similar process that is time-dependent: varying t, all the random variables $\mathbf{X}_t$ have the same law; all the moments, if they exist, are constant in time; the distribution of $\mathbf{X}(t_1)$ and $\mathbf{X}(t_2)$ depends only on the difference $\tau = t_2 - t_1$ (time lag), i.e,

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2) = P(\mathbf{x}_1, \mathbf{x}_2; \tau).$$

A conceptual map of main kinds of stochastic processes that we will discuss in the remainder of this Chapter is presented in Fig. 9.13.

9.5 How to Leave the Past Behind: Markov Processes

The most simple kind of stochastic process is the Purely Random Process in which there are no correlations. From Eq. (9.25):

$$\begin{aligned} P(1,2) &= P(1)P(2) \qquad (9.28)\\ P(1,2,3) &= P(1)P(2)P(3)\\ P(1,2,3,4) &= P(1)P(2)P(3)P_1(3)\\ &\ldots \end{aligned}$$

One such process can be obtained for example by repeated coin tossing. The complete independence property can be written explicitly as:

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \ldots) = \prod_i P(\mathbf{x}_i, t_i), \tag{9.29}$$

the uppercase greek letter $\prod$ indicates a product of factors, e.g., for $i = 1, 2, 3$, $P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \mathbf{x}_3, t_3) = P(\mathbf{x}_1, t_1)P(\mathbf{x}_2, t_2)P(\mathbf{x}_3, t_3)$.

Equation 9.29 means that the value of $\mathbf{X}$ at time t is completely independent of its values in the past (or future). A special case occurs when the $P(\mathbf{x}_i, t_i)$ are independent of t, so that the same probability law governs the process at all times. Thus, a completely memoryless stochastic process is composed by a set of **independent and identically distributed** (**i.i.d**) RVs. Put simply, a series of i.i.d. RVs is a series of samples where individual samples are "independent" of each other and are generated from the same probability distribution ("identically distributed").

More realistically, we know that most processes in nature, present some correlations between consecutive values. For example, the direction of the following gaze shift is likely to be positively correlated with the direction of current gaze shift. A step towards a more realistic description consists then of assuming that the next value of each RV in the process depends explicitly on the current one (but not explicitly on any other previous to that). An intuitive example is the *simple random walk*, which is briefly discussed in Box 7, and can be modelled by the simple equation

$$x_t = x_{t-1} + \xi_t, \tag{9.30}$$

where the noise term ξ_t is sampled, at each step t, from a suitable distribution $P(\xi)$.

Box 7: Random walks

Random walks (RW) are a special kind of stochastic process and can be used, as we will see, to model the dynamics of many complex systems. A particle moving in a field, an animal foraging, and indeed the "wandering" eye can be conceived as examples of random walkers.

In general, RWs exhibit what is called serial correlation, conditional independence for fairly small values of correlation length t_{corr} (cfr., Box 6, Eq. 9.20), and a simple stochastic historical dependence. For instance, a simple additive 1−dimensional random walk has the form:

$$x_t = x_{t-1} + \xi_t, \text{ where } \xi_t \sim P(\xi) \tag{9.31}$$

In the above formulation, time t proceeds in discrete steps. ξ_t is a RV drawn i.i.d. from a distribution $P(\xi)$, called the *noise* or *fluctuation* distribution. Thus, the differences in sequential observations $x_t - x_{t-1} = \xi_t \sim P(\xi)$ are i.i.d. We have here **independent displacements**.

However, the observations themselves are not independent, since (9.31) encodes the generative process, or evolution law, $x_{t-1} \to x_t$ where x_t explicitly depends on x_{t-1}, but not on earlier $x_{t-2}, x_{t-3}, x_{t-4}, \ldots$. Thus Eq. (9.31) represents a Markov process.

Conventionally, fluctuations are Gaussian distributed with mean μ and variance σ^2, that is, $\xi \sim \mathcal{N}(\mu, \sigma^2)$, as this makes mathematical analysis considerably simpler. In this case by simply extending to two dimensions Eq. 9.31,

$$\begin{aligned} x_t &= x_{t-1} + \xi_{x,t} \\ y_t &= y_{t-1} + \xi_{y,t}, \end{aligned} \tag{9.32}$$

the simulation of a simple Brownian RW can be obtained (see Fig. 9.14).

However, any probability distribution, for instance, a Laplace (exponential tails) or double-Pareto distribution (power-law tails), also works.

Note that if we iterate Eq. (9.30) for a number of steps $t = 1, 2, 3, \ldots$ and collect the output sequence of the equation, that is $x_1, x_2, x_3, \ldots$, we obtain a single trajectory/path of the walk, which is one possible realisation of the underlying stochastic process. This corresponds to consider one horizontal slice of Fig. 9.12, that is a (discrete) **time series**. It is worth mentioning that there exist a vast literature on time series analysis, which can be exploited in neuroscience data analysis and more generally in other fields. Indeed, the term "time series" more generally refers to data that can be represented as a sequence. This includes for example financial data in which

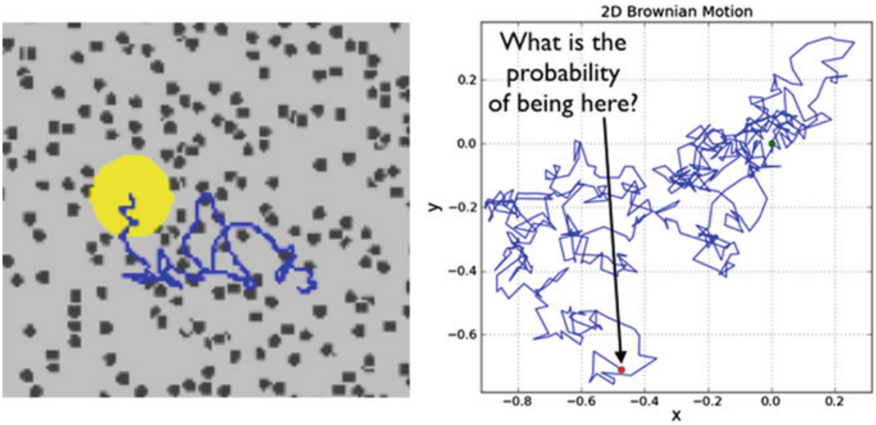


Fig. 9.14 Motion of a Brownian particle. Left: The physical mechanism of the displacement (in blue): the bigger particle performs a **Brownian motion** (Bm) as a result of the collisions with small particles (figure is not to scale). Right: Sample path of the Bm performed by the bigger particle. Here the fundamental question is: What is the probability of the particle being at location $\mathbf{x} = (x, y)$ at time t? (cfr. Box 9)

the sequence index indicates time, as in our case, but also genetic data (e.g. ACATGC . . .) where the sequence index has no temporal meaning. Some of the methods that have been developed in this research area might be useful for eye movement modelling and analysis. We have no place here to further discuss this issue, however Box 8 provides some "pointers" for the reader.

Box 8: As time goes by: time series analysis

The random walk model summarised by Eq. 9.30, can be seen as a special case of the model

$$x_t = \alpha_1 x_{t-1} + \xi_t, \tag{9.33}$$

with model parameter $\alpha_1 = 1$.

In **time series** analysis, under the assumption that ξ_t is sampled from a Gaussian distribution mean zero and variance σ^2, the model specified by Eq. 9.33 is known as an autoregressive or **AR model** of order 1, abbreviated to AR(1).

The AR(1) model, in turn is a special case of an autoregressive process of order p, denoted as AR(p):

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p} + \xi_t, \tag{9.34}$$

with model parameters $\alpha_p \neq 0$. Note that such model is a regression of x_t on past terms from the same series; hence the use of the term "autoregressive".

Considering again the RW of Eq. 9.30. One can substitute the term x_{t-1}, that by using the same equation can be calculated as $x_{t-1} = x_{t-2} + \xi_{t-1}$; thus,

$$x_t = x_{t-2} + \xi_{t-1} + \xi_t. \tag{9.35}$$

Continuing and substituting for x_{t-2}, followed by x_{t-3} and so on (a process known as "back substitution") gives

$$x_t = \xi_1 + \xi_1 + \cdots + \xi_{t-1} + \xi_t, \tag{9.36}$$

where x_t is written as the sum of the current noise term ξ_t and the past noise terms.

This result can be generalised by writing x_t as the linear combination of the current white noise term and the q most recent past noise terms

$$x_t = \xi_t + \beta_1 \xi_{t-1} + \beta_2 \xi_{t-2} + \cdots + \beta_q \xi_{t-q}. \tag{9.37}$$

This defines a moving average or **MA model** of order q, shortly MA(q).

Putting all together, we can write the general expression:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \cdots + \alpha_p x_{t-p} + \xi_t + \beta_1 \xi_{t-1} + \beta_2 \xi_{t-2} + \cdots + \beta_q \xi_{t-q}. \tag{9.38}$$

The time series is said to follow an autoregressive moving average or **ARMA model** of order (p, q), denoted ARMA(p, q).

By expanding on the former, a great deal of models can be conceived. Also, a variety of methods, algorithms, and related software, is at hand for estimating the model parameters from time series data. Cowpertwait and Metcalfe (2009) provide a thorough introduction with R language examples, for R fans. The book edited by Barber et al. (2011) offers a comprehensive picture of modern time series techniques, specifically those based on Bayesian probabilistic modelling. Time series modelling is a fast-growing trend in neuroscience data analysis, which is addressed in-depth by Ozaki (2012).

Going back to stochastic processes, if a process has no memory beyond the last transition then it is called a *Markov process* and the transition probability enjoys the property:

$$P(\mathbf{x}_n, t_n | \mathbf{x}_{n-1}, t_{n-1}; \ldots; \mathbf{x}_1, t_1) = P(\mathbf{x}_n, t_n | \mathbf{x}_{n-1}, t_{n-1}) \tag{9.39}$$

with $t_1 < t_2 < \cdots < t_n$.

A Markov process is fully determined by the two densities $P(\mathbf{x}_1, t_1)$ and $P(\mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1)$; the whole hierarchy can be reconstructed from them. For example, from Eq. (9.25) using the Markov property $P(3 \mid 1, 2) = P(3 \mid 2)$:

$$P(1, 2, 3) = P(1)P(2 \mid 1)P(3 \mid 2). \tag{9.40}$$

The factorisation of the joint pdf can thus be explicitly written in full notation as

$$P(\mathbf{x}_n, t_n; \mathbf{x}_{n-1}, t_{n-1}; \ldots; \mathbf{x}_1, t_1) = P(\mathbf{x}_1, t_1) \prod_{i=2}^{n} P(\mathbf{x}_i, t_i \mid \mathbf{x}_{i-1}, t_{i-1}), \tag{9.41}$$

with the propagator $P(\mathbf{x}_{i+1}, t_{i+1} \mid \mathbf{x}_i, t_i)$ carrying the system forward in time, beginning with the initial distribution $P(\mathbf{x}_1, t_1)$.

A well known example of Markov process is the **Wiener-Lévy process** describing the position of a Brownian particle (Fig. 9.14).

The fact that a Markov process is fully determined by $P(1)$ and $P(2 \mid 1)$ does not mean that such two functions can be chosen arbitrarily, for they must also obey two important identities.

The first one is Eq. (9.26) that in explicit form reads:

$$P(\mathbf{x}_2, t_2) = \int_{\mathbf{x}_1} P(\mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1) P(\mathbf{x}_1, t_1) d\mathbf{x}_1. \tag{9.42}$$

This equation simply constructs the one time probabilities in the future t_2 of t_1, given the conditional probability $P(\mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1)$.

The second property can be obtained by marginalising the joint pdf $P(\mathbf{x}_3, t_3, \mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1)$ with respect to $\mathbf{x}_2$ and by using the definition of conditional density under the Markov property:

$$P(\mathbf{x}_3, t_3 \mid \mathbf{x}_1, t_1) = \int_{\mathbf{x}_2} P(\mathbf{x}_3, t_3 \mid \mathbf{x}_2, t_2) P(\mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1) d\mathbf{x}_2, \tag{9.43}$$

Equation (9.43) is known as the **Chapman-Kolmogorov Equation** (C-K equation, from now on). It is "just" a statement saying that to move from position $\mathbf{x}_1$ to $\mathbf{x}_3$ you just need to average out all possible intermediate positions $\mathbf{x}_2$ or, more precisely, by marginalisation over the nuisance variable $\mathbf{x}_2$ (Fig. 9.15).

Such equation is a consistency equation for the conditional probabilities of a Markov process and the starting point for deriving the equations of motion for Markov processes. Aside from providing a consistency check, the real importance of the C-K equation is that it enables us to build up the conditional probability densities over the "long" time interval $[t_1, t_3]$ from those over the "short" intervals $[t_1, t_2]$ and $[t_2, t_3]$.

The C-K equation is a rather complex nonlinear functional equation relating all conditional probabilities $P(\mathbf{x}_i, t_i \mid \mathbf{x}_j, t_j)$ to each other. Its solution would give us a complete description of any Markov process, but unfortunately, no general solution to this equation is known: in other terms, it expresses the Markov character of the process, but containing no information about any particular Markov process.

The idea of forgetting the past so to use the present state for determining the next one might seem an oversimplified assumption when dealing, for instance, with eye movements performed by an observer engaged in some overt attention task. However, this conclusion may not in fact be an oversimplification. This was discussed by Horowitz and Wolfe (1998), and will be detailed in the following subsection.

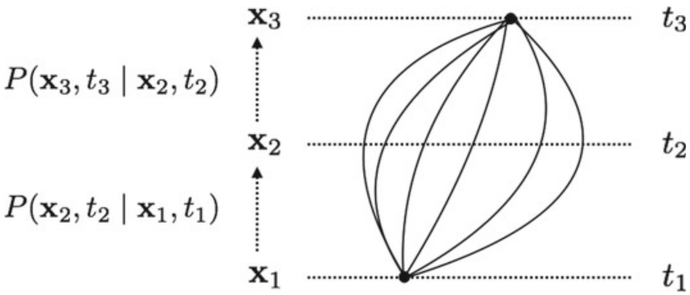


Fig. 9.15 The Chapman-Kolmogorov equation at work: the probability of transition from the event $(\mathbf{x}_1, t_1)$ to $(\mathbf{x}_3, t_3)$ is broken into a subprocess from $(\mathbf{x}_1, t_1)$ to an intermediate, nuisance event $(\mathbf{x}_2, t_2)$ (which is not observed in practice) and then from $(\mathbf{x}_2, t_2)$ to $(\mathbf{x}_3, t_3)$, by considering all the paths from $\mathbf{x}_1$ to $\mathbf{x}_3$

9.5.1 Case Study: The Horowitz and Wolfe Hypothesis of Amnesic Visual Search

Serial and parallel theories of visual search have in common the memory-driven assumption that efficient search is based on accumulating information about the contents of the scene over the course of the trial.

Horowitz and Wolfe in their seminal Nature paper (Horowitz & Wolfe, 1998) tested the hypothesis whether visual search relies on memory-driven mechanisms. They designed their stimuli so that, during a trial, the scene would be constantly changing, yet the meaning of the scene (as defined by the required response) would remain constant. They asked human observers to search for a letter “T” among letters “L”. This search demands visual attention and normally proceeds at a rate of 20–30 ms per item. In the critical condition, they randomly relocated all letters every 111 ms. This made it impossible for the subjects to keep track of the progress of the search. Nevertheless, the efficiency of the search was unchanged.

On the basis of achieved results they proposed that visual search processes are “amnesic”: they act on neural representations that are continually rewritten and have no permanent existence beyond the time span of visual persistence.

In other terms, the visual system does not accumulate information about object identity over time during a search episode. Instead, the visual system seems to exist in a sort of eternal present. Observers are remarkably oblivious to dramatic scene changes when the moment of change is obscured by a brief flicker or an intervening object.

Interestingly enough, they claim that an amnesic visual system may be a handicap only in the laboratory. The structure of the world makes it unnecessary to build fully elaborated visual representations in the head. Amnesia can be an efficient strategy for a visual system operating in the real world.

Box 9: The hall of fame of Markov processes

The most famous Markov process is the Wiener-Lévy process describing the position of a Brownian particle. Brownian particles can be conceived as a bodies of microscopically-visible size suspended in a liquid, performing movements of such magnitude that they can be easily observed in a microscope, on account of the molecular motions of heat (Einstein, 1906). Figure 9.14 shows an example of the 2-dimensional motion of one such particle.

A probabilistic description of the random walk of the Brownian particle must answer the question: What is the probability $P(\mathbf{x}, t)$ of the particle being at location $\mathbf{x} = (x, y)$ at time t?

In the 1−dimensional case, the probability $P(x, t)$ and its evolution law are defined for $-\infty < x < \infty$, $t > 0$ by the densities

$$P(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right), \tag{9.44}$$

$$P(x_2, t_2 \mid x_1, t_1) = \frac{1}{\sqrt{4\pi D(t_2 - t_1)}} \exp\left(-\frac{(x_2 - x_1)^2}{4D(t_2 - t_1)}\right). \tag{9.45}$$

that satisfy the Chapman-Kolmogorov equation. In both equations, D denotes a **diffusion coefficient**. The diffusion concept has deep roots in statistical physics: indeed, Einstein was the first to show in his seminal work on Brownian motion (Einstein, 1906) that the coefficient D captured the average or mean squared displacement in time of a moving Brownian particle ("[...] a process of diffusion, which is to be looked upon as a result of the irregular movement of the particles produced by the thermal molecular movement", Einstein, 1906). We will further discuss this important concept in Sect. 9.5.3.

9.5.2 *Stationary Markov Processes and Markov Chains*

Recall that for stationary Markov processes the transition probability $P(\mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1)$ only depends on the time interval. For this case one can introduce the special notation

$$P(\mathbf{x}_2, t_2 \mid \mathbf{x}_1, t_1) = T_\tau(\mathbf{x}_2 \mid \mathbf{x}_1) \text{ with } \tau = t_2 - t_1. \tag{9.46}$$

The Chapman-Kolmogorov equation then becomes

$$T_{\tau+\tau'}(\mathbf{x}_3 \mid \mathbf{x}_1) = \int_{\mathbf{x}_2} T_{\tau'}(\mathbf{x}_3 \mid \mathbf{x}_2) T_\tau(\mathbf{x}_2 \mid \mathbf{x}_1) d\mathbf{x}_2. \tag{9.47}$$

If one reads the integral as the product of two matrices or integral kernels, then

$$T_{\tau+\tau'} = T_{\tau'} T_\tau \quad (\tau, \tau' > 0) \tag{9.48}$$

A simple but important class of stationary Markov processes are the **Markov chains** defined by the following properties:

1. the state space of $\mathbf{x}$ is a *discrete set of states*;
2. the *time variable is discrete*;

In this case the dynamics can be represented as the PGM in Fig. 9.16. The PGM shows that the joint distribution for a sequence of observations $P(\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$ can be written as the product:

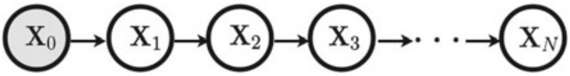


Fig. 9.16 The PGM of a Markov chain: given the prior or initial condition $P(\mathbf{x}_1)$, the behaviour of the system is determined by the conditional probability $P(\mathbf{x}_t \mid \mathbf{x}_{t-1})$

$$P(\mathbf{x}_1)P(\mathbf{x}_2 \mid \mathbf{x}_1) \cdots P(\mathbf{x}_N \mid \mathbf{x}_{N-1}) = P(\mathbf{x}_1)\prod_{t=2}^{N} P(\mathbf{x}_t \mid \mathbf{x}_{t-1}) \tag{9.49}$$

This is also known as an **observable Markov process**.

A **finite Markov chain** is one whose range consists of a finite number of N states. In this case the first probability distribution is an N component vector. The transition probability $T_\tau(\mathbf{x}_2 \mid \mathbf{x}_1)$ is an $N \times N$ matrix.

Thus, the C-K equation, by using the form in Eq. 9.48, leads to the matrix equation

$$T_\tau = (T_1)^\tau \tag{9.50}$$

Hence the study of finite Markov chains amounts to investigating the powers and the properties of the $N \times N$ transition matrix: this is a *stochastic matrix* whose elements are nonnegative and each row adds up to unity (i.e., they represent transition probabilities). One seminal application of Markov chains to scan paths has been provided by Ellis and Stark (1986).

9.5.2.1 Case Study: Modelling Gaze Shifts as Observable Finite Markov Chains

Ellis and Stark pioneered the use of Markov analysis for characterising scan paths (Ellis & Stark, 1986) in an attempt to go beyond visual inspection of the eye movement traces and application of a subjective test for similarity of such traces. In particular, they challenged the assumption of what they defined "apparent randomness", that many studies at the time were supporting in terms of either simple random or stratified random sampling (Ellis & Stark, 1986). To this end (see Fig. 9.17), they defined regions of interest (ROI) defined on the viewed picture, each ROI denoting a *state* into which the fixations can be located. By postulating that the transitions from one state to another have certain probabilities, they effectively described the generating process for these sequences of fixations as Markov processes. This way, they were able to estimate the marginal probabilities of viewing a point of interest i, i.e., $P(X = s_i)$, and the conditional probability of viewing a point of interest j given a previous viewing of a point of interest i, i.e., $T(X = s_j \mid X = s_i)$, where s_i, s_j are states in the state-space S (see Fig. 9.17).

By comparing expected frequency of transitions according to random sampling models with observed transition frequencies, they were able to assess the statisti-

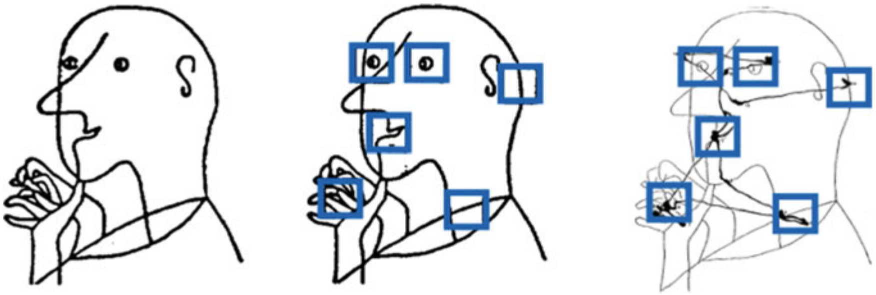


Fig. 9.17 Markov analysis of eye movements made by a subject viewing for the first time a drawing adapted from the Swiss artist Klee (left). Centre: ROIs superimposed on the drawing, defining the states of the Markov chain: $S = \{s_1 =$ "left eye", $s_2 =$ "right eye", $s_3 =$ "nose", $s_4 =$ "mouth", $s_5 =$ "hand", $s_6 =$ "neck"$\}$. Right: saccades represented as state transitions in the state-space. Modified after (Ellis & Stark, 1986; Hacisalihzade, Stark, & Allen, 1992)

cally significant differences that occurred (subject-by-subject basis with a chi-square goodness-of-fit test on the entire distribution of observed and expected transitions). Thus, they concluded that "there is evidence that something other than stratified random sampling is taking place during the scanning" (Ellis & Stark, 1986). In a further study (Hacisalihzade et al., 1992), examples have been provided for exploiting the observable Markov chain as a generative machine apt to sample simulated scan paths, once the transition matrix has been estimated / learned from data.

9.5.3 Levels of Representation of the Dynamics of a Stochastic Process

Up to this point, we have laid down the basis for handling eye movements in the framework of stochastic processes (cfr., Sect. 9.4 and previous subsections of the present one). Now we are ready for embracing a broader perspective.

Let us go back to Fig. 9.11 showing an ensemble of scan paths recorded from different observers while viewing the same image (in a similar vein, we could take into account an ensemble of scan paths recorded from the same observer in different trials). Our assumption is that such ensemble represents the outcome of a stochastic process and modelling/analysis should confront with the fundamental question: What is the probability $P(\mathbf{x}, t)$ of gazing at location $\mathbf{x}$ at time t?

There are three different levels to represent and to deal with such question; for grasping the idea it is useful to take a physicist perspective. Consider each trajectory (scan path) as the trajectory of a particle (or a random walker). Then, Fig. 9.11 provides a snapshot of the evolution of a many-particle system. In this view, the probability $P(\mathbf{x}, t)$ can be interpreted as the density $\rho(\mathbf{x}, t)$ (number of particles per unit length, area or volume) at point $\mathbf{x}$ at time t. In fact the density $\rho(\mathbf{x}, t)$ can be recovered by multiplying the probability density $P(\mathbf{x}, t)$ by the number of particles.

The finest grain of representation of a many-particle system is the individual particle, where each stochastic trajectory becomes the basic unit of the probabilistic description of the system. This is the **microscopic level**. In its modern form, it was first proposed by the french physicist Paul Langevin, giving rise to the notion of random walks where the single walker dynamics is governed by both regular and stochastic forces (which resulted in a new mathematical field of **stochastic differential equation**, briefly SDE, cfr. Box 11). At this level, $P(\mathbf{x}, t)$ can be be obtained by considering the collective statistical behaviour as given by the individual simulation of many individual particles (technically a Monte Carlo simulation, see Box 4 and refer to Fig. 9.21 for an actual example).

In the opposite way, we could straightforwardly consider, in the large scale limit, the equations governing the evolution of the space-time probability density $P(\mathbf{x}, t)$ of the particles. Albert Einstein basically followed this path when he derived the diffusion equation for Brownian particles (Einstein, 1905). This coarse-grained representation is the **macroscopic level** description.

A useful analogy for visualising both levels is provided by structure formation on roads such as jam formation in freeway traffic. At the microscopic scale, one can study the motion of an individual vehicle, taking into account many peculiarities, such as motivated driver behaviour and physical constraints. On the macroscopic scale, one can directly address phase formation phenomena collectively displayed by the car ensemble.

How do we relate the macroscopic and the microscopic levels? The crucial link is provided by the intermediate **mesoscopic level** of description. Pushing on with the traffic analogy, instead of following the motion of each vehicle (microscopic level), a stochastic cluster (mesoscopic) of congested cars is considered by starting to average or integrating microscopic fundamental laws. Then, further coarse-graining (and related approximations) allow to reach the macroscopic description where single vehicle behaviour has no place.

In comparison with the microscopic approach based on SDEs, the mesoscopic description does not allow to get individual realisations of the process but yet keeps the whole amount of statistical information of the underlying microscopic process. Technically it consists in finding integral or integro-differential equations for the probability that governs the evolution of the system. In the picture we have so far outlined the mesoscopic level is represented by the Chapman-Kolmogorov equation. At the microscopic level we actually consider the dynamics of the particle as governed by a regular force plus a fluctuating, random force due to the incessant molecular impacts. The C-K equation coarse-grains the landscape stating that the probability of a particle being at a point $\mathbf{x} + \Delta\mathbf{x}$ at time $t + \Delta t$ is given by the probabilities of all possible "pushes" $\Delta\mathbf{x}$, multiplied by the probability of being at $\mathbf{x}$ at time t. This assumption is based on the independence of the "push" of any previous history of the motion; it is only necessary to know the initial position of the particle at time t not at any previous time. This is the Markov postulate, and the C-K equation is the central dynamical equation for all Markov processes.

Summing up, the passage from a microscopic description to a macroscopic one, can be envisioned as a coarse-graining operation. At the mesoscopic level an appropriately chosen coarse-graining of the observation time-scale, permits the physical process to be described as Markovian, and such coarse-graining allows to switch from the individual particle to a density of particles $\rho(\mathbf{x}, t)$. Subsequently, under the same coarse-graining, the C-K equation can be reduced to a master equation or a diffusion equation describing the evolution of the system at the macroscopic level. More details are given in the following subsections.

9.5.3.1 The Microscopic Level

The microscopic description of a system amounts to writing down the evolution equations or **differential equations** (see Box 10) that describe the fine-grained dynamics of individual trajectories: e.g., the path of a Brownian particle or the scan path of an eye-tracked observer. A simple form of such equations is the following:

$$\overbrace{\frac{dx}{dt}}^{\text{state-space rate of change}} = \overbrace{a(x,t)}^{\text{deterministic comp.}} + \overbrace{b(x,t)\xi(t)}^{\text{stochastic comp.}}, \tag{9.51}$$

which we call the **Langevin equation**, in analogy with the well known equation that in statistical physics describes the time evolution of the velocity of a Brownian particle. In Eq. (9.51) the drift term $a(x,t)$ represents the deterministic component of the process; the diffusive component $b(x,t)\xi(t)$ is the stochastic component, $\xi(t)$ being the “noise” sampled from some probability density, i.e. $\xi(t) \sim P(\xi)$, usually a zero-mean Normal distribution.

Box 10: Dynamical systems and differential equations

A system that changes with time is called a **dynamical system**. A dynamical system consists of a **space of states** and entails a law of motion between states, or a **dynamical law**. The deterministic component of Langevin equation (9.51)

$$\frac{dx(t)}{dt} = a(x(t), t) \tag{9.52}$$

is one such law, the variable $x(t)$ being the variable that, moment to moment, takes values in the state space of positions. Equation (9.52) is a **differential equation** describing the rate of change of state-space variable x.

In simple terms, a dynamical law is a rule that tells us the next state given the current state. This can be more readily appreciated if we recall the definition of derivative given in Box 1, but avoiding the shrinking operation ($\lim_{\Delta t \to 0}$), i.e. we approximate the derivative as a discrete difference

$$\frac{dx(t)}{dt} \approx \frac{x(t+\Delta t) - x(t)}{\Delta t}.$$

By assuming for simplicity a unit time step, i.e., $\Delta t = 1$ and substituting in Eq. (9.52):

$$\overbrace{x(t+1)}^{\text{next state}} = \overbrace{x(t)}^{\text{current state}} + a(x(t), t) \tag{9.53}$$

Equation (9.53) is the discrete-time version of the differential equation (9.52), namely a finite-difference equation. The model in discrete time emphasises the predictive properties of the law: indeed, with the scientific method we seek to make predictions about phenomena that are subject to change. Caveat: we should always be cautious about how predictable the world is, even in classical physics. Certainly, predicting the future requires a perfect knowledge of the dynamical laws governing the world but at the same time entails the ability to know the initial conditions with almost perfect precision. However, perfect predictability is not achievable, simply because we are limited in our resolving power. There are cases in which the tiniest differences in the initial conditions (the starting state), leads to large eventual differences in outcomes. This phenomenon is called *chaos*.

The law formalised in Eqs. (9.52) or (9.53) are deterministic. In stark contrast, the Langevin equation (9.51) "corrupts" the deterministic law of motion with the "noise" introduced by the RV $\xi(t)$. Thus, the eventual outcome is not deterministic but stochastic (though it may be predictable in probability). Langevin equation is but one example of **stochastic differential equation** (SDE).

Equation (9.51) is an SDE, which in a more formal way can be written in the Itô form of Eq. (9.54) as detailed in Box 11.

Box 11: Stochastic Differential Equations

The Langevin equation written in the form (9.51) poses some formal problems. Since $\xi(t)$ is noise it consists of a set of points that in some cases can be even uncorrelated. As a consequence $\xi(t)$ is often non-differentiable. Thus, $x(t)$ should be non-differentiable too, so that the left hand side of (9.51) is incoherent from this point of view. To overcome this problem, the 1−dimensional Langevin equation is usually presented in the mathematically sound form:

$$dx(t) = a(x(t), t)dt + b(x(t), t)\xi(t)dt = a(x(t), t)dt + b(x(t), t)dW(t) \tag{9.54}$$

with $W(t) = \int_0^t \xi(t')dt'$, so that the integration of the stochastic component $\int b(x, t)dW(t)$ can be performed according to the rules of stochastic calculus (in the Itô or Stratonovich approach Higham, 2001). Throughout this chapter we shall use with a certain liberality both forms (9.51) and (9.54) at our convenience. Thus, a stochastic quantity $x(t)$ obeys an Itô SDE written as in (9.54), if for all t and t_0,

$$x(t) = x(t_0) + \int_{t_0}^{t} a(x(t'), t')dt' + \int_{t_0}^{t} b(x(t'), t')dW(t') \tag{9.55}$$

A discretised version of the SDE can be obtained by taking a mesh of points t_i (Fig. 9.19)

$$t_0 < t_1 < t_2 < \cdots < t_{n-1} < t_n = t$$

and writing the equation as

$$x_{i+1} = x_i + a(x_i, t_i)\Delta t_i + b(x_i, t_i)\Delta W_i \tag{9.56}$$

Here, $x_i = x(t_i)$ and

$$\Delta t_i = t_{i+1} - t_i, \tag{9.57}$$

$$\Delta W_i = W(t_{i+1}) - W(t_i) \propto \sqrt{\Delta t_i}\xi_i. \tag{9.58}$$

The approximate procedure for solving the equation is to calculate x_{i+1} from the knowledge of x_i by adding a deterministic term $a(x_i, t_i)\Delta t_i$ and a stochastic term $b(x_i, t_i)\Delta W_i$, which contains the element ΔW_i, namely the increment of the Wiener process. The solution is then formally constructed by letting the mesh size go to zero. The method of constructing a solution outlined above is called the Cauchy-Euler method, and can be used to generate simulations. By construction the time development of $x(t)$ for $t > t_0$ is independent of $x(t_0)$ for $t > t_0$ provided $x(t_0)$ is known. Hence, $x(t)$ is a Markov process.

For an intuitive, Matlab based, thorough introduction to SDEs see Higham (2001).

Concretely, the construction of a trajectory (a solution) can be performed by refining the intuitive discretisation approach presented in Box 10. Eq. (9.51) is discretised as in Eq. (9.56) by executing a sequence of drift and diffusion steps as illustrated in Fig. 9.19.

In continuous time a 2-dimensional random motion of a particle, with stochastic position/state $\mathbf{x}(t)$, under the influence of an external force field (Siegert & Friedrich, 2001), can be described by the Langevin stochastic equation

$$d\mathbf{x}(t) = \mathbf{A}(\mathbf{x}, t)dt + \mathbf{B}(\mathbf{x}, t)\boldsymbol{\xi}dt. \tag{9.59}$$

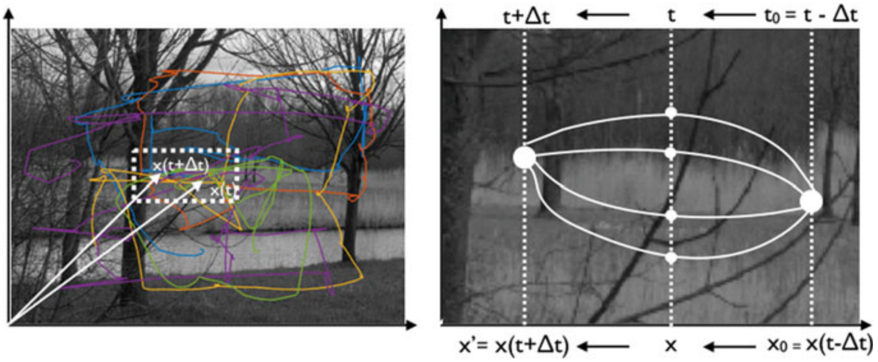
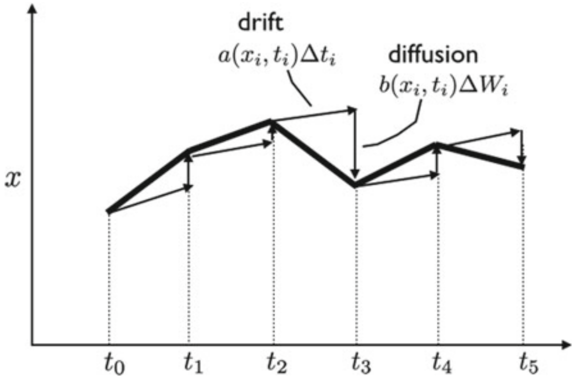


Fig. 9.18 Left panel: the ensemble of scan paths that was presented in Fig. 9.11. At the microscopic level, each trajectory can be seen as the output of the simulation of Langevin equation. Right panel: a close-up of the window selected from the image in the left panel (dotted line). The mesoscopic representation used by the C-K equation for computing the conditional probability of shifting the gaze from $(\mathbf{x}_0, t_0)$ to $(\mathbf{x}', t + \Delta t)$. Here the individual trajectories are replaced by the particle densities around such points. Δt is a small time interval

Fig. 9.19 The Cauchy-Euler procedure for constructing an approximate solution of the Langevin SDE in the Itô form (cfr. Box 11)



As in the one-dimensional case, the trajectory of $\mathbf{x}$ is determined by a deterministic part $\mathbf{A}$, the drift, and a stochastic part $\mathbf{B}(\mathbf{x}, t)\boldsymbol{\xi} dt$, where $\boldsymbol{\xi}$ is a random vector and $\mathbf{B}$ is a diffusion factor.

By simulating from Langevin-like equations, for suitable choice of parameters $\mathbf{A}$, $\mathbf{B}$ and noise distribution $P(\boldsymbol{\xi})$ (which is likely to be non Gaussian as we will see) it is possible to obtain trajectories that are similar to the individual trajectories shown in the left panel of Fig. 9.18.

9.5.3.2 The Mesoscopic Level

If we select an appropriate spatio-temporal scale we can coarse-grain our description by summarising single particle dynamic behaviours in terms of variations on particle densities at different points, say $\mathbf{x}_0$ and $\mathbf{x}'$, as illustrated in the right panel of Fig. 9.18. This density based description to an actual image the abstract C-K construction we previously presented in Fig. 9.15. In this case, by considering the state $\mathbf{x} = (n, m)$ as indexing a discrete grid of image pixels and, admittedly, to avoid burdening mathematical details, we can rewrite C-K equation (9.43) in discrete form:

$$P(\mathbf{x}', t + \Delta t \mid \mathbf{x}_0, t_0) = \sum_{\mathbf{x}} P(\mathbf{x}', t + \Delta t \mid \mathbf{x}, t) P(\mathbf{x}, t \mid \mathbf{x}_0, t_0). \tag{9.60}$$

Now, the microscopic individual trajectories are replaced on a coarse-grain scale by the particle densities around such points and single particle dynamics is summarised by the Markov-based transition probability $P(\mathbf{x}', t + \Delta t \mid \mathbf{x}_0, t_0)$. Note that even in the discrete case where the C-K equation has an intuitive form, it is equally a complicated equation that a function of four parameters namely has to fulfil.

However, from this level it is possible to compute the evolution of particle densities in a larger scale limit, where we consider the coarse-grained dynamics of the overall pdf of the many-particle system, as detailed below.

9.5.3.3 The Macroscopic Level

There are two possible macroscopic limits from mesoscopic equations: the macroscopic limit in time or in space. When we consider the macroscopic limit in time of the C-K equation, we obtain the **Master equation**; when we consider the macroscopic limit both in time and in the state space of the C-K equation, we obtain the famous **Fokker-Planck equation** (F-P).

To illustrate how this works we present a simple derivation of the Master equation starting from the discrete C-K equation (9.60). First of all, we assume that Δt is a small time interval. Then we characterise the short time properties of the conditional probability $P(\mathbf{x}', t + \Delta t \mid \mathbf{x}, t)$, where $\mathbf{x}$ denote a number of intermediate states between the initial state $\mathbf{x}_0$ and the ending state $\mathbf{x}'$ as usually considered in the C-K construction (cfr. Fig. 9.18, right).

We also assume that a transition from $\mathbf{x}$ to state $\mathbf{x}'$ is proportional to time when time gets small and that such transition occurs at a transition rate which we denote $w(\mathbf{x}' \mid \mathbf{x})$ (density variation per unit time). Thus:

$$P(\mathbf{x}', t + \Delta t \mid \mathbf{x}, t) \approx \Delta t \times w(\mathbf{x}' \mid \mathbf{x}), \tag{9.61}$$

which obviously holds only for $\mathbf{x}' \neq \mathbf{x}$ since $w(\mathbf{x} \mid \mathbf{x}) = 0$.

However, to be complete we must also consider the probability that no state transition occurs in the small time interval:

$$Q(\mathbf{x}) = 1 - \Delta t \sum_{\mathbf{x}' \neq \mathbf{x}} w(\mathbf{x}' \mid \mathbf{x}). \tag{9.62}$$

Eventually,

$$P(\mathbf{x}', t + \Delta t \mid \mathbf{x}, t) \approx \Delta t \times w(\mathbf{x}' \mid \mathbf{x}) + Q(\mathbf{x})\delta_{\mathbf{x}',\mathbf{x}}, \tag{9.63}$$

where $\delta_{\mathbf{x}',\mathbf{x}}$ is the Kroenecker symbol: $\delta_{\mathbf{x}',\mathbf{x}} = 0$ when $\mathbf{x}' \neq \mathbf{x}$, and $\delta_{\mathbf{x}',\mathbf{x}} = 1$ when $\mathbf{x}' = \mathbf{x}$.

If we plug the right hand side of Eq. (9.63) in C-K equation (9.60), after some (tedious and unrelevant) algebra we obtain the following result:

$$\begin{aligned} P(\mathbf{x}', t + \Delta t \mid \mathbf{x}_0, t_0) = P(\mathbf{x}', t \mid \mathbf{x}_0, t_0) + \\ \Delta t \sum_{\mathbf{x}} \left[w(\mathbf{x}' \mid \mathbf{x}) P(\mathbf{x}, t \mid \mathbf{x}_0, t_0) - w(\mathbf{x} \mid \mathbf{x}') P(\mathbf{x}', t \mid \mathbf{x}_0, t_0) \right]. \end{aligned} \tag{9.64}$$

Note that in the limit $\Delta t \to 0$, we can write

$$\lim_{\Delta t \to 0} \frac{P(\mathbf{x}', t + \Delta t \mid \mathbf{x}_0, t_0) - P(\mathbf{x}', t \mid \mathbf{x}_0, t_0)}{\Delta t} = \frac{\partial P(\mathbf{x}', t \mid \mathbf{x}_0, t_0)}{\partial t}$$

This is the definition of a *partial derivative* of first order with respect to time, denoted by the symbol $\frac{\partial}{\partial t}$.[7]

Using such definition in Eq. (9.65), in the limit $\Delta t \to 0$:

$$\frac{\partial P(\mathbf{x}', t \mid \mathbf{x}_0, t_0)}{\partial t} = \sum_{\mathbf{x}} \left[w(\mathbf{x}' \mid \mathbf{x}) P(\mathbf{x}, t \mid \mathbf{x}_0, t_0) - w(\mathbf{x} \mid \mathbf{x}') P(\mathbf{x}', t \mid \mathbf{x}_0, t_0) \right]. \tag{9.65}$$

This equation can be further simplified by using the marginalisation rule

$$\sum_{\mathbf{x}_0} P(\mathbf{x}', t \mid \mathbf{x}_0, t_0) P(\mathbf{x}_0, t_0) = P(\mathbf{x}_0, t_0)$$

Multiplying equation (9.65) by $P(\mathbf{x}_0, t_0)$ and summing over $\mathbf{x}_0$, we eventually obtain

$$\frac{\partial P(\mathbf{x}', t)}{\partial t} = \sum_{\mathbf{x}} \left[w(\mathbf{x}' \mid \mathbf{x}) P(\mathbf{x}, t) - w(\mathbf{x} \mid \mathbf{x}') P(\mathbf{x}', t) \right]. \tag{9.66}$$

This is our final Master equation and has a very simple interpretation. The density $P(\mathbf{x}', t)$ is the probability of the ensemble of observers (the many particle system) of being in (gazing at) state $\mathbf{x}'$ at time t. How does this density change over time? The system can be in any state $\mathbf{x} \neq \mathbf{x}'$ and move into the state $\mathbf{x}'$. The system is in state $\mathbf{x}$

---

[7] If we have a function of more than one variable, e.g., $f(x, y, z, \ldots)$, we can calculate the derivative with respect to one of those variables, with the others kept fixed. Thus, if we want to compute $\frac{\partial f(x,y,z,\ldots)}{\partial x}$, we define the increment $\Delta f = f([x + \Delta x], y, z, \ldots) - f(x, y, z, \ldots)$ and we construct the partial derivative as in the simple derivative case as $\frac{\partial f(x,y,z,\ldots)}{\partial x} = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$. By the same method we can obtain the partial derivative with respect to any of the other variables.

with probability $P(\mathbf{x}, t)$ and from any of these states $\mathbf{x}$ it moves to the state $\mathbf{x}'$ with probability $w(\mathbf{x}' \mid \mathbf{x})P(\mathbf{x}, t)$. But simmetrically, it could also already be in state $\mathbf{x}'$ and move to one of the other states $\mathbf{x}$ with probability $w(\mathbf{x} \mid \mathbf{x}')P(\mathbf{x}', t)$.

In other terms the Master equation is a probability flux balance equation, best understood if $P(\mathbf{x}', t)$ is interpreted as a particle density: the rate of change in that density (derivative) is the difference of what comes in and what goes out.

Technically speaking the Master equation (9.66) is a stochastic **partial differential equation** (PDE) defining the "law of motion" of the density $P(\mathbf{x}', t)$ in probability space and it has been obtained by taking the limit in time of the C-K equation.

When we consider the macroscopic limit both in time and in the state space of the C-K equation, we obtain another PDF, namely the **Fokker-Planck** (F-P) equation.We omit the formal derivation because beyond the scope of this Chapter.

In the simple 1-dimensional case, the F-P equation for diffusive processes is the following:

$$\frac{\partial P(x,t)}{\partial t} = -\frac{\partial}{\partial x}[a(x,t)P(x,t)] + \frac{1}{2}\frac{\partial^2}{\partial x^2}[b(x,t)^2 P(x,t)] \tag{9.67}$$

The symbols $\frac{\partial}{\partial x}$ and $\frac{\partial^2}{\partial x^2} = \frac{\partial}{\partial x}(\frac{\partial}{\partial x})$ denote partial derivatives with respect to space of first and second-order, respectively.

What is important to note is that there is a formal link between the microscopic description provided by the Langevin equation (cfr., the 1-dimensional case of Eq. 9.51) and the macroscopic description addressed by the F-P equation (9.67), which is established via $a(x, t)$ and $b(x, t)$. The term $a(x, t)$ represents a **drift** which is related to the average deviation of the process

$$a(x,t) = \lim_{\Delta t \to 0} \frac{\langle \Delta x \rangle}{\Delta t}. \tag{9.68}$$

over a small time interval Δt; the bracket operator $\langle\rangle$ denotes the average or expectation value of any function $f(X)$.[8] Δx is a deviation or displacement in state-space.

The term $b^2(x, t)$ represents a **diffusion** term, which is related to the mean square deviation of the process:

$$b^2(x,t) = \lim_{\Delta t \to 0} \frac{\langle (\Delta x)^2 \rangle}{\Delta t}. \tag{9.69}$$

Eventually, at the macroscopic level, by knowing the evolution of $P(x, t)$ in time, one can obtain statistical "observables" as the moments, correlations, etc. These obviously lack microscopic details from the underlying stochastic process, which for some specific purposes may be important.

We will turn now to the fundamental example of the Wiener process to make clear the connections between the macroscopic and microscopic levels of description.

---

[8] This is physicists' preferred notation, which you are likely to most frequently run into when dealing with these problems. In other more mathematically inclined papers and books you will find the expectation notation $E[f(X)]$ or $Ef(X)$.

9.5.3.4 Example: The Wiener Process

Recall again the most famous Markov process: the Wiener process describing Brownian motion (Bm), cfr., Box 9. The SDE defining the motion of a particle undergoing 1-dimensional Brownian motion can be obtained by setting to zero the drift component $a(x, t)$ and letting $b(x, t) = \sqrt{2D}$, where D is the diffusion coefficient. Thus:

$$dx = \sqrt{2D}dW(t) \tag{9.70}$$

By using Eqs. (9.56) and (9.58) introduced in Box 11, the discretised version of the Wiener process (9.70) over a small but finite time interval $\Delta t = T/N$, N being the discrete number of integration steps, can be written as

$$x_{i+1} = x_i + \sqrt{2D}\Delta W_i = x_i + \sqrt{2D\Delta t_i}\xi_i \tag{9.71}$$

with ξ sampled from a zero-mean Gaussian distribution of unit variance $\mathcal{N}(0, 1)$

Equation 9.71 shows that the system describes a refined version of the simple additive random walk. Once again, it is worth noting that since the $\xi(t)$ are sampled i.i.d, then the differences in sequential observations are i.i.d, namely, $x_{i+1} - x_i = \Delta x_i$, rather than the observations themselves. In fact, if we compute the auto-correlation function of the $\{x(t)\}$ time series, it exhibits a slower decay —differently from the white noise process—, which shows how this simple random walk exhibits memory (Fig. 9.20).

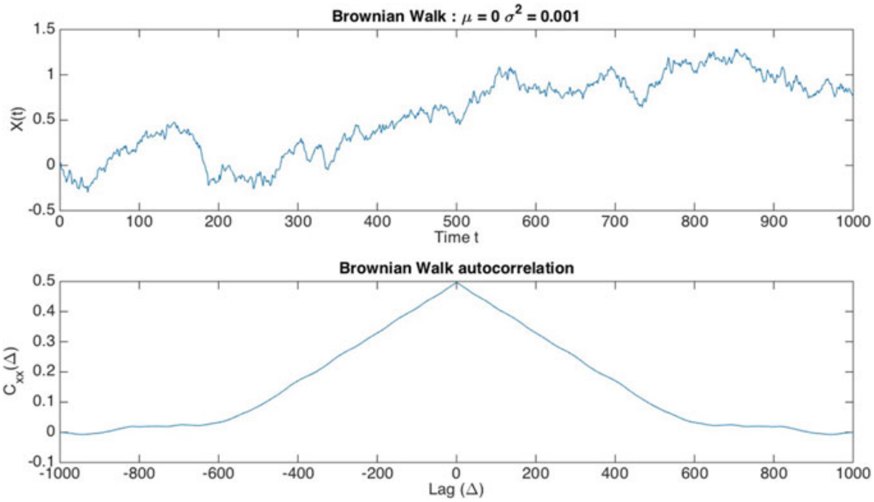


Fig. 9.20 One dimensional Brownian motion. Top: the random walk process; bottom: the autocorrelation of the process (cfr. Box 6)

Let us simulate the Brownian motion of a large number of particles, say 10^5 particles. Resorting to a Monte Carlo approach, we can obtain this result by running in parallel 10^5 random walks, each walk being obtained by iterating Eq. (9.71). Figure 9.21 (top) shows an example of 20 such trajectories. In probabilistic terms each trajectory is a realisation, a sample of the stochastic process $\{X(t)\}$.

We may be interested in gaining some statistical insight of the collective behaviour of all such random walkers. This can be obtained by considering the the dynamics of the pdf $P(x, t)$ describing the probability of finding a particle at position x at time t. Empirically, we can estimate $P(x, t)$ at any time t by computing the density of particles occurring within a certain bin $(x - \delta, x + \delta)$ centred on x, that is by computing the histogram $h(x, t)$ and normalising it with respect to the total number of particles. This procedure is shown in the bottom of Fig. 9.21: the empirical pdf has a nice bell shape, i.e. it is a Normal distribution, which spreads as time increases.

This insight can be given a formal justification by resorting to the macroscopic level of description of process dynamics as provided by F-P equation (9.67). By setting again $a(x, t) = 0$ and $b(x, t) = \sqrt{2D}$:

$$\frac{\partial P(x,t)}{\partial t} = D \frac{\partial^2 P(x,t)}{\partial x^2} \tag{9.72}$$

This is the well-known **heat** or **diffusion equation**. Thus, the pdf $P(x, t)$ of finding a particle at position x at time t evolves in time according to the diffusion equation when the underlying microscopic dynamics is such that the particle position corresponds to a Wiener process.

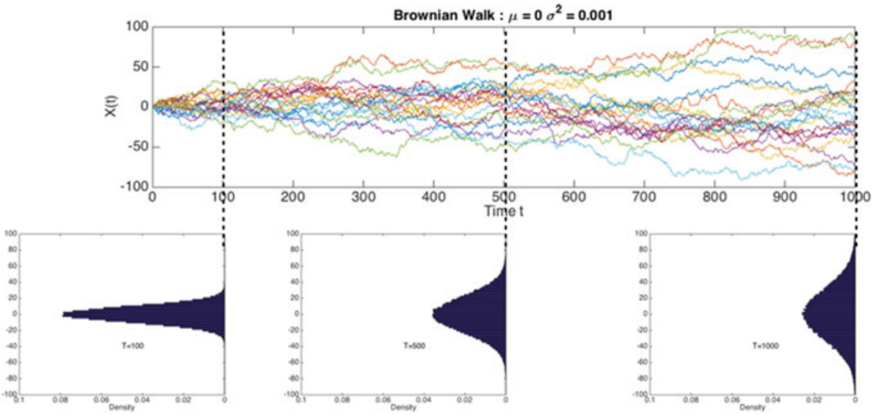


Fig. 9.21 A Monte Carlo simulation (cfr. Box 4) of the macroscopic dynamics of $P(x, t)$ in the case of Brownian motion. Top: The simulation of individual trajectories of 10^5 random walkers: only 20 are shown for visualisation purposes. Bottom: The distributions (histograms) of the walkers, after $T = 100$, $T = 500$ and $T = 1000$ time steps. The distribution initially concentrated at a point takes later the Gaussian form, whose width grows in time as $t^{1/2}$. This kind of diffusion is called the *normal diffusion*

The solution to the heat Eq. (9.72) is the time-dependent Gaussian pdf, as anticipated in Box 9:

$$P(x,t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left(-\frac{x^2}{4Dt}\right), \tag{9.73}$$

By comparing the Gaussian pdf variance $\sigma^2 = 4Dt$ to the definition of $b^2(x,t)$ given in Eq. (9.69), we can set the following correspondences:

$$\sigma^2 = 2Dt = b^2 t \approx \langle x^2 \rangle \tag{9.74}$$

In other terms for Bm, the average square deviation of the walk, and thus the spread of the Gaussian, grows linearly with time, as it can be intuitively appreciated from Fig. 9.21.

More precisely, define the **Mean Square Displacement** (MSD) of a walk that starts at position x_0 at time t_0:

$$MSD = \langle |x - x_0|^2 \rangle, \tag{9.75}$$

which is, quoting Einstein, the square of the displacement in the direction of the x-axis "that a particle experiences on the average" (Einstein, 1906). Here x_0 denotes the initial position. In the case of Brownian motion, Einstein (1906) was the first to show that:

$$MSD = 2Dt \tag{9.76}$$

Note that $\langle |x - x_0|^2 \rangle = \langle x^2 \rangle + x_0^2 - 2x_0 \langle x \rangle$. Hence, when the initial position is at $x_0 = 0$, $MSD = \langle x^2 \rangle \propto t$.

Equation (9.76) is sometimes more generally written in terms of the **Hurst exponent** H

$$MSD = kt^{2H} \tag{9.77}$$

with $H = \frac{1}{2}$ for Bm. This is useful for characterising different kinds of diffusions, like hyperdiffusion or subdiffusion, as discussed in Box 12.

Box 12: The Hurst exponent: the Swiss army knife of diffusion processes (without SDE pain)

The Hurst exponent, H, is related to the signal correlation behaviour and it allows the detection of the long-range dependences. In general, properties of Gaussian diffusion may be expressed in terms of the MSD of x and its relation to time:

$$MSD = \langle |x(t) - x(0)|^2 \rangle = kt^{2H} \tag{9.78}$$

When $H = 0.5$, MSD is linear in time:

$$MSD = kt, \tag{9.79}$$

which exemplifies the ordinary condition of Bm, the derivative of Bm being additive white Gaussian noise. When $H > 0.5$, increments are positively correlated, i.e. the random walk shows the tendency to continue to move in the current direction. This behaviour is called **persistence**. In this case, MSD increases nonlinearly with respect to time, indicative of **hyperdiffusion**. In particular, for $H = 1$,

$$MSD = kt^2. \tag{9.80}$$

In this case diffusion follows correlated **fractional Brownian motion** (fBm), whose derivative is fractional Gaussian noise.

In the case $H < 0.5$, the random walk generates negatively correlated increments and is **anti-persistent**.

It is important to note that for $H \neq 0.5$, the increments are not independent, thus the fBm is a Gaussian process but it is not a Markov process.

Interesting work has been reported in the recent literature on the use of the H exponent to analyse eye movements, e.g. by Engbert and colleagues on random walk analysis of fixational eye movements (Engbert, 2006). For such purposes the scaling exponent H can be estimated as follows (Engbert, 2006). Consider a time series of gaze positions of length N, $\mathbf{x}_1, \ldots, \mathbf{x}_N$. Define the displacement estimator

$$\overline{\delta^2(\Delta_m)} = \frac{1}{N-m} \sum_{i=1}^{N-m} \|\mathbf{x}_{i+m} - \mathbf{x}_i\|^2, \tag{9.81}$$

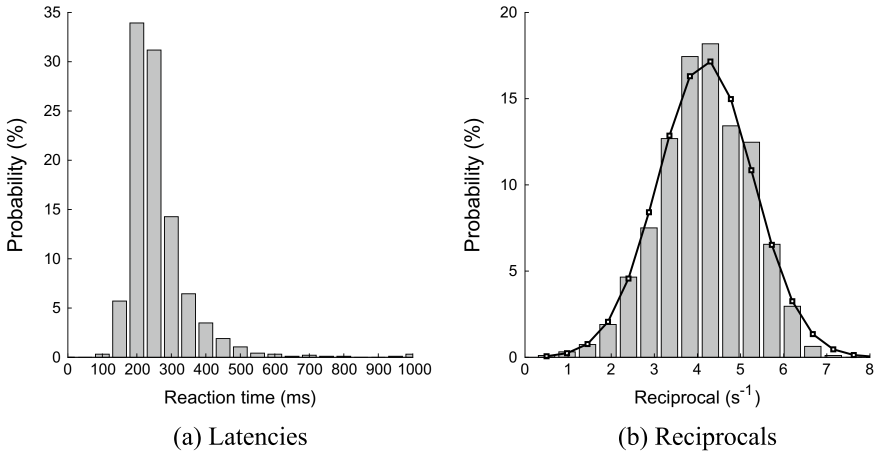
namely the **time averaged MSD**. By recalling that $MSD \approx t^{2H}$, the Hurst exponent H can be obtained by calculating the slope in the plot of $\log \overline{\delta^2(\Delta_m)}$ as a function of $\log \Delta_m$, where $\Delta_m = mT_0$ is the *time lag*, T_0 (ms) being the sampling time interval and $m = 1, 2, 3, \ldots$.

More recently, Engbert has proposed more sophisticated estimation framework based on the Bayesian approach (Makarava, Bettenbühl, Engbert, & Holschneider, 2012).

A last, but important remark: the equivalence

$$\overline{\delta^2(\Delta)} \approx \langle x(\Delta)^2 \rangle \tag{9.82}$$

holds when the process is **ergodic**, where ensemble averages and long-time averages are equivalent in the limit of long measurement times.



(a) Latencies (b) Reciprocals

Fig. 9.22 Empirical distributions (histograms) of latencies (left panel) and reciprocals (right panel). The latencies/reaction times themselves are typically skewed, with a tail towards long reaction times. When plotted as a function of reciprocal reaction time (inverse of the reaction times or promptness), the distribution looses its skewness, and it turns out to obey the most fundamental of all stochastic laws, the normal or Gaussian distribution. The reciprocal histogram can thus be approximated by a normal distribution (the black line superimposed to the histogram, with the same mean and standard deviation of actual reciprocals) (Carpenter & Williams, 1995; Noorani & Carpenter, 2016)

9.5.3.5 Case Study: From Random Walks to Saccade Latency

A saccade represents the output of a decision, a choice of where to look, and reaction time, or latency, can be regarded as an experimental "window" into decision processes. Typical latencies are around 200 ms (see Fig. 9.22a) whereas the sensorimotor components of a saccade are only around 60 ms: this suggests that reaction time is composed of more than the simple sum of times of for sensory input and motor output, and such extra time (also called neural procrastination) reflects the time taken for the brain to choose a response. Put simply, reaction time is a useful indicator of decision time. In experimental paradigms, reaction time varies stochastically between one trial and the next, despite standardized experimental conditions. Furthermore, distribution of reaction times is typically skewed, with a tail towards long reaction times (Fig. 9.22a). However, if we take the reciprocal of the latencies and plot these in a similar fashion, the resulting distribution appears Gaussian (Fig. 9.22b).

Box 13: The probit function

When confronting with reaction times and in particular with the well known Carpenter's LATER model (Carpenter & Williams, 1995; Noorani & Carpenter, 2016), one tipically has to deal with reciprobit plots, a tool that sometimes students find difficult to conceptualise. Indeed, the subtle probit function lies behind such graphs.

Recall from your elementary probability and statistics class that a quantile function returns the value x such that $F_X(x) = P(X \leq x) = p$ where $0 < p < 1$ is a probability value and $F_X(\cdot)$ is the **cumulative distribution function** (CDF) of random variable X. In simple words, it returns a threshold value x below which random draws from the given CDF would fall p percent of the time.

The probit function $probit(\cdot)$ is exactly the quantile function associated with the standardised normal distribution $\mathcal{N}(0, 1)$ and the standard CDF $\Phi(\cdot)$, formally

$$probit(p) = \Phi^{-1}(p). \tag{9.83}$$

where $\Phi^{-1}(\cdot)$ is the inverse CDF of the standard normal. Clearly, by the above definition, the following properties hold:

$$\Phi(probit(p)) = \Phi(\Phi^{-1}(p)) = p$$

and

$$probit(\Phi(z)) = \Phi^{-1}(\Phi(z)) = z$$

where $z \sim \mathcal{N}(0, 1)$ is a standardised random variable (the famous "z-score") sampled from the standard normal.

Unfortunately, the standard CDF Φ and its inverse Φ^{-1} are not available in closed form, and computation requires careful use of numerical procedures (unless we do not exploit good old tables). To such end, $\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-t^2}{2}} dt$ is related to the error function $erf(x) = \frac{1}{\sqrt{\pi}} \int_{-x}^{x} e^{-t^2} dt$ as $\Phi(z) = \frac{1}{2}\left[1 + erf\left(\frac{z}{\sqrt{2}}\right)\right]$. However, we need not worry too much about these mathematical details since numerical calculations of the error and the inverse error functions are widely available in software for statistics and probability modelling.

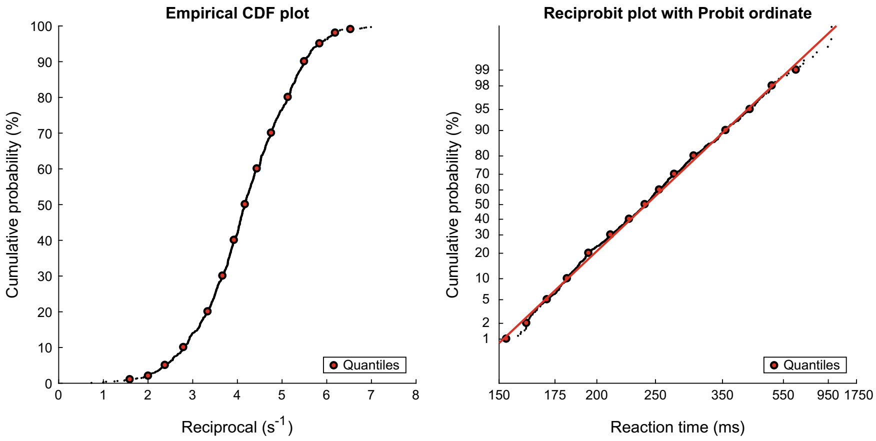
Eventually, given a probability value $\Phi(z) = p$ from the normal CDF, the probit of such value can be computed in terms of the inverse error function $erf^{-1}(\cdot)$ as

$$probit(p) = \sqrt{2} erf^{-1}(2p - 1), \tag{9.84}$$

which, for instance, in Matlab code boils down to the simple statement

```
prob_p = sqrt(2) * erfinv(2 * p -1)
```

This is all we need for producing beautiful reciprobit plots (see Fig. 9.23b).



(a) Reciprocal ECDF (b) Reciprobit plot

Fig. 9.23 Empirical CDF of reciprocals (left panel) and reciprobit plot (right panel) of the same data presented in Fig. 9.22. In the cumulative probability of inverse reaction time, every dot is an actual individual data point, while red markers indicate quantiles at probabilities $p = 0.01, 0.02, 0.05, 0.1, 0.2, 0.3, \ldots, 0.7, 0.8, 0.9, 0.95, 0.98, 0.99$. In the right panel, the same data as in the left panel plotted as a cumulative histogram but using a probit scale. Note that the latency uses a reciprocal scale resulting in a reciprobit graph

Histograms, however, are in some sense problematic since their shape depends on the bin size and they have an effectively arbitrary vertical scale. As an alternative, cumulative histograms or empirical Cumulative Distribution Function (ECDF) are normalized (running from 0 to 1 or 100%), can represent all data without binning (continuous), thus facilitating comparisons between data sets. The ECDF $\widetilde{F_N}(x)$, is an unbiased estimator of the CDF, the Cumulative Distribution Function $F_X(x)$, of a RV X (cfr. Box 13 for refreshing these basic concepts). In practice, given N data points x_i, $i = 1, \ldots, N$ the computation of $\widetilde{F_N}(x)$ boils down to two steps: (1) count the number of data less than or equal to x; divide the number found in (1) by the total number of data in the sample. The ECDF of inverse reaction times previously presented in Fig. 9.22 is illustrated in Fig. 9.23a.

An even better result can be achieved by using a non-linear probit scale (an inverse error function transformation, see Box 13, Eq. 9.84) for the vertical frequency axis, rather than a linear scale (Carpenter & Williams, 1995; Noorani & Carpenter, 2016. Such transformation stretches the ends of the ordinate axis in such a way as to generate a straight line if the data is indeed Gaussian, as it can be seen in Fig. 9.23b. Since the latency uses a reciprocal scale, this is known as a **reciprobit plot** (because it combines a *reci*procal and a *probit* scale). Such plot provides at a glance the visual impression of the two parameters describing the normal distribution of reciprocals, the mean μ_r and the variance σ_r^2: μ_r corresponds to the median of the plot (where the line intersects $p = 50\%$), and the standard deviation σ_r to its slope (a steep line has a small variance, a shallow one a greater one).

The basic idea that reciprocal latencies follow a Gaussian or normal distribution and that these reciprocals have equal variability around a mean value μ_r, as captured by the representations we have introduced above, lies at the heart of the LATER model. The LATER model (“Linear Approach to Threshold with Ergodic Rate” (Carpenter & Williams, 1995; but see Noorani and Carpenter (2016) for a recent review) is one of the simplest, and yet one of the most elegant and powerful models of reaction time distributions in decision tasks: it is assumed that some decision signal is accumulated over time at a constant rate of rise r until a threshold is reached, at which point a response is triggered (Fig. 9.24, left panel). Crucially, the rate r at which such decision signal accumulates is normally distributed across trials. In mathematical terms, the model is easily specified. If

1. the response is triggered when the evidence - starting from a resting level S_0 - reaches a threshold level S_T, and
2. evidence accumulates at a constant rate r which, across trials, follows a normal distribution, $\mathcal{N}(\mu_r, \sigma_r^2)$,

then the response latency T is determined by:

$$T = \frac{S_T - S_0}{r}. \tag{9.85}$$

If one further assumes that both S_0 and S_T are relatively constant across trials, then the distribution of the times is the reciprocal of a normal distribution:

$$\frac{1}{T} = \mathcal{N}\left(\frac{\mu_r}{S_T - S_0}, \left(\frac{\sigma_r}{S_T - S_0}\right)^2\right), \tag{9.86}$$

which Carpenter terms the **Recinormal distribution** (Carpenter & Williams, 1995; Noorani & Carpenter, 2016).

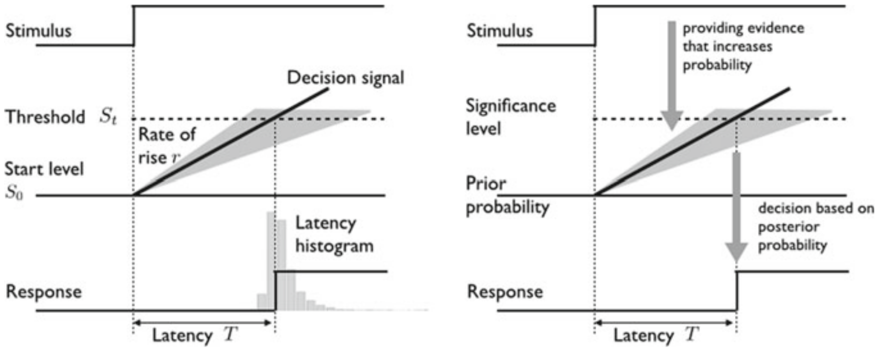


Fig. 9.24 The LATER model (Carpenter & Williams, 1995). Left: the essential model. Right: LATER as an ideal Bayesian decision-maker (see Box 14, for a formal discussion)

The stochasticity behind the LATER model can be described at the *macroscopic level*, in terms of full probability distribution functions, by interpreting LATER as an optimal model of Bayesian decision making, see Box 14 and Fig. 9.24 (right).

The corresponding description at the microscopic level explains the LATER model in terms of a random walk model. In this perspective, we consider the function $S(t)$, as the evidence accumulated at time t starting from prior level $S(0) = S_0$ in the process of reaching the threshold S_T. The random "trajectory" in time of $S(t)$ is that of a drifting gaussian random walker as shown in Fig. 9.25. Indeed, it can be formally shown that the average accumulation of evidence $<S(t)>$ with mean rate r is described by a normal distribution centred at a mean $S_0 + rt$ having variance at time t equal to $\sigma_r^2 t^2$ (Moscoso del Prado Martin, 2008). Then the microscopic behaviour of the random walker $S(t)$ is described by a Langevin-type SDE of drift r and diffusion coefficient $\sigma_r \sqrt{2t}$:

$$dS(t) = rdt + \sigma_r \sqrt{2t} dW(t), \tag{9.87}$$

where $W(t)$ is the standard Wiener process with linear drift. Thus LATER can be considered a non-linear version of the Drift Diffusion Model (Ratcliff & McKoon, 2008) of decision making. Note how the random walk model provides the microscopic dynamics of how the long-tail skewed distribution of reaction times is originated (Fig. 9.25).

One general implication of all this is that the large random variation observed in latencies is not the result of noise at the input, as has commonly been supposed, but represents a gratuitous, "deliberate" randomising device (summarised by the random term $dW(t)$ in Eq. 9.87). It has been surmised that its purpose is to prevent the generation of undesirably stereotyped behaviour - a roulette-wheel within the brain, that one may or may not care to think of as the basis for the sense of free will and creativity (Noorani & Carpenter, 2016).

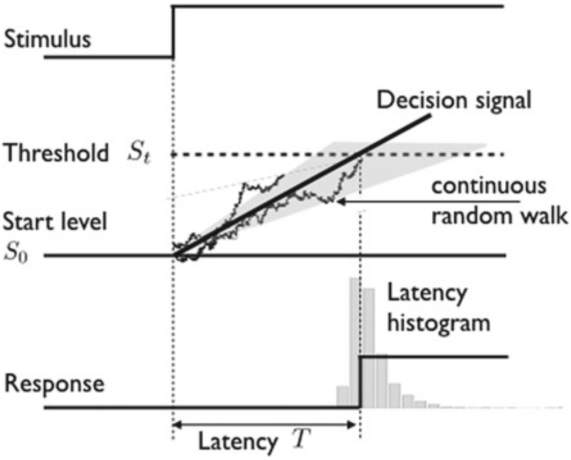


Fig. 9.25 The LATER model as a diffusion model of decision making based on continuous random walks generated through the Langevin-type SDE of Eq. 9.87. Two trajectories only are shown for graphical clarity

Eventually, it is worth pointing out that the predictions of latencies from random walk models and from the original LATER model are very similar, and in fact can only really be distinguished when there are a vast number of data points for latencies.

Box 14: LATER as a Bayesian decision-maker

Classic, **frequentistic hypothesis testing** considers the **likelihood ratio** $\frac{P(\mathcal{D}|\mathcal{H})}{1-P(\mathcal{D}|\mathcal{H})}$, where $P(\mathcal{D} \mid \mathcal{H})$ is the likelihood of the hypothesis $\mathcal{H}$ being tested, e.g., $\mathcal{H} =$ “the stimulus is present”, and $\mathcal{D}$ is the evidence provided by the stimulus. In **Bayesian hypothesis testing**, posterior and prior probabilities need to be taken into account. Denote $P(\mathcal{H})$ the prior probability of the hypothesis $\mathcal{H}$ and $P(\mathcal{H} \mid \mathcal{D})$ the posterior probability of the hypothesis given the evidence $\mathcal{D}$.

At the macroscopic description level, LATER can be directly interpreted as an optimal Bayesian model of hypothesis testing as sketched in the right panel of Fig. 9.24.

To such end, simply rewrite Eq. (9.85) as

$$S(T) = S(0) + rT \tag{9.88}$$

Then, by using Bayes’ rule, we can rewrite the LATER parameters in Eq. (9.88) in terms of **log-odds** (namely, the log-ratio of the probability that an event will happen to the probability that the event will not happen):

$$S(T) = \log \frac{P(\mathcal{H} \mid \mathcal{D})}{1 - P(\mathcal{H} \mid \mathcal{D})} = \overbrace{\log \frac{P(\mathcal{H})}{1 - P(\mathcal{H})}}^{S(0)} + \int_0^T \overbrace{\log \frac{P(\mathcal{D} \mid \mathcal{H})}{1 - P(\mathcal{D} \mid \mathcal{H})}}^{r} dt = S(0) + rT, \tag{9.89}$$

where the starting level $S(0)$ denotes the **log-prior odds** $\log \frac{P(\mathcal{H})}{1-P(\mathcal{H})}$ and the rate of information intake r is the **log Bayes factor** of the stimulus.

Thus, the accumulated evidence $S(T)$ is an optimal estimate of the *log-posterior odds* $\log \frac{P(\mathcal{H}|\mathcal{D})}{1-P(\mathcal{H}|\mathcal{D})}$ (a.k.a., the logit of the posterior probability) of the hypothesis being tested.

9.5.4 Walking on the Safe Side: The Central Limit Theorem

Once more, consider Eq. (9.71), namely the discretised version of the Wiener process. We rewrite here - with step index $n = i + 1$ - for the reader’s convenience:

$$x_n = x_{n-1} + \sqrt{2D}\Delta W_{n-1}. \tag{9.90}$$

It is easy to see that by repeated substitution and by assuming the initial condition $x_0 = 0$, after n integration steps:

$$\begin{aligned} x_n &= x_{n-1} + \sqrt{2D}\Delta W_{n-1} \\ &= x_{n-2} + \sqrt{2D}\Delta W_{n-2} + \sqrt{2D}\Delta W_{n-1} \\ &= x_{n-3} + \sqrt{2D}\Delta W_{n-3} + \sqrt{2D}\Delta W_{n-2} + \sqrt{2D}\Delta W_{n-1} \\ &= \dots \\ &= \sum_{i=0}^{n-1} \sqrt{2D}\Delta W_i \end{aligned} \tag{9.91}$$

Note that by definition of Brownian motion, the Wiener increment ΔW_i is zero-mean Gaussian distributed with variance $\sigma^2 = \Delta t_i$, that is $\Delta W_i \sim \mathcal{N}(0, \Delta t_i)$. By using elementary properties of Gaussian RVs, the increment ΔW_i can be computed by (i) sampling $\xi_i \sim \mathcal{N}(0, 1)$ and (ii) multiplying ξ_i by the standard deviation $\sigma = \sqrt{\Delta t_i}$, namely:

$$\Delta W_i = \sqrt{\Delta t_i}\xi_i.$$

Then, Eq. (9.91) can be simply written as

$$x_n = \sum_{i=0}^{n-1} \sqrt{2D\Delta t_i}\xi_i. \tag{9.92}$$

This result is nothing but the answer to our fundamental question, "What is the probability $P(x_n)$ of being at point x_n after n steps?" As before, the final position x_n is a zero-mean Gaussian random variable with variance $\sigma^2 = 2Dt_n$ (assuming initial time $t_0 = 0$). Thus, $P(x_n) = \mathcal{N}(0, 2Dt_n)$.

In short, by exploiting the simple property that the sum of independent Gaussian RVs (here the particle displacements) is a Gaussian RV, we have derived (in discrete form) the result discussed in Box 9: the probability $P(\mathbf{x}, t)$ of a Brownian particle being at location $\mathbf{x}$ at time t is a Gaussian distribution $\mathcal{N}(0, 2Dt)$. Indeed, the Monte Carlo simulation presented in Fig. 9.21, which illustrates both the microscopic trajectories of a large number of Brownian particles and the macroscopic evolution of the solution of the diffusion equation, is a procedure based on Eq. (9.92): (1) find the cumulative sums up to x_n for a large number of walkers, and (2) compute their histogram to approximate $P(\mathbf{x}, t)$.

Though intuitive, this view is an overly simplified picture of the whole story. To see why, consider a very simple kind of discrete random walker, an instantiation of the basic model presented at the very beginning in Box 7, Eq. (9.31).

At fixed time intervals Δt the walker either jumps to the right with probability p or to the left with probability $q = 1 - p$. The intuition is that for $p = \frac{1}{2}$, at each step the walker is tossing a fair coin to make the left/right decision.

The microscopic behaviour of the walker is governed by the following discrete equation:

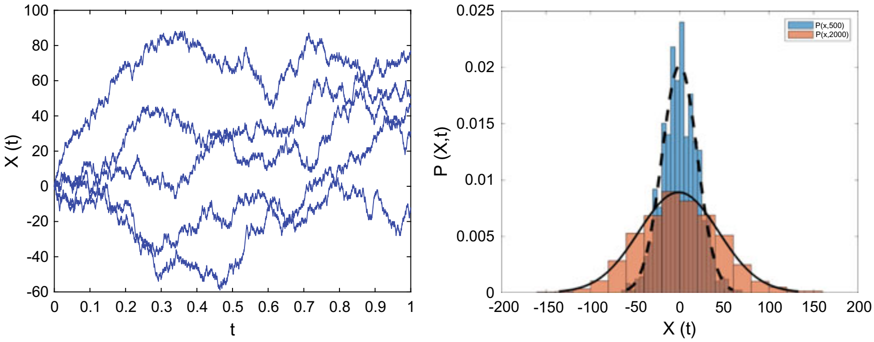
$$x_n = x_{n-1} + (2\xi_n - 1), \;\; n = 1, 2, \ldots, \tag{9.93}$$

where x_0 is the initial position, e.g. $x_0 = 0$. The discrete RV $\xi_n \in \{0, 1\}$ stands for the coin toss at step n, and it is sampled from the Bernoulli distribution, namely $\xi_n \sim Bern(p) = p^{\xi_n} q^{1-\xi_n}$. In other terms $\{\xi_n\}_{n \geq 1}$ is a Bernoulli process.

The dynamics behind Eq. (9.93) is simple: at each step, if $\xi_n = 0$, then Eq. (9.93) gives $x_n = x_{n-1} + 1$ (right step); otherwise, $x_n = x_{n-1} - 1$ (left step). Figure 9.26a shows five trajectories of the simple walker drawn by iterating Eq. (9.93) 2000 steps. For such a long sequence, trajectories look similar to those produced by a Gaussian random walker (Brownian particle). This becomes evident by inspecting the macroscopic behaviour as illustrated in Fig. 9.26b. The collective behaviour of 1000 walkers at a step n is obtained by computing for each walker the sum of a number of i.i.d displacements/increments up to step n, as in Eq. (9.92), then the (empirical) distribution $P(x_n)$ is computed. The figure shows two "snapshots" of the evolution of $P(x_n)$ at $n = 500$ and $n = 2000$: it is easy to see that its behaviour is that of a Gaussian distribution spreading in time, much like the Brownian diffusion behaviour shown in Fig. 9.21. To better visualise such trend, we also plot the Gaussian distributions fitted at the same time steps.

This result is not surprising: the Gaussian or Normal distribution correctly describes an amazing variety of phenomena. Most important for our purposes, the bell-shaped curve appears in nature ubiquitously due to the wide applicability of the **Central Limit Theorem** (CLT).

The key idea behind the classical CLT (but see Box 15 for more details) is the following. If we sum a large number n of RVs X_i that are



(a) Microscopic behaviour (b) Macroscopic behaviour

Fig. 9.26 Microscopic and macroscopic behaviour of the simple random walker. Panel 9.26a shows five among 1000 trajectories simulated via Eq. (9.93), for 2000 iteration steps. Panel 9.26b: the macroscopic behaviour of the many-walker system captured by computing the empirical distributions $P(x_n)$ at steps $n = 500$ and $n = 2000$. Both distributions are overlaid with the Gaussian pdfs fitted at the same time steps (dotted and continuous line, respectively)

1. *statistically independent* and
2. *identically distributed* (i.i.d), and that
3. have a *finite variance*,

the distribution $P(S_n)$ for the sum $S_n = \sum_{i=1}^{n} X_i$ converges to a Gaussian distribution with mean $n\mu$ and variance $n\sigma^2$, namely, $P(S_n) \approx \mathcal{N}(n\mu, n\sigma^2)$.

The CLT theorem generalises the thorough intuition we initially gained by exploiting the Gaussian nature of a Brownian random walk. If a trajectory consists of a set of independent displacements with finite variance, then the total distance covered (the sum of these displacements) follows a Gaussian law whose variance is proportional to the number of these displacements (and, in consequence, proportional to time).

If the three conditions required by the CLT are fulfilled, then the MSD will behave at large times like $\langle x(t)^2 \rangle \propto t$ (no matter how complicated the motion pattern is) in the limit $t \to \infty$. Such scaling law $\langle x(t)^2 \rangle \propto t$, as previously stated, is characteristic of the diffusion equation, but it also arises asymptotically in many other cases. This is not a coincidence but a direct consequence of the baseline CLT.

A tenet of this Chapter, following Paul and Baschnagel (2013), is exploiting the classical CLT as the pivotal concept to distinguish stochastic behaviours that occur within its limits and behaviours that violate its limits. Examples of the latter are considered in the following Section.

Box 15: The Gaussian bell tolls for thee: The Central Limit Theorem

In a nutshell, the simplest form of the CLT states that if one sums together many **independent and identically distributed** (i.i.d) random variables with **finite variance**, then the probability density of the sum will be close to a Gaussian.

Formally, consider n i.i.d RVs $X_1, X_2, \ldots, X_n$, with mean

$$\langle X_i \rangle = \mu$$

and finite variance

$$Var(X_i) = \sigma^2 < \infty$$

for all $i = 1, \ldots, n$. Consider then the RV S_n, which is the sum

$$S_n = X_1 + X_2 + \cdots + X_n. \tag{9.94}$$

Then, as $n \to \infty$, the distribution of the normalised sum $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ converges to the standard Gaussian distribution. More precisely, for $-\infty < a < \infty$

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq a\right) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-\frac{x^2}{2}} dx, \tag{9.95}$$

where $\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} = \mathcal{N}(0, 1)$ is the standard Normal PDF with zero mean and unit variance.

The two most common manifestations of the CLT are the following. First, as $n \to \infty$, the sum S_n "tends" to a Gaussian variable with distribution $\mathcal{N}(n\mu, n\sigma^2)$. Second, the same holds for $\frac{S_n}{n}$, where

$$\frac{S_n}{n} = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i,$$

which is nothing but the empirical mean or sample average. Namely, $\frac{S_n}{n}$ "tends" to the distribution $\mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$.

If these issues might seem too abstract or of limited interest for you, recall that you have been trained in your daily lab work to collect results from several repeated measurements, to take mean values and to eventually estimate confidence intervals by using tables from the Gaussian distribution. Your data do not necessarily follow a Gaussian distribution, yet you are not worried about. Any problem where the final output results from the average over a set of identical and independent variables with finite variance leads to a Gaussian PDF. More often than not, the CLT is your unconscious safety belt.

This is the "baseline" or classical CLT (a.k.a. the Lindeberg-Lévy CLT). However, it is worth mentioning that the CLT comes in various stronger and weaker forms (Schuster, 2016). Under certain conditions, the RVs X_i are not required to be identically distributed. In such case we have different variances, $Var(X_i) = \sigma_i^2$ but the CLT still holds if the contribution of any individual random variable to the overall variance $\sigma_n^2 = \sum_{i=1}^{n} \sigma_i^2$ is arbitrarily small for $n \to \infty$ (Lindeberg's condition).

What happens if the X_i are originated from a distribution whose variance is not finite? It is indeed the case that there are experimental time series (e.g., saccades in free viewing, go back to Fig. 9.4), for which the distribution of independent increments exhibit much fatter tails than the normal, and sometimes considerable skewness as well. In Fig. 9.8 we have seen one such example: the Cauchy distribution with a power-law tail (see Box 16). The Cauchy distribution does not have a finite variance. The Generalized CLT - due to Gnedenko, Kolmogorov and Lévy - states that the sum of a number of RV with a power-law tail will tend to a non Gaussian α-stable distribution (stable Lévy noise) as $n \to \infty$ (Kolmogorov & Gnedenko, 1954; Paul, 1954). We will see in Sect. 9.6.2, that such kind of heavy-tailed distributions play an important role in the modelling of saccadic eye movements.

9.6 Walking on the Wild Side: Eye Movements Beyond the CLT

In spite of the nice behaviour of RWs patrolled by the CLT, when dealing with eye movements most interesting cases happen when the CLT is violated:

(i) Violation of independency: Long-range correlations are present, so once the random walker decides moving in one direction it keeps on doing the same for a long time (this will lead to superdiffusion) or, alternatively, once it stops it remains resting for an arbitrarily long time (then subdiffusion will emerge)
(ii) Violation of identity: motion consists of non-identical displacements that become gradually shorter (subdiffusion) or longer (superdiffusion) probably because of external constraints, and non stationarity arises.
(iii) Violation of moment finiteness: The displacements forming the trajectory can be fitted to a PDF with non-finite mean or variance, so as a result arbitrarily large displacements are likely with a certain frequency (long tail distributions)

When one of the three conditions is violated then the process is said to exhibit **anomalous diffusion**. A simple way to define anomalous diffusion or an anomalous random walk is when $\langle x(t)^2 \rangle$ does not increase linearly with time. In such case, the corresponding MSD shows a power-law behaviour

$$MSD = kt^{\gamma}, \tag{9.96}$$

with $\gamma \neq 1$.

9.6.1 *A First Violation: i.i.d Denied*

One way to anomalous diffusion is by introducing "memory" effects in the process. This gives rise to **long-range** power–law autocorrelations in the underlying noise that drives the random walk. Long-range memory effects violate the condition of independent random variables.

One intuitive example is the **self-avoiding walk** (SAW). In this process the random walker has to keep track of the whole history of his path while he moves along, since he is not allowed to visit a site twice. Sites already visited therefore act like a repulsive potential for the continuation of the walk; Fig. 9.27 illustrates a 2-dimensional SAW. Intuitively, this "long–range repulsive interaction" along the path should make the overall displacement grow stronger with increasing t than in the case of the Bm (Paul & Baschnagel, 2013). **Fractional Brownian motion** (fBm) is another example. It was introduced by Mandelbrot and van Ness (1968) to account for processes obeying a scaling law of the functional form $MSD = kt^{2H}$, with $0 < H < 1, \quad H \neq \frac{1}{2}$, where $H = \frac{1}{2}$ is the special case of Bm (cfr. Box 12).

It is described by the time-varying pdf

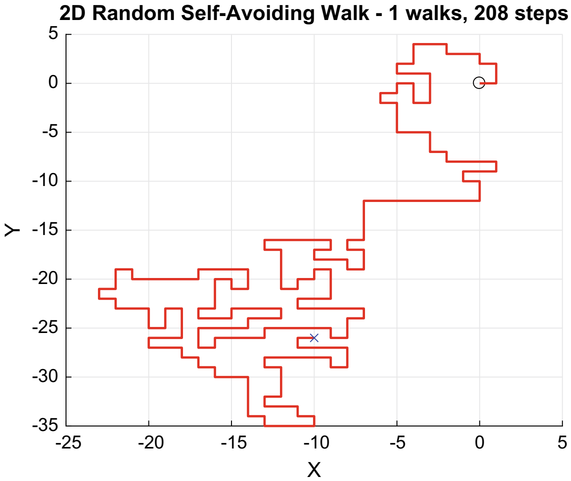


Fig. 9.27 A single trajectory of a 2-dimensional Self-Avoiding Walk (SAW). Whereas a random walk can intersect its path arbitrarily often, a SAW may occupy each site only once. This leads to an increase in the distance between the start and the end, compared to the classic random walk. Self-avoidance has been proposed as the key mechanism driving the drift observed in fixation tasks (Engbert et al., 2011)

$$P(x,t) = \frac{1}{\sqrt{4\pi D t^{2H}}} \exp\left[-\frac{x^2}{4Dt^{2H}}\right]. \tag{9.97}$$

fBm has been used as a mathematical reference for random-walk analysis of fixational eye movements (FEMs) and for studying their correlations across time (e.g., Engbert et al., 2011). Also, properties of persistence/antipersistence have been exploited, for instance, in analysing optokinetic nystagmus (OKN, Trillenberg, Gross, & Shelhamer, 2001).

9.6.1.1 Case Study: Random Walk Analysis of Microsaccades

In a number of studies, (e.g., Engbert, 2006, Engbert et al., 2011; Makarava et al., 2012) have shown that a typical trajectory generated by the eyes during FEMs exhibits clear features of a random walk.

For instance, on a short time scale (2–20 ms), the RW is persistent, whereas on a long time scale (100–800 ms) it exhibits anti-persistent behaviour. Thus, they observed a time-scale separation with two qualitatively different types of motion. On the short time scale, drift produces persistence and this tendency is increased by the presence of microsaccades. On the long time scale, the anti–persistent behaviour is specifically created by microsaccades. Since the persistent behaviour on the short time scale helps to prevent perceptual fading and the anti-persistent behaviour on the long time scale is error-correcting and prevents loss of fixation, they concluded that microsaccade are optimal motor acts to contribute to visual perception

A more recent model of FEMs has also incorporated self-avoidance as the key mechanism driving drifts observed in fixation tasks (Engbert et al., 2011). The Self-Avoiding walk model encodes history by treating space as a lattice and recording the number of visits to each site: the SAW proceeds by choosing the least-visited neighbour at each step. A SAW simulation is provided in Fig. 9.27. The model proposed in Engbert et al. (2011) also includes a confining potential to keep the random walk near the origin, which is needed for the long-time subdiffusive nature of fixation tasks, as well as a mechanism for triggering microsaccades when occupying highly-visited sites.

9.6.1.2 Case Study: Optokinetic Nystagmus

OKN is a reflexive eye movement with target-following slow phases (SP) alternating with oppositely directed fast phases (FP). For a quick grasp of this kind of eye movement, a video is worth a thousand words, e.g., visit https://www.youtube.com/watch?v=KSJksSA6Q-A. Trillenberg et al. (2001) by measuring FP beginning and ending positions, amplitudes, and intervals and SP amplitudes and velocities, tried to predict future values of each parameter on the basis of past values, using state-space representation of the sequence (time-delay embedding) and local second-order approximation of trajectories. Since predictability is an indication of determinism, this approach allows to investigate the relative contributions of random and deterministic dynamics in OKN. FP beginning and ending positions showed good predictability, but SP velocity was less predictable. FP and SP amplitudes and FP intervals had little or no predictability. FP beginnings and endings were as predictable as randomised versions that retain linear auto-correlation; this is typical of random walks. Predictability of FP intervals did not change under random rearrangement, which also is a characteristic of a random process. They concluded that there is undoubtedly a gross level of deterministic behaviour in OKN. Yet within this range, there is apparently significant random behaviour, with a small amount of predictability. The random behaviour has overlaid on it a form of long-term correlation in the form of anti-persistence. This mixture of dynamics is intriguing and provides a challenge for mathematical modelling efforts, though the physiological meaning of these dynamics is open to conjecture.

9.6.2 A Second Violation: Loosing Your Moments

Even in the absence of correlations, a mechanism for disrupting convergence to Brownian motion in the long time limit is using power–law tailed distributions in the random walk steps (i.e., power–law distributed noise rather than Wiener or similar noise). **Lévy flights** (LFs) are one such mechanism. LFs are stochastic processes characterised by the occurrence of extremely long jumps, so that their trajectories are not continuous anymore. The length of these jumps is distributed according to a Lévy stable statistics with a power–law tail and divergence of the second moment. This peculiar property strongly contradicts the ordinary Bm, for which all the moments of the particle coordinate are finite.

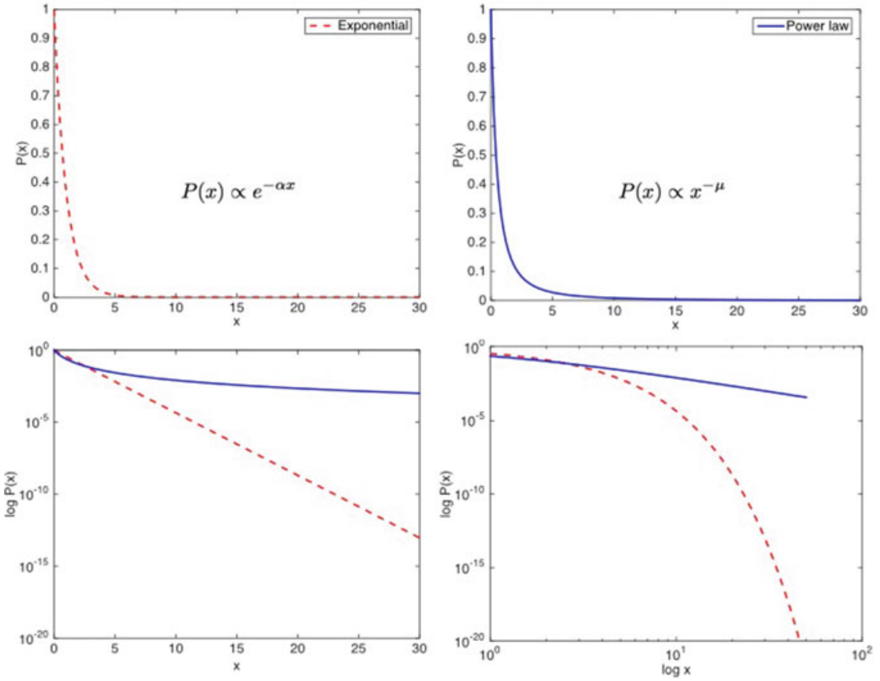


Fig. 9.28 Exponential distribution $P(x) \propto e^{-\alpha x}$ (top-left) vs. power–law distribution $P(x) \propto x^{-\mu}$ (top-right): pdf shapes look apparently similar. However, the exponential pdf is represented as a straight line on a semilog graph of $\log P(x)$ versus x (bottom-left), whilst the power-law shapes as a straight line on a log-log graph (bottom-right), a signature of the heavy-tail behaviour

For a random walker who takes steps of size l according to a probability density function

$$P(l) \approx l^{-\mu}, \tag{9.98}$$

the resulting type of diffusion depends on the value of μ. In particular:

(a) $\mu > 3$: the CLT guarantees convergence to normal diffusion and Brownian regime holds;
(b) $\mu \to 1$: the ballistic motion limit is reached;
(c) $1 < \mu < 3$: superdiffusive behaviour occurs.

LFs arise in the super diffusive regime, when the jump size distribution has a power–law tail with $\mu < 3$. As discussed in Box 16, for such values of the power–law exponent, the RVs can have diverging variance. The necessary and sufficient conditions of the classical CLT do not hold in this case. Lévy flight patterns comprise sequences of randomly orientated straight-line movements. Frequently occurring but relatively short straight-line movement randomly alternate with more occasionally occurring longer movements, which in turn are punctuated by even rarer, even longer

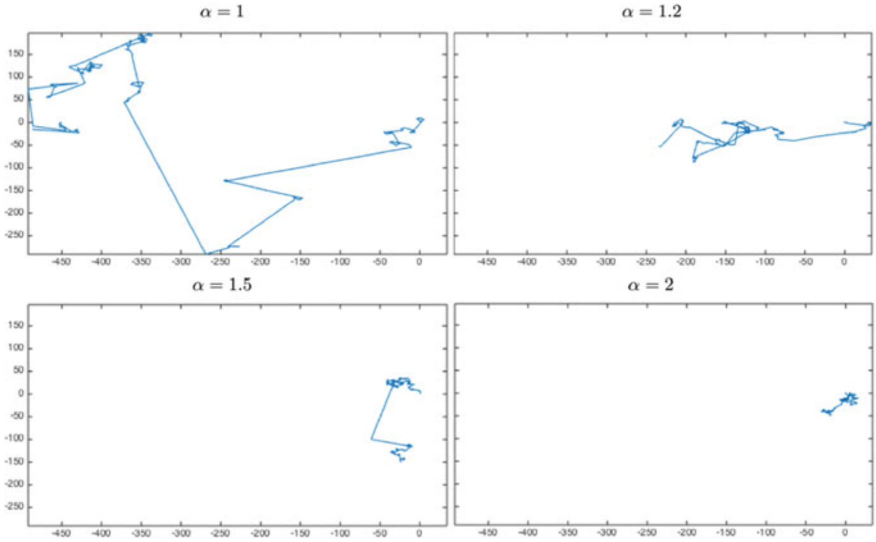


Fig. 9.29 Different 2-dimensional α-stable motions obtained by sampling the "noise" component $\boldsymbol{\xi} \sim f(\boldsymbol{\xi}; \alpha, \beta, \gamma, \delta)$ in Eq. (9.102) for different values of the characteristic index parameter α. The plots shown in the four panels – left to right, top to bottom –, have been generated via $\alpha = 1$, $\alpha = 1.2$, $\alpha = 1.5$, $\alpha = 2$, respectively. The same number of discrete steps (#500) has been fixed for all the examples, but note how the "scale" of the exploration restricts as $\alpha \to 2$, eventually reaching the limit case $\alpha = 2$ where classic Bm is generated (bottom right plot)

movements, and so on with this pattern repeated at all scales. Some examples of LF patterns are provided in Fig. 9.29. As a consequence, the straight-line movements have no characteristic scale, and LFs are said to be **scale-free**.

Box 16: Power–law distribution

Many empirical quantities cluster around a typical value: speeds of cars on a highway, the temperature in Freiburg at noon in February, etc. Distributions of these quantities place a negligible amount of probability far from the typical value, making the typical value representative of most observations. In short, the underlying processes that generate these distributions fall into the general class well-described by the CLT.

Not all distributions fit this pattern, however, and in some cases the deviation is not a defect or problem, but rather an indication of interesting underlying complexity in the generating process. Complex social, biological and technological systems give rise to countless example of "non-normal" and **heavy-tailed** distributions. A power-law distribution is one such kind of probability distribution (see Newman, 2005 for a nice review). When the probability of measuring a particular value of some quantity varies inversely as a power of that value, the quantity is said to follow a power law, also known as the **Pareto distribution**. There are several ways to define them mathematically, one way, for a continuous random variable is the following (Newman, 2005):

$$P(x) = Cx^{-\mu}, x \geq x_{min}, \tag{9.99}$$

where $C = (\mu - 1)x_{min}^{\mu-1}$. Note that this expression only makes sense for $\mu > 1$, which is indeed a requirement for a power–law form to normalize.

Power–law distributions have many interesting mathematical properties. Many of these come from the extreme right-skewness of the distributions and the fact that only the first $(\mu - 1)$ moments of a power-law distribution exist; all the rest are infinite. In general, the k-th moment is defined as

$$\langle x^k \rangle = \int_{x_{min}}^{\infty} x^k P(x) dx = C \int_{x_{min}}^{\infty} x^{-\mu+k} dx = x_{min}^k \left(\frac{\mu - 1}{\mu - 1 - k} \right), \tag{9.100}$$

with $\mu > k + 1$. Thus, when $1 < \mu < 2$, the first moment (the mean or average) is infinite, along with all the higher moments. When $2 < \mu < 3$, the first moment is finite, but the second (the variance) and higher moments are infinite! In contrast, all the moments of the vast majority of other pdfs are finite.

Another interesting property of power-law distributions is **scale invariance**. If we compare the densities at $P(x)$ and at some $P(cx)$, where c is some constant, they are always proportional, i.e. $P(cx) \propto P(x)$. This behaviour shows that the relative likelihood between small and large events is the same, no matter what choice of "small" we make. That is, the density "scales." It is easy to see that if we take the logarithm of both sides of Eq. (9.99):

$$\ln P(x) = \ln C - \mu \ln x, \tag{9.101}$$

the rescaling $x \rightarrow cx$ simply shifts the power-law up or down on a logarithmic scale. The result shows another well-known property of a power-law distribution: it appears as a straight line on a $\log - \log$ plot, as opposed to the strongly curved behaviour of an exponential distribution (see Fig. 9.28).

Inspiring analyses of eye movements and visual search in terms of power–law behaviour and power spectra have been conducted by Aks, Zelinsky, and Sprott (2002) suggesting that our oculomotor system may produce a complex and self-organising search pattern providing maximum coverage with minimal effort.

At the microscopic level, the simulation of individual LF trajectories do not require complex calculations to execute. They are a Markovian process and can be easily obtained from Eq. (9.59), by setting $\mathbf{A} = 0$

$$d\mathbf{x}(t) = \mathbf{B}(\mathbf{x}, t)\boldsymbol{\xi}(t)dt = \mathbf{B}(\mathbf{x}, t)d\mathbf{L}_{\alpha}(t) \tag{9.102}$$

The form of this equation is that of the Wiener process, however in this case the stochastic increment $d\mathbf{L}_\alpha(t) = \boldsymbol{\xi}(t)dt$ is sampled from an α-**stable distribution** $f(\boldsymbol{\xi}; \alpha, \beta, \gamma, \delta)$ (cfr. Box 17):

$$\boldsymbol{\xi}(t) \sim f(\boldsymbol{\xi}; \alpha, \beta, \gamma, \delta). \tag{9.103}$$

In other terms, $d\mathbf{L}_\alpha(t)$ in the context of Eq. (9.102) represents an infinitesimal Lévy motion.

The macroscopic description of the pdfs for particles undergoing a Lévy flight can be modeled using a generalised version of the Fokker-Planck equation. The equation requires the use of fractional derivatives and we will not discuss it here since really beyond the scope of an introductory chapter.

By discretising and iterating Eq. (9.102), over a large number of trials a Lévy flight will be distributed much farther from its starting position than a Brownian random walk of the same length (see again Fig. 9.29). Indeed, the MSD of a Brownian walker has a linear dependence on time whereas that of a Lévy flier grows faster and depends on time raised to some power > 1. This result gives a precise meaning to their characterisation as "super-diffusive." The probability density function for the position of the walker converges to a Lévy α-stable distribution with Lévy index $\alpha = \mu - 1$, with $0 < \alpha \leq 2$ (with the special case $\alpha = 2$ corresponding to normal diffusion). The Lévy index is an important feature. For instance, it has been shown that it is suitable to characterise the variation or activity of fMRI signals of different networks in the brain under resting state condition. Visual and salience networks seem to present a definite Lévy motion-like behaviour of activity, whereas areas from the cerebellum exhibit Brownian motion (Costa, Boccignone, Cauda, & Ferraro, 2016).

The Hurst exponent H, the characteristic index α, and the power–law exponent μ are related as follows:

$$H = \frac{1}{\alpha} = \frac{1}{\mu - 1}. \tag{9.104}$$

Thus, rephrasing the conditions that have been discussed for the μ exponent, for $\alpha < 2$ one cannot define the MSD because it diverges. Instead, one can study moments of order lower than α because they do not diverge. Nevertheless, one can define some "empirical" width, such as half widths at half maximum, and show that a *pseudo*-MSD grows as $\approx t^{\frac{1}{\alpha}}$ for Lévy flights.

Box 17: Stable distributions

The family of α-stable distributions (Gnedenko & Kolmogórov, 1954) form a four-parameter family of continuous probability densities, say $f(\boldsymbol{\xi}; \alpha, \beta, \gamma, \delta)$. The parameters are the skewness β (measure of asymmetry), the scale γ (width of the distribution) and the location δ and, most important,the **characteristic**

exponent α, or index of the distribution that specifies the asymptotic behaviour of the distribution. The relevance of α derives from the fact that the pdf of jump length scales, asymptotically, as $l^{-1-\alpha}$. Thus, relatively long jumps are more likely when α is small. By sampling $\boldsymbol{x} \sim f(\boldsymbol{x}; \alpha, \beta, \gamma, \delta)$, for $\alpha \geq 2$ the usual Bm occurs; if $\alpha < 2$, the distribution of lengths is "broad" and the so called Lévy flights take place.

One example of α-stable motions generated for varying the α index is illustrated in Fig. 9.29.

A random variable X is said to have a stable distribution if: (1) the parameters of its probability density function $f(x; \alpha, \beta, \gamma, \delta)$ are in the following ranges $\alpha \in (0; 2]$, $\beta \in [-1; 1]$, $\gamma > 0$, $\delta \in \mathbb{R}$ and (2) if its characteristic function $E\left[\exp(itx)\right] = \int_{\mathbb{R}} \exp(itx) dF(x)$, F being the CDF, can be written as

$$E\left[\exp(itx)\right] = \begin{cases} \exp\left[-|\gamma t|^{\alpha}\left(1 - i\beta\frac{t}{|t|}) \tan(\frac{\pi\alpha}{2}\right) + i\delta t\right] \\ \exp\left[-|\gamma t|\left(1 + i\beta\frac{2}{\pi}\frac{t}{|t|}\ln|t|\right) + i\delta t\right] \end{cases}$$

The first expression holds if $\alpha \neq 1$, the second if $\alpha = 1$.

Special cases of stable distributions whose pdf can be written analytically, are given for $\alpha = 2$, the **Normal distribution** with

$$f(x; 2, 0, \frac{\sigma}{\sqrt{2}}, \mu) = \mathcal{N}(x; \mu, \sigma^2), \tag{9.105}$$

for $\alpha = 1$, the **Cauchy** or **Lorentz distribution**

$$f(x; 1, 0, \gamma, \delta) = \frac{1}{\pi\gamma}\left[\frac{\gamma^2}{(x-\delta)^2 + \gamma^2}\right], \tag{9.106}$$

and for $\alpha = 0.5$, the **Lévy distribution**

$$f(x; 0.5, 1, \gamma, \delta) = \sqrt{\frac{\gamma}{2\pi}}\frac{\exp -(\frac{\gamma}{2(x-\delta)})}{(x-\delta)^{3/2}}. \tag{9.107}$$

For all other cases, only the characteristic function is available in closed form, and numerical approximation techniques must be adopted for both sampling and parameter estimation (Chambers, Mallows, & Stuck, 1976; Koutrouvelis, 1980; Nolan, 1997). A very nice and simple to use Matlab package for parameter inference and computation of α-stable distributions is freely downloadable at Mark Veilette's homepage http://math.bu.edu/people/mveillet/html/alphastablepub.html.

Some examples of α-stable pdfs and related **complementary cumulative distribution function** (CCDF) are given in Fig. 9.30. The use of the CCDF, or upper tail, of jump lengths is the standard convention in the literature, for the

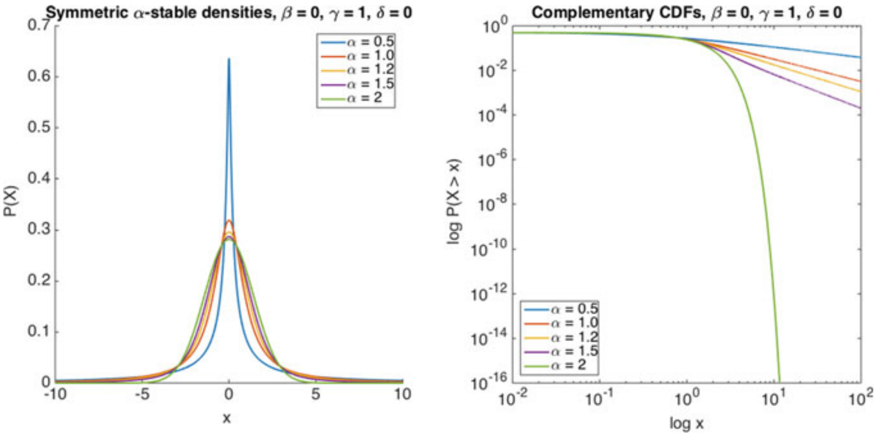


Fig. 9.30 Plots of symmetric α-stable distributions (left) and their complementary CDF (CCDF) on $\log - \log$ axes (right) for different values of the characteristic index parameter $\alpha = 0.5, 1, 1.2, 1.5, 2$. The CCDF shows the rapid fall-off of the tail of the Gaussian case ($\alpha = 2$) as opposed to the power–law tail behaviour of actual Lévy flights $\alpha < 2$

sake of a more precise description of the tail behaviour, i.e. the laws governing the probability of large shifts. The CCDF is defined as $\overline{F}_X(x) = P(X > x) = 1 - F_X(x)$, where $F_X(\cdot)$ is the CDF (cfr., Box 13).

9.6.2.1 Case Study: The Lévy Flight of Saccades

Brockmann and Geisel (2000) have assumed a power–law dependence in the tail of the saccade amplitude distribution, for which they found empirical support in free viewing of natural scenes. Minimisation of the time needed to scan the entire visual space then led them to predict that eye movement trajectories behave as Lévy flights, as opposed to more common diffusive random walks, which would result from a Gaussian amplitude distribution. But in order to obtain simulated eye trajectories that look like their observed scan paths, an empirical determination of a salience field for the correspondingly viewed scene is still needed. Brockmann and Geisel derived that salience field from the spatial distribution of fixations made by observers throughout the scene (a picture of a party). As to the amplitude distribution they considered the Cauchy distribution (Eq. 9.106).

The stochastic assumptions of saccade generation made by Brockmann and Geisel (2000) involve a Markovian process, consistent with an interpretation of visual search originally proposed by Horowitz and Wolfe (1998). However, the predictions and results of the Brockmann and Geisel model do not change substantially if those assumptions are relaxed so as to allow a sufficiently rapidly decaying correlation in the saccade sequences.

Further evidence and characterisation of Lévy-like diffusion in eye movements associated with spoken-language comprehension have recently been provided by Stephen, Mirman, Magnuson, and Dixon (2009).

9.6.2.2 Case Study: The Microsaccade Conundrum

Martinez-Conde et al. (2013) and Otero-Millan, Macknik, Langston, and Martinez-Conde (2013) have put forward the proposal that microsaccades and saccades are the same type of eye movement (the "continuum hypothesis"). The microsaccade–saccade continuum is sustained by evidence that saccades of all sizes share a common generator.

In this respect, a straightforward hypothesis on the function microsaccades is that they help to scan fine details of an object during fixation. This hypothesis would imply that fixational eye movements represent a search process. According to this analogy, the statistics of microsaccades can be compared to other types of random searches, namely inspection saccades during free picture viewing (Brockmann & Geisel, 2000).

Given these assumptions, Engbert (2006) checked whether the amplitude distribution of microsaccades and saccades follows a similar law.

To investigate the distribution of microsaccade amplitude in a data set of 20,000 microsaccades, they analysed the tail of the distribution on a double logarithmic scale. They obtained a power–law decay of the tail with exponent $\mu = 4.41$, which would reject the hypothesis of a Lévy flight for microsaccades (requiring $\mu < 3$), if compared with numerical results obtained by Brockmann and Geisel (2000). In turn, this apparently would also lead to reject the "continuum hypothesis". However there are many subtleties that should be taken into account in order to fairly compare two such different analyses (e.g., sampling rate and discretisation of the eye tracking raw data) before reaching a conclusion. But at least this nice piece of work is useful to show how advanced statistical methods for eye movement analysis can address big questions in the field.

9.6.3 The Foraging Perspective

Consider Fig. 9.31: *prima facie*, is seems to illustrate a bizarre jest. However, from eye-movements studies (Canosa, 2009; Dorr et al., 2010; Over, Hooge, Vlaskamp, & Erkelens, 2007; Tatler, Baddeley, & Vincent, 2006; 2008; Tatler & Vincent, 2009; Tatler et al., 2011), there is evidence that eye movement trajectories and their statistics are strikingly similar, with respect to the resulting movement patterns and their statistics to those exhibited by foraging animals, (Codling, Plank, & Benhamou, 2008; Plank & James, 2008; Reynolds, 2008; Viswanathan, Raposo, & da Luz, 2008).

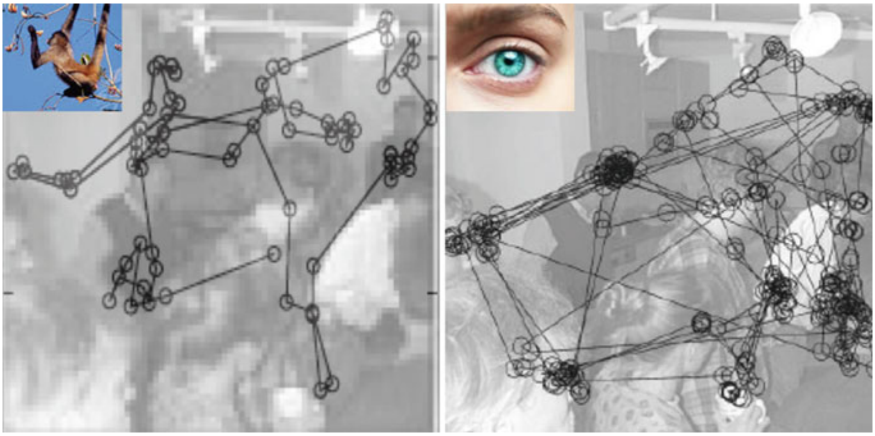


Fig. 9.31 Monkey or human: can you tell the difference? The left image has been obtained by superimposing a typical trajectory of spider monkeys foraging in the forest of the Mexican Yucatan, as derived from Ramos-Fernandez et al. (2004), on the "party picture" used in Brockmann and Geisel (2000). The right image is an actual human scan path (modified after Brockmann & Geisel, 2000)

In other terms, eye movements and animal foraging address in some way a similar problem (Brockmann & Geisel, 2000). Under the foraging metaphor, the eye (and the brain modules controlling the eye behaviour) is the forager, the input visual representation $\mathcal{D}$ is the foraging landscape. Points attracting fixations are foraging sites (in the case of static images) or moving preys (time-varying scenes); gaze shifts occur due to local exploration moves, prey pursuit and long relocation from one site to another.

An intriguing issue is whether the foraging theory underpinning the proposed analyses just provides a useful computational theory metaphor, or constitutes a more substantial ground. Interestingly enough, Hills (2006) has argued that what was once foraging in a physical space for tangible resources became, over evolutionary time, foraging in cognitive space for information related to those resources. Adaptations that were selected for during ancestral times are, still adaptive now for foraging on the internet or in a supermarket, or for goal-directed deployment of visual attention (Wolfe, 2013). In these terms, the foraging approach may set a broader perspective for discussing fundamental themes in eye movement behaviour, e.g., the "continuum hypothesis" of Martinez-Conde et al. (2013) and Otero-Millan et al. (2013).

Building on this rationale, gaze shift models have been proposed coping with different levels of visual representation complexity (Boccignone & Ferraro, 2004, 2013a, b, 2014; Clavelli et al., 2014; Napoletano et al., 2015) and eye movement data analyses have been performed in terms of foraging efficiency (Cain et al., 2012; Wolfe, 2013).

More formally, rewrite the 2-dimensional Langevin equation (9.59), interpreting the deterministic component $\mathbf{A}(\mathbf{x}, t)$ as an external force field due to a potential

$V(\mathbf{x}, t)$ (Boccignone & Ferraro, 2004) (see Fig. 9.32), that is $\mathbf{A}(\mathbf{x}, t) = -\nabla V(\mathbf{x}, t)$, where the “del” (or “nabla”) symbol ∇ denotes the gradient operator.[9]

Then,

$$d\mathbf{x}(t) = -\nabla V(\mathbf{x}, t)dt + \mathbf{B}(\mathbf{x}, t)d\mathbf{L}_\alpha(t). \tag{9.108}$$

Equation (9.108) now provides a microscopic description (trajectory) of a RW biased by an external force field.

We can thus generalise to two dimensions the discretisation method used to obtain the 1−dimensional Eq. (9.55) so to gain an operative definition of the SDE (9.108)

$$\overbrace{\mathbf{x}_{i+1}}^{\text{new gaze location}} = \overbrace{\mathbf{x}_i}^{\text{current gaze location}} - \overbrace{\nabla V(\mathbf{x_i}, t_i)\Delta t_i}^{\text{external force}} + \overbrace{\mathbf{B}(\mathbf{x}_i, t_i)(\Delta t_i)^{\frac{1}{\alpha}}\boldsymbol{\xi}_i}^{\text{Lévy motion}} \tag{9.109}$$

which makes clear that next gaze position is obtained by shifting from current gaze position following a Lévy displacement that is constrained by the external potential field. The external potential summaries the informative properties of the “visual landscape” of the forager.

For instance in Boccignone and Ferraro (2004) Eq. (9.109) was used as a *generative model* of eye movements. In that case, the external potential was taken as a function of the salience field and $\boldsymbol{\xi}_i$ was sampled from a Cauchy distribution. Each sampled gaze shift was then accepted according to a Metropolis-like algorithm (cfr. Box 4), governed by a “temperature” parameter suitable to tune the randomness of the visual exploration (i.e., the attitude of the forager to to frequently engage in longer relocations/saccade rather then keeping with local/fixational exploration). One example is provided in Fig. 9.32.[10]

As to the parameters of Eq. (9.109), it is worth noting that established numerical techniques are available for fitting such parameters from real data (see, e.g. 2001). Once the parameters have been learned, the generative capabilities of Eq. (9.109) can be straightforwardly used to Monte Carlo simulate gaze shifts whose characteristics can then be compare with human data. For instance in a recent paper by Liberati et al. (2017), it is shown that scan paths of children with typical development (TD) and Autism Spectrum Disorder (ASD) can be characterised by eye movements geometrically equivalent to Lévy flights, but with a different degree of randomness, which can be captured by a temperature parameter as in Boccignone and Ferraro (2004).

As previously discussed, the heavy-tailed distributions of gaze shift amplitudes are close to those characterising the foraging behaviour of many animal species. Lévy flights have been used to model optimal searches of foraging animals, namely their moment-to-moment relocations/flights used to sample the perceived habitat

---

[9] A salience map, and thus the potential field V derived from salience, varies in space (as shown in Fig. 9.32). The map of such variation, namely the rate of change of V in any spatial direction, is captured by the vector field ∇V. To keep things simple, think of ∇ as a “vector” of components $(\frac{\partial}{\partial x}, \frac{\partial}{\partial y})$. When ∇ is applied to the field V, i.e. $\nabla V = (\frac{\partial V}{\partial x}, \frac{\partial V}{\partial y})$, the gradient of V is obtained.

[10] Matlab software for the simulation is freely downloadable at http://www.mathworks.com/matlabcentral/fileexchange/38512.

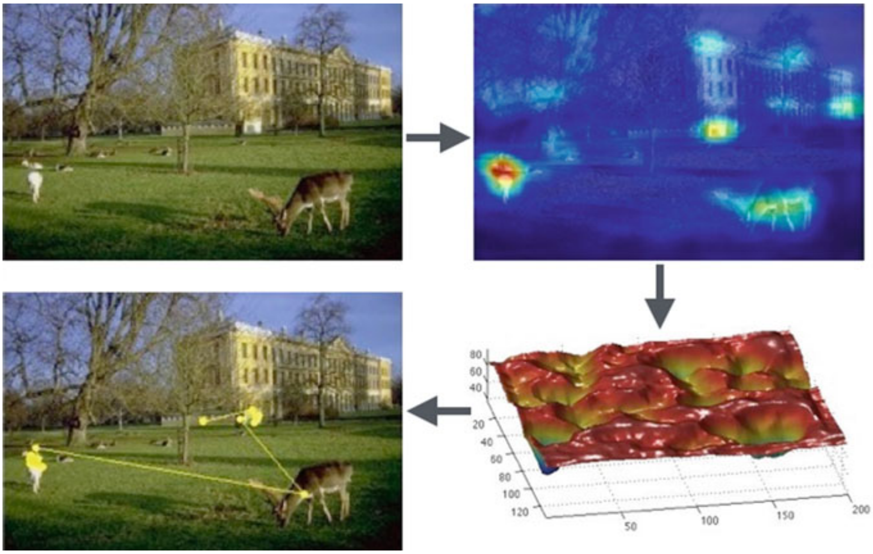


Fig. 9.32 The Lévy model used as a generative model. Top left: the original image. Top right: the salience map. Bottom right: the potential $V(\mathbf{x_i}, t_i)$ computed from saliency: *potential wells* represent informative regions that can attract gaze. Bottom, left: the final scan path superimposed on the original image

(Viswanathan et al., 2008). However, the general applicability of Lévy flights in ecology and biological sciences is still open to debate. In complex environments, optimal searches are likely to result from a *composite strategy*, in which Brownian and Lévy motions can be adopted depending on the structure of the landscape in which the organism moves (Plank & James, 2008). Lévy flights are best suited for the location of randomly, sparsely distributed patches and Brownian motion gives the best results for the location of densely but random distributed within-patch resources (Reynolds, 2008).

Also it is possible to compose the strategy as a hybrid strategy in which Brownian-like motion is adopted for local exploration (e.g., FEMs and small saccades) and Lévy displacements are exploited for long relocations (medium / long saccades) A preliminary attempt towards such a composite strategy for modelling gaze shift mechanisms has been presented in Boccignone and Ferraro (2011). However, that approach only conjectured a simple binary switch between a Gaussian and a Cauchy-like walk. In Boccignone and Ferraro (2014) the approach was generalised to handle observers watching videos and thus accounting for multiple kinds of shifts: FEMs, saccade and smooth pursuit (cfr. Fig. 9.33). To this end, Eq. (9.109) is reformulated as a 2-dimensional dynamical system in which the stochastic part is driven by one-of-K possible types of α-stable motion $\boldsymbol{\xi}_i^k$.

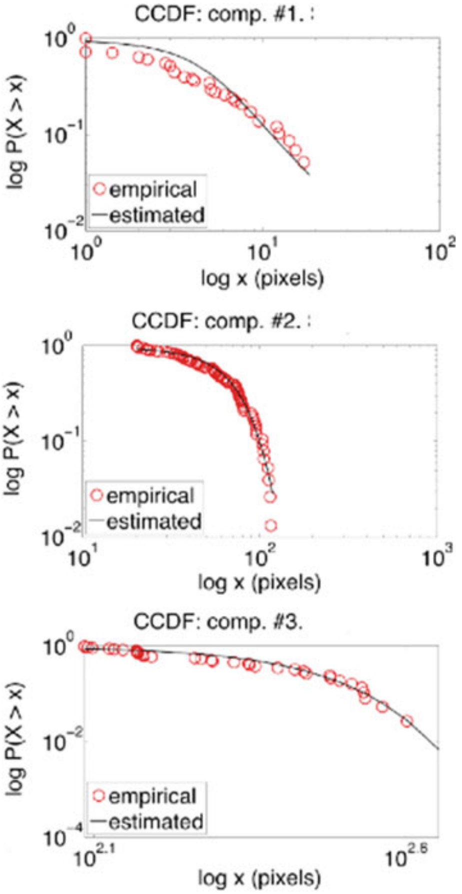


Fig. 9.33 Analysis of gaze shift dynamics from a video. The different components were automatically separated by using a clustering procedure based on the Variational Expectation-Maximization algorithm (see Boccignone & Ferraro, 2014 for details). Then, each component was fitted by an α-stable distribution. Fitting results for one eye-tracked subject are shown in terms of double log plot of the CCDF. From top to bottom: first component accounting for smooth-pursuit and FEMs motions; the medium saccade component; the long saccade component

In the Ecological Sampling model (Boccignone & Ferraro, 2014) the switch from one motion type to the other was bottom-up determined as a multinomial choice biased by the complexity of the sampled visual landscape. More recently this idea was extended to a top-down and task-dependent probabilistic choice in the framework of Bayesian Decision theory (Clavelli et al., 2014). Bayesian Decision theory has gained currency in modelling active sensing behaviour (Yang, Wolpert, & Lengyel, 2016): though minimal, from a theoretical standpoint, a gaze shift action can indeed be considered as the result of a decision-making process (either conscious or unconscious).

9.7 From Patterns of Movement to Patterns of the Mind: Unveiling Observer's Hidden States

The last point we are addressing in this Chapter is: can we infer target hidden states of observer's mind by analysing his eye movement trajectories? Or, in foraging terms, can we say something on forager's internal state by observing his foraging patterns?

Generally speaking, the hidden states that we may target could be, for example, the task the observer is accomplishing, his expertise, his emotional state (but, also, a certain pathology affecting a group of patients as opposed to a control group).

More formally, if $\mathbf{T}$ denotes the target internal state (or a set of states) and $\mathbf{X}$ the visible eye movement behaviour, e.g., $\mathbf{X} = \{\mathbf{x}_F(1), \mathbf{x}_F(2), \ldots\}$, (or, alternatively, the sequence of gaze, amplitudes, directions and durations), one can assume a generative process $\mathbf{T} \rightarrow \mathbf{X}$, where the observer's hidden state shapes the kind of eye trajectories.

In probabilistic terms the generative process can be captured by the simple PGM sketched in Fig. 9.34, which factorises the joint pdf $P(\mathbf{T}, \mathbf{X})$ as $P(\mathbf{T}, \mathbf{X}) = P(\mathbf{X} \mid \mathbf{T})P(\mathbf{T})$ (product rule).

This way, anything we can infer on the hidden state given the observable behaviour is obtained by "inverting the arrow", i.e., by applying Bayes' rule:

$$P(\mathbf{T} \mid \mathbf{X}) = \frac{P(\mathbf{X} \mid \mathbf{T})P(\mathbf{T})}{P(\mathbf{X})}. \tag{9.110}$$

Once the posterior has been computed, then we use decision theory to determine output $\mathbf{T} = t$ for each new input $\mathbf{X} = \mathbf{x}$.

Note that this is a very general formulation of the problem, which actually may entail a large number of solutions, and the PGM shown in Fig. 9.34 could be further specified/specialised in number of ways. Also, to keep the description simple, we have omitted the set of parameters involved by the actual specification of the pdfs in Eq. (9.110). Clearly, before Eq. (9.110) can be put into work, such parameters are to be specified and fitted, or, adopting a more modern term, learned. To this end, a huge amount of **Machine Learning** (ML) techniques are today available (see Box 18, for ML basic terminology in a nutshell, and (Bishop, 2006; Murphy, 2012) for an in-depth presentation).

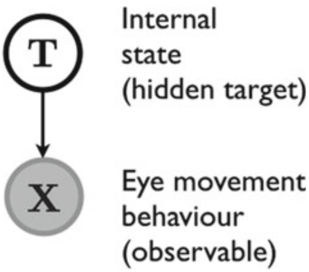


Fig. 9.34 The PGM specifying the generative process $\mathbf{T} \rightarrow \mathbf{X}$ through the joint pdf factorisation: $P(\mathbf{T}, \mathbf{X}) = P(\mathbf{X} \mid \mathbf{T})P(\mathbf{T})$. The shaded node denotes that RV $\mathbf{X}$ is observable

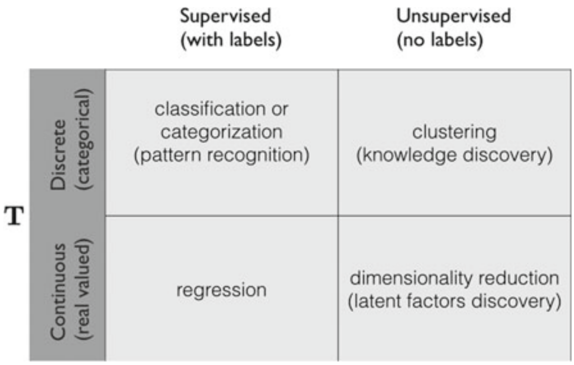
Keeping to such general level, from a methodological standpoint there are at least three distinct approaches to cope with the inverse inference problem posed by Eq. (9.110). In decreasing order of complexity:

1. First solve the inference problem of determining the likelihood function $P(\mathbf{X} \mid \mathbf{T})$ and the prior probabilities $P(\mathbf{X})$. Then use Bayes' theorem in the form given in Eq. (9.110). Equivalently, we can model the joint distribution $P(\mathbf{T}, \mathbf{X})$ directly and then normalise to obtain the posterior probabilities $P(\mathbf{T}, \mathbf{X})$. Approaches that explicitly or implicitly model the distribution of inputs as well as outputs are known as *generative models*, because by sampling from them it is possible to generate synthetic data points in the input space. The popular **Naive Bayes** and Linear Discriminant Analysis methods are very simple instances (though effective in many practical cases) of a generative approach; **Hidden Markov Models** (HMM) Bishop (2006) and Murphy (2012) for modelling time series provide an appealing example of a generative approach that has been often exploited for eye movement analysis.
2. Solve straightforwardly the inference problem of determining the posterior probabilities $P(\mathbf{T}, \mathbf{X})$, and then subsequently use decision theory to assign each new $\mathbf{X} = \mathbf{x}$ to an output target. Approaches that model the posterior probabilities directly are called *discriminative models*. **Logistic regression** is one notable and classic example, **Conditional Random Fields** (CRFs) for modelling time series are more sophisticated one (Bishop, 2006; Murphy, 2012).
3. Find a function $f : \mathbf{X} \rightarrow \mathbf{T}$, called a *discriminant function*, which maps each input $\mathbf{X} = \mathbf{x}$ directly onto a class label. For instance, in the case of two-class problems, e.g., distinguishing experts from novice observers, $\mathbf{T}$ might be binary valued and such that $f = 0$ represents class $\mathbf{T} = t_1$ and $f = 1$ represents class $\mathbf{T} = t_2$. In this case, probabilities play no role. Many popular artificial neural nets or modern methods such as the **Support Vector Machine** (SVM) for regression and classification (a baseline technique in ML) implement this approach (Bishop, 2006; Murphy, 2012).

Note that, since many applications require a posterior class probability, methods based on discriminant functions can be "transformed" into discriminative ones in order to gain an output in probabilistic form. For instance, the output $f(\cdot)$ of a binary SVM classifier can be fed into a sigmoid function, to approximate the posterior (e.g., $P(\mathbf{t} \mid \mathbf{X}) \approx \frac{1}{1+\exp(Af+B)}$, where A, B are parameters that can be determined via regularised maximum likelihood).

Clearly, the generative approach is in principle the most appealing one. However it should be recalled that apart from simple cases such as Naïve Bayes (see Box 19), the normalisation of the joint pdf can be a hard task. Referring again to Eq. (9.110), calculating the normalisation factor $P(\mathbf{X})$ requires the marginalisation $P(\mathbf{X}) = \sum_{\mathbf{T}} P(\mathbf{T}, \mathbf{X})$ ($P(\mathbf{X}) = \int P(\mathbf{T}, \mathbf{X}) d\mathbf{T}$ when RVs are continuous), which, in real cases, is hardly computable. Thus, complex approximation techniques such as Monte Carlo or Variational Bayes are to be taken into account (Bishop, 2006; Murphy, 2012).

Fig. 9.35 The main problems Machine Learning is addressing



T		Supervised (with labels)	Unsupervised (no labels)
	Discrete (categorical)	classification or categorization (pattern recognition)	clustering (knowledge discovery)
	Continuous (real valued)	regression	dimensionality reduction (latent factors discovery)

Box 18: Machine Learning in a nutshell

In the inferential process defined by Eq. (9.110), **T** represents the output of the process and **X** the given input. In Statistical Machine Learning (briefly, ML) terminology **X** is usually shaped as a random vector of features (or attributes, or covariates), $\mathbf{X} = \{\mathbf{X}_i\}_{i=1}^N$, where i is a suitable index. For example when i is a time index, then $\{\mathbf{x}_1, \mathbf{x}_2\}$ is the realisation of a stochastic process. ML does not relate to a specific problem thus $\mathbf{X}_i$ could be a complex structured object, such as an image, a sentence, an email message, a graph, etc.

The form of the output or response variable **T** can be either discrete (categorical, nominal) or continuous (real-valued). One example of the first type, is when **T** can take the label of one of two tasks given to the observer, e.g., $\mathbf{T} = t_k$ where $t_1 =$ "look for people" and $t_2 =$ "look for cars". Another example, is **T** taking values over the discrete set of basic emotion ("fear", "disgust", "joy", etc.). As opposed to this latter example, we could try instead, to infer from eye movements a continuous affect state, so that **T** is taking values t_i in the real valued space of valence and arousal.

From a practical standpoint, when using a ML approach to analyse our eye tracking data we can be in one of these two conditions (Fig. 9.35):

1. **supervised learning**: we know where input $\mathbf{x}_i$ comes from (e.g., $\mathbf{x}_i$ was measured while the observer was scanning a happy face); more formally $\mathbf{x}_i$ is paired with target value or label t_i, thus we have a training set $\mathcal{D} = \{(\mathbf{x}_i, t_i)\}_{i=1}^N$;
2. **unsupervised learning**: we have no labels, and our dataset is represented by the bare input data $\mathcal{D} = \{(\mathbf{x}_i)\}_{i=1}^N$

Thus, in the supervised setting the goal is to learn a mapping from input **X** to output **T**, given a labeled set of input-output pairs. When **T** is discrete the problem is known as **classification** or pattern recognition; when **T** is real-valued, we are performing **regression**.

In the unsupervised setting, we have no labels available, thus the goal is to find "interesting patterns" in the data. This is sometimes named knowledge discovery or data mining. It is a much less well-defined problem, since we are not told what kinds of patterns to look for, and there is no obvious error metric to use (unlike supervised learning, where we can compare our prediction for a given x to the observed value). When **T** is discrete, the problem is known as **clustering**. When **T** is real-valued we are typically in the case of **dimensionality reduction**. The latter is used when dealing with high dimensional data: it is often useful to reduce the dimensionality by projecting the data to a lower dimensional subspace which captures the "essence" of the data. Indeed, although the input data may appear high dimensional, there may only be a small number of degrees of variability, corresponding to **latent factors** (Principal Component Analysis or Factor Analysis being well known examples).

Statistical Machine Learning is nowadays a broad and mathematically sophisticated field. Two excellent and up-to-date textbooks are those by Bishop (2006) and Murphy (2012).

9.7.1 Inverting Yarbus to Infer the Task

We now assume that the target internal state **T** of the observer stems from a given visual task, and for simplicity we straightforwardly use **T** to denote the task. In other terms, the task **T** is the hidden state of interest.

The seminal experiment by Yarbus (1967) studied the effect of the visual task on the trajectories of eye movements **X**. On the basis of our introductory discussion, we know that in probabilistic terms the effect $\mathbf{T} \rightarrow \mathbf{X}$, or forward mapping, is formally captured by the likelihood function $P(\mathbf{X} \mid \mathbf{T})$. Also, we know that Eq. (9.110) is an application of Bayes' rule, to compute the posterior probability of the task **T** after having observed eye movement trajectories **X**. Summing up, the inference of the task requires the computation of the inverse mapping $\mathbf{T} \leftarrow \mathbf{X}$, which is easily understood as an "inverse Yarbus" process (Haji-Abolhassani & Clark, 2013).

Clearly, as previously mentioned, there are several ways of "inverting Yarbus".

9.7.1.1 Case Study: Inverting Yarbus via Naïve Bayes

In a recent study (Henderson et al., 2013), Henderson et al. considered four tasks: scene search, scene memorisation, reading, and pseudo-reading. Task inference was achieved by classifying the observers' task by implementing Eq. (9.110) as the baseline Naïve Bayes' (NB) classifier. Namely they addressed two problems: (i) whether the task associated with a trial could be identified using training from other trials within the same experimental session (within-session classification); (ii) whether the task performed in one session could be identified based on training from a session conducted on a different day (cross-session classification). Twelve members of the University of South Carolina community participated in the experiment. A dedicated classifier was trained for each observer, thus the baseline NB has proved to be sufficient. NB classifiers were trained on a feature vector **X** of dimension 8, i.e., eight eye movement features capturing eye movement patterns for each trial: the mean and standard deviation of fixation duration, the mean and standard deviation of saccade amplitude, the number of fixations per trial, and the three parameters μ, σ, and τ quantifying the shape of the fixation duration distribution with an ex-Gaussian distribution, which is known to change for different eye-movement tasks (cfr., Henderson et al., 2013) for details.

Box 19: Naïve Bayes

The Naïve Bayes algorithm is a classification algorithm based on Bayes rule, that assumes the attributes $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2, \ldots\}$ are all conditionally independent of one another given $\mathbf{T}$. Consider, for example, the two feature case, where $\mathbf{X} = \{\mathbf{X}_1, \mathbf{X}_2\}$, then

$$P(\mathbf{X} \mid \mathbf{T}) = P(\mathbf{X}_1, \mathbf{X}_2 \mid \mathbf{T}) = P(\mathbf{X}_1 \mid \mathbf{X}_2, \mathbf{T}) P(\mathbf{X}_2 \mid \mathbf{T}) = P(\mathbf{X}_1 \mid \mathbf{T}) P(\mathbf{X}_2 \mid \mathbf{T}) \tag{9.111}$$

Thus, if **X** contains n attributes: $P(\mathbf{X} \mid \mathbf{T}) = \prod_{i=1}^{n} P(\mathbf{X}_i \mid \mathbf{T})$. This way Eq. 9.110 can be written as

$$P(\mathbf{T} = t_k \mid \mathbf{X}) = \frac{P(\mathbf{T} = t_k) \prod_{i=1}^{n} P(\mathbf{X}_i \mid \mathbf{T} = t_k)}{\sum_j P(\mathbf{T} = t_j) \prod_{i=1}^{n} P(\mathbf{X}_i \mid \mathbf{T} = t_j)}. \tag{9.112}$$

If we are interested only in the most probable value of **T**, then we have the Naïve Bayes classification/decision rule: $\mathbf{T} \leftarrow \arg\max_{t_k} P(\mathbf{T} = t_k \mid \mathbf{X})$, which simplifies Eq. (9.112) to the following (because the denominator does not depend on t_k):

$$\mathbf{T} \leftarrow \arg\max_{t_k} P(\mathbf{T} = t_k) \prod_{i=1}^{n} P(\mathbf{X}_i \mid \mathbf{T} = t_k) \tag{9.113}$$

9.7.1.2 Case Study: Inverting Yarbus via HMM

In (Haji-Abolhassani & Clark, 2013) Haji-Abolhassani and Clark present a study in which Eq. (9.110) is shaped to explicitly account for the gaze shift sequence as a stochastic process. They present different experiments and models, but here, for clarity sake, we will consider a basic condition and a baseline model so to capture the rationale behind their approach and, also, to compare with the Ellis and Stark model presented in Sect. 9.5.2.1. Recall that in that case the transition matrix A_{kj}

was directly estimated by "counting" the percentage of transitions from one point of interest to another (more formally, via Maximum Likelihood estimation). Thus, the model was an observable Markov model. In the inverse Yarbus setting, it can be represented as the Markov chain conditioned on task **T** that is depicted in Fig. 9.37 (top panel, solution 1). Different tasks are likely to give rise to different transition matrices. In (Haji-Abolhassani & Clark, 2013) visual tasks considered in the simplest experiment were counting red bars, green bars, blue bars, horizontal bars, vertical bars, or characters.

Even if we leave apart subtle issues such as the dissociation between the centre of gaze and the covert focus of attention (Haji-Abolhassani & Clark, 2013), it is very unlikely for a saccade to land exactly on the chosen point of interest (objects or salient locations). The fixation locations may undershoot or overshoot the targets due to oculomotor properties of human eyes or the noisiness of the eye tracker. To account for this problem, in terms of an observable Markov model, a "practical" viable solution is to relax the exact point of interest condition to a more flexible region of interest surrounding the point. This indeed was the solution adopted by Stark and colleagues (Ellis & Stark, 1986; Hacisalihzade et al., 1992), depicted in Fig. 9.17, and which we now recall in Fig. 9.36.

As a more principled alternative (see, Fig. 9.36) one can assume that the exact points of interest, correspond to "hidden" targets or states: when, under a given task, one such target is chosen at time t, say $\mathbf{z}_t$, the corresponding actual fixation $\mathbf{x}_t$ will be generated by adding some noise ϵ (e.g., distributed according to a zero mean Gaussian pdf), i.e. $\mathbf{x}_t = \mathbf{z}_t + \epsilon$. In other terms, we are assuming that $P(\mathbf{x}_t \mid \mathbf{z}_t) = \mathcal{N}(\mathbf{x}_t; \mathbf{z}_t, \Sigma)$. That is, when $\mathbf{z}_t$ is chosen, the actual observation is obtained by sampling from a Gaussian distribution $\mathcal{N}(\mathbf{x}_t; \mathbf{z}_t, \Sigma)$ centred on the true target $\mathbf{z}_t$, where the inverse of the covariance matrix Σ will define the precision of target shooting.

An HMM can be explicitly conditioned on task **T** generalising to the the Dynamic Bayesian Network depicted in Fig. 9.37 (bottom panel, solution 2). The problem of learning such DBN can be further simplified by learning a separate HMM for a given task $\mathbf{T} = t_k$. This way each task will be implicitly defined through the set of parameters defining the corresponding HMM, $\mathbf{T} = t_k \Longleftrightarrow \Theta = \Theta_k$ (cfr, Fig. 9.38).

Eventually, task inference is performed by choosing the HMM providing the higher likelihood for the input observation $\mathbf{x}_{new}$.

A nice study using HMM in the specific context of face exploration modelling has been presented by Coutrot, Binetti, Harrison, Mareschal, and Johnston (2016). In this study, similarly to the example outlined in Fig. 9.36, each hidden state represents specific parts of the face that are likely to be fixated; the actual distribution of eye positions (emission density) is modelled as a 2-D Gaussian distribution. Coutrot, Hsiao, and Chan (2017) have recently a Matlab toolbox freely available to the community to support scan path modeling and classification with HMMs.[11] A more complex example of how to exploitmore general DBNs - namely, an Input-Output Coupled

---

[11] http://antoinecoutrot.magix.net/public/code.html.

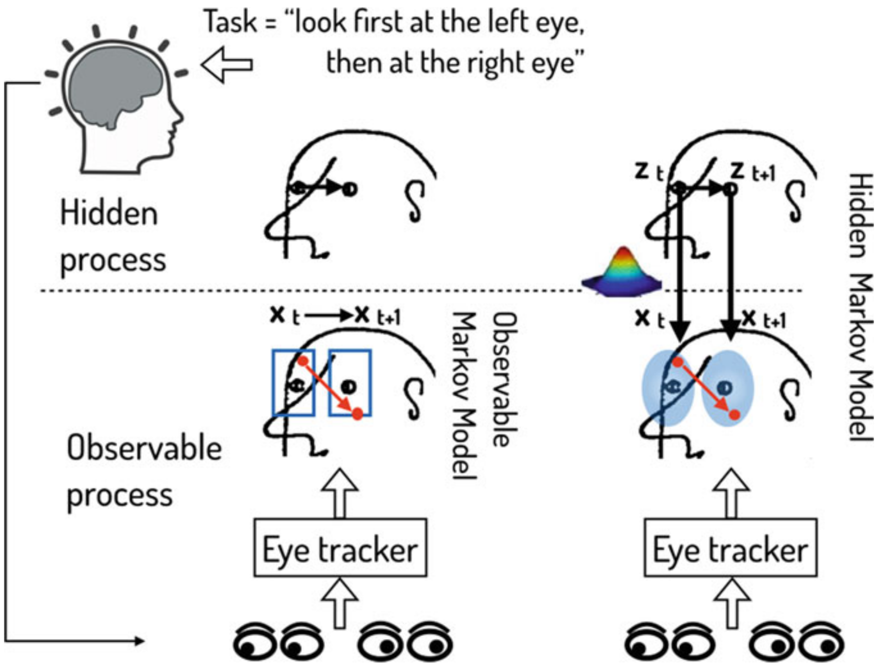


Fig. 9.36 Observable Markov Model vs. HMM. We assume the task "look first at the left eye, then at the right eye" of Klee's drawing presented in Fig. 9.17. The noise-free state space is here $S = \{s_1 = \text{"left eye"}, s_2 = \text{"right eye"}\}$, where, ideally, s_1, s_2 are exactly centred on the eyes. Thus, the "perfect shift" of the mind's eye would be $\mathbf{z}_t = s_1 \rightarrow \mathbf{z}_{t+1} = s_2$. However, as in Plato's cave, we can only observe the noisy and variable eye-tracked shift $\mathbf{x}_t, \rightarrow \mathbf{x}_{t+1}$ as a surrogate. The observable Markov Model simplifies the analysis by considering the two ROIs as a coarse-grained representation of the hidden states s_1, s_2 and assumes all $\mathbf{x}_t$'s falling within the ROI as equivalent. On the contrary, the HMM allows the hidden shift $\mathbf{z}_t \rightarrow \mathbf{z}_{t+1}$ to be part of the model, and the visible shift $\mathbf{x}_t \rightarrow \mathbf{x}_{t+1}$ to be nothing but its sampled noisy realisation (e.g., Gaussian). See text for details

HMM - for dynamically intertwining eye movements and hand actions in a drawing task is provided in Coen-Cagli, Coraggio, Napoletano, and Boccignone (2008) and Coen-Cagli et al. (2009).

9.7.2 *Assessing Cognitive Impairments and Expertise*

Equation (9.110), can be used beyond the important issue of task classification. More generally, the value of $\mathbf{T}$ can represent a label ℓ to identify groups of observers that exhibit different eye movement behaviour with respect to a given task. In these circumstances, Eq. (9.110) formalises the probability that one observer belongs to one group. The posterior can then be used for classification (e.g., via the arg max decision rule).

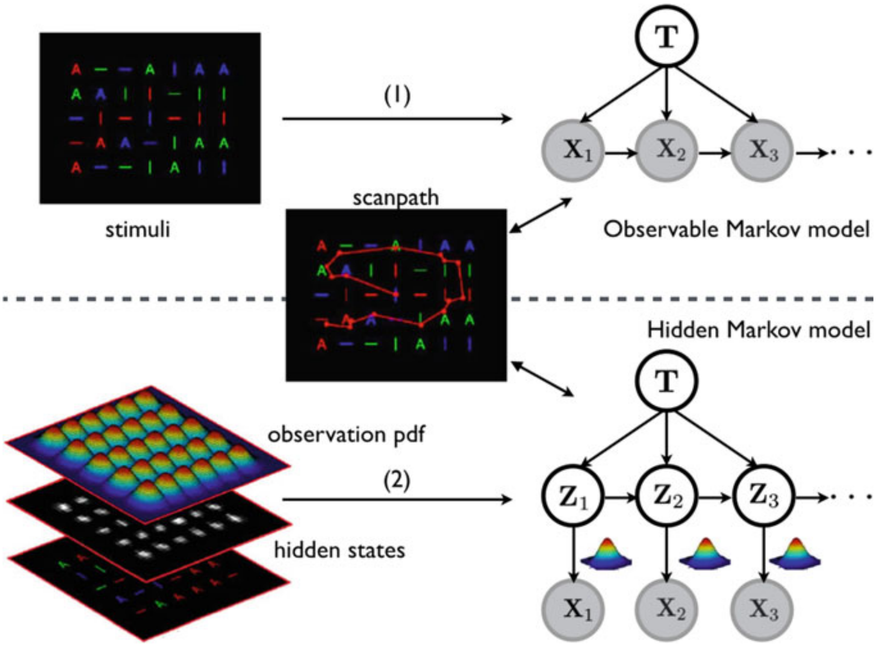


Fig. 9.37 Two solutions for the inverse Yarbus given a time series of observations: (1) the task-conditioned observable Markov model (Ellis & Stark, 1986; Hacisalihzade et al., 1992) (top panel) and the task-conditioned hidden Markov model (Haji-Abolhassani & Clark, 2013) (bottom panel)

9.7.2.1 Case Study: Assessing Cognitive Impairments

On the rationale that patients with mild cognitive impairment (MCI) often progress to Alzheimer's disease (AD), Lagun, Manzanares, Zola, Buffalo, and Agichtein (2011) applied ML methods to analyse and exploit the information contained in the characteristics of eye movement exhibited by healthy and impaired subjects during the viewing of stimuli in the Visual Paired Comparison (VPC) task for the detection of memory impairment associated with MCI. The VPC assessment proceeds in two steps. During the familiarisation phase, subjects are presented with two identical visual stimuli, side by side, on a computer screen. Eye tracked subjects are allowed to look at the pictures for a specified amount of time. During the test phase, subjects are presented with pictures of the old stimulus and a novel stimulus, side by side. Control subjects typically spend 70% of the time during the test phase looking at the novel stimulus, which indicates that they have a memory for the repeated, and now less interesting, stimulus. In contrast, age-matched MCI patients did not spend more time looking at the novel stimulus than the repeated stimulus.

Data analysis was conducted via supervised classification (two class/label problem, **T** taking values in $\ell = \{\text{“impaired”}, \text{“control”}\}$), by exploiting standard techniques, namely Naïve Bayes, Logistic Regression, and the Support Vector Machine.

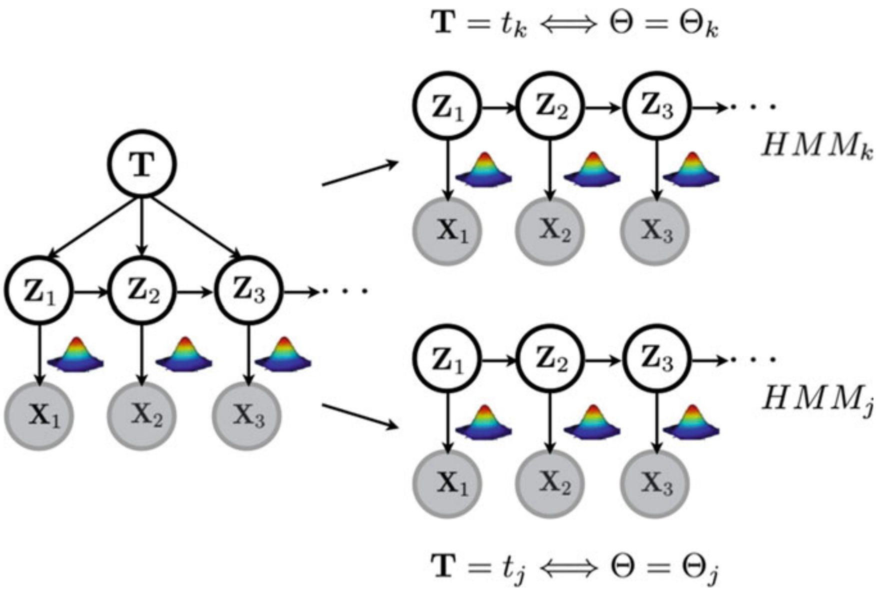


Fig. 9.38 For K tasks, the problem of learning the parameters for the DBN on the left is simplified to learning K simple HMM parameters. In the learning stage, for each task $\mathbf{T} = t_k$ a specific set of parameter Θ_k is from observations, obtaining the k-th HMM. Task inference is performed by choosing the HMM providing the higher likelihood for the input observation $\mathbf{x}_{new}$

They first trained the classification models on the multidimensional representation $\mathbf{X}$ of eye movements from a sample of the impaired and control subjects, $\mathcal{D} = \{\mathbf{x}_{train}, \ell\}$ and then used the model to predict the status of new subjects based on their eye movement characteristics, i.e, $P(\mathbf{T} \mid \mathbf{X} = \mathbf{x}_{new})$. The results showed that eye movement characteristics including fixation duration, saccade length and direction, and re-fixation patterns (gaze position re-visits on previously seen parts of the stimuli) can be used to automatically distinguish impaired and normal subjects. In this study the SVM classifier outperformed the other techniques.

Beyond the specific issue addressed by Lagun et al., it is worth looking at their paper (2011) because it provides a gentle introduction to the Naïve Bayes, Logistic Regression, and SVM algorithms.

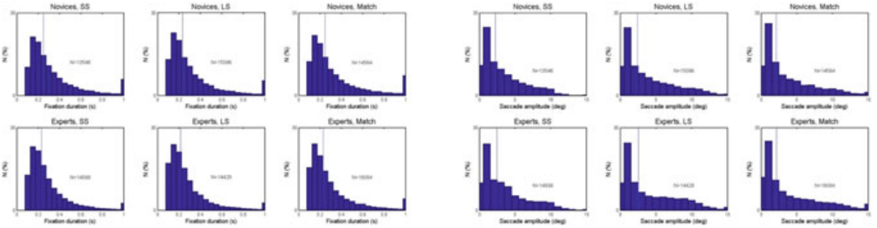
9.7.2.2 Case Study: Classifying Billiard Player Expertise

The study presented in Boccignone et al. (2014) analysed the oculomotor behaviour of individual observers engaged in a visual task, with the aim of classifying them as experts or novices (two class/label problem, $\mathbf{T}$ taking values in $\ell = \{$“expert”, “novice”$\}$). To this end, various visual stimuli and tasks were administered to 42 subjects, half novices and half expert billiard players. Stimuli were a portion of a real match, video-recorded from the top, containing several shots of variable length and complexity, as well as a number of ad-hoc individual shots, also video-recorded from the top in a real setting. The match stimulus was associated to a free-viewing observation condition, while for the individual shots, which were occluded in the final part of the trajectory, observers were asked to predict the outcome of the shot, which placed implicitly a significant constraint on the deployment of visuospatial attention, and, consequently, on the overt scan path.

The input $\mathbf{X}$ was obtained as follows. For each observer, given the sequence of fixations $\{\mathbf{x}_t\}_{t=1}^{N_T}$, where the vector $\mathbf{x}_t$ represents the fixation position (coordinates) at time t, the amplitude and direction of each gaze shift were computed: $\{l_t, \theta_t\}_{t=1}$. Third feature was the fixation duration $\{f_t\}_{t=1}^{N_T}$.

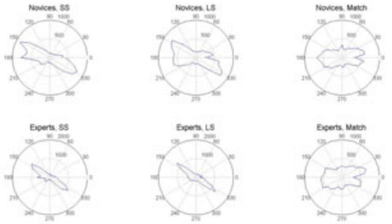
The random sample $\{l_t, \theta_t, f_t\}_{t=1}^{N_T}$ was summarised through the empirical distribution functions (histograms), that is the random vectors $\mathbf{x}^l = \left[x_1^l \dots x_D^l\right]^T$, $\mathbf{x}^\theta = \left[x_1^\theta \dots x_D^\theta\right]^T$ and $\mathbf{x}^f = \left[x_1^f \dots x_D^f\right]^T$, respectively, where the vector dimension D represents the number of bins of the histogram. The feature vector $\mathbf{x}^s$ is thus a summary of the behaviour of a single observer with respect to a particular feature space or source of information $s = 1, \dots S$, here $S = 3$. Thus, eventually, $\mathbf{X} = \{\mathbf{x}^s\}_{s=1}^{S}$.

From Fig. 9.39, note that differences between experts and novices are barely noticeable in terms of features. Clearly, when addressing a scenario in which individual observers are classified as belonging to one or another population, and differ-



(a) Fixation duration

(b) Gaze shift amplitude



(c) Gaze shift direction

Fig. 9.39 Empirical distributions (histograms) of $\{l_t, \theta_t, f_t\}_{t=1}^{N_T}$ used to classify expertise. Top panels (9.39a), fixation duration; middle panels (9.39b), gaze shift amplitude; bottom panels (9.39c), gaze shift direction. Vertical solid lines, median values. SS = Short Shots, LS = Long Shots. Modified after (Boccignone et al., 2014) ences between features are so subtle, more sophisticated ML tools are needed. On this basis, each feature space s was treated as independent and mapped to a specific kernel space (either linear or Gaussian, Murphy, 2012). Then, the posterior $P(\mathbf{T} \mid \mathbf{X})$ was rewritten as $P(t_n|\mathbf{x}_n^1, \ldots, \mathbf{x}_n^S) = P(t_n|\mathbf{W}, \mathbf{k}_n^\beta)$, where the term on the r.h.s is the Multinomial probit likelihood. Here, $\mathbf{W} \in \mathbb{R}^{N \times C}$ is the matrix of model parameters; the variable $\mathbf{k}_n^\beta$ is a row of the kernel matrix $\mathbf{K}^\beta \in \mathbb{R}^{N \times N}$ - whose elements are the $K^\beta(\mathbf{x}_i, \mathbf{x}_j)$, i.e. the different kernels - and it expresses how related, based on the selected kernel function, observation $\mathbf{x}_n$ is to the others of the training set. This way (cfr. Fig. 9.40), sources can be combined within a composite kernel space level and classified through a Relevance Vector Machine (RVM), namely a multiple-kernel RVM (Damoulas & Girolami, 2009; Psorakis, Damoulas, & Girolami, 2010).

Discussing in detail this solution is out of the scope of this chapter (see Boccignone et al., 2014, of a short presentation, and Bishop, 2006 for more details). However, just to give some hints, RVMs can be considered the Bayesian counterpart of SVMs. They are Bayesian sparse machines, that is they employ sparse Bayesian learning via an appropriate prior formulation. Not only do they overcome some of the limitations affecting SVMs but also they achieve sparser solutions (and hence they are faster at test time) than SVM. Indeed, by combining only three basic parameters of visual exploration, the overall classification accuracy, expressed as percent correct and averaged across stimulus types and oculomotor features, scored a respectable 78%. More interesting is to consider the best performance for each stimulus type, which testifies the achievement of the classifier, and which depends on the features used. The best performance ranged between 81.90% and 88.09–1.852 to 2.399 in terms of d', which is an interesting result, especially considering the naturalistic, unconstrained viewing condition.

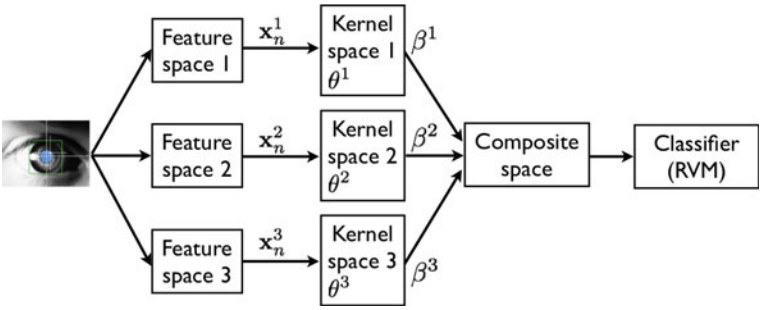


Fig. 9.40 Data analysis in multiple-kernel representation. The fixation sequence is represented in different feature spaces $s = 1, \ldots, S$; each feature $\mathbf{x}^s$ is then separately mapped in a kernel space, each space being generated via kernel K^s of parameters θ^s. The separate kernel spaces are then combined in a composite kernel space, which is eventually used for classification. Modified after (Boccignone et al., 2014)

9.8 A Final Note on the Use of Machine Learning in Visual Attention Modelling

For the sake of completeness, it is worth mentioning that modern Statistical Machine Learning techniques such as those described in the previous Section, are currently adopted even at the earliest stages of visual attention modelling, markedly in salience computation.

Going back to Eqs. 9.11 and 9.15 (Sect. 9.3), as we already pointed out, in the case the prior $P(\mathbf{L})$ is assumed to be uniform (no spatial bias, no preferred locations), then $P(\mathbf{L} = 1 \mid \mathbf{F}) \simeq P(\mathbf{F} \mid \mathbf{L} = 1)$, where $\mathbf{L} = 1$ is now a binary RV (1 or 0) simply denoting location (x, y) as salient/non salient. In such case, the likelihood function $P(\mathbf{F} \mid \mathbf{L} = 1)$ can be determined in many ways; e.g., nonparametric kernel density estimation has been addressed (Seo & Milanfar, 2009), where center/surround local regression kernels are exploited for computing $\mathbf{F}$.

More generally, taking into account the ratio $f(\mathbf{L}) = \frac{P(\mathbf{L}=1|\mathbf{F})}{P(\mathbf{L}=0|\mathbf{F})}$ (or, commonly, the log-ratio) casts the saliency detection problem in a classification problem, in particular a discriminative one, for which a variety of learning techniques are readily available. Kienzle, Wichmann, Franz, and Schölkopf (2006) pioneered this approach by learning the saliency discriminant function $f(\mathbf{L})$ directly from human eye tracking data using an SVM. Their approach has paved the way to a relevant number of works: from Judd, Ehinger, Durand, and Torralba (2009) – where a linear SVM is trained from human fixation data using a set of low, middle and high-level features to define salient locations–, to most recent ones that wholeheartedly endorse ML trends. Henceforth, methods have been proposed relying on sparse representation of “feature words” (atoms) encoded in salient and non-salient dictionaries; these are either learned from local image patches (Lang, Liu, Yu, & Yan, 2012; Yan, Zhu, Liu, & Liu, 2010) or from eye tracking data of training images (Jiang, Xu, Ye, & Wang, 2015). Graph-based learning is one other trend, from the seminal work of Harel, Koch, and Perona (2007) to Yu, Zhao, Tian, and Tan (2014) (see the latter, for a brief review of this field). Crucially, for the research practice, data-driven learning methods allow to contend with large scale dynamic datasets. SVMs are used by Mathe and Sminchisescu (2015) in the vein of Kienzle et al. (2006) and Judd et al. (2009), but they remarkably exploit state-of-the art computer vision datasets annotated with human eye movements collected under the ecological constraints of a visual action recognition task.

As a general comment on (discriminative) ML-based methods, on the one hand it is embraceable the criticism by Borji and Itti (2013), who surmise that these techniques make “models data-dependent, thus influencing fair model comparison, slow, and to some extent, black-box.” But on the other hand, one important lesson of these approaches lies in that they provides a data-driven way of deriving the most relevant visual features as optimal predictors. The learned patterns can shape receptive fields (filters) that have equivalent or superior predictive power when compared against hand-crafted (and sometimes more complicated) models (Kienzle, Franz, Schölkopf, & Wichmann, 2009). Certainly, this lesson is at the base of the current exponentially growth of methods based on *deep learning techniques* (LeCun, Bengio, & Hinton, 2015), in particular Convolutional Neural Networks (CNN, cfr. Srinivas et al., 2016 for a focused review), where the computed features seem to outperform, at least from an engineering perspective, most of, if not all, the state-of-the art features conceived in computer vision.

Again, CNNs, as commonly exploited in the current practice, bring no significant conceptual novelty as to the use of Eq. 9.11: fixation prediction is formulated as a supervised binary classification problem (in some case, regression is addressed, Wang, Wang, & Ji, 2016). For example, in (Vig, Dorr, & Cox, 2014) a linear SVM is used for learning the saliency discriminant function $f(\mathbf{L})$ after a large-scale search for optimal features $\mathbf{F}$. Similarly, Shen and Zhao (2014) detect salient region via linear SVM fed with features computed from multi-layer sparse network model. Work described in Lin, Kong, Wang, and Zhuang (2014) uses the simple normalization step (Itti et al., 1998 to approximate $P(\mathbf{L} = 1 \mid \mathbf{F})$, where in Kruthiventi, Ayush, and Babu (2015) the last 1×1 convolutional layer of a fully convolutional net is exploited. Cogent here is the outstanding performance of CNN in learning and representing features that correlate well with eye fixations, like objects, faces, context.

Clearly, one problem is the enormous amount of training data necessary to train these networks, and the engineering expertise required, which makes them difficult to apply for predicting saliency. However, Kümmerer, Theis, and Bethge (2014) by exploiting the well known network from Krizhevsky, Sutskever, and Hinton (2012) as starting point, have given evidence that deep CNN trained on computer vision tasks like object detection boost saliency prediction. The network described in Krizhevsky et al. (2012) has been optimised for object recognition using a massive dataset consisting of more than one million images, and results reported (Kümmerer et al., 2014) on static pictures are impressive when compared to state-of-the-art methods, even to previous CNN-based proposals (Vig et al., 2014).

9.9 Suggested Readings

To explore beyond the contents of this Chapter, we recommend the following. A brief and clear introduction to stochastic processes (from a physicist's point of view) can be found in a few chapters of Huang's "*Introduction to statistical physics*" (Huang, 2001). A comprehensive treatment of stochastic processes is given in Gardiner's "*Stochastic Methods: A Handbook for the Natural and Social Sciences*" (Gardiner, 2009); it is a great starting point if you are looking for specific information about a specific stochastic process. One of the finest books on stochastic processes is van Kampen's classic "*Stochastic processes in physics and chemistry*" (Van Kampen, 2001); difficult reading, but well–worth the effort. A modern treatment of the subject is provided in Paul and Baschnagel "*Stochastic Processes – From Physics to Finance*" (Paul & Baschnagel, 2013), with a clear discussion of what happens beyond the Central Limit Theorem.

A beautiful bridge between stochastic processes and foraging is outlined in Méndez, Campos, and Bartumeus, "*Stochastic Foundations in Movement Ecology: Anomalous Diffusion, Front Propagation and Random Searches*" (Méndez, Campos, & Bartumeus, 2014). However, if one wants to skip more technical details, an affordable, easy to read introduction to foraging and Lévy flights is "*The physics of foraging*" by Viswanathan, Da Luz, Raposo, and Stanley (2011).

Eventually, for what concerns Statistical Machine Learning, which is nowadays a vast field, a thorough and simple introduction is provided by Rogers and Girolami (2011). A deeper insight can be gained by reading Bishop's textbook (Bishop, 2006). The most comprehensive and up-to-date textbook is that by Murphy (2012). If some of the readers are daring to surf the big wave of deep learning, then the book by Goodfellow, Bengio, and Courville (2016) provides the vital outfit.

9.10 Questions Students Should Be Able to Answer

1. What are the main reasons for considering the sequence of gaze shifts to be the observable outcome of a stochastic process?
2. If you were to set up a probabilistic model of gaze shifts, which factors would you consider to design a prior distribution $P(\mathbf{x})$ on gaze position $\mathbf{x}$?
3. What kind of information is provided by a Probabilistic Graphical Model?
4. You have conducted an eye tracking experiment where you recorded the first ten fixations for each subject, say $\mathbf{x}_i$, $i = 1, 2, \ldots, 10$ at times t_i, $i = 1, 2, \ldots, 10$. You devise two possible models:

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \ldots) = \prod_{i=1}^{10} P(\mathbf{x}_i, t_i),$$

$$P(\mathbf{x}_1, t_1; \mathbf{x}_2, t_2; \ldots) = P(\mathbf{x}_1, t_1) \prod_{i=2}^{10} P(\mathbf{x}_i, t_i \mid \mathbf{x}_{i-1}, t_{i-1}).$$

   What are the assumptions behind the two models? What are the pros and cons of each model?
5. You set up an eye tracking experiment where a control group and a patient group observe pictures displaying a number of objects of interest. At the end of the experiment you define objects as Areas of Interest (AOIs) and compute the observation transition frequency between AOIs in each picture for each subject of the two groups. What is the probabilistic model behind such statistics?
6. You are simulating a 2-dimensional stochastic process according to the following equations:

$$x_t = \xi_{x,t}$$
$$y_t = \xi_{y,t},$$

where x_t, y_t denote spatial coordinates and the random variable $\xi_{x,t}$ is sampled from a 1-dimensional Gaussian distribution $\mathcal{N}(\mu_x, \sigma^2)$, i.e. $\xi_{x,t} \sim \mathcal{N}(\mu, \sigma^2)$, with $\mu_x = 10, \sigma = 1$, while $\xi_{y,t} \sim \mathcal{N}(\mu_y, \sigma^2)$, with $\mu_y = 10$. What kind of 2-dimensional pattern is likely to be drawn if you run the simulation for $t = 1, 2, \ldots 100$ iterations?

7. You are studying the central fixation bias in scene viewing (i.e., the marked tendency to fixate the center of the screen when viewing scenes on computer monitors). The resolution of the CRT monitor where stimuli are presented is 1280 × 1020 pixels. What could be a possible microscopic level description of such bias? What a possible macroscopic description? (*Hint*: Reconsider Question 5)
8. Assume that you are able to fit the empirical distribution of some experimental data with the law $P(x) \approx x^{-\mu}$. Which kind of information cold you infer from such result?
9. Repeat the simulation proposed in Question 5, but now assume to sample $\xi_{x,t}$ and $\xi_{y,t}$ from a Cauchy distribution $\frac{1}{\pi\gamma}\left[\frac{\gamma^2}{(\xi-\delta)^2+\gamma^2}\right]$ with $\delta = 10$, $\gamma = 1$. What kind of 2-dimensional pattern do you expect? Would the histogram of the sequence of step amplitudes have a bell shape? Could we still consider this kind of random walk a Markovian walk?
10. You plan an eye tracking experiment with n subjects to distinguish experts from non experts when viewing five paintings. After recording the data, you define a number N_k of AOIs on paintings (for simplicity, assume an equal number for all paintings). Then, for each subject and each painting, you measure l_k and t_k, namely the number of fixations and the average fixation time, respectively, in k-th AOI, $k = 1, 2, \ldots, N_k$. How would you proceed to make the desired discrimination?

References

Aks, D. J., Zelinsky, G. J., & Sprott, J. C. (2002). Memory across eye-movements: 1/f dynamic in visual search. *Nonlinear Dynamics, Psychology, and Life Sciences*, *6*(1), 1–25.

Bachelier, L. (1900). *Théorie de la spéculation*. Gauthier-Villars.

Barber, D., Cemgil, A. T., & Chiappa, S. (2011). *Bayesian time series models*. Cambridge: Cambridge University Press.

Baronchelli, A., & Radicchi, F. (2013). Lévy flights in human behavior and cognition. *Chaos, Solitons & Fractals*, *56*, 101–105.

Begum, M., Karray, F., Mann, G., & Gosine, R. (2010). A probabilistic model of overt visual attention for cognitive robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, *40*(5), 1305–1318.

Bishop, C. M. (2006). *Pattern recognition and machine learning (Information science and statistics)*. New York Inc, Secaucus, NJ: Springer.

Boccignone, G., & Ferraro, M. (2004). Modelling gaze shift as a constrained random walk. *Physica A: Statistical Mechanics and its Applications*, *331*(1–2), 207–218.

Boccignone, G., & Ferraro, M. (2011). The active sampling of gaze-shifts. In G. Maino & G. Foresti (Eds.), *Image analysis and processing ICIAP 2011, Lecture Notes in Computer Science* (Vol. 6978, pp. 187–196). Berlin/Heidelberg: Springer.

Boccignone, G., & Ferraro, M. (2013a). Feed and fly control of visual scanpaths for foveation image processing. *Annals of Telecommunications, 68* (3–4), 201–217.

Boccignone, G., & Ferraro, M. (2013b). Gaze shift behavior on video as composite information foraging. *Signal Processing: Image Communication*, *28*(8), 949–966.

Boccignone, G., & Ferraro, M. (2014). Ecological sampling of gaze shifts. *IEEE Transactions on Cybernetics*, *44*(2), 266–279.

Boccignone, G., Ferraro, M., & Caelli, T. (2001). An information-theoretic approach to active vision. In *Proceedings 11th International Conference on Image Analysis and Processing, (ICIAP)* (pp. 340–345). New York, NY: IEEE Press.

Boccignone, G., Ferraro, M., Crespi, S., Robino, C., & de'Sperati, C. (2014). Detecting expert's eye using a multiple-kernel relevance vector machine. *Journal of Eye Movement Research, 7*(2), 1–15.

Boccignone, G., Marcelli, A., Napoletano, P., Di Fiore, G., Iacovoni, G., & Morsa, S. (2008). Bayesian integration of face and low-level cues for foveated video coding. *IEEE Transactions on Circuits and Systems for Video Technology*, *18*(12), 1727–1740.

Borji, A., & Itti, L. (2013). State-of-the-art in visual attention modeling. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(1), 185–207.

Borji, A., Sihite, D.N., & Itti, L. (2012). An object-based Bayesian framework for top-down visual attention. In *Twenty-Sixth AAAI Conference on Artificial Intelligence*.

Brockmann, D., & Geisel, T. (2000). The ecology of gaze shifts. *Neurocomputing*, *32*(1), 643–650.

Bundesen, C. (1998). A computational theory of visual attention. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1373), 1271–1281.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (2012). A bayesian optimal foraging model of human visual search. *Psychological Science*, *23*(9), 1047–1054.

Canosa, R. (2009). Real-world vision: Selective perception and task. *ACM Transactions on Applied Perception*, *6*(2), 11.

Carpenter, R., & Williams, M. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature*, *377*(6544), 59–62.

Cerf, M., Frady, E., & Koch, C. (2009). Faces and text attract gaze independent of the task: Experimental data and computer model. *Journal of Vision, 9*(12).

Cerf, M., Harel, J., Einhäuser, W., & Koch, C. (2008). Predicting human gaze using low-level saliency combined with face detection. *Advances in Neural Information Processing Systems*, *20*.

Chambers, J., Mallows, C., & Stuck, B. (1976). A method for simulating stable random variables. *Journal of the American Statistical Association*, *71*(354), 340–344.

Chernyak, D. A., & Stark, L. W. (2001). Top–down guided eye movements. *IEEE Transactions on Systems Man Cybernetics - B, 31*, 514–522.

Chikkerur, S., Serre, T., Tan, C., & Poggio, T. (2010). What and where: A Bayesian inference theory of attention. *Vision Research*, *50*(22), 2233–2247.

Clavelli, A., Karatzas, D., Lladós, J., Ferraro, M., & Boccignone, G. (2014). Modelling task-dependent eye guidance to objects in pictures. *Cognitive Computation*, *6*(3), 558–584.

Codling, E., Plank, M., & Benhamou, S. (2008). Random walk models in biology. *Journal of the Royal Society Interface*, *5*(25), 813.

Coen-Cagli, R., Coraggio, P., Napoletano, P., & Boccignone, G. (2008). What the draughtsman's hand tells the draughtsman's eye: A sensorimotor account of drawing. *International Journal of Pattern Recognition and Artificial Intelligence*, *22*(05), 1015–1029.

Coen-Cagli, R., Coraggio, P., Napoletano, P., Schwartz, O., Ferraro, M., & Boccignone, G. (2009). Visuomotor characterization of eye movements in a drawing task. *Vision Research*, *49*(8), 810–818.

Costa, T., Boccignone, G., Cauda, F., & Ferraro, M. (2016). The foraging brain: Evidence of levy dynamics in brain networks. *PloS one, 11*(9), e0161,702.

Coutrot, A., Binetti, N., Harrison, C., Mareschal, I., & Johnston, A. (2016). Face exploration dynamics differentiate men and women. *Journal of Vision*, *16*(14), 16–16.

Coutrot, A., Hsiao, J. H., & Chan, A. B. (2017). Scanpath modeling and classification with hidden markov models. *Behavior Research Methods*. https://doi.org/10.3758/s13428-017-0876-8.

Cowpertwait, P. S., & Metcalfe, A. V. (2009). *Introductory time series with R*. Dordrecht: Springer.

Damoulas, T., & Girolami, M. A. (2009). Combining feature spaces for classification. *Pattern Recognition*, *42*(11), 2671–2683.

deCroon, G., Postma, E., & van den Herik, H. J. (2011). Adaptive gaze control for object detection. *Cognitive Computation*, *3*, 264–278.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*(1), 193–222.

Doob, J. L. (1942). The brownian movement and stochastic equations. *Annals of Mathematics,* pp. 351–369.

Dorr, M., Martinetz, T., Gegenfurtner, K., & Barth, E. (2010). Variability of eye movements when viewing dynamic natural scenes. *Journal of Vision, 10*, (10).

Dubkov, A. A., Spagnolo, B., & Uchaikin, V. V. (2008). Lévy flight superdiffusion: An introduction. *International Journal of Bifurcation and Chaos*, *18*(09), 2649–2672.

Einhäuser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14). https://doi.org/10.1167/8.14.18, http://www.journalofvision.org/content/8/14/18.abstract.

Einstein, A. (1905). On the motion required by the molecular kinetic theory of heat of small particles suspended in a stationary liquid. *Annalen der Physik*, *17*, 549–560.

Einstein, A. (1906). Zur theorie der brownschen bewegung. *Annalen der Physik, 324*(2), 371–381.

Elazary, L., & Itti, L. (2010). A bayesian model for efficient visual search and recognition. *Vision Research*, *50*(14), 1338–1352.

Ellis, S., & Stark, L. (1986). Statistical dependency in visual scanning. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *28*(4), 421–438.

Engbert, R. (2006). Microsaccades: A microcosm for research on oculomotor control, attention, and visual perception. *Progress in Brain Research*, *154*, 177–192.

Engbert, R., Mergenthaler, K., Sinn, P., & Pikovsky, A. (2011). An integrated model of fixational eye movements and microsaccades. *Proceedings of the National Academy of Sciences*, *108*(39), E765–E770.

Feng, G. (2006). Eye movements as time-series random variables: A stochastic model of eye movement control in reading. *Cognitive Systems Research*, *7*(1), 70–95.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*(2).

Frintrop, S., Rome, E., & Christensen, H. (2010). Computational visual attention systems and their cognitive foundations: A survey. *ACM Transactions on Applied Perception, 7*(1), 6.

Gardiner, C. (2009). *Stochastic methods: A handbook for the natural and social sciences. Springer series in synergetics*. Berlin, Heidelberg: Springer.

Gnedenko, B., & Kolmogórov, A. (1954). *Limit distributions for sums of independent random variables*. Addison-Wesley Pub. Co.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. http://www.deeplearningbook.org

Hacisalihzade, S., Stark, L., & Allen, J. (1992). Visual perception and sequences of eye movement fixations: A stochastic modeling approach. *IEEE Transactions on Systems, Man, and Cybernetics, 22*(3), 474–481.

Haji-Abolhassani, A., & Clark, J. J. (2013). A computational model for task inference in visual search. *Journal of Vision*, *13*(3), 29.

Harel, J., Koch, C., & Perona, P. (2007). Graph-based visual saliency. In *Advances in neural information processing systems* (Vol. 19, pp. 545–552). Cambridge, MA: MIT Press.

Heinke, D., & Backhaus, A. (2011). Modelling visual search with the selective attention for identification model (VS-SAIM): A novel explanation for visual search asymmetries. *Cognitive Computation*, *3*(1), 185–205.

Heinke, D., & Humphreys, G. W. (2003). Attention, spatial representation, and visual neglect: Simulating emergent attention and spatial memory in the selective attention for identification model (SAIM). *Psychological Review*, *110*(1), 29.

Heinke, D., & Humphreys, G. W. (2005). Computational models of visual selective attention: A review. *Connectionist Models in Cognitive Psychology*, *1*(4), 273–312.

Henderson, J. M., Shinkareva, S. V., Wang, J., Luke, S. G., & Olejarczyk, J. (2013). Predicting cognitive state from eye movements. *PLoS ONE, 8*(5), e64,937.

Higham, D. (2001). An algorithmic introduction to numerical simulation of stochastic differential equations. *SIAM Review,* pp. 525–546.

Hills, T. T. (2006). Animal foraging and the evolution of goal-directed cognition. *Cognitive Science*, *30*(1), 3–41.

Ho Phuoc, T., Guérin-Dugué, A., & Guyader, N. (2009). A computational saliency model integrating saccade programming. In *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing* (pp. 57–64). Porto, Portugal.

Horowitz, T., & Wolfe, J. (1998). Visual search has no memory. *Nature*, *394*(6693), 575–577.

Huang, K. (2001). *Introduction to statistical physics*. Boca Raton, FL: CRC Press.

Humphreys, G. W., & Muller, H. J. (1993). Search via recursive rejection (SERR): A connectionist model of visual search. *Cognitive Psychology*, *25*(1), 43–110.

Insua, D., Ruggeri, F., & Wiper, M. (2012). *Bayesian analysis of stochastic process models*. Hoboken, NJ: Wiley.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*, 1254–1259.

Jarrow, R., & Protter, P. (2004). A short history of stochastic integration and mathematical finance: The early years, 1880–1970. Lecture Notes-Monograph Series, pp. 75–91.

Jaynes, E. T. (2003). *Probability theory: The logic of science*. New York, NY: Cambridge University Press.

Jetley, S., Murray, N., & Vig, E. (2016). End-to-end saliency mapping via probability distribution prediction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5753–5761).

Jiang, L., Xu, M., Ye, Z., & Wang, Z. (2015). Image saliency detection with sparse representation of learnt texture atoms. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* (pp. 54–62).

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). Learning to predict where humans look. In *IEEE 12th International conference on Computer Vision* (pp. 2106–2113). New York, NY: IEEE.

Keech, T., & Resca, L. (2010). Eye movements in active visual search: A computable phenomenological model. *Attention, Perception, & Psychophysics*, *72*(2), 285–307.

Kienzle, W., Franz, M. O., Schölkopf, B., & Wichmann, F. A. (2009). Center-surround patterns emerge as optimal predictors for human saccade targets. *Journal of Vision*, *9*(5), 7–7.

Kienzle, W., Wichmann, F. A., Franz, M. O., & Schölkopf, B. (2006). A nonparametric approach to bottom-up visual saliency. In *Advances in neural information processing systems* (pp. 689–696).

Kimura, A., Pang, D., Takeuchi, T., Yamato, J., & Kashino, K. (2008). Dynamic Markov random fields for stochastic modeling of visual attention. In *Proceedings of ICPR '08* (pp. 1–5). New York, NY: IEEE.

Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: Towards the underlying neural circuitry. *Human Neurobiology*, *4*(4), 219–27.

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: Principles and techniques*. Cambridge, MA: MIT press.

Kolmogorov, A., & Gnedenko, B. (1954). *Limit distributions for sums of independent random variables*. Cambridge, MA: Addison-Wesley.

Kolmogorov, A. N. (1941). Dissipation of energy in isotropic turbulence. *Doklady Akademii Nauk SSSR*, *32*, 325–327.

Koutrouvelis, I. (1980). Regression-type estimation of the parameters of stable laws. *Journal of the American Statistical Association,* pp. 918–928.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, & K. Q. Weinberger (Eds.), *Advances in Neural Information Processing Systems* (Vol. 25, pp. 1097–1105). Curran Associates, Inc.

Kruthiventi, S. S., Ayush, K., & Babu, R. V. (2015). Deepfix: A fully convolutional neural network for predicting human eye fixations. arXiv preprint arXiv:1510.02927.

Kümmerer, M., Theis, L., & Bethge, M. (2014). Deep gaze I: Boosting saliency prediction with feature maps trained on imagenet. arXiv preprint arXiv:1411.1045.

Lagun, D., Manzanares, C., Zola, S. M., Buffalo, E. A., & Agichtein, E. (2011). Detecting cognitive impairment by eye movement analysis using automatic classification algorithms. *Journal of Neuroscience Methods*, *201*(1), 196–203.

Laing, C., & Lord, G. J. (2010). *Stochastic methods in neuroscience*. Oxford: Oxford University Press.

Lang, C., Liu, G., Yu, J., & Yan, S. (2012). Saliency detection by multitask sparsity pursuit. *IEEE Transactions on Image Processing, 21*(3), 1327–1338.

Langevin, P. (1908). Sur la théorie du mouvement brownien. *Comptes-Rendus de Académie Sciences Paris, 146*(530–533), 530.

Le Meur, O., & Coutrot, A. (2016). Introducing context-dependent and spatially-variant viewing biases in saccadic models. *Vision Research*, *121*, 72–84.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444.

Liberati, A., Fadda, R., Doneddu, G., Congiu, S., Javarone, M. A., Striano, T., & Chessa, A. (2017). A statistical physics perspective to understand social visual attention in autism spectrum disorder. *Perception, 46*(8), 889–913.

Lin, Y., Kong, S., Wang, D., & Zhuang, Y. (2014). Saliency detection within a deep convolutional architecture. In *Workshops at the Twenty-Eighth AAAI Conference on Artificial Intelligence*.

Logan, G. D. (1996). The code theory of visual attention: An integration of space-based and object-based attention. *Psychological Review*, *103*(4), 603.

MacKay, D. (2002). *Information theory. Inference and learning algorithms*. Cambridge, MA: Cambridge University Press.

Makarava, N., Bettenbühl, M., Engbert, R., & Holschneider, M. (2012). Bayesian estimation of the scaling parameter of fixational eye movements. *EPL,* **100**(4), 40,003.

Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business*, *36*(4), 394–419.

Mandelbrot, B. B., & Van Ness, J. W. (1968). Fractional brownian motions, fractional noises and applications. *SIAM Review*, *10*(4), 422–437.

Mantegna, R. N., Stanley, H. E., et al. (2000). *An introduction to econophysics: Correlations and complexity in finance*. Cambridge, MA: Cambridge University Press.

Marat, S., Rahman, A., Pellerin, D., Guyader, N., & Houzet, D. (2013). Improving visual saliency by adding Ôface feature mapÕand Ôcenter biaÕ. *Cognitive Computation*, *5*(1), 63–75.

Marr, D. (1982). *Vision*. New York, NY: W.H. Freeman.

Martinez-Conde, S., Otero-Millan, J., & Macknik, S. L. (2013). The impact of microsaccades on vision: Towards a unified theory of saccadic function. *Nature Reviews Neuroscience*, *14*(2), 83–96.

Mathe, S., Sminchisescu, C. (2013). Action from still image dataset and inverse optimal control to learn task specific visual scanpaths. In *Advances in neural information processing systems* (pp. 1923–1931).

Mathe, S., & Sminchisescu, C. (2015). Actions in the eye: Dynamic gaze datasets and learnt saliency models for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(7), 1408–1424.

Méndez, V., Campos, D., & Bartumeus, F. (2014). *Stochastic foundations in movement ecology: Anomalous diffusion. Front propagation and random searches. Springer series in synergetics*. Berlin, Heidelberg: Springer.

Meyer, P. A. (2009). Stochastic processes from 1950 to the present. *Electronic Journal for History of Probability and Statistics*, *5*(1), 1–42.

Mozer, M. C. (1987). Early parallel processing in reading: A connectionist approach. Lawrence Erlbaum Associates, Inc.

Murphy, K. P. (2012). *Machine learning: A probabilistic perspective*. Cambridge, MA: MIT press.

Najemnik, J., & Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391.

Napoletano, P., Boccignone, G., & Tisato, F. (2015). Attentive monitoring of multiple video streams driven by a Bayesian foraging strategy. *IEEE Transactions on Image Processing*, *24*(11), 3266–3281.

Nelson, E. (1967). *Dynamical theories of Brownian motion*. Princeton, NJ: Princeton University Press.

Newman, M. E. (2005). Power laws, pareto distributions and zipf's law. *Contemporary Physics*, *46*(5), 323–351.

Nolan, J. (1997). Numerical calculation of stable densities and distribution functions. *Communications in Statistics-Stochastic Models*, *13*(4), 759–774.

Noorani, I., & Carpenter, R. (2016). The LATER model of reaction time and decision. *Neuroscience & Biobehavioral Reviews*, *64*, 229–251.

Osborne, M. F. (1959). Brownian motion in the stock market. *Operations Research*, *7*(2), 145–173.

Otero-Millan, J., Macknik, S. L., Langston, R. E., & Martinez-Conde, S. (2013). An oculomotor continuum from exploration to fixation. *Proceedings of the National Academy of Sciences*, *110*(15), 6175–6180.

Over, E., Hooge, I., Vlaskamp, B., & Erkelens, C. (2007). Coarse-to-fine eye movement strategy in visual search. *Vision Research*, *47*, 2272–2280.

Ozaki, T. (2012). *Time series modeling of neuroscience data*. CRC Press.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research*, *40*(10), 1227–1268.

Papoulis, A., & Pillai, S. U. (2002). *Probability, random variables, and stochastic processes*. New York, NY: McGraw-Hill.

Paul, L. (1954). *Théorie de l'addition des variables aléatoires*. Paris: Gauthiers-Villars.

Paul, W., & Baschnagel, J. (2013). *Stochastic processes: From physics to finance*. Berlin, Heidelberg: Springer International Publishing.

Phaf, R. H., Van der Heijden, A., & Hudson, P. T. (1990). Slam: A connectionist model for attention in visual selection tasks. *Cognitive Psychology*, *22*(3), 273–341.

Plank, M., & James, A. (2008). Optimal foraging: Lévy pattern or process? *Journal of The Royal Society Interface*, *5*(26), 1077.

Moscoso del Prado Martin, F. (2008). A theory of reaction time distributions. http://cogprints.org/6310/1/recinormal.pdf

Psorakis, I., Damoulas, T., & Girolami, M. A. (2010). Multiclass relevance vector machines: Sparsity and accuracy. *IEEE Transactions on Neural Networks*, *21*(10), 1588–1598.

Ramos-Fernandez, G., Mateos, J., Miramontes, O., Cocho, G., Larralde, H., & Ayala-Orozco, B. (2004). Lévy walk patterns in the foraging movements of spider monkeys (Ateles geoffroyi). *Behavioral Ecology and Sociobiology*, *55*(3), 223–230.

Rao, R. P., Zelinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (2002). Eye movements in iconic visual search. *Vision Research*, *42*(11), 1447–1463.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Rensink, R. (2000). The dynamic representation of scenes. *Visual Cognition*, *1*(3), 17–42.

Reynolds, A. (2008). How many animals really do the Lévy walk? Comment. *Ecology*, *89*(8), 2347–2351.

Reynolds, A. (2008). Optimal random Lévy-loop searching: New insights into the searching behaviours of central-place foragers. *EPL (Europhysics Letters), 82*, 20,001.

Richardson, L. F. (1926). Atmospheric diffusion shown on a distance-neighbour graph. *Proceedings of the Royal Society of London. Series A, 110*(756), 709–737.

Rogers, S., & Girolami, M. (2011). *A first course in machine learning*. Boca Raton, FL: CRC Press.

Rutishauser, U., & Koch, C. (2007). Probabilistic modeling of eye movement data during conjunction search via feature-based attention. *Journal of Vision, 7*(6).

Schinckus, C. (2013). How physicists made stable lévy processes physically plausible. *Brazilian Journal of Physics*, *43*(4), 281–293.

Scholl, B. (2001). Objects and attention: The state of the art. *Cognition*, *80*(1–2), 1–46.

Schuster, P. (2016). *Stochasticity in processes*. Berlin: Springer.

Schütz, A., Braun, D., & Gegenfurtner, K. (2011). Eye movements and perception: A selective review. *Journal of Vision, 11*(5).

Seo, H., & Milanfar, P. (2009). Static and space-time visual saliency detection by self-resemblance. *Journal of Vision*, *9*(12), 1–27.

Shen, C., & Zhao, Q. (2014). Learning to predict eye fixations for semantic contents using multi-layer sparse network. *Neurocomputing*, *138*, 61–68.

Siegert, S., & Friedrich, R. (2001). Modeling of nonlinear Lévy processes by data analysis. *Physical Review E, 64*(4), 041,107.

Srinivas, S., Sarvadevabhatla, R. K., Mopuri, K. R., Prabhu, N., Kruthiventi, S., & Radhakrishnan, V. B. (2016). A taxonomy of deep convolutional neural nets for computer vision. *Frontiers in Robotics and AI, 2*(36). https://doi.org/10.3389/frobt.2015.00036, http://www.frontiersin.org/vision_systems_theory,_tools_and_applications/10.3389/frobt.2015.00036/abstract

Stephen, D., Mirman, D., Magnuson, J., & Dixon, J. (2009). Lévy-like diffusion in eye movements during spoken-language comprehension. *Physical Review E, 79*(5), 056,114.

Stigler, G. J. (1964). Public regulation of the securities markets. *The Journal of Business*, *37*(2), 117–142.

Sun, Y., Fisher, R., Wang, F., & Gomes, H. M. (2008). A computer vision model for visual-object-based attention and eye movements. *Computer Vision and Image Understanding*, *112*(2), 126–142.

Tatler, B. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of Vision, 7*(14).

Tatler, B., Baddeley, R., & Vincent, B. (2006). The long and the short of it: Spatial statistics at fixation vary with saccade amplitude and task. *Vision Research*, *46*(12), 1857–1862.

Tatler, B., Hayhoe, M., Land, M., & Ballard, D. (2011). Eye guidance in natural vision: Reinterpreting salience. *Journal of Vision, 11*(5).

Tatler, B., & Vincent, B. (2008). Systematic tendencies in scene viewing. *Journal of Eye Movement Research*, *2*(2), 1–18.

Tatler, B., & Vincent, B. (2009). The prominence of behavioural biases in eye guidance. *Visual Cognition*, *17*(6–7), 1029–1054.

Torralba, A. (2003). Contextual priming for object detection. *International Journal of Computer Vision*, *53*, 153–167.

Treisman, A. (1998). Feature binding, attention and object perception. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 353*(1373), 1295–1306.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136.

Trillenberg, P., Gross, C., & Shelhamer, M. (2001). Random walks, random sequences, and nonlinear dynamics in human optokinetic nystagmus. *Journal of Applied Physiology*, *91*(4), 1750–1759.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the brownian motion. *Physical Review*, *36*(5), 823.

Van Der Linde, I., Rajashekar, U., Bovik, A. C., & Cormack, L. K. (2009). Doves: A database of visual eye movements. *Spatial Vision*, *22*(2), 161–177.

Van Kampen, N. G. (2001). *Stochastic processes in physics and chemistry*. Amsterdam, NL: North Holland.

Vig, E., Dorr, M., Cox, D. (2014). Large-scale optimization of hierarchical features for saliency prediction in natural images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 2798–2805).

Viswanathan, G., Raposo, E., & da Luz, M. (2008). Lévy flights and superdiffusion in the context of biological encounters and random searches. *Physics of Life Reviews*, *5*(3), 133–150.

Viswanathan, G. M., Da Luz, M. G., Raposo, E. P., & Stanley, H. E. (2011). *The physics of foraging: An introduction to random searches and biological encounters*. Cambridge, MA: Cambridge University Press.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks*, *19*(9), 1395–1407.

Wang, K., Wang, S., & Ji, Q. (2016). Deep eye fixation map learning for calibration-free eye gaze tracking. In *Proceedings of the Ninth Biennial ACM Symposium on Eye Tracking Research & Applications* (pp. 47–55). New York, NY: ACM.

Wiener, N. (1930). Generalized harmonic analysis. *Acta Mathematica*, *55*(1), 117–258.

Wischnewski, M., Belardinelli, A., Schneider, W., & Steil, J. (2010). Where to look next? Combining static and dynamic proto-objects in a TVA-based model of visual attention. *Cognitive Computation*, *2*(4), 326–343.

Wolfe, J. M. (1994). Guided search 2.0 a revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.

Wolfe, J. M. (2013). When is it time to move to the next raspberry bush? Foraging rules in human visual search. *Journal of Vision*, *13*(3), 10.

Yan, J., Zhu, M., Liu, H., & Liu, Y. (2010). Visual saliency detection via sparsity pursuit. *Signal Processing Letters, IEEE*, *17*(8), 739–742.

Yang, S. C. H., Wolpert, D. M., & Lengyel, M. (2016). Theoretical perspectives on active sensing. *Current Opinion in Behavioral Sciences*, *11*, 100–108.

Yarbus, A. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Yu, J. G., Zhao, J., Tian, J., & Tan, Y. (2014). Maximal entropy random walk for region-based visual saliency. *IEEE Transactions on Cybernetics*, *44*(9), 1661–1672.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review*, *115*(4), 787.

Chapter 10
Temporal Methods for Eye Movement Analysis

Walter F. Bischof, Nicola C. Anderson and Alan Kingstone

Contents

W. F. Bischof (✉) · N. C. Anderson · A. Kingstone
Department of Psychology, University of British Columbia, Vancouver, BC, Canada
e-mail: wfb@ualberta.ca

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_10

Abstract This chapter is aimed at introducing the reader to current methods for the spatial and temporal analysis of eye movements. There are four main parts. In the first part of the chapter, we introduce the relation between attention and eye movements and then review the foundations of attention research, social attention and the effect of stimulus saliency on attention. In the second part, we first review traditional characterizations of eye movements, specifically measures of fixations and saccades, and then we review the major methods for the spatial analysis of eye movements, including heat maps and area-of-interest analyses. In the third part, we introduce a new temporal analysis of eye movements, recurrence quantification analysis. We review the basic methods and measures of this analysis, and we discuss several applications. In the last part of the chapter, we provide an overview of methods for the comparison of scanpaths and show that recurrence quantification analysis can be generalized successfully to scanpath comparisons.

With the information provided in this chapter, readers will be able to understand

- The relationship between attention and eye movements
- The different methods for characterizing eye movements in the spatial and temporal domain
- The fundamentals of recurrence quantification analysis for characterizing eye movement dynamics
- The major scanpath comparison methods.

Abbreviations

The following abbreviations are used in this chapter

AOI	Area of interest
CORM	Center of recurrence mass
DET	Determinism
LAM	Laminarity
LDA	Linear discriminant analysis
REC	Recurrence
RQA	Recurrence quantification analysis

10.1 Introduction: Attention and Eye Movements

Human attention is the mechanism that enables one to filter out some sensory information in the environment and sharpen the perception of other information, much as one disregards the surrounding text on a page while reading specific words (Broadbent, 1971). Typically, attentional orienting is accompanied by fast jerky eye movements—called saccades—that rapidly acquire detailed information from the environment, such as those being performed while reading the sentences in this chapter (see e.g. chapter by Hyönä & Kaakinen, this volume). These shifts in attention are called *overt* because they involve reorienting the eyes, a behaviour that can be observed by others (Findlay & Gilchrist, 2003). Interestingly, though, eye movements are not needed for shifts of attention to occur. For instance, when you reach the end of this sentence, stare at the period and then identify the word that is one line below it while keeping your eyes on the period. These shifts in attention in the absence of eye movements are called *covert* because they do not require reorienting of the eyes and thus cannot be directly observed by others (Posner, 1978). In general we look at things that interest us, and therefore overt attention and covert attention align. Therefore the focus of discussion in this chapter concerns attention, as measured by overt shifts in attention, specifically, those that are executed as saccadic eye movements.

Both forms of orienting, covert and overt, can be activated *reflexively* by external events in the environment (this is sometimes referred to as exogenous orienting) and *voluntarily* by internal goals and expectations (also called endogenous orienting) (Kingstone, 1992). Consider for example that in reading this sentence you are moving your overt attention voluntarily from one word to the next.

Alternatively, imagine that you are driving a car when a ball suddenly bounces into the road ahead of you. Initially, your attention is captured by the ball (overt reflexive orienting) but you do not track the ball for long. Rather, you look back to the area that the ball came from (overt voluntary orienting) to see if a child is about to run into the road after the ball.

In everyday life there are often competing demands on reflexive and volitional attention, and in order to respond appropriately to these competing demands, one needs to be able to coordinate attentional processes. For example, in the driving scenario described above, you would need to be able to break your reflexive attention away from the ball to search in a volitional manner for a child.

In the lab, two classic paradigms used to isolate these two processes, and their coordination, are the prosaccade task and the anti-saccade task (see Fig. 10.1 and chapter by Pierce et al., this volume). Saccadic latency is shorter when saccades are directed towards a target (prosaccades) than away from it (anti-saccades). This difference is typically attributed to the fact that different processes are involved in prosaccade and antisaccade generation. Prosaccades are reflexive (exogenous) responses triggered by the onset of a stimulus. Antisaccades require two processes: the inhibition of a prosaccade and volitional (endogenous) programming of a saccade in the opposite direction (Olk & Kingstone, 2003).

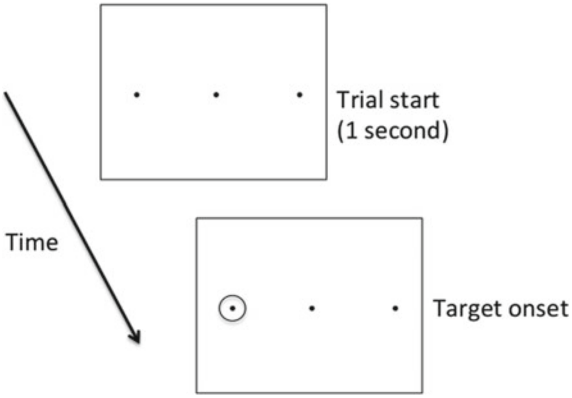


Fig. 10.1 Example trial sequence in a prosaccade and antisaccade task. Participants are presented with a central fixation dot flanked by two possible target locations. After a delay of delay of 1 s, a target circle is presented on the left or right. In a prosaccade task the participant is to fixate the target location as quickly as possible. In an antisaccade task the participant is to fixate the location that is mirror opposite to where the target appeared. Thus, in the trial that is illustrated, if a participant was performing a prosaccade task a left eye movement would be correct, and if a participant was performing an antisaccade task a right eye movement would be correct

Difficulty in coordinating attentional processes can be a major source of disability in people with neurologic disorders (see Müri et al., this volume). For example, patients with frontal lobe lesions, caused by a stroke or a closed head injury, may perform normally when looking reflexively to a peripheral light that appears suddenly, as in the prosaccade task. But when asked to look away from the stimulus light, as in the antisaccade task, these patients are extremely slow to respond, and often incorrectly continue to make reflexive eye movements towards the light. It is as if the driver in our scenario above could not stop tracking the ball and thus could not begin to search for a child. Thus it is believed that frontal brain systems are crucial for generating voluntary shifts of attention and inhibiting reflexive ones.

10.2 Historical Annotations

10.2.1 Attention Research: Assumptions of Process Stability and Control

The study of human attention can be segmented into three historical stages. The first stage occurs in the late 1950s and is characterized by a rapid scientific progression propelled by the methods of traditional psychophysics and experimental psychology. The second stage appears in the mid-1970s and is driven by computational analyses that heralded the arrival of cognitive science. The third phase, which began in the mid-1980s, incorporated evidence from neuropsychology and animal neurophysiology, and more recently, brain imaging, and is subsumed by the field of cognitive neuroscience.

Each of these historical stages are grounded on two basic research assumptions. One is that human attention is controlled by processes that are stable across different situations, meaning that, for example, the processes that are studied in the lab are the same as the processes that are expressed in the real world. Second one can maximize research power by exerting experimental control and minimizing all variability in a situation save for the factor of interest. Below we describe briefly why these two assumptions may very well be invalid insofar as they fail to shed light on attention in real world situations, and how the field has responded by conducting far more complex studies that occur in real life or at least better approximate real-world situations. These more sophisticated studies have in turn demanded more advanced analytical tools, and these analytical methods—in particular, the recurrence quantification analysis—are the main focus of the present chapter. Before turning to these analyses, however, a brief historical review is warranted and presented.

10.2.2 Social Attention and Stimulus Saliency

While the assumptions of process stability and situational control are commonly held and applied in studies of attention, adopting them comes with a degree of risk. The assumption of stability for example eliminates any need or obligation by the scientist to confirm that the factors being manipulated and measured in the lab actually express themselves in the real world. The field does of course check routinely that the effects being measured are stable within the lab environment by demanding that results in the lab be replicable. Unfortunately a result that is stable within a controlled laboratory environment does not necessarily mean that it is stable outside the lab. Indeed there are many examples within the field of human attention indicating that even minimal changes within a laboratory situation will compromise the replicability of the effect (e.g., Chica, Klein, Rafal, & Hopfinger, 2010; Hunt, Chapman, & Kingstone, 2008; Soto, Morein-Zamir, & Kingstone, 2005).

It has been proposed in much greater detail elsewhere (e.g., Kingstone, Smilek, & Eastwood, 2008; Risko & Kingstone, 2011) that an impoverished highly controlled experimental situation is unlikely to inform the field about the attentional processes as they are expressed in everyday real-life situations. It stands to reason then that by increasing situational complexity and reducing experimental control one will begin to better approximate the mechanisms that operate in everyday life.

This approach can be illustrated by first considering that the prevailing eye movement model of Itti and Koch (2000) assumes that where people look is determined by a ‘winner take all’ visual saliency map. This saliency map is generated from basic stimulus features, such as luminance, contrast and colour. These features are claimed to be combined in a biologically plausible way (based on the workings of the visual cortex) to represent the most interesting or ‘salient’ regions of a display, image or

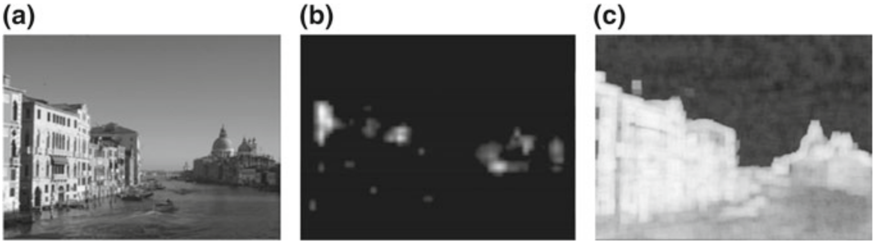


Fig. 10.2 **a** Example image; **b** Salience map as computed by the Saliency Toolbox in MATLAB (Walther & Koch, 2006); **c** Salience map as computed by the Attention based on Information Maximization (AIM) model (Bruce & Tsotsos, 2009)

video (see Fig. 10.2). Despite the mounting evidence that this model has, at best, minimal construct validity in an very narrow band of situations (e.g., Nyström & Holmqvist, 2008; Tatler, Baddeley, & Gilchrist, 2005), it is still prominent in the literature.

What is striking about this model, besides its vast popularity, is how the scenes that have been used to test the model have rarely contained any people. The real world is much more than images of landscapes, buildings, and empty rooms. The real world contains people, and much of it operates in service of the needs of people.

The recent eye tracking work of Birmingham and colleagues (Birmingham, Bischof, & Kingstone, 2008a, b; Levy, Foulsham, & Kingstone, 2012) has revealed that people are extraordinarily interested in people, in particular, the eyes of people, even when they are embedded in complex scenes. It does not appear to matter very much where the people are in the scenes, what they are wearing, or even how tiny they are represented in a scene. If there is a person somewhere in a photo, then participants are going to look at them quickly, and often—especially their eyes.

From the perspective of the saliency model, these results are not expected because often the people in the scenes are very small and not at all visually salient (Birmingham, Bischof, & Kingstone, 2009). And yet, observers quickly, consistently, and repeatedly seek them out. Thus, there seemed to be a profound bias to search out people, and in particular the eyes of individuals, in complex social scenes.

10.2.3 Social Attention in the Real World

The research conducted using social scenes, of course, pertain to simple static scenes (i.e. photographs) of people. In real life people move about, they look at each other, and they talk to one another. What happens in such a situation? Foulsham, Cheng, Tracy, Henrich, and Kingstone (2010) asked precisely this question. In their study participants watched videos of different groups of three individuals sitting around a table discussing a hypothetical situation regarding the most important items that they would take to the moon while having their eye movements tracked. Foulsham et al. (2010) found that despite the fact that the individuals in these videos moved, talked and interacted with one another, there remained a tremendous consistency in the participants' looking behavior. Specifically, participants fixated primarily on the eyes of the people in the video (see also Cheng, Tracy, Foulsham, Kingstone, & Henrich, 2013). Thus, even in this dynamic social context, participants' looking behavior evidenced a clear bias to attend to the eyes of others. Furthermore, as with Birmingham, Bischof, and Kingstone (2009), these findings cannot be explained in terms of basic low-level stimulus saliency, in this case, features like visual motion and sound onsets. Foulsham and Sanderson (2013) and Coutrot and Guyader (2014) both investigated whether looks to the faces and eyes of individuals engaged in conversation were significantly affected by changes in visual saliency, or whether the audio is present or absent. In both studies, participants again view complex, dynamic scenes featuring conversation while their eye movements were recorded. Their results indicated that the addition of an audio track increased looks to the faces and eyes of the talkers, and also resulted in greater synchrony between observers when they looked at the speakers (Foulsham & Sanderson, 2013). Critically, however, whether sound was present or not, and independent of changes in low-level visual saliency (Coutrot & Guyader, 2014), people spent most of their time looking at the faces and eyes of the individuals in the videos.

10.2.4 Summary

To summarize, it has been found that attention paradigms that are conducted in isolation using simple non-social, carefully controlled visual scenes, a model that assumes that people look at the most salient items can explain some eye movement behaviour. However, when contrary to the classic research approach of simplification and control, participants are shown a wide variety of photos containing people (the pictures are all different) and the behaviour is unconstrained (participants are free to look wherever they wished), then it is discovered that people are primarily interested in looking at the people in the scenes, especially their eyes. These findings persisted when stimulus complexity is further increased by introducing videos that involved people moving and talking. Finally, these data provided a strong test of the Itti-Koch saliency model of human looking behaviour and found that it could not account for such behaviour in these more natural and complex displays.

To address these and similar criticisms, the Itti-Koch saliency model has been revised regularly by incorporating additional features, such as depth and motion, by adding top-down mechanisms, such as face detectors, or contextual guidance (Anderson, Donk, & Meeter, 2016; Torralba, Oliva, Castelhano, & Henderson, 2006). Finally, some recent work has used machine learning to learn an optimal set of bottom-up features (e.g. Vig, Dorr, & Cox, 2014). These extensions are further discussed in the chapter by Foulsham (this volume). We also encourage the reader to consult the Saliency Benchmark website http://saliency.mit.edu for a review of recent saliency models and their performance.

10.3 Traditional Characterizations of Eye Movements

So far we have reviewed fundamental characteristics of attention with a focus on eye movements, and some of the basic paradigms used in attention research. We began with very simple tasks for the study of reflexive and volitional attention using, for example, the prosaccade and antisaccade tasks. We concluded that although these studies are useful and important, highly controlled experimental situations may not fully inform us about the attentional processes expressed in everyday real-life situations (see Sect. 10.3.2). We then reviewed studies employing more complex tasks such as viewing of static and dynamic images depicting complex social scenes.

In the remainder of this chapter, we focus on the measurement and description of eye movements. First, we briefly review some of the traditional basic eye movement measures and note that they are simply not able to capture and reflect the complex patterns of spatial and temporal eye movement behaviours being produced. Second, we review popular methods for assessing spatial and temporal characteristics of eye movements that are more suitable for describing eye movements in viewings complex scenes. Third, we introduce recurrence quantification analysis, a method that is well suited to the description of the temporal characteristic of eye movements in real-world situations. Finally, we review several methods of comparing sequences of eye movements.

10.3.1 Fundamental Measures for Eye Movements

From a psychological perspective, the most important eye movement events are fixations (see also the chapter by Alexander & Martinez-Conde, this volume) and saccades, as described earlier in the chapter. A saccade is a rapid, ballistic motion of the eyes from one point to another, while a fixation is the brief (around 200 ms) pause between saccades during which the most visual information is gleaned. Fixations and saccades are extracted from the raw eye-tracking data that is recorded using specialized eye-tracking equipment by applying an algorithm, or series of algorithms to the raw data. Both fixations and saccades can be described in the spatial and the temporal domains. One of the more important spatial fixation measures is the variability of fixations, or where exactly people looked, which can be used for assessing the consistency of eye movements across different observers or across repeated presentations for the same observer. The variability can be measured by determining the variance or the range of fixation positions, and it can be measured with respect to the whole stimulus area or with respect to particular regions of interest. The temporal fixation measures are based on the duration of fixations and include, for example, the average fixation duration, the distribution of fixation durations, or the total fixation duration for fixations within different regions of interest.

The spatial measures of saccadic eye movements include amplitude and direction of saccades. The former refers to the size of saccades and is typically measured by the average saccade amplitude or the distribution of saccade amplitudes for each experimental condition. The direction distribution of saccades describes how often saccades were made in each direction. The temporal measures include saccade duration, that is how much time the saccades take on average, and saccade rate, that is how frequently saccades are made. Finally, spatio-temporal saccade measures include, for example, saccade velocity and acceleration (for further details, see Holmqvist et al., 2011).

In practice, eye movement studies often use combinations of spatial and temporal fixation and saccade measures for assessing differences between experimental groups or conditions. While these measurements are undoubtedly important for characterizing eye movement behaviour, it is often difficult to explain why certain differences exist between experimental conditions or experimental groups, except in very specific circumstances. For example, the average saccade amplitudes may differ between two groups of participants, but without an analysis of eye movement dynamics and without a model of saccade generation, it would be difficult to explain why this would be the case. This, in turn, puts limits on the usefulness of these measures for characterizing eye movement behaviour with complex stimuli. For this reason, we introduce spatial and temporal eye movement measures that are suitable for such stimuli.

10.3.2 Spatial Analysis of Eye Movements

This section presents an overview of the predominant methods for visualizing and analyzing the spatial distribution of fixations. Figure 10.3a shows a scene of three people playing cards, with over 2000 fixations produced by 21 participants while viewing this scene. Each red dot represents one or more fixations at that image location. From the distribution of fixations, one can see that a large proportion of fixations landed on the persons in the scene, and in particular on their faces. The fixation plot of Fig. 10.3a can be visualized in different ways that make areas with high fixation densities more explicit.

In the gridded heat map (see Fig. 10.3b), the stimulus area is partitioned in a rectangular grid of square cells, in this case of 16 by 12 cells. The grey level of each cell is proportional to the number of fixations that landed within the cell. The gridded heat map makes it easy to see the image areas with the most fixations; here the two faces on the left, the card deck on the table and the picture on the wall. The visualization of fixations using gridded heat maps has a clear advantage over raw fixation maps (as in Fig. 10.3a), and it allows to make statistical comparisons of fixation frequency counts between different groups or experimental conditions fairly easy. On the negative side, this technique introduces artificial boundaries between cells that have no relation to the scene content.

Alternatively, the fixations can be visualized using a smooth heat map. In this method, each fixation is replaced by a 2D-Gaussian with a pre-defined standard deviation, and these Gaussians are added together, resulting in a smooth fixation

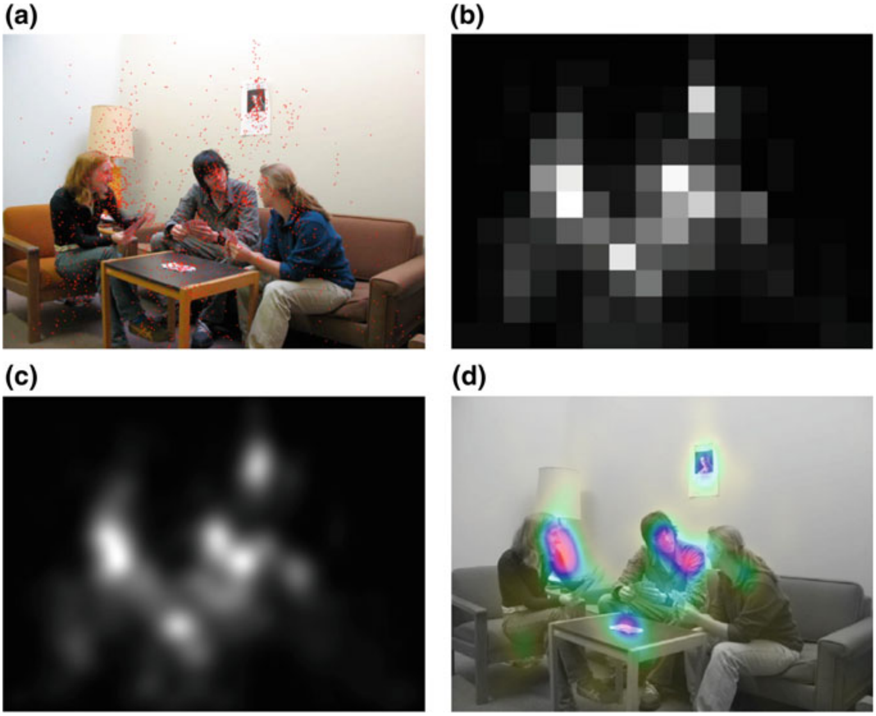


Fig. 10.3 **a** Image of a social scene with fixations overlaid as red dots. **b** Gridded heat map with 16 × 12 cells. Grey level of each cell is proportional to the number of fixations that landed within the cell. **c** Smooth heat map, where fixations have been replaced by 2D-Gaussians. The grey level at is point is proportional to the height of the heat map. **d** Variation of smooth heat map where only the peaks are shown in colour with the color proportional to the height of the heat map, whereas the original image is shown in areas of low heat map values

map, which is often referred to as a heat map. In the case of Fig. 10.3c, the standard deviation was chosen to be 20 pixels (with an image size of 800 by 600 pixels). There are, however, no hard and fast guidelines regarding the best choice of the standard deviation. This can often make it difficult to compare heatmaps across different experimental conditions. The heat map can also be visualized by assigning a grey level proportional to the height of the heat map (Fig. 10.3c) or by assigning a Gaussian standard deviation based on fixation duration. Again, the heat map makes it clear which image areas were fixated the most. An interesting variation of heatmaps was created by Woodford (2014), where the heat map is overlayed over the original image, but only the peaks are shown in colour with the color proportional to the height of the heat map, whereas in regions with low values, the original image is shown (see Fig. 10.3d).

10.3.2.1 Limitations of Heat Maps

Heat maps are good visualization tools and can give a quick overview of fixation patterns for a large number of participants and to easily locate fixation hotspots such as the faces in Fig. 10.3. For several reasons, however, the analysis and the interpretation of heat maps are difficult. First, heat maps are specific to a particular stimulus layout, and hence can be used only to compare heat maps of fixation patterns obtained for the same images (or images with identical spatial layouts). Second, fixation hotspots may not be due to participants fixating a particular scene point, but could be due to the fact that participants tend to fixate the center area of stimuli more frequently than peripheral areas. This central fixation bias is illustrated in Fig. 10.4, which shows the heat map obtained from the fixations of 21 participants looking at 10 different images. The (red) hotspots in the center indicate a strong bias towards fixating central image areas independent of particular scene contents (Tatler, 2007).

Box 1: Comparison of Heat Maps

Heat maps must be analyzed statistically to establish differences and similarities obtained under different experimental conditions, as well as to establish whether certain fixation hotspots are significantly higher than the rest of the heat map (for a detailed overview of heat map comparisons see Le Meur & Bacchino, 2013). The similarity of heat maps can, for example, be measured using simple correlation between two heat maps, measured over all image positions. Second, one can compare the two heat maps using the Kullback-Leibler divergence, a non-symmetric measure of the difference between two probability distributions (Kullback & Leibler, 1951). Third, the heatmaps can be compared using ROC analysis (Green & Swets, 1966) where the ROC curve

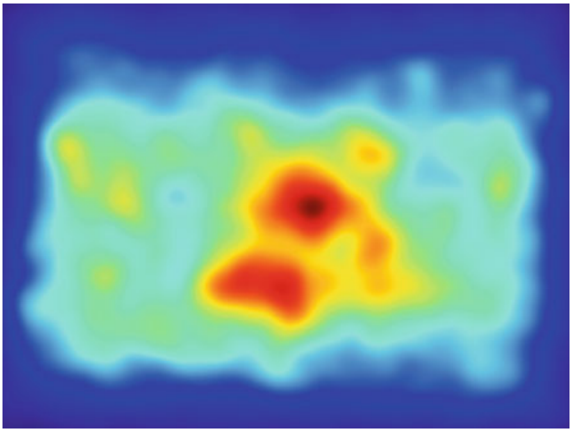


Fig. 10.4 Heat map obtained from fixations by 21 participants looking at 10 different images. Red areas indicate regions with high fixation counts and blue indicates areas with low fixation counts

expresses the correspondence between the two maps (for more details see Judd, Ehinger, Durand, & Torralba, 2009; Le Meur & Baccino, 2013).

The three methods are unproblematic for gridded heat maps because the number of fixations is independent between different cells of the grid. For smooth heat maps, however, the comparison is more complicated due to the fact that independence between nearby positions is lost due to the spatial smoothing with the Gaussian filter. This is a situation similar to the one in the analysis of fMRI results, where a smoothing is applied to the raw data (Friston, Jezzard, & Turner, 1994; Friston et al., 1995). For eye movements, Caldara and Miellet (2011) has proposed a pixelwise comparison of heat maps that takes this dependence into account. In general, it is best to rely on bootstrapping methods (Efron & Tibshirani, 1993) for the statistical analysis of heat maps.

The methods presented in this section visualize the spatial distribution of fixations of many participants in an image, and they are easy to understand. Their statistical analysis is well understood, but their interpretation is a bit more difficult. They express the spatial distribution of fixations in image coordinates without reference to specific scene content, so the analysis with respect to scene content requires additional work.

10.3.3 Area of Interest Analysis of Eye Movements

In order to take the content of a visual scene into account, a second group of eye movement measures is concerned with analyzing the proportion and duration of fixations in pre-defined areas of interest (AOIs). In a visual search display, one might define AOIs for each target and each distractor. In images of more complex scenes, for example, in work with social scene stimuli (e.g., Birmingham et al., 2008a, b), the AOIs can include the bodies, faces and eye regions of the persons in the scenes (see Fig. 10.5). An AOI analysis consists of counting all fixations or determining the fixation proportions or average fixation durations (dwell time) for the eye movements within each AOI.

10.3.3.1 Advantages of AOI Analyses

AOI analyses have proven useful for assessing eye movement patterns with simple and complex images. Fixation frequencies or dwell times on different AOIs can be compared between different experimental groups or conditions and even between different stimuli provided they have the same image structure, i.e. they contain the same types of regions (for example, people, faces or eyes in social scenes). Assuming a uniform distribution of fixations across images, the number of fixations or dwell time in an AOI is proportional to the AOI area. If fixation counts or durations need

Fig. 10.5 Areas of interest of the social scene. The color of each area has been chosen randomly

to be compared between different AOIs, then it is advisable to normalize the number or duration of fixations by the AOI area and obtain the number of fixations or dwell time per unit area in each AOI.

The definition of AOIs is relatively straight-forward with static images and many eye tracking systems provide software for defining static AOIs. With dynamic stimuli, the definition of AOIs becomes more difficult. Example of such cases include movies as stimuli or with eye tracking in real-world experiments, where participants are wearing a mobile eye tracker and fixations are projected onto the recorded scenes by the eye tracking software. This is illustrated in Fig. 10.6, where the rider changes position and size in every movie frame, leading to an enormous effort for defining the AOI rider for long frame sequences. One possible solution to this problem relies on so-called keyframing. In this technique, the relevant AOIs are defined only in certain frames (key frames), and the computer then interpolates the AOI boundaries for all the frames in between (Igarashi, Moscovich, & Hughes, 2005). With this technique, the effort of encoding AOIs is substantially reduced. Under certain circumstances, in particular with computer-generated stimuli, AOIs can be automatically generated with relatively little effort. In mobile eye tracking it is also possible to utilize markers (unique patterns printed on a piece of paper) to define regions of interest in the real environment. These markers can then be used to translate world-view fixation

Fig. 10.6 Three frames from the opening scene of the movie "The Good, the Bad and the Ugly". The size and location of the rider change in each frame, increasing the effort of defining areas of interest

coordinates to fixation coordinates based on the area around a marker instead (e.g., Krassner, Patera, & Bulling, 2014)

10.3.3.2 Disadvantages of AOI Analyses

The AOI analysis is useful for the comparison of fixation patterns for identical or structurally similar stimuli with a common set of elements. It can, however, be difficult to use AOIs to derive meaningful results if highly diverse stimuli, e.g., when a mixture of social scenes, landscapes and artificial stimuli is being used. In this case, there is no obvious way of comparing AOIs across the set of stimuli.

The sections on the spatial analysis of fixation patterns using heat maps and areas of interest have introduced several analysis methods that are very popular in the eye movement literature for analyzing the spatial allocation of attention in images. On the downside, the methods do not capture the dynamics of eye movement behaviour, unless one analyzes their development over time. This is, however, only possible with a very coarse temporal resolution. Methods more suitable for the temporal analysis of eye movements are discussed in the next section.

10.3.4 Temporal Analysis of Eye Movements

Temporal eye movement analysis is concerned with the dynamic aspects of eye movements, the temporal sequence of fixations and saccades as they unfold over time (i.e., the scanpath). In the context of AOI analysis, it is concerned with the transitions between different AOIs and typically focuses on AOI transitions matrices. For example, with a set of stimuli derived from an image of two people sitting at a table, the regions head1, torso1, head2, torso2, table, and chairs, could yield a transition matrix between the AOI regions as shown in Table 10.1.

Each row in Table 10.1 indicates the probability of saccades from the AOI on the left to each of the AOIs. For example, if a fixation at time t is in the region head2, then the next fixation at time t + 1 is in the region head3 with a probability of 0.3, in the region torso1 with a probability of 0.1, and so on. The AOI transitions in

Table 10.1 Example of a transition matrix for six AOI regions

	head1	torso1	head2	torso2	chairs	table
head1	0.1	0.4	0.3	0.1	0.1	0.0
torso1	0.2	0.2	0.1	0.2	0.2	0.1
head2	0.3	0.1	0.3	0.3	0.0	0.0
torso2	0.1	0.1	0.4	0.2	0.1	0.1
chairs	0.1	0.3	0.1	0.3	0.1	0.1
table	0.3	0.1	0.3	0.1	0.2	0.0

Table 10.1 characterize the dynamic sequence of eye movements for a given experimental condition, and comparisons to other conditions can be made by assessing the similarity of the transition matrices. It is important to point out, however, that such transition matrices capture only the overall characteristics of the transitions, but not, for example, their change over time.

In specific applications, the transition matrices have been modeled using Markov and hidden Markov models (Boccignone & Ferraro, 2014; Holmqvist et al., 2011; Stark & Ellis, 1981), for example, in face recognition (Chuk, Chan, & Hsiao, 2014) or participants solving items from Raven's Advanced Progressive Matrices Test (Hayes, Petrov, & Sederberg, 2011; see also Boccignone's chapter in this volume). As pointed out earlier, such AOI-based analyses are feasible only when the class of stimuli is sufficiently restricted, e.g. contain a common set of regions. Otherwise, the transition matrices cannot be compared across different stimuli. In contrast, we introduce in the next section a method for characterizing scanpaths obtained with related as well as unrelated images.

10.3.5 Summary

In this section we have introduced the analysis of eye movements and reviewed some common spatial and temporal methods of analysing fixations. We have discussed spatial techniques such as using heatmaps to determine where people look as well as AOI analyses that include information about the scene content. We have briefly discussed the use of transition matrices to analyze the dynamic aspects of using AOI's.

An important group of other dynamic fixation measures are the scanpath analyses that were developed in reading research. In these studies, scanpath events, such as backtracking, regression, look-ahead and return have a direct interpretation. In scene viewing, however, scanpath measures are more difficult to interpret, and there is little consensus on good measures for scanpath comparison. Almost no measures have been developed for directly quantifying a single scanpath in an image. In the sections below, we describe one recently developed method for quantifying a single scanpath, then we describe several techniques for comparing similarities between scanpaths.

10.4 Recurrence Quantification Analysis (RQA)

Recurrence analysis has been used successfully as a tool for describing complex dynamic systems, for example, climatological data (Marwan & Kurths, 2002), electrocardiograms (Webber & Zbilut, 2005), or postural fluctuations (Pellecchia & Shockley, 2005; Riley & Clark, 2003), which are inadequately characterized by standard methods in time series analysis (e.g., Box, Jenkins, & Reinsel, 2008). It has also been used for describing the interplay between dynamic systems in cross-recurrence analysis, e.g., for analyzing the postural synchronization of two persons (Shockley, Baker, Richardson, & Fowler, 2007; Shockley, Santana, & Fowler, 2003; Shockley & Turvey, 2005). The fundamental idea of recurrence analysis is to analyze the temporal pattern of repeated (recurrent) events, for example, the same tidal height in tide analysis, the same waves in the ventricular cycle in electrocardiogram analysis, the same postural relation in the analysis of postural synchronization, or the same fixated locations in an image. Here we describe a simplified version of recurrence analysis based on categorical data. This simplified analysis is ideal because it allows for direct interpretation of the various recurrence measures as they apply to categorical fixation data.

10.4.1 Categorical Recurrence Quantification

Richardson, Dale and colleagues have generalized recurrence analysis to the analysis of categorical data and have used it for analyzing the coordination of gaze patterns between individuals (e.g., Cherubini, Nüssli, & Dillenbourg, 2010; Dale, Kirkham, & Richardson, 2011; Dale, Warlaumont, & Richardson, 2011; Richardson & Dale, 2005; Richardson, Dale, & Tomlinson, 2009; Shockley, Richardson, & Dale, 2009). For example, Dale, Warlaumont & Richardson, 2010, quantified the coordination between a speaker and a listener's eye movements as they viewed actors on a screen (see Fig. 10.7).

In this experiment, one person (left) was talking about a particular television character while the other person (right) listened in. Pictures of characters including the one discussed were displayed in a 2 × 3 grid in front of the speaker and the listener. The eye movements were categorized into a grid corresponding to the six photos of the television characters. The eye movements generated by speaker and listener during a 60 s period are shown in a series of numbers 1–6 corresponding to the six grid locations. Dale et al. (2010) used a cross-recurrence analysis of these fixations and were able to show that the listener tended to follow the same fixation patterns as the speaker, with a delay of approximately 2 s.

10.4.2 Generalized Recurrence Quantification

Dale et al.'s cross-recurrence analysis can provide an overall measure of similarity across two eye movement sequences (i.e., a form of scanpath comparison that we will discuss later in this chapter). Recently, we have introduced a generalized form of Dale et al.'s categorical recurrence analysis to characterize gaze patterns of a single observer (Anderson, Bischof, Laidlaw, Risko, & Kingstone, 2013), and we were able to show that it is a very useful tool for encoding general characteristics of fixation sequences. The essential idea is to consider each fixation as one in a series throughout

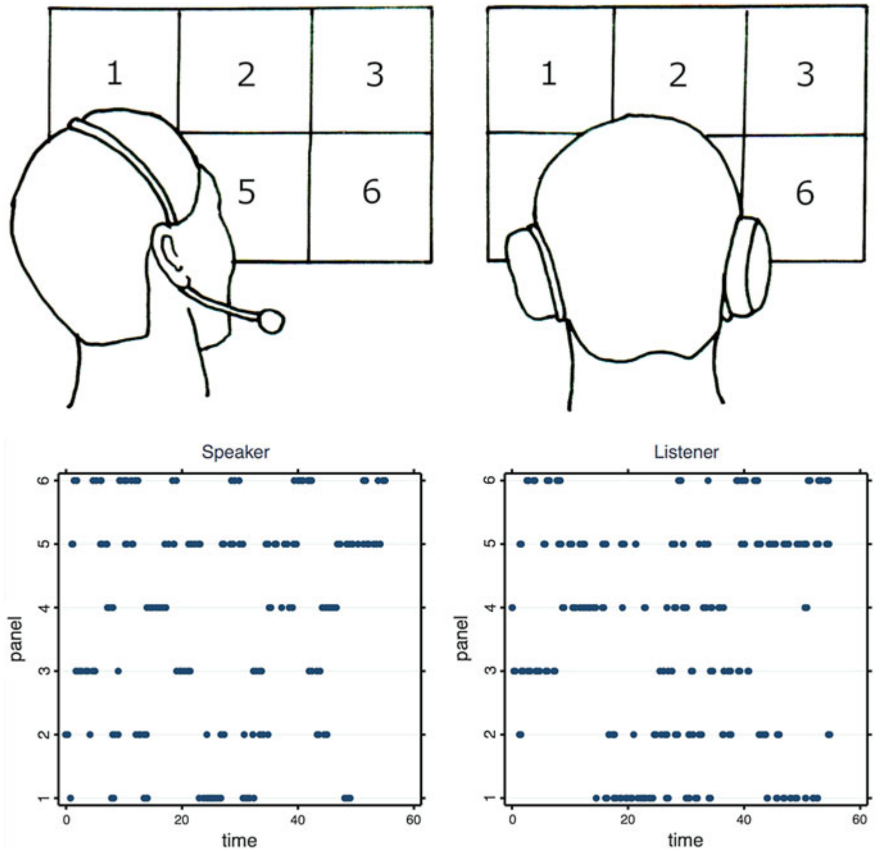


Fig. 10.7 Dale, Warlaumont, and Richardson's (2011) experiment. Upper panel: The left person is speaking about a 2 × 3 grid displaying television characters while the right person listens in. Lower panel: Each interval of the one-minute experiment produces a numeric value representing the panel that was fixated. Using these data, a cross-recurrence analysis compares the panels fixated by speaker and listener over different time lags. Figure adapted from Dale et al. (2011)

the image. Recurrence analysis quantifies when one fixation might overlap in space with another (one person looking at the same place twice or more times in the course of looking at a scene). In the following, we first review the fundamentals of recurrence quantification analysis (RQA) for use with categorical eye movement data, with specific consideration of fixation sequences. Then we describe and interpret the main RQA measures.

10.4.3 Formal Definition of Recurrence

Consider a sequence of N fixations f_i, $i = 1,\ldots, N$, with each f_i characterized by its spatial coordinates. Two fixations are considered recurrent, if they are close together. “Closeness” can be defined in several ways, but in general, one can define recurrence r_{ij} as

$$r_{ij} = \begin{cases} 1, & d\left(f_i, f_j\right) \le \rho \\ 0, & otherwise \end{cases} \tag{10.1}$$

where d is some distance metric (usually Euclidian distance), and ρ is a given radius, i.e., two fixations are considered recurrent if they are within a distance ρ of each other. Guidelines for selecting a value for ρ are introduced further below.

10.4.4 Recurrence Plot

Recurrence can be represented in a recurrence diagram, which plots recurrences of a fixation sequence over all possible time lags. The essential starting point of a recurrence analysis is drawing this plot. While it is not strictly a necessary step, all of the recurrence measures are based on patterns that emerge from this plot. The plot is drawn as follows: If fixations i and j are recurrent (i.e. if $r_{ij} = 1$), then a dot is plotted at position i,j. All fixations are recurrent with themselves (since $d(f_i, f_i) = 0$), hence all elements on the major diagonal—the line of incidence—are recurring. Furthermore, since distance metrics are symmetric (i.e., $d(f_i, f_j) = d(f_j, f_i)$) recurrence plots are also symmetric. This is illustrated in Fig. 10.8. Figure 10.8a shows a landscape image with a scanpath consisting of 30 fixations, with repeated fixations mainly in the cloud formation. The recurrence plot for the fixation scanpath is shown in Fig. 10.8b, with each recurrence indicated by a red dot. A recurrence plot can be generated for each sequence of fixations, e.g. for each experimental trial.

10.4.5 Recurrence Quantification Measures

The recurrence diagram provides a useful visual representation of the recurrence patterns for a fixation sequence, but it must be complemented by a recurrence quantification analysis for comparison across different fixation sequences, e.g., across different trials, participants and experimental conditions. Here, we describe a subset of RQA measures, those that are particularly useful for the analysis of fixation sequences (see Webber & Zbilut, 2005 and Marwan & Kurths, 2002 for a complete list of measures). All of these measures describe certain patterns that emerge in the

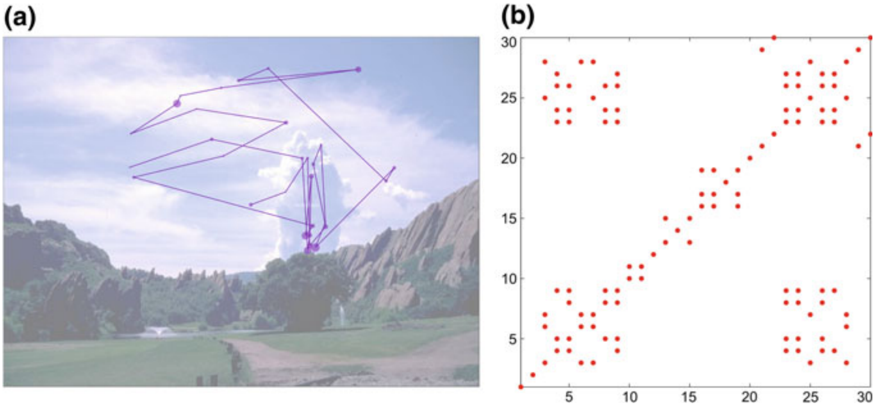


Fig. 10.8 **a** Image of a landscape with the scanpath produced by one participant overlaid (the size of the circle at each fixation point represents the duration of the fixation); **b** recurrence plot of the scanpath in **a**

recurrence plot. For example, the recurrence measure itself is simply a percentage of the total number of possible recurrent points. Determinism is the percentage of recurrent points that form a diagonal line on the plot, while laminarity is the percentage of vertical and horizontal lines. Given the symmetry of the recurrence diagram, these quantitative measures are usually extracted from the upper triangle of the recurrence diagram, excluding the line of incidence, which does not add any additional information (recall that the line of incidence indicates that each fixation is recurrent with itself).

First, we give some useful definitions: Given a fixation sequence of length N, f_i, i = 1,…, N, let R be the sum of recurrences in the upper triangle of the recurrence diagram, i.e., $R = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r_{ij}$. Let D_L be the set of diagonals lines, H_L the set of horizontal, and V_L the set of vertical lines, all in the upper triangle, and all with a length of at least L, and let |·| denote cardinality.

10.4.5.1 Recurrence

The **recurrence** measure is defined as

$$\mathrm{REC} = 100 \frac{2R}{N(N-1)} \tag{10.2}$$

For a sequence of N fixations Recurrence represents the percentage of recurrent fixations (i.e. it indicates how often an observer re-fixates previously fixated image positions). As fixations are plotted sequentially, the larger the distance between a recurrent point and the main diagonal, the larger the time interval (in number of fixations) between the original fixation and the re-fixation.

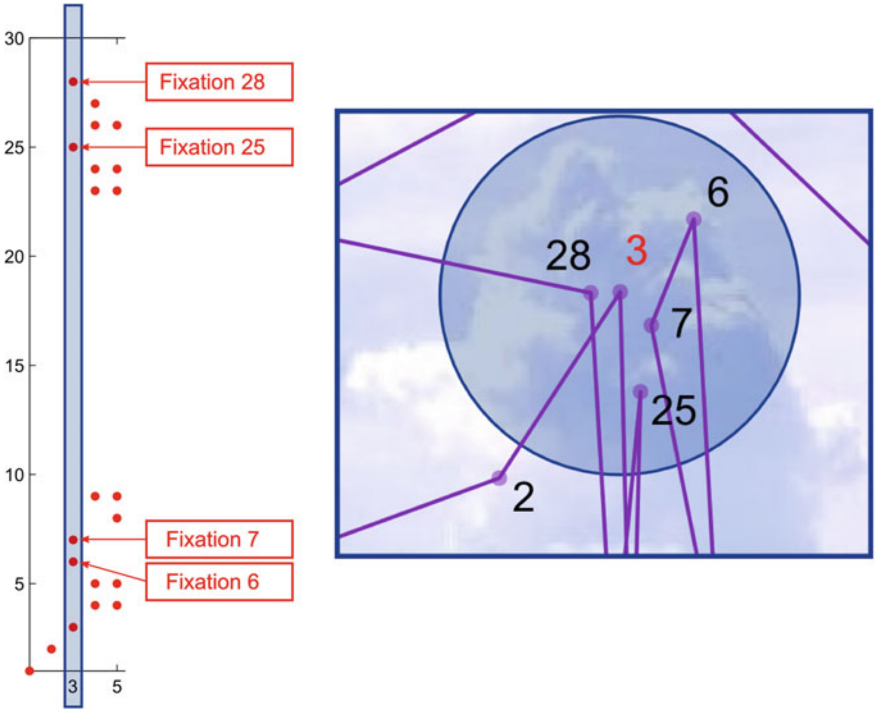


Fig. 10.9 Detailed view of the recurrence plot and the fixation plot shown in Fig. 10.8. Fixations 6, 7, 25 and 28 are within the radius ρ of Fixation 3, as indicated by the blue circle, producing the shown recurrences in the recurrence plot

Figure 10.9 provides an illustration of recurrence. Fixations 6, 7, 25 and 28 in Fig. 10.9 are all close-by, more specifically within a radius of ρ of fixation 3, hence these fixations are recurrent.

10.4.5.2 Determinism

The **determinism** measure is defined as

$$\mathrm{DET} = 100\frac{|D_L|}{R} \tag{10.3}$$

Determinism measures the proportion of recurrent points forming diagonal lines in the recurrence plot and represents repeating fixation sequences in the recurrence diagram. This may represent two areas of a scene where one fixation is more likely to follow another. For example, if a person looks first at one person, then another in that same order twice in a trial. It is a small section of a scanpath repeated within a trial. The minimum line length of diagonal line elements is typically set to L =

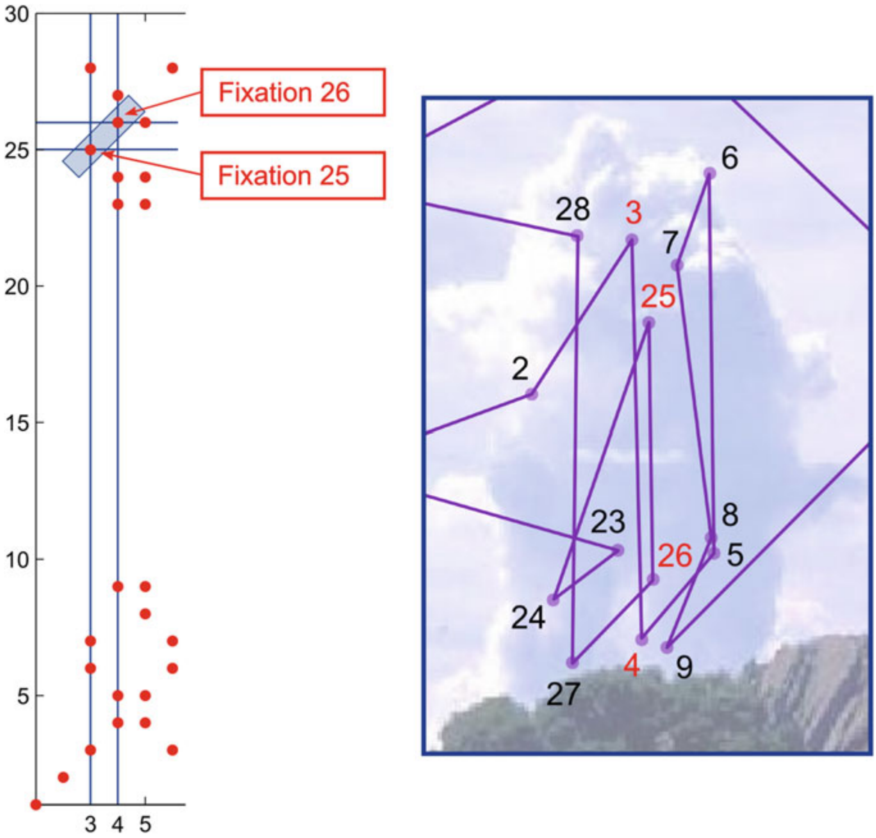


Fig. 10.10 Detailed view of the recurrence plot and the fixation plot shown in Fig. 10.8. Fixations 3 and 4 as well as fixations 25 and 26 follow the same path. This creates a diagonal line on the recurrence plot and defines a deterministic recurrence

2. The length of the diagonal line element reflects the number of fixations making up the repeated scanpath, and the distance from the diagonal reflects the time (in numbers of fixations) since the scanpath was first followed.

Figure 10.10 provides an illustration of determinism. Fixations 3 and 4 as well as fixations 25 and 26 follow the same path, defining a deterministic recurrence.

10.4.5.3 Laminarity

The **laminarity** measure is defined as

$$\mathrm{LAM} = 100\frac{|H_L| + |V_L|}{2R} \tag{10.4}$$

Vertical lines represent areas that were fixated first in a single fixation and then re-scanned in detail over consecutive fixations at a later time (e.g., several fixations later), and horizontal lines represent areas that were first scanned in detail and then re-fixated briefly later in time. Again, we set the minimum line lengths of vertical and horizontal lines to $L = 2$. We find that the recurrence diagrams sometimes contain recurrence clusters (with horizontal and vertical lines), indicating detailed scanning of an area and nearby locations. Laminarity indicates that specific areas of a scene are repeatedly fixated, for example, when an observer returns to an interesting area of the scene to scan it in more detail.

An example of laminarity is shown with fixations in Fig. 10.9. The position fixated in fixation 3 is refixated in a more detailed inspection in fixations 6 and 7.

10.4.5.4 Center of Recurrence Mass

The **center of recurrence mass** (corm) is defined as the distance of the center of gravity of recurrent points from the line of incidence, normalized such that the maximum possible value is 100.

$$\mathrm{CORM} = 100\frac{\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}(j-i)r_{ij}}{(N-1)R} \tag{10.5}$$

This measure indicates approximately where in time most of the recurrent points are situated. Small corm values indicate that re-fixations tend to occur close in time, i.e. most recurrent points are close to the line of incidence. For example, if an observer sequentially scans three particular areas of a scene in detail and never returns to those areas later in the trial, most of the recurrent points would fall close to the line of incidence. This would be represented by a small corm value. In contrast, large corm values indicate that re-fixations tend to occur widely separated in time, i.e. most recurrent points are close to the upper left and lower right corners of the recurrence diagram. This occurs, for example, when an observer re-fixates only one scene area, once at the beginning and once at the end of the fixation sequence, but not in between.

Box 2: RQA and Fixation Duration

Fixation duration can be an important indicator of processing during fixation (Holmqvist et al., 2011, pp. 377ff). RQA can be generalized to take fixation durations into account (Anderson et al., 2013).

Given a fixation sequence f_i, i = 1,…, N, and the associated vector of fixation durations t_i, i = 1,…, N, one can redefine recurrence r_{ij}^t as

$$r_{ij}^t = \begin{cases} t_i + t_j, & d(f_i, f_j) \le \rho \\ 0, & otherwise \end{cases} \tag{10.6}$$

with the (Euclidian) distance metric d and the radius ρ. With the modified recurrence definition of r^t_{ij}, the RQA measures have to be renormalized. Let $R = \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} r^t_{ij}$, and $T = \sum_{i=1}^{N} t_i$. Then the revised definitions for *REC, DET, LAM and CORM* are as follows.

$$REC^t = 100 \frac{2R^t}{(N-1)T}$$

$$DET^t = \frac{100}{R^t} \sum_{(i,j) \in D_L} r^t_{ij}$$

$$LAM^t = \frac{100}{2R^t} \left(\sum_{(i,j) \in H_L} r^t_{ij} + \sum_{(i,j) \in V_L} r^t_{ij} \right)$$

$$CORM^t = 100 \frac{\sum_{i=1}^{N-1} \sum_{j=i+1}^{N} (j-i) r^t_{ij}}{(N-1)^2 T}$$

In summary, the recurrence and corm measures capture the global temporal structure of fixation sequences. They measure how many times given scene areas are re-fixated (recurrence) and whether these re-fixations occur close in time or widely separated in a trial (corm). In contrast, determinism and laminarity measure the finer temporal structure of fixation sequences. Specifically, they indicate sequences of fixations that are repeated (determinism) and points at which detailed inspections of an image area are occurring (laminarity). These measures can then be compared across different types of images, experimental contexts and participants to assess the dynamic structure of eye movements.

10.4.6 *Selection of the Recurrence Radius*

As explained in Sect. 10.5.3, two fixations are considered recurrent if they are within a distance ρ of each other, with the radius ρ being a free parameter. The number of recurrences is directly related to the radius. As the radius ρ approaches zero, (off-diagonal) recurrences approach zero, and as ρ approaches the image size, recurrences approach 100%. The dependence of recurrence on radius leads to the obvious question of how an appropriate radius for recurrence analysis should be selected.

Webber and Zbilut (2005) suggested several guidelines for selecting the proper radius, including the selection of a radius such that percentage of recurrences remains low, for example about 1–2%. In the case of eye movements, one can apply more content-oriented criteria. For example, fixations can be considered as recurring if their foveal areas overlap, using a radius size of 1–2° of visual angle. This is discussed further by Anderson et al. (2013).

10.4.7 Statistical Analyses of Recurrence Measures

In this section, we discuss the distribution of RQA measures as well as their correlations. This is important when RQA measures are used to compare and discriminate between experimental conditions and groups. Figure 10.11 shows the histograms of the measures recurrence, determinism, laminarity and corm obtained with 104 participants viewing 1872 images, each for a duration of 10 s. As the histograms show, the RQA measures are distributed more or less symmetrically, with the exception of Recurrence, permitting the use of analyses of variance for their statistical analysis.

Figure 10.12 illustrates scatter diagrams and correlations between RQA measures, obtained for the same group of participants and the same images as in Fig. 10.11. As the Figure shows, the correlations between RQA measures vary substantially. While the correlation between laminarity and corm is essentially zero, other correlations, e.g. between recurrence and laminarity are relatively high. Such dependencies between the RQA measures are not surprising, given that they are defined and selected to be easily interpretable, rather than being independent of each other. Indeed, laminarity and determinism are very closely related to recurrence as they are themselves percentages of recurrent points.

For this reason, it is useful to consider several or all RQA measures for comparing and discriminating different groups of participants or different experimental condi-

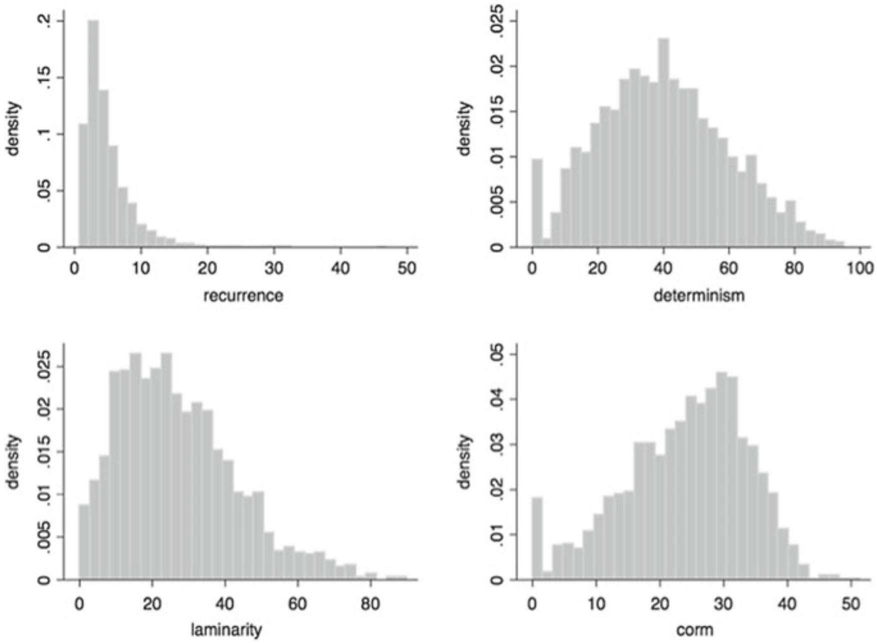


Fig. 10.11 Histograms of the measures recurrence, determinism, laminarity and corm obtained with 104 participants viewing 1872 images, each for a duration of 10 s

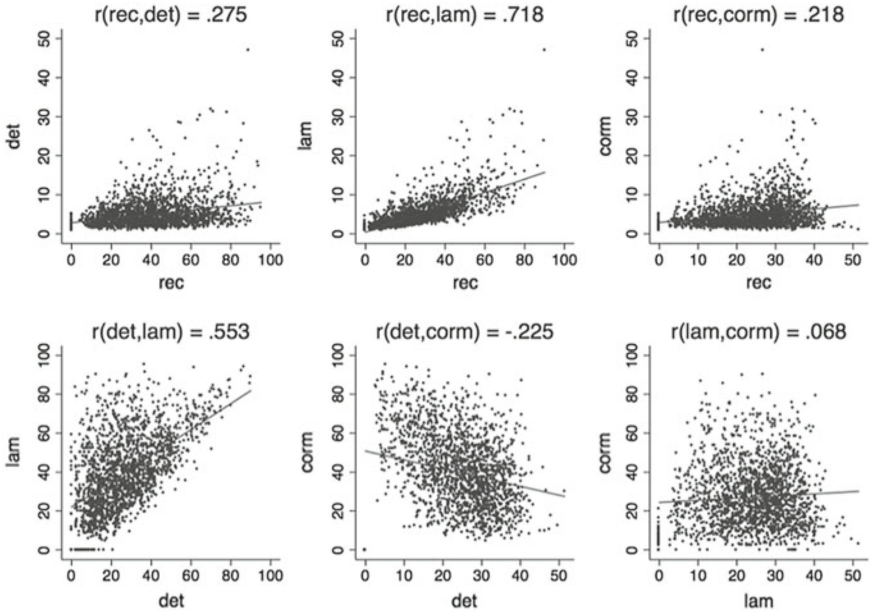


Fig. 10.12 Correlations between RQA measures. Each panel shows a scatterplot for a pair of measures, the regression line and the correlation value

tions. In other words, it is advisable to take all RQA measures into account and use discriminant analyses for distinguishing different scanpath patterns, a multivariate analysis of variance, analysis of variance, or canonical correlation in regression-type analysis. This is further described below.

10.4.8 Discrimination of Gaze Patterns Using RQA

The previous sections showed that RQA is a useful method for discriminating eye movement patterns based purely on dynamic characteristics. Each of the RQA measures used (recurrence, determinism, laminarity and corm) can discriminate to a certain extent between experimental groups or different groups of participants. At the same time, however, the measures are not independent, as the correlations between RQA measures in the previous section showed. To elaborate on this aspect, we evaluated how well the RQA measures discriminate between eye movement patterns of participants who viewed social scenes under natural viewing conditions (Full-View condition) and participants who view the same scenes through a gaze-contingent window of limited spatial extent (Restricted-View condition).

Figure 10.13 shows histograms for the measures recurrence, determinism, laminarity and corm, for the Full-View and the Restricted-View conditions. For clarity of presentation, the histograms were smoothed using a Gaussian smoothing filter with half-width = 2% or $\sigma = 0.849\%$ of the total range. An inspection of the four histograms shows that each measure can discriminate to some extent between the two groups, but, at the same time, there is a substantial overlap that varies between RQA measures. Consequently, classification accuracy of the discrimination between the two groups varies accordingly, with accuracy for the recurrence measure at 75%, for the determinism measure of 88.6%, for laminarity at 77.7%, and for corm at 70.8%. In other words, each of the measures can discriminate between the two experimental groups, but discrimination performance is far from perfect.

Discrimination between the eye movement patterns of the two groups can be improved substantially using combinations of RQA measures. This is possible because the measures are somewhat, but definitely not perfectly correlated. For example, using the measures recurrence + laminarity + corm improves discrimination accuracy to 85.5%, and the discrimination accuracy using all four measures is 94.5%, a substantial improvement over the use of single measures. The results show

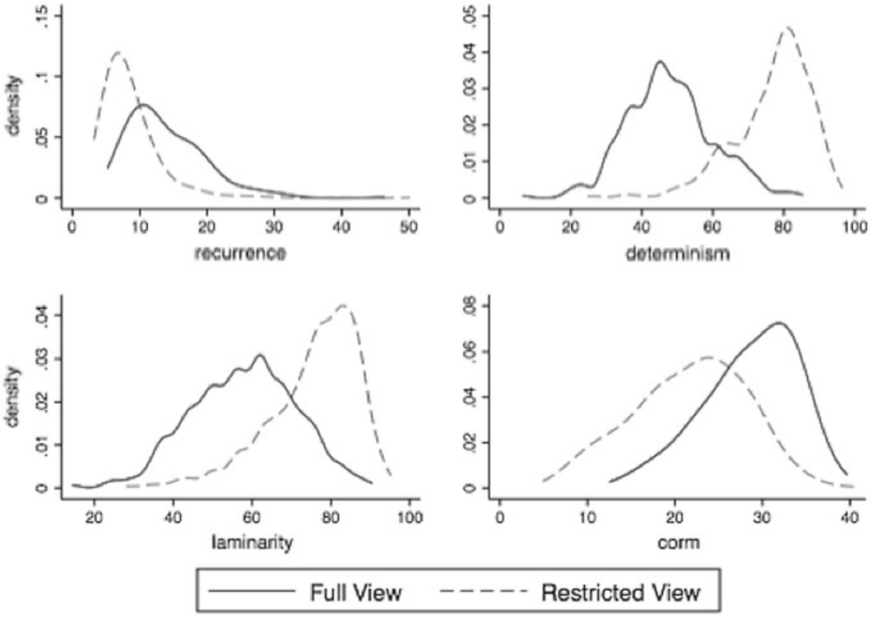


Fig. 10.13 Smoothed histograms for the measures recurrence, determinism, laminarity and corm for the full-view (natural viewing) condition and the Restricted-View (viewing through a small, gaze-contingent window). The histograms were smoothed with a Gaussian filter with half-width = 2%

clearly that the RQA measures are sensitive to, and useful for discriminating gaze patterns under Full-View and Restricted-View conditions.

Box 3: Potential Limitations of the RQA Method

1. *RQA can measure the characteristics of re-fixations, but it cannot measure idiosyncratic eye movement characteristics, such as saccadic distances, angles, and so on.*

The purpose of RQA is to capture the dynamic aspects of scanpaths and should be used **in addition** to other measures (e.g. fixation and saccade characteristics), not replace them.

2. *RQA analyzes scanpaths under the assumption that there is always recurrence.*

This is not quite correct. First, the recurrence measure can always be computed and is zero when there is no recurrence. It is, however, correct that the other measures, determinism, laminarity, and corm are undefined if there is no recurrence. Second, one can always compute the RQA measures with larger recurrence radii ρ (see Eq. 10.1) and possibly with a whole range of recurrence radii. As the recurrence radius ρ increases, so does the number of recurrences (see Sect. 10.5.4 and Anderson et al., 2013).

3. *RQA cannot be applied if there is a single fixation.*

For this extreme case, it is indeed true that RQA cannot be applied, but the same is true for many other measures including those describing saccade characteristics.

4. *What are the computational limitations of RQA?*

The number of cells in the recurrence plot increases with the square of the number of fixations n in a scan path. Consequently, computing RQA measures becomes computationally more and more expensive as n increases. In our own work, we have rarely used RQA for scanpaths with more than 100 fixations.

10.4.9 Summary

In this section we have presented a new technique, RQA, for quantifying the characteristics of a single scanpath. This is a useful measure that compliments other spatial and temporal measures of eye movement characteristics. We have described recurrence in general, several other useful measures associated with RQA, and how to use RQA to discriminate gaze patterns across groups of participants or stimuli. In the next section we will describe methods to compare between two scanpaths, a technique that is useful for describing how two eye movement patterns are similar to each other.

10.5 Comparison of Scanpaths

Eye movements unfold over time, hence one can also examine the inter-relationship between sequences of eye movements. In his seminal work, Yarbus (1967) noticed that observers displayed similar scan patterns in viewings of Repin's painting "The Unexpected Visitor" and concluded that "observers differ in the way they think and, therefore, differ also to some extent in the way they look at things" (p. 192). A brief inspection of these scan patterns reveals that they are complex and non-random, and they contain sequences of repeated fixations. Noton and Stark (1971) noticed that observers tend to show similar scan patterns during encoding and later recognition of images. According to their "Scanpath Theory", the sequence of fixations during the first viewing of a stimulus is stored in memory as a spatial model, and stimulus recognition is facilitated through observers following the same scanpath during repeated exposures to the same image. These early observations were made informally by visual inspection, but later research has aimed at quantifying the similarity of scanpaths, for the same observer at different time points or when solving different tasks, or between different subjects.

In the following sections, we describe the scanpath comparison methods that have been introduced in the literature (see Anderson, Anderson, Kingston, & Bischof, 2014). In each case, we give a short description, and the reader is advised to consult the original publications for further details.

10.5.1 Edit Distance

One successful way for comparing scanpaths is based on the string edit distance (Bunke, 1992; Levenshtein, 1966; Wagner & Fischer, 1974), which is used to measure the dissimilarity of character strings. In this method, a sequence of transformations (insertions, deletions, and substitutions), is used to transform one string into the other and their similarity is represented as the number of transformation steps between the two strings. This method has been adapted for comparing the similarity of scanpaths (Brandt & Stark, 1997; Foulsham & Kingstone, 2013; Foulsham & Underwood, 2008; Harding & Bloj, 2010; Underwood, Foulsham, & Humphrey, 2009). To achieve this, a grid is overlaid on an image, and each cell in the grid is assigned a unique character. Fixation sequences are then transformed into a sequence of characters by replacing the fixation with the character corresponding to the grid cell a fixation falls in. With this approach, scanpaths are being represented by strings of characters, and the dissimilarity of two scanpaths can then be represented by the number of transformations required to convert the string corresponding to the first scanpath to the string corresponding to the second scanpath.

The string edit distance method has been very popular in early scanpath comparison work (e.g., Brandt & Stark, 1997) and has been used subsequently in a variety of experimental contexts (e.g., Harding & Bloj, 2010; Underwood et al., 2009). This is an advantage for researchers wishing to directly compare results to these earlier studies. The main advantage of the string edit measure, however, lies in the fact that it captures the intuitive notion of the distance between two scanpaths (i.e., their dissimilarity) in a simple and straightforward way.

Several criticisms have been raised against the use of edit distance for scanpath comparisons. First, as described in Sect. 10.4.2 for gridded heatmaps, the grid is defined independently of image content. It may thus be too coarse in regions of interest while being too fine in other regions. Second, two fixations may be considered different even when they are close together, namely if they fall on either side of a grid line. Some variants of the string edit distance have been developed to address these problems. For instance, assigning characters to pre-defined areas of interest allow the researcher to add semantic information to the quantization process (Josephson & Holmes, 2002; West, Haake, Rozanski, & Karn, 2006), but the definition of regions of interest can be time-consuming.

10.5.2 ScanMatch

Cristino, Mathôt, Theeuwes, and Gilchrist (2010) proposed a generalized scanpath comparison method that addresses many of the deficiencies of the string edit distance method. Their generalization aligns eye movement sequences based on the Needleman-Wunsch algorithm (Needleman & Wunsch, 1970), which is used in bioinformatics to compare DNA sequences. In their method, scanpaths are spatially and temporally binned and then recoded to create a sequence of letters that retains fixation location, duration, and sequence information. The two character sequences are compared by maximizing the similarity score computed from a substitution matrix, which in turn provides the score for all letter pair substitutions, and includes a penalty for gaps. Critically, the substitution matrix can encode information about the relationship between specific regions of interest, thus providing the opportunity to include semantic information in the similarity measure.

A major advantage of the ScanMatch method is that it can take into account spatial, temporal and sequential similarity in the comparison of scanpaths. In addition, semantic information can be easily added using the substitution matrix. One disadvantage of this method is that it suffers from the quantization issues inherent to any measure using grids or regions of interest.

10.5.3 Sample-Based Measures

Shepherd, Steckenfinger, Hasson, and Ghasanfar (2010) introduced several measures for assessing the similarity of two scanpaths, which are described in the following subsections. For each of the measures, the eye positions are first resampled uniformly in time (at 60 Hz), and truncated to the length of the shorter sequence. These measures introduced below are sample-based in the sense that they do not require pre-processing of eye-tracking data into discrete fixation-saccade sequences, as is usually the case in eye movement analyses.

10.5.3.1 Fixation Overlap

Fixation overlap is a measure of the similarity of two scanpaths in space and over time. To this effect, two gaze samples are considered overlapping if they are within a predefined radius. This measure is extremely sensitive to differences in absolute timing between two scanpaths, but is slightly less sensitive to differences in position (due to the use of the radius). Given these sensitivities, it is reasonable to expect this measure to perform similarly to the ScanMatch measure, which is also sensitive to the spatial and temporal similarities between two scanpaths.

10.5.3.2 Temporal Correlation

Shepherd et al. (2010) also introduced temporal correlation [see also Hasson, Yang, Vallines, Heeger and Rubin (2008)] as a measure of the similarity between scanpaths. For two scanpaths f and g, the temporal correlation is defined as the average of the correlation between the x-coordinates of f and g and between the y-coordinates of f and g.

This measure is very sensitive to temporal and spatial differences between the two scanpaths. The sensitivity to temporal differences can be advantageous when timing is important, e.g., when the stimuli change over time, such as in videos. The correlation measure is also sensitive to small differences in fixation positions, given that there is no spatial quantization of the fixations. A significant advantage of this method is its use of the straightforward and readily interpretable correlation analysis. This measure is more sensitive to similarities in position than the fixation overlap method, while also taking sequential information into account. However, this strong spatial-temporal sensitivity may be less robust to noisy data than other measures that are grid-based or rely on a radius-based definition of fixation proximity.

10.5.3.3 Gaze Shift

Shepherd and colleagues' (Shepherd et al., 2010) gaze shift measure assesses how similar the saccade times and amplitudes are between two scanpaths. Gaze shift is computed as the correlation between the absolute values of the first derivative of each scanpath and is computed in the same manner as the temporal correlation, but using the first derivative instead of the position.

To smooth the scanpaths and compute their derivatives, each scanpath is convolved with the derivative of a Gaussian filter. Gaze shift is sensitive to the amplitude of the saccade as well as its temporal location, and it reflects how similar two scanpaths are in terms of the sequence of large and small saccades. This captures some aspects of a global viewing strategy, as subjects who produce small saccades within a localized region have very different scanpaths than subjects who produce large saccades within the entire visible area. This is also useful for comparing dynamic stimuli (e.g., video) to assess how subjects respond to temporal changes in the scene.

10.5.4 Linear Distance

Mannan, Ruddock, and Wooding (1995) analyzed the overall similarity of two scanpaths by computing the linear distances between the fixations of the first scanpath and the nearest neighbour fixations of the second scanpath, as well as the linear distances between the fixations of the second scanpath and the nearest neighbour fixations of the first scanpath. These distances are averaged and normalized against randomly generated scanpath sequences. Mannan et al.'s method was further developed by Mathot, Cristino, Gilchrist, and Theeuwes (2012).

The major advantage of the linear distance method is that it does not need to be quantized like, for example, the string edit distance method. It simply compares each fixation of one scanpath with the fixations of another scanpath in terms of their spatial similarity. However, by comparing only nearest neighbour fixations in terms of distance, this method ignores sequential information. To address some of these issues, Mannan et al.'s (1995) method was modified by Henderson, Brockmole, Castelhano, and Mack (2007) to enforce a one-to-one mapping between two scanpaths, provided that they have the same length. The results for the two methods are very similar (Foulsham & Underwood, 2008), which is likely due to the fact that Mannan et al. average the distances from the first to the second and from the second to the first scanpath, hence clusters of fixations in one scanpath are averaged out.

10.5.5 Scasim

Scasim is a scanpath comparison technique developed initially for use in analysing eye movement patterns while reading (von der Malsburg & Vasishth, 2011). It uses similar logic to the string edit and Levenshtein distance metrics, however, it does not require discretization of fixations into regions of interest and it takes fixation duration into account. It compares both the duration and spatial location between fixations by adding or subtracting durations dependent on their distance apart. One unique advantage of Scasim is that it allows the user to specify what sort of cost spatial distance between fixations might have. While by default, this is related to the drop-off of visual acuity from the fovea, this cost can be changed depending on whether spatial distance is more or less important to the hypothesis in question.

10.5.6 MultiMatch

Recently, Jarodzka, Holmqvist, and Nyström (2010), Dewhurst et al. (2012) and Foulsham et al. (2012) introduced the MultiMatch method for comparing scanpaths. The MultiMatch methods consists of five separate measures that capture the similarity between different characteristics of scanpaths, namely shape, direction, length, position and duration. Computation of each MultiMatch measure begins with scanpath simplification, which involves combining iteratively successive fixations if they are within a given distance or within a given directional threshold of each other. This simplification process aids in reducing the complexity of the scanpaths while preserving the spatial and temporal structure.

Following this simplification, scanpaths are aligned based on their shape using a dynamic programming approach. The alignment is computed by optimizing the vector difference between the scanpaths (note, however, that scanpaths may be aligned on any number of dimensions in MultiMatch). This alignment reduces the comparison's sensitivity to small temporal or spatial temporal variations and allows the algorithm to find the best possible match between the pair of scanpaths. All subsequent similarity measures are computed on these simplified, aligned scanpaths.

10.5.6.1 Vector Similarity

The MultiMatch **vector similarity** measure is computed as the vector difference between aligned saccade pairs, normalized by the screen diagonal and averaged over scanpaths. This measure is sensitive to spatial differences in fixation positions without relying on pre-defined quantization. It is a measure of the overall similarity in shape between two fixation-saccade sequences.

10.5.6.2 Length Similarity

MultiMatch **length similarity** is computed as the absolute difference in the amplitude of aligned saccade vectors, normalized by the screen diagonal and averaged over scanpaths. This measure is sensitive to saccade amplitude only, not to the direction, location or the duration of the fixations.

10.5.6.3 Direction Similarity

MultiMatch **direction similarity** is computed as the angular difference between aligned saccades, normalized by p and averaged over scanpaths. This measure is sensitive to saccade direction only, but not to amplitude or absolute fixation location.

10.5.6.4 Position Similarity

MultiMatch **position similarity** is computed as the Euclidean distances between aligned fixations, normalized by the screen diagonal and averaged over scanpaths. This measure is sensitive to both saccade amplitudes and directions.

10.5.6.5 Duration Similarity

MultiMatch **duration similarity** is computed as the absolute difference in fixation durations of aligned fixations, normalized by the maximum duration and averaged over scanpaths. This measure is insensitive to fixation position or saccade amplitude.

The main advantage of the MultiMatch method is that it provides several measures to choose from for assessing scanpath similarity, and each measure on its own captures a unique component of scanpath similarity. Given the multiplicity of measures, it remains, however, difficult to assess which measure, or which set of measures, is most applicable in a given scenario. Furthermore, because each scanpath is initially simplified it is also not clear how robust each measure is to scanpath variations.

10.5.7 Cross-Recurrence Quantification Analysis

The recurrence quantification analysis introduced in Sect. 10.5 can be generalized for the comparison of scanpaths in a cross-recurrence analysis. To this effect, we have generalized the RQA measures for the comparison of scanpaths, and we now introduce these generalized measures.

Consider two fixation sequences f and g that have the same lengths. For sequences of unequal length, the longer sequence is truncated. Within these sequences, two fixations f_i and g_j are cross-recurrent if they match or are close together, i.e., if their distance is below a given threshold. One can define cross-recurrence c_{ij} as

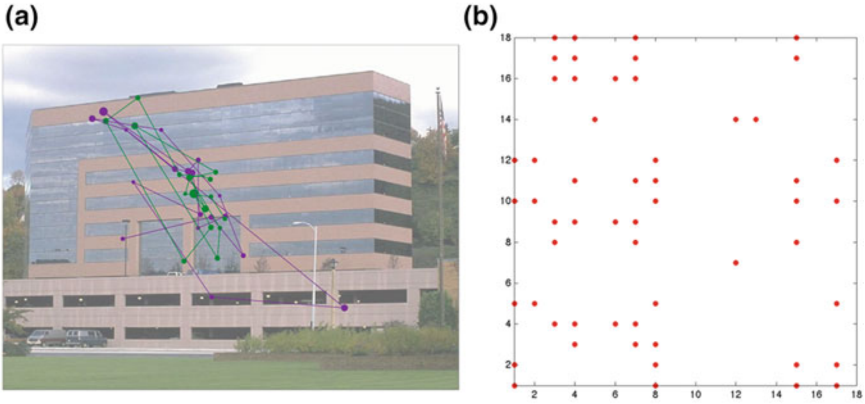


Fig. 10.14 **a** Image of building with two scanpaths (purple and orange) produced by the same participant. **b** Corresponding cross-recurrence diagram

$$c_{ij} = \begin{cases} 1, & d(f_i, g_j) \le \rho \\ 0, & otherwise \end{cases} \tag{10.7}$$

where d is distance metric and ρ is a given radius, as in the definition of recurrence. Cross-recurrence can be represented in a cross-recurrence diagram, which plots cross-recurrences of the fixation sequences over all possible time lags. If fixations f_i and g_j are recurrent (i.e. if $c_{ij} = 1$), then a dot is plotted at position i,j.

This is illustrated in Fig. 10.14, which shows an example from a study in which participants looked at many images, each for about 10 s, and later saw the same images again among many new ones. The black lines in Fig. 10.14a indicate the first scanpath, the red lines the second scanpath, and the black circles indicate cross-recurrences, i.e., fixations of one scanpath that were close to fixations of the other scanpath.

The resulting cross-recurrence diagram is shown in Fig. 10.14b, in which the fixations of the first scanpath are shown along the x-axis and the fixations of the second scanpath along the y-axis. In contrast to the recurrence diagram shown in Fig. 10.8, the cross-recurrence diagram is not symmetric. Second, there is no line of incidence in the cross-recurrence diagram. For these reasons, the RQA measures have to be generalized for the application in cross-recurrence analysis.

10.5.7.1 Cross-Recurrence

The **cross-recurrence** measure of two fixation sequences represents the percentage of cross-recurrent fixations, i.e. the percentage of fixations that match between the two fixation sequences. The more similar two fixation sequences, the higher the number of cross-recurrent points on the plot. It is invariant to differences in the order of the fixations as fixations are considered recurrent only if they overlap in position.

Given that cross-recurrence quantifies similarity in position only, it is most similar to the linear distance measure and the MultiMatch position measure.

10.5.7.2 Determinism

The **determinism** measure encodes the percentage of cross-recurrent points that form diagonal lines in the cross-recurrence plot and represents the percentage of fixation trajectories common to both fixation sequences. That is, determinism quantifies the overlap of a specific sequence of fixations, preserving their sequential information. An advantage of this measure is that it provides unique information about the type of similarity between two scanpaths. Although two scanpaths may be quite dissimilar in their overall shape or fixation positions, this measure can show whether certain smaller sequences of those scanpaths are shared.

10.5.7.3 Laminarity

The **laminarity** measure is a measure of repeated fixations on a particular region that are common to both scanpaths. Laminarity is closely related to determinism. If both laminarity and determinism are high, then in both scanpaths fixations tend to cluster on one or a few particular locations and remain there across several fixations. If laminarity is high, but determinism is low, then it quantifies the number of locations that were fixated in detail in one of the fixation sequences, but only fixated briefly in the other fixation sequence. It is a measure of the clustering of fixations across two sequences.

10.5.7.4 Center of Recurrence Mass

The **center of recurrence mass** (corm) is defined as the distance of the center of gravity of recurrences from the main diagonal in a recurrence plot. The corm measure indicates the dominant lag of cross-recurrences. Small corm values indicate that the same fixations in both fixation sequences tend to occur close in time, whereas large corm values indicate that cross-recurrences tend to occur with either a large positive or negative lag. This is a measure of whether one scanpath leads (with positive lag) or follows (with negative lag) its paired scanpath. Their overall similarity in shape or position may be different, but offset, such that one sequence proceeds in a particular trajectory, and the other follows the same trajectory only later on in time (e.g., a few fixations later). If there is no specific prediction about whether one scanpath leads or follows the other, the absolute value of the corm value can be used rather than averaging over positive and negative values.

In summary, cross-recurrence has been shown to be a natural extension of recurrence. While recurrence is used to characterize individual scanpaths, cross-recurrence is used to characterize the similarity of two different scanpaths. One major advantage of recurrence and cross-recurrence analysis lies in the fact that the same measures, recurrence, determinism, laminarity and corm can be applied to both situations. As a note of caution, however, it must be emphasized that the measures have to be interpreted differently in the two cases.

10.5.8 Summary

In this section, we have reviewed several common methods for assessing the similarity between scanpaths, and by extension, differences between scanpaths. These various methods all have their strengths and weaknesses, but all provide unique information regarding the similarity between two scanpath sequences. Comparing scanpaths is useful, as like RQA (described in Sect. 10.5), they preserve and quantify the temporal characteristics of eye movement behavior. These methods can be used for comparing the similarities and differences in eye movements between or among different observers.

10.6 General Summary

In the preceding sections, we first reviewed traditional measures for characterizing eye movements, starting with basic fixation and saccade measures. We then considered spatial eye movement analyses with a focus on heat maps and area of interest analyses. Finally, we examined popular temporal analyses of eye movements. Taken together, the reviews showed that the spatial analyses were not capturing the dynamic characteristics of eye movements, whereas the temporal analyses were applicable only in restricted circumstances.

We introduced recurrence quantification analysis (RQA) as a method for measuring the dynamic characteristics of eye movements. Although the analysis may appear to some as rather complicated, RQA can readily and accurately be conceptualized as an analysis of the temporal pattern of refixations. Critically, the RQA measures we introduced (recurrence, determinism, laminarity, and corm) capture important and interpretable aspects of this pattern. We showed for instance that RQA is suitable for discriminating eye movement patterns independent of the spatial structure of stimuli by focusing on temporal aspects exclusively. For this reason, it is ideally suited for the analysis of eye movements in more complex and dynamic situations.

Finally, we discussed current approaches to the comparison of scanpaths (edit distance, sample-based measures, linear distance, Scasim, ScanMatch, and MultiMatch), and showed that a simple generalization of RQA, that is, cross-RQA is well suited for the spatial comparison of scanpaths (i.e., how closely the individual fixations of two scanpaths overlap) while also capturing aspects of their temporal nature (e.g., their sequential information). Collectively then, RQA and cross-RQA, represents two powerful analysis techniques for future studies of eye movement behaviour, both within controlled laboratory settings and less controlled, more complex, natural environments. More generally. as eye movement analysis moves into the real world, techniques that capture the temporal nature of eye movements will no doubt prove useful in quantifying this more complex behavior.

10.7 Suggested Readings

Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. Behavior research methods, 45(3), 842–856.

- *This paper introduces recurrence quantification analysis for the analysis of eye movements. Much of the material in section 6 is based on this paper.*

Anderson, N. C., Anderson, F., Kingston, A., & Bischof, W. F. (2014). A comparison of scanpath comparison methods. Behavior Research Methods, https://doi.org/10.3758/s13428-014-0550-3.

- *This paper reviews and compares most of the recent scanpath comparison methods. Much of the material in section 7 is based on this paper.*

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011b). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. International Journal of Bifurcation and Chaos, 21, 1153–1161. https://doi.org/10.1142/s0218127411028970.

- *Recurrence analysis can be generalized to categorical data. This paper introduces categorical cross-recurrence analysis for analyzing the coordination of gaze patterns between individuals (see section 6).*

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). Eye tracking: A comprehensive guide to methods and measures. Oxford, U.K.: Oxford University Press.

- *This handbook is a comprehensive guide to the methods and measures for eye tracking. We recommend that researchers in the area of eye movement analysis consult this handbook. A new edition is scheduled for 2016 or 2017.*

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. Physics Letters A, 302, 299–307.

- *This paper introduces recurrence analysis as a tool for describing complex dynamic systems. The works of Dale et al. (2011b) and of Anderson et al.* (2013) *are applications of, and extensions to this paper.*

10.8 Questions to Students

a. What are the fundamental differences between grid-based and heat map methods for the spatial analysis of eye movements?
b. What are the characteristics of stimuli for which area-of-interest analyses are useful, and when are they less useful or not at all useful?
c. What are the fundamental advantages of recurrence analysis over the other spatial and temporal methods presented in this chapter?
d. What aspect of the recurrence patterns are captured by the measures recurrence, determinism, laminarity and corm?

Acknowledgements The authors acknowledge the support of the Natural Sciences and Engineering Research Council of Canada.

Bibliography

Anderson, N. C., Anderson, F., Kingston, A., & Bischof, W. F. (2014). A comparison of scanpath comparison methods. *Behavior Research Methods*, https://doi.org/10.3758/s13428-014-0550-3.

Anderson, N. C., Bischof, W. F., Laidlaw, K. E., Risko, E. F., & Kingstone, A. (2013). Recurrence quantification analysis of eye movements. *Behavior Research Methods, 45*(3), 842–856.

Anderson, N. C., Donk, M., & Meeter, M. (2016). The influence of a scene preview on eye movement behavior in natural scenes. *Psychonomic Bulletin & Review*, https://doi.org/10.3758/s13423-016-1035-4.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008a). Gaze selection in complex social scenes. *Visual Cognition, 16,* 341–355. https://doi.org/10.1080/13506280701434532.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2008b). Social attention and real-world scenes: The roles of action, competition and social content. *Quarterly Journal of Experimental Psychology, 61,* 986–998. https://doi.org/10.1080/17470210701410375.

Birmingham, E., Bischof, W. F., & Kingstone, A. (2009). Saliency does not account for fixations to eyes within social scenes. *Vision Research, 49,* 2992–3000.

Boccignone, G., & Ferraro, M. (2014). Ecological sampling of gaze shifts. *IEEE Transactions on Cybernetics, 44*(2), 266–279. https://doi.org/10.1109/TCYB.2013.2253460.

Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2008). *Time series analysis* (4th ed.). New York, NY: Wiley.

Brandt, S. A., & Stark, L. W. (1997). Spontaneous eye movements during visual imagery reflect the content of the visual scene. *Journal of Cognitive Neuroscience, 9,* 27–38. https://doi.org/10.1162/jocn.1997.9.1.27.

Broadbent, D. E. (1971). *Decision and stress*. London: Academic Press.

Bruce, N. D. B., & Tsotsos, J. K. (2009). Saliency, attention, and visual search: An information theoretic approach. Journal of Vision, 9(3), 5. https://doi.org/10.1167/9.3.5.

Bunke, H. (1992). Recent advances in string matching. In H. Bunke (Ed.), *Advances in structural and syntactic pattern recognition* (pp. 107–116). Singapore: World Scientific.

Caldara, R., & Miellet, S. (2011). iMap: A Novel Method for Statistical Fixation Mapping of Eye Movement data. *Behavior Research Methods, 43*(3), 864–878.

Cheng, J. T., Tracy, J. L., Foulsham, T., Kingstone, A., & Henrich, J. (2013). Two ways to the top: Evidence that dominance and prestige are distinct yet viable avenues to social rank and influence. *Journal of Personality and Social Psychology, 104*(1), 103.

Cherubini, M., Nüssli, M.-A., & Dillenbourg, P. (2010). This is it! Indicating and looking in collaborative work at distance. *Journal of Eye Movement Research, 3*(3), 1–20.

Chica, A. B., Klein, R. M., Rafal, R. D., & Hopfinger, J. B. (2010). Endogenous saccade preparation does not produce inhibition of return: Failure to replicate Rafal, Calabresi, Brennan, & Sciolto (1989). *Journal of Experimental Psychology: Human Perception and Performance, 36*(5), 1193.

Chuk, T., Chan, A. B., & Hsiao, J. H. (2014). Understanding eye movements in face recognition using hidden Markov models. *Journal of Vision, 14*(11):8, 1–14.

Coutrot, A., & Guyader, N. (2014). How saliency, faces, and sound influence gaze in dynamic social scenes. *Journal of Vision, 14*(8), 5.

Cristino, F., Mathôt, S., Theeuwes, J., & Gilchrist, I. D. (2010). ScanMatch: A novel method for comparing fixation sequences. *Behavior Research Methods, 42,* 692–700. https://doi.org/10.3758/BRM.42.3.692.

Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011a). The dynamics of reference and shared visual attention. *Frontiers in Psychology, 2*(355), 1–11. https://doi.org/10.3389/fpsyg.2011.00355.

Dale, R., Warlaumont, A. S., & Richardson, D. C. (2011b). Nominal cross recurrence as a generalized lag sequential analysis for behavioral streams. *International Journal of Bifurcation and Chaos, 21,* 1153–1161. https://doi.org/10.1142/S0218127411028970.

Dewhurst, R., Nyström, M., Jarodzka, H., Foulsham, T., Johansson, R., & Holmqvist, K. (2012). It depends on how you look at it: Scanpath comparison in multiple dimensions with multimatch, a vector-based approach. *Behavior Research Methods, 44*(4), 1079–1100.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.

Findlay, J. M., & Gilchrist, I. D. (2003). *Active vision: The psychology of looking and seeing*. Oxford: Oxford University Press.

Foulsham, T., Cheng, J. T., Tracy, J. L., Henrich, J., & Kingstone, A. (2010). Gaze allocation in a dynamic situation: Effects of social status and speaking. *Cognition, 117*(3), 319–331.

Foulsham, T., & Kingstone, A. (2013). Fixation-dependent memory for natural scenes: An experimental test of scanpath theory. *Journal of Experimental Psychology: General*. https://doi.org/10.1037/a0028227.

Foulsham, T., & Sanderson, L. A. (2013). Look who's talking? Sound changes gaze behaviour in a dynamic social scene. *Visual Cognition, 21*(7), 922–944.

Foulsham, T., & Underwood, G. (2008). What can saliency models predict about eye movements? Spatial and sequential aspects of fixations during encoding and recognition. *Journal of Vision, 8*(2)(6), 1–17. https://doi.org/10.1167/8.2.6.

Foulsham, T., Dewhurst, R., Nystrom, M., Jarodzka, H., Johansson, R., Underwood, G., & Holmqvist, K. (2012). Comparing scanpaths during scene encoding and recognition: A multidimensional approach. *Journal of Eye Movement Research, 5*(4), 1–14.

Friston, K. J., Holmes, A. P., Poline, J.-B., Grasby, P. J., Williams, S. C. R., Frackowiak, R. S. J., et al. (1995). Analysis of fMRI time-series revisited. *Neuroimage, 2,* 45–53.

Friston, K. J., Jezzard, P., & Turner, R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping, 1,* 153–171.

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.

Harding, G., & Bloj, M. (2010). Real and predicted influence of image manipulations on eye movements during scene recognition. *Journal of Vision, 10*(2)(8), 1–17. https://doi.org/10.1167/10.2.8.

Hasson, U., Yang, E., Vallines, I., Heeger, D. and Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *Journal of Neuroscience, 28*, 2539–2550.

Hayes, T. R., Petrov, A. A., & Sederberg, P. B. (2011). A novel method for analyzing sequential eye movements reveals strategic influence on Raven's Advanced Progressive Matrices. *Journal of Vision, 11,* 1–11.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. L. (2007). Visual saliency does not account for eye movements during visual search in real-world scenes. In R. van Gompel,

M. Fischer, W. Murray, & R. W. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Amsterdam, The Netherlands: Elsevier.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, U.K.: Oxford University Press.

Hunt, A. R., Chapman, C. S., & Kingstone, A. (2008). Taking a long look at action and time perception. *Journal of Experimental Psychology: Human Perception and Performance, 34*(1), 125.

Igarashi, T., Moscovich, T., & Hughes, J. F. (2005). Spatial keyframing for performance-driven animatoin. In *Proceedings of the 2005 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, (pp. 107–115). New York: ACM.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40,* 1489–1506.

Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). A *vector-based, multidimensional scanpath similarity measure*. Paper presented at the Proceedings of the 2010 Symposium on Eye-Tracking Research and Applications.

Josephson, S., & Holmes, M. E. (2002). Visual attention to repeated images on the World-Wide Web: Another look at scanpath theory. *Behavior Research Methods, Instruments, & Computers, 34*(4), 539–548.

Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009). *Learning to predict where humans look*. Paper presented at the IEEE International Conference on Computer Vision (ICCV). Kyoto, Japan.

Kingstone, A. (1992). Combining expectancies. *The Quarterly Journal of Experimental Psychology, 44A,* 69–104.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive ethology: A new approach for studying human cognition. *British Journal of Psychology, 99*(3), 317–340.

Krassner, M., Patera, W., Bulling, A. (2014) Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. *Computer Vision and Pattern Recognition*. arXiv: 1405.0006.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics, 22*(1), 79–86. https://doi.org/10.1214/aoms/1177729694.

Le Meur, O., & Baccino, T. (2013). Methods for comparing scanpaths and saliency maps: Strengths and weaknesses. *Behavior Research Methods, 45,* 251–266.

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions, and reversals. *Soviet Physics Doklady, 10*, 707–710.

Levy, J., Foulsham, T., & Kingstone, A. (2012). Monsters are people too. *Biology letters, Feb 23;9*(1):20120850. https://doi.org/10.1098/rsbl.2012.0850.

Mannan, S., Ruddock, K. H., & Wooding, D. S. (1995). Automatic control of saccadic eye movements made in visual inspection of briefly presented 2-D images. *Spatial Vision, 9,* 363–386. https://doi.org/10.1163/156856895X00052.

Marwan, N., & Kurths, J. (2002). Nonlinear analysis of bivariate data with cross recurrence plots. *Physics Letters A, 302,* 299–307.

Mathôt, S., Cristino, F., Gilchrist, I. D., & Theeuwes, J. (2012). A simple way to estimate similarity between pairs of eye movement sequences. *Journal of Eye Movement Research, 5*(1).

Needleman, S. B., & Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *Journal of Molecular Biology, 48*(3), 443–453.

Noton, D., & Stark, L. (1971). Scanpaths in eye movements during pattern perception. *Science, 171,* 308–311. https://doi.org/10.1126/science.171.3968.308.

Nyström, M., & Holmqvist, K. (2008). Semantic override of low-level features in image viewing - both initially and overall. *Journal of Eye Movement Research, 2*, 1–11.

Olk, B., & Kingstone, A. (2003). Why are antisaccades slower than prosaccades? A novel finding using a new paradigm. *NeuroReport, 14,* 151–155.

Pellecchia, G. L., & Shockley, K. (2005). Cross recurrence quantification of interpersonal postural activity. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods* *for the behavioral sciences* (pp. 95–141). Washington, DC: National Science Foundation, Directorate for Social, Behavioral & Economic Sciences. Retrieved August 25, 2011, from www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale, NJ: Erlbaum.

Richardson, D. C., & Dale, R. (2005). Looking to understand: The coupling between speakers' and listeners' eye movements and its relationship to discourse comprehension. *Cognitive Science, 29,* 1045–1060.

Richardson, D. C., Dale, R., & Tomlinson, J. M. (2009). Conversation, gaze coordination, and beliefs about visual context. *Cognitive Science, 33,* 1468–1482. https://doi.org/10.1111/j.1551-6709.2009.01057.x.

Riley, M. A., & Clark, S. (2003). Recurrence analysis of human postural sway during the sensory organization test. *Neuroscience Letters, 342,* 45–48.

Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics, 73*(2), 291–296.

Shepherd, S. V., Steckenfinger, S. A., Hasson, U., & Ghasanfar, A. A. (2010). Human–monkey gaze correlations reveal convergent and divergent patterns of movie viewing. *Current Biology, 20,* 649–656.

Shockley, K., Baker, A. A., Richardson, M. J., & Fowler, C. A. (2007). Articulatory constraints on interpersonal postural coordination. *Journal of Experimental Psychology: Human Perception and Performance, 33*(1), 201.

Shockley, K., Richardson, D. C., & Dale, R. (2009). Conversation and coordinative structures. *Topics in Cognitive Science*, *1*, 305–319.

Shockley, K., Santana, M.-V., & Fowler, C. A. (2003). Mutual interpersonal postural constraints are involved in cooperative conversation. *Journal of Experimental Psychology: Human Perception and Performance, 29*(2), 326–332. https://doi.org/10.1037/0096-1523.29.2.326.

Shockley, K., & Turvey, M. T. (2005). Encoding and retrieval during bimanual rhythmic coordination. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 31*(5), 980.

Soto-Faraco, S., Morein-Zamir, S., & Kingstone, A. (2005). On audiovisual spatial synergy: The fragility of the phenomenon. *Perception and Psychophysics, 67*(3), 444–457.

Stark, L. W. & Ellis, S. R. (1981) Scanpaths revisited: cognitive models direct active looking. In D. F. Fisher et al., (Eds.) *Eye movements: cognition and visual perception*, (pp. 193–226). Hillsdale, NJ: Lawrence Erlbaum Associates.

Tatler, B. W. (2007). The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions. *Journal of vision*, *7*(14), 4–4.

Tatler, B. W., Baddeley, R. J., & Gilchrist, I. D. (2005) Visual correlates of fixation selection: effects of scale and time. *Vision Research, 45*(5), 643–659

Torralba, A., Oliva, A., Castelhano, M. S., & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-worldscenes: The role of global features in object search. *Psychological Review, 113*(4), 766–786.

Underwood, G., Foulsham, T., & Humphrey, K. (2009). Saliency and scan patterns in the inspection of real-world scenes: Eye movements during encoding and recognition. *Visual Cognition, 17,* 812–834.

Vig, E., Dorr, M., & Cox, D. (2014). Large-scale optimization of hierarchical features for saliency prediction in natural images. In The IEEE Conference on Pattern Recognition, Columbus Ohio, June 24–27.

von der Malsburg, T., & Vasishth, S. (2011). What is the scanpath signature of syntactic reanalysis? *Journal of Memory and Language, 65*(2), 109–127.

Wagner, R. A., & Fischer, M. J. (1974). The string-to-string correction problem. *Journal of the Association for Computing Machinery, 21,* 168–178.

Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural Networks, 19*, 1395–1407.

Webber, C. L., Jr., & Zbilut, P. J. (2005). Recurrence quantification analysis of nonlinear dynamical systems. In M. A. Riley & G. C. Van Orden (Eds.), *Tutorials in contemporary nonlinear methods for the behavioral sciences* (pp. 26–94). Washington, DC: National Science Foundation, Directorate for Social, Behavioral & Economic Sciences. Retrieved August 25, 2011, from www.nsf.gov/sbe/bcs/pac/nmbs/nmbs.jsp.

West, J. M., Haake, A. R., Rozanski, E. P., & Karn, K. S. (2006). *eyePatterns: software for identifying patterns and similarities across fixation sequences.* Paper presented at the Proceedings of the 2006 Symposium on Eye Tracking Research & Applications, (pp. 149–154). San Diego, California, March 27–29, 2006.

Woodford, O. (2014). SC - powerful image rendering. https://www.mathworks.com/matlabcentral/fileexchange/16233-sc-powerful-image-rendering. MATLAB Central File Exchange.

Yarbus, A. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Chapter 11
Pupillometry

Bruno Laeng and Dag Alnaes

Contents

Abstract The eye pupil is an organ optimally designed by evolution to control the amount of light entering the eye so as to obtain sharp visual images and allow monitoring of as much as possible of the visual field. The opposing pupillary movements of

---

B. Laeng (✉) · D. Alnaes
Department of Psychology, University of Oslo, Oslo, Norway
e-mail: bruno.laeng@psykologi.uio.no

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_11

constrictions and dilations optimize some of these goals while at the same time may reduce the effectiveness of others. The pupils are controlled by subcortical anatomical structures of the brain that allow rapid adjustments of the pupil (in particular to strong light). However, higher-level cortical mechanisms in the brain can affect in a top-down manner pupil size, possibly to prepare the eye as for forthcoming visual events. Today it has become relatively easy and inexpensive to measure pupil size with current infrared eye-trackers. Yet, advancements in equipment and data analysis are still needed in order to apply pupillometry to situations outside the laboratory where luminance conditions cannot be controlled. Psychological studies have clearly implicated pupil responses as reliable indexes of motivational and emotional states. The most promising application of pupillometry stands in its ability to index robustly and reliably the level of cognitive workload or "mental effort" engaged during a task. Recent studies in neuroscience have clarified that the pupil reflects the activity of the *Locus coeruleus*, which is the brain's hub of the norepinephrine system and whose activity has the effect of "energizing" the whole brain at a particular moment in time. Hence pupillary responses provide a window on attentional processes that are relevant for learning and memory.

11.1 Introduction and Learning Objectives

> We know today that every vivid mental event, every physical effort, every impulse of will, each activation of attention, and especially each affect causes pupillary dilation just like every sensory stimulus that flows to the brain from the periphery.
>
> **Oswald Bumke** (Die Pupillenstörungen bei Geistes-und Nervenkrankheiten, 1904)

11.1.1 The Eye Pupil: An Introduction

If we look closely to an animal's eye while it moves about and looks at different objects, we will notice that the size of the pupil (i.e., the black circle or hole at the center of the eye inside the colored iris) is changing continuously (Denton, 1956). These changes in size are mainly adjustments of the eye opening to the continuous variations in light conditions or luminance. The term 'pupil' apparently derives from Latin *pupilla*, meaning 'small doll', and referring to the experience of a mirror-like reflection of the observer's face when closely looking at someone else's eye (Janisse, 1977).

The eye pupil allows light to enter inside the eye and travel through the watery fluid inside the eye to stimulate the retina (a neural layer of light-sensitive photoreceptors) in the internal surface of the eye. The eye is a biological optical device (Campbell & Gregory, 1960) that can adjust to focus light optimally for visual acuity (like the diaphragm of a camera also does). When the external illumination is high, the pupil is made as small as possible to reduce the effects of optical aberrations due to the eye's spherical shape and watery fluid contained within. That is, when light goes through a semi-transparent medium (e.g., a glass prism) the light rays do not continue in a straight line but can become diffuse, causing blurriness in the projected image on the retina. In contrast, at low level illumination, visual acuity may be more limited by the total quanta of light energy (i.e., the number of photons) that can reach the retinal photoreceptors (Denton & Pirenne, 1952) and, consequently, it can be useful to increase the amount of input despite the distortions caused by optical aberrations and blurriness. Thus, the pupil adjusts by either increasing its depth of field and visual acuity by constriction (albeit that reduces the amount of light signals entering the eye) or by dilating the pupil to achieve the needed amount of light stimulation while at the same time decreasing vision's depth of field and sharpness of vision (McDougal & Gamlin, 2008). Although some animals can apparently voluntarily constrict each pupil independently of the other and also of the stimulating light's actual intensity (e.g., geckos; Denton, 1956), humans cannot voluntarily control the size of the pupils (Laeng & Sulutvedt, 2014). Pupil diameters normally change in a consensual, binocular, manner even when light is shown onto a single pupil (though in the latter case the degree of constriction is a bit lower than when both eyes are stimulated).

11.1.2 This Chapter's Learning Objectives

You will learn about:

1. The function of the eye opening or pupil and what parts of the brain control constrictions and dilations
2. How to design a pupillometric study by controlling potential nuisance factors when using visual stimuli.
3. How pupillometry has contributed so far to a variety of fields within psychology.
4. The relationship between mental effort or cognitive workload and changes in pupil diameter.
5. How pupillometry provides a window into the functioning of neural networks that are closely related to attention and cognitive resources.

11.2 Historical Annotations

According to the most authoritative review on pupillometry by Irene Loewenfeld (1999), scientific knowledge that pupil size may change despite no change in the light conditions is quite ancient. The oldest report about pupillary changes seems to have appeared some centuries ago, written by the Italian scientist Fontana (1765). However, there can be found even earlier literary references to the eyes as being the *occulus animi index*, i.e., "the eye as the index of the mind" and the English Paediatrician Thomas Phaer wrote in 1545 (in the *Regimen sanitatis Salernitanum*): "The eyes…are the wyndowes of the mynde, for both ioye & ange…are seene… through them." Understandably, the pupils have attracted the interest of poets who typically referred to them as the "windows of the soul"; e.g. the French poet Guillaume de Salluste Du Bartas (1544–1590), translated in English by fellow poet Joshua Sylvester (1563–1618) with the verses: "*These lovely lamps, these windowes of the soule*."

Always in Italy, during the Renaissance (i.e., from the 14th to the 17th century), it was common lore that romantic or sexual interest would be reflected in dilated pupils (Simms, 1967). It is well known that, at the time, women could use the extract of the plant *Belladonna* to (artificially) enlarge their pupils in order to look both "interested" and "interesting" (indeed, 'bella-donna' translates from the Italian as "pretty woman"; see Fig. 11.1).

Neurology was the first modern discipline to discover the usefulness of monitoring changes in pupil size as a way to identify specific pathological processes within the nervous system. Doctors looked for large-scale changes in pupil diameter, either at rest or when the eye was stimulated by light (i.e., by checking the presence of the pupil light reflex), which could be easily observed with a "naked eye" (i.e., qualitatively and without using specialized equipment; Beatty & Lucero-Wagner, 2000). In addition, with the widespread use of narcotics and other substance of abuse within modern societies, it became a custom of both doctors and police to examine the size of the pupil to establish if a person was under the influence of some specific drug. For example, alcohol or opioids (e.g., morphine) constrict the eye pupil (even to a "pinpoint"; Murray et al., 1983), whereas other drugs, such as LSD, MDMA, mescaline, psilocybin mushrooms, cocaine and amphetamine can cause pupil dilation. Nowadays, in several countries, a quick evaluation of pupil size is routinely used by police officers to detect drugs' use (Richman, McAndrew, et al., 2004).

The practice of measuring quantitatively the pupil developed within the neurosciences based on the initial "pupillography" work of Lowenstein, which started in the 1920 and was continued by his student Loewenfeld; this work mainly focused on clinical aspects of the pupillary responses, charting in detail the pupillary light reflex as well as the dilation reflex. The clinical research of the physiologists Lowenstein and Loewenfeld (1958) was built upon the earlier research of Johann Friedrich

Fig. 11.1 The eyes of the Mona Lisa, in the famous "portrait" by Leonardo da Vinci, look rather dilated (about 50% of the area of her irises), which may be attributable to either being partially in the shade or as being "attracted" to the painter

Horner (1869) and Douglas Argyll Robertson (1869), who both gave names to neurological syndromes that are expressed in dramatic changes in pupil behavior. Indeed, the study of the pupillary responses has a long tradition within medicine, in particular neurology, ophthalmology, and physiology (e.g., Wilhelm, Wilhelm, & Lüdtke, 1999; Wilhelm, Wilhelm, Moro, & Barbur, 2002).

The discovery that small-scale pupil changes (not so easily visible with the naked eye at a typical distance of social interaction) occurred also during mental work was first reported within the field of neurology with a study by Schiff (1875) where he measured dilations of the pupil in individuals engaged in mental arithmetic. According to Loewenfeld (1999), in the 1870s, Schiff had labeled the pupil as the body's "finest esthesiometer" (i.e., as a gauge of the presence of challenges to the mind and body, like pain, or mental and physical exertion). This effect on the pupil of cognitive processing was therefore "re-discovered" in the early 1960 by the psychologist Eckard Hess at University of Chicago (Hess & Polt, 1964). Indeed, the pupillometry method in psychology firmly established itself within 'cognitive' psychology after the appearance of three seminal studies (Hess & Polt, 1960, 1964; Kahneman & Beatty, 1966), all of them published in the prestigious journal *Science*.

Beside his strong contributions with pupillometry to cognitive psychology, Eckard Hess became a strong advocate for the application of pupillometry in several domains of research and business, especially marketing and advertisement. Hess (1965) was also the first to use pupillometry as a way to probe political feelings or ideology, by showing photographs of American presidential candidates during election times, as well changes in attitude (after exposure to political propaganda). As pointed out by Irene Loewenfeld (1999), the main authority on pupil studies in the medical field: "Most people perusing their Sunday papers during the 1960 and 1970s sooner or later learned about E. H. Hess's fascinating discovery" […] "A veritable avalanche of publicity followed, as some of the largest advertising companies in the country vied with *Marplan* (a research subsidiary of the advertising conglomerate *Interpublic*, who had Hess under contract) to use this marvelous new method that could detect the customer's true likes and dislikes without his conscious cooperation, and indeed even against his will." Hess summarized his psychological and marketing research in the book *The tell-tale eye: How your eyes reveal hidden thoughts and emotions* (1975).

Some of the earliest researchers, like Heinrich Harms (1937), had already observed the usefulness of pupillometry in uncovering what an individual consciously perceives at a particular moment. He employed "binocular rivalry" situations, where stimuli differing in brightness are separately but simultaneously presented to each eye. Harms found that the pupil size closely reflect which of the stimuli dominates awareness, since it will constrict more when the observer reports being aware of the brighter stimulus relatively to when the darker stimulus dominates. Recent research has confirmed by use of bistable figures or stimuli simultaneously competing for attention the ability of the pupil to signal the locus of focusing of the mind on either a specific object or particular perspective of it (Laeng, Sirois, & Gredebäck, 2012).

There have been since the inception of pupillometry within medicine and psychology several technical changes or improvements in the recording of pupil data.

Many of the studies conducted before the 1990s were necessarily based on analog photographic and cinematographic recordings. In some of the older studies, the method involved infrared film running at 2 Hz and measuring the diameter of the pupil was tediously made by hand, after projecting a greatly enlarged image of the iris onto a large surface and measuring diameter with an ordinary ruler (Beatty & Lucero-Wagner, 2000). For example, Denton (1956) describes the measuring of an animal's pupil by taking infrared still photos of the pupil and projecting it on graph paper and counting the squares. All early experimenters (e.g., Bender, 1933; Gradle & Ackerman, 1932) analyzed pupil size by measuring it from videocamera's film projections and in order to achieve enough contrast and improve the identification of the pupil's borders they used exclusively participants with light-colored irises. Given such a slow analysis method, it is not surprising that so many studies before the 1990s tended to have few trials and small sample sizes (for example, Hess & Polt's, 1964, seminal study was based on 5 participants' responses to just 4 trials). These may also be some of the reasons why there has been something like a 30-year gap before pupillometry has come to be extensively used again in psychological and neuroscience research (Laeng, Sirois, & Gredebäck, 2012). The modern infrared technology, originally used in clinical application of pupillometry, has later greatly benefitted from computerized systems for the control of the camera as well as the development of software specialized for the automatic analysis of eye-data. Nowadays, it has become straightforward, less time expensive, and very precise to measure pupil size by use of such computerized infrared-based "eye-trackers".

11.2.1 The Anatomy of the Pupil

Based on the signals from three types of photoreceptors in the retina of each eye, i.e. the three cones (each sensitive to different ranges of colors) and the rods and melanopsin receptors which are sensitive to the intensity of light stimulation (Gamlin et al., 2007; Kawasaki & Kardon, 2007), the brain in turn controls changes in eye pupil size by sending simultaneously excitatory and inhibitory commands to two opposing smooth muscles of the iris (see Fig. 11.2). These are the radial muscles of the *dilator pupillae*, innervated by the sympathetic branch of the autonomous nervous system (ANS), and the circular muscles of the *sphincter (constrictor) pupillae*, innervated by the parasympathetic branch of the ANS. For example, a pupillary *constriction* or *miosis* can result from a stimulation of the constrictor or an inhibition of the dilator, whereas a pupillary *dilation* or *mydriasis* results from the inverse pattern of excitatory and inhibitory stimulation of the constrictor and the dilator.

In "standard" or comfortable light conditions, the average size of the pupil is about 3 mm and typically ranges from 2 to 8 mm on average (Watson & Yellott, 2012). In dim light or darkness, the pupil enlarges to allow more light energy to enter the eye up to an average size of about 7 mm with a standard deviation (from this average) of about 0.9 mm (MacLachlan & Howland, 2002). In complete darkness, after an adaptation period of some minutes, the pupils of an alert young (age <35) subject do not measure

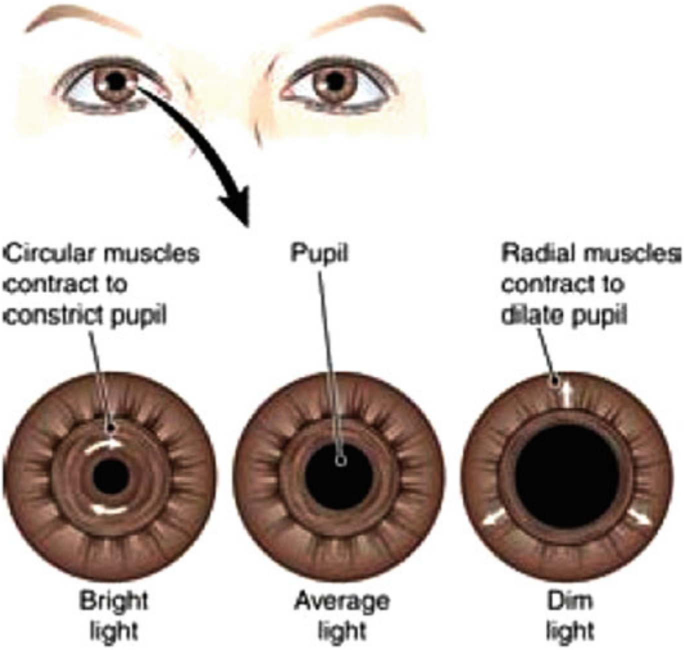


Fig. 11.2 Different muscles controlling the pupil size become active in relation to the amount of light entering the eye

less than 7 mm (with some spontaneous fluctuations that do not exceed 0.5 mm). However, when pupils are measured in darkness in the same but tired subjects, the pupils will be smaller and begin to oscillate (Lavie, 1979; Lowenstein, Feinberg, & Loewenfeld, 1963; Warga, Lüdtke, Wilhelm, & Wilhelm, 2009). Pupillary changes that are non-luminance mediated but are driven by psychological factors (like mental work or emotional states, as it will be discussed later in depth) are typically much more modest than the ones described above in reaction to light and are rarely greater than 0.5 mm (Beatty & Lucero-Wagoner, 2000). This happens regardless of whether pupil changes are evoked by visual or non-visual stimuli (Loewenfeld, 1958), for example when listening to musical stimuli (e.g., Damsma & van Rijn, 2017).

11.2.2 *The Physiology of Pupillary Changes*

Sudden increases in illumination (e.g., a camera's flash), trigger the so-called "pupillary light reflex" (PLR; see Fig. 11.3), a rapid constriction of the eye pupil (Loewenfeld, 1999). The biological and physiological functions of the PLR seem very clearly "designed" by evolution, in a way reminiscent to the human design of the lens' opening of a camera that adjusts to light levels so to obtain optimal photo images.

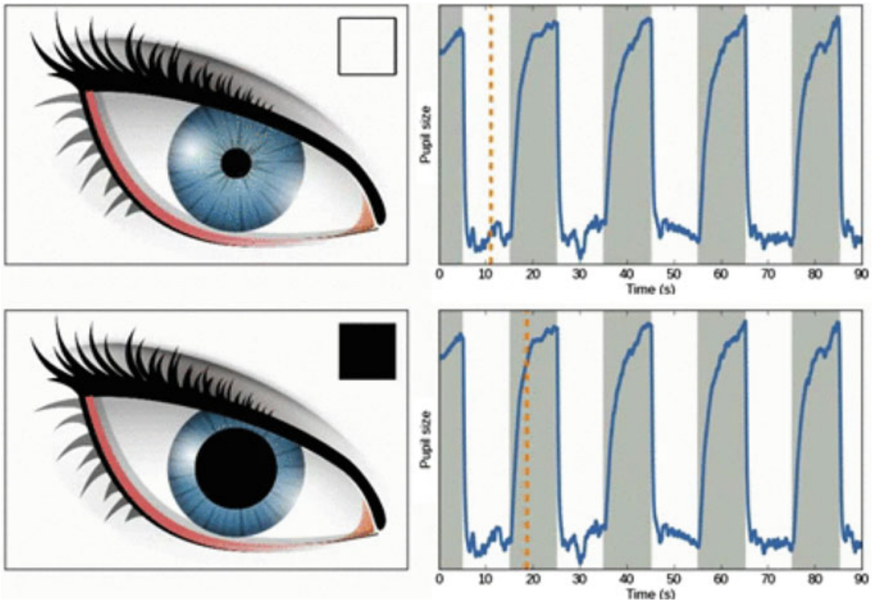


Fig. 11.3 Changes in pupil size when one switches luminance in a room to be very bright (top) versus in complete darkness (bottom). The dotted line indicates a particular moment during the time period in which the room is bright or dark and the blue curve indicates the corresponding size of the pupil at each moment

When the external illumination is high, pupils are made small to increase focus and acuity by reducing the effects of the eye's optical aberrations, like diffraction and stray light; as it happens with the camera images, pupillary constrictions improve contrast and depth of field of vision. In contrast, at low level illumination, visual acuity is limited by the total 'quanta' (i.e., photons) of light energy that can reach the photoreceptors and the amount of input needs to be increased by enlarging the pupils, despite the resulting optical distortions and loss of contrast acuity. The total number of quanta per second that fall on the retina is directly proportional to the area of the pupil (Cornsweet, 1970). Hence the pupil diameter should affect the visibility of a visual scene by enhancing the perceived light intensity (for example, if the pupil diameter increases 50% from time$_1$ to time$_2$ then twice the amount of the quanta radiating from the image will be able to pass and the light in the image would be twice more intense).

Another advantage of dilating the pupil is that when the pupil is wider, visual acuity for dynamic or moving stimuli in the periphery of the visual field are better than when the pupil is constricted. In other words, when the pupil is large a visual stimulus can be effectively detected and identified within a wider area on the retina (Ueda, Nawa, Yukawa, Taketani, & Hara, 2006). Typically, the pupil enlarges rapidly when an animal is startled (for example by a sudden and loud sound) and large pupils are one of the features of the expression of fear. One possibility is that pupil dilations to startling stimuli play an adaptive role by increasing the visual ability to monitor a wider visual area and spot more effectively potential threats entering the field of vision (cf. Lee, Susskind, & Anderson, 2013; Susskind et al., 2008).

In contrast, the pupil can optimize sight by automatically constricting when distance from the eye to the observed object changes from far to near (e.g., less than 6 m); this phenomenon is called "pupillary near response" (PNR) and it increases the depth field of view while reducing defocus across the depth planes. Nota Bene: the same phenomenon happens in a photo camera: when decreasing the size of the diaphragm one can obtain a photo image where objects look sharp at different levels of depth. However, the camera will need a longer exposure (i.e. more light stimulation) onto the photocells of a digital camera (or—for the older cameras—to allow the film's chemicals to be affected by light). Constrictions of the pupils are also commonly thought to function as a protective mechanism to avoid that excessively saturating light energy could overstimulate and ultimately damage the retinal photoreceptors. According to Denton (1956), the protective function of the constriction of the iris seems however poorly fulfilled in humans, since the minimal area of the pupil is only one-sixteenth of the maximal size.

The pupil constriction latency increases with age (e.g., Feinberg & Podolak, 1965) as well as average pupil size (e.g., Winn, Whitaker, Elliot, & Phillips, 1994). Since a reduced pupil size in adults has the effect of increasing the depth of field of vision, this partially compensates for the progressive reduction of the eye lenses' accommodative range with age. Differently from adults, children have elastic lenses that can rapidly accommodate. There exist also weak sex differences in the PLR and, in a dark-adapted condition, females have significantly larger relative constriction amplitudes than males (Fan et al., 2009).

Note also that pupil dilations cannot be provoked voluntarily but can occur in an indirect manner by mentally imaging an object or event that would normally evoke a spontaneous pupillary dilation (Laeng & Sulutvedt, 2014); for example, engaging in imagery of a sexual nature (Whipple, Ogden, & Komisaruk, 1992) or simply imagining being in a dark room (Laeng & Sulutvedt, 2014). Even constrictions may be provoked indirectly (e.g., by imagining looking at something very bright, like the sun; cf. Binda, Pereverzeva, & Murray, 2013). Similarly, engaging in arithmetic computations (Hess & Polt, 1964) or simply recalling your own phone number (Beatty & Khaneman, 1966) dilates the pupil.

Figure 11.4 illustrates the neural network that controls pupillary dilations and constrictions. Pupillary dilations evoked by psychologically relevant stimuli and not by changes in light are typically the result of a neural inhibitory mechanism by the *Locus coeruleus* (LC) on the parasympathetic oculomotor complex (e.g. Steinhauer, Siegle, Condray, & Pless, 2004) or Edinger-Westphal nucleus. The LC is the brain's hub of the noradrenergic or norepinephrine (NE) neuromodulatory system (Samuels & Szabadi, 2008; Szabadi, 2012; Wilhelm et al., 1999). It is a small brainstem nucleus located bilaterally in the rostral pons and the only source of noradrenergic innervation in the neocortex (i.e., the outer layer of the brain). The current total count of processing units within the human LC is at a small number of 22,000–51,000 for each lateral nucleus (Mouton, Pakkenberg, Gundersen, & Price, 1994). Despite the

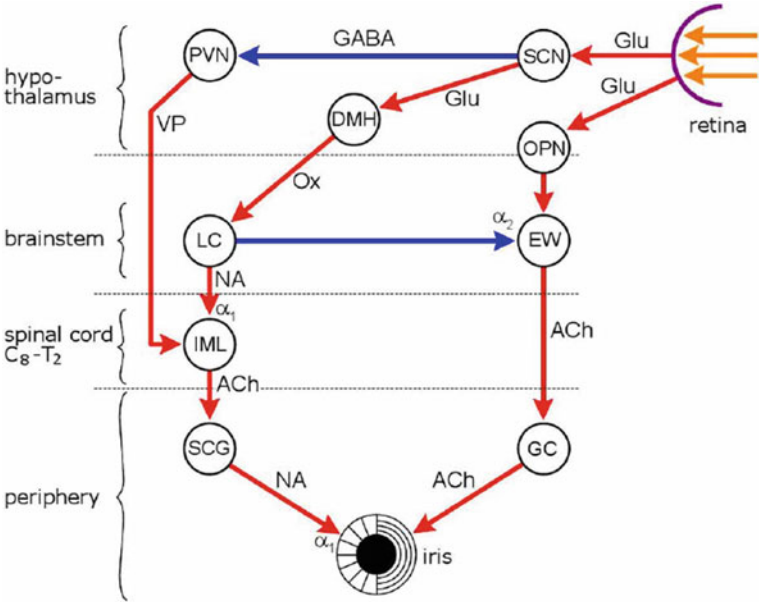


Fig. 11.4 The neural network controlling the diameter of the pupil and the neurotransmitters responsible for either excitatory, in red, or inhibitory connections, in blue (from Szabadi, 2012). The PLR is a parasympathetic reflex: light that stimulates cells of the retina will engage the neural chain OPN → EW → GC, leading to pupil constriction, while inhibiting sympathetic activity via an inhibitory output from the SCN to the PVN. Abbreviations: *SCN* suprachiasmatic nucleus; *PVN* paraventricular nucleus; *DMH* dorsomedial hypothalamus; autonomic premotor nuclei: *OPN* olivary pretectal nucleus; *LC Locus coeruleus*; parasympathetic nucleus/ganglion: *EW* Edinger Westphal nucleus; *GC* ganglion ciliare; *sympathetic nucleus/ganglion: IML* intermedio-lateral column of spinal cord; *SCG* superior cervical ganglion. Neurotransmitters: *Glu* glutamate; *GABA* γ-aminobutyric acid; *VP* vasopressin; *Ox* orexin; *ACh* acetylcholine; *NA* noradrenaline. Adrenoceptors: $\alpha 1$ excitatory and $\alpha 2$ inhibitory

scarcity of these cells, all of the neurotransmitter/modulator norepinephrine (NE) to the cortex, cerebellum and hippocampus (i.e., neural structures that are implicated with high-level cognition and which, all together, include several billions of neurons) originate from neurons within these small nuclei. The LC also sends excitatory projections to neurons in the brainstem and spinal cord that help to maintain muscle tone; in fact, during REM sleep, when the muscle tone is lowest, the activity of the LC is also lowest (Gottesmann, 2011).

The LC was earlier known among psychologists for its role in clinical syndromes like depression, panic disorder, and anxiety (e.g., Einhäuser, Koch, & Carter, 2010; Svensson, 1987). It is activated by stress and it can increase NE secretion through the hypothalamic-pituitary-adrenal axis. By altering the activity of prefrontal cortex, the LC can have large repercussions on cognitive function (Bouret & Sara, 2004) as well as on emotion and motivation via its connections to the nucleus accumbens and amygdala (both crucial to the so-called "reward" and "valence" systems of the brain;

Bouret, Duvel, Onat, & Sara, 2003). Most relevantly for cognitive psychology, the LC has particularly dense projection to cortical regions involved in attentional processes, e.g. parietal cortex, pulvinar nucleus of the thalamus and the superior colliculus (Foote & Morrison, 1987). Hence, the LC is a very important neural structure for understanding the effects on the pupil of attention and states of consciousness, as well as other psychological variables that typically engage attention towards external stimuli, like emotions, or internally, like memory and decision-making. As we will discuss later, several studies that have recorded activity directly from neurons of the LC in monkeys and those neuroimaging studies with humans that have monitored activity in the brainstem converge in indicating the LC as the major player in pupillary responses to psychologically relevant events.

11.3 Measuring the Eye Pupil

11.3.1 Pupillometry Method

Many research labs have developed within the last 50 years their own pupillometry apparatus with custom-made programming scripts (e.g., Bernhardt, Dabbs, & Riad, 1996; McConkie, Kerr, Reddix, & Zola, 1988). However, within the last 25 years, the interest for the measurement of both eye movements and the pupil has spread into various areas of research beside psychology per se, like marketing research (Arch, 1979; Blackwell, Miniard, & Engel, 2001; King, 1972; Krugman, 1965) and applications to industry and design (e.g., Carbon, Hutzler, & Minge, 2006; Laeng, Suegami, & Aminihajibashi, 2016; Schwalm, Keinath, & Zimmer, 2008). Various companies located in the USA, Germany and Sweden have made commercially available several eye-tracking systems that, despite being mainly designed to quantitatively measure eye movements (saccades) and eye fixations, can provide detailed information about the diameters of the pupil of both eyes. Using such commercially-available "eye-trackers" (i.e., computerized systems receiving input from a video camera sensitive to infrared light; Silverman, Mooney, & Shepherd, 1992), dynamic changes in the size of the pupil can be accurately and non-invasively measured. Infrared eye-tracker are therefore capable of tracking the movements of the pupils in space, like its direction (also in depth), and changes in its size. Infrared illumination allowed precise measurements of both brown and blue eyes and in virtual darkness (since the human eye is "blind" to the infrared spectrum).

11.3.2 Experimental Set-Up for Eye Tracking

Current remote eye-trackers provide to users the raw data about pupil diameters (both horizontal and vertical) as samples (coded by sample frequency) with values typically given in number of pixels of the eye camera. Some eye-trackers can also report pupil diameter in millimeters, after a simple point-calibration routine, by also keeping track of head movements and distance in relation to the calibrated points. Pupillometry with eye-tracking glasses (e.g., the SMI® ET-glasses) provide pupil data from high-resolution sensors, which are free of head movement and distance artifacts. Researchers may often use only the horizontal diameter since the vertical diameter is thought to be more affected by eyelid closure (e.g., during eye blinks). However, both the horizontal and vertical diameters are subject to the optical perspective based on the direction of gaze. If one observer looks, for example, to one of the corners of the computer screen, this may result in underestimations of both the horizontal and vertical pupil diameter due to foreshortening artifacts. Different types of commercially available eye-trackers may introduce measurements distortions depending on the eye position (Brisson et al., 2013; Gagl, Hawelka, & Huzler, 2011). Given that most psychological experiments are based on viewing single patterns centered on the screen and that the pupil responses may be (a) averaged over both eyes (when using binocular eye-tracking system) as well as (b) averaged over several fixations, which are typically distributed over different locations of the display, it seems unlikely that the above mentioned distortions would introduce more than a small amount of random noise in the data of most cognitive studies.

A typical analysis based on raw data output (see Fig. 11.5 for a graphic display

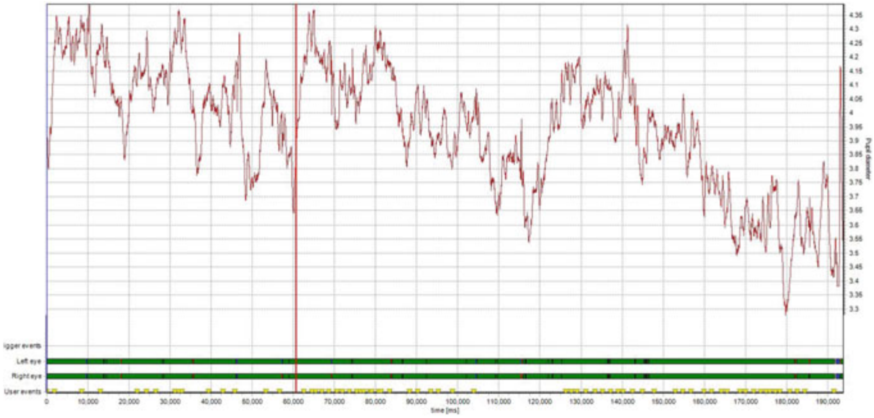


Fig. 11.5 The raw data of one pupil in mm from one participant (unpublished data from the authors' laboratory; display generated with SMI software *BeGaze*) displayed over time (in milliseconds). The participant listened to a musical piece while looking to a blank grayish screen; hence, pupillary changes reflect internal, mental, processes related to listening to music and not to external visual factors (Laeng, Eidet, Sulutvedt, & Panksepp, 2016). The small red bars within the horizontal green bars indicate when the eyes moved from one location of the screen to another (i.e., saccades). The yellow bars at the bottom indicate key presses done by the participant that indicated those moments during the song that were particularly liked so that the pupillary events can be related to moments during the music and emotional state. The thick vertical red line indicates a period in which the eyelids were shortly shut (i.e. an eye blink occurred)

of raw data) may proceed as follows: Data are pre-processed by down-sampling the raw data, especially when using high-speed eye-trackers (e.g., from 240 to 40 Hz, by including—for example—every sixth sample point). Some eye-trackers allow users to select in advance the sampling rate (e.g., 60 Hz are sufficient for capturing the pupillary movements triggered by cognitive processes). The relevant data can be copied into columns of a spread sheet showing, most essentially, the clock time (in milliseconds or smaller fractions) of the sample, an event file number or name (e.g., identifying a picture, movie, or sound file that was used during the trial), and the diameters of the pupil in the X-dimension (in pixels) and in the Y-dimension (in pixels). The X and Y diameters per sample point are typically averaged into a single combined pupil diameter, also to compensate for the distortions due to optical perspective, as discussed above. The average pupil diameter in number of pixels can be later converted in pupil diameter in millimetres (if these are not already available), which can be done by dividing all pupil diameter values by a conversion factor that varies with distance in cm from the screen and with the resolution of the video camera. To obtain the conversion factor, one method consists in placing a circle of known size (e.g., 5 mm) in front of the eye tracker's camera at the same distance of the participant's eyes. The value obtained in pixels is then divided by the known actual size of the 'pupil' in millimeters to yield the conversion factor.

11.3.3 Key Factors in the Experiment and the Removal of Artifacts

Eye blinks, which necessarily obscure the pupil from the camera, are excluded from the raw data by identifying sample points that contain the value 0, so as to replace these with an empty value in data sampling by replacing (filling) the gaps by means of interpolation methods, using the two ends of the recordings around the missing data period. However, some analysis softwares made available by current eye-tracking companies already implements algorithms that automatically remove eye blinks and/or can compute pupil diameters exclusively during eye fixations (e.g., a standard criterion for a fixation is the average pupil within consecutive samples whenever the eye remains within an area of 100 pixels radius for at least 80 ms). The maximum dilation velocity of a human pupil is known to be 0.702 mm/s and the maximum constriction velocity of a human pupil is 3.204 mm/s. Hence, one can use an algorithm that compares two points of a trial and check whether the slope between these two points is steeper than could physiologically be expected. The distance between these points can be chosen to be 250 ms so as to effectively remove short, physiologically impossible, fluctuations in pupil diameter as well as relatively long periods of erroneous tracking (up to 500 ms). Moreover, the diameter of human pupils normally vary from 1.5 to 8 mm (McDougal & Gamlin, 2008) and pupil sizes smaller or larger than these values are likely to be measurement artefacts. Outliers can also be deleted within each trial; e.g., values outside a cut-off equal to mean pupil size ±3 SD (in each trial). One should also be aware that rapid changes in measured pupil size can often be caused by the tracking of objects other than the pupil (e.g., the participant's eyelashes, especially in individuals wearing make-up around the eyes, like mascara). Trials containing less than 50% of the data after the cleaning steps are typically excluded from analysis. Given the possibility of ceiling effects when the baseline pupil is close to the maximal dilation value (e.g., when testing in a darkened room), it may be advisable to run experiments in a room with constant (e.g., windowless) but moderate illumination (Goldwater, 1972; Porter & Troscianko, 2003). In fact, pupillary dilations during cognitive demanding tasks may be mainly mediated by parasympathetic inhibitory modulation on the constrictor muscles of the pupil from the cortex to the midbrain nuclei, so that an already dark-adapted pupil may show minimal dilations due to the cognitive effects (Steinhauer et al., 2004).

Given that spontaneous fluctuations in pupil diameter are of approximately the same magnitude of changes triggered by cognitive and affective processes, it is important in psychological experiments to measure pupillary changes with event-related designs, where the pupil size is averaged over several task-evoked events and then compared to baseline measurements immediately preceding each event. Indeed, it is a well-accepted principle in psychophysiology that the magnitude of any physiological response is partially a function of the initial level (Goldwater, 1972), often as a negative correlation between the pre-stimulus level and the post-stimulus response. Thus, a good experimental design would require that, immediately before the presentation of any static picture, a baseline measurement is taken by showing another image, commonly a blank gray field of matched average luminance to the following picture. Preferably one could use a "mosaic" or pixelated version of the actual stimulus (e.g., Bernhardt et al., 1996; Wu et al., 2011; see Fig. 11.6), which preserves the same distribution and intensity of colors, without revealing the content of the following image. Such event-related baseline measurements typically vary between 2 s and 200 ms durations. The pupil diameter measured during the visual presentation of the baseline can be used as a control condition in statistical analyses or, as it is sometime done in neuroimaging research, as a "rest" condition that can be subtracted from the pupil diameter measured in the event related condition. Box 1 illustrates such "subtraction method" in pupillometry.

Changes in visual information that do not imply changes in luminance, i.e. isoluminant images, can lead nevertheless to systematic pupillary constrictions (e.g., Barbur, 1991, 1995; Barbur, Harlow, & Sahraie, 1992; Gamlin, Zhang, Harlow, & Barbur, 1998; Porter & Troscianko, 2003). For example, color changes like "chromatic flicker" (Kahn & Clynes, 1969; Drew, Sayres, Watanabe, & Shimojo, 2001) or a shift from monocular (seen with one eye) to stereoscopic stimuli (seen with both eyes and therefore in 3D) can result in reductions of the pupil diameters in the absence of any increase in light energy stimulating the eye (Li & Sun, 2004). These small and rapid constrictions are thought to result from transient weakening of the steady-state inhibitory signal from the cortex to the subcortical nuclei controlling the size of the pupil (Barbur, 2004). Thus, in many pupillometry experiments that present visual stimuli, when switching from looking to a neutral blank field or fixation cross to an actual stimulus (e.g., a face), a striking transient "drop" in pupillary diameter

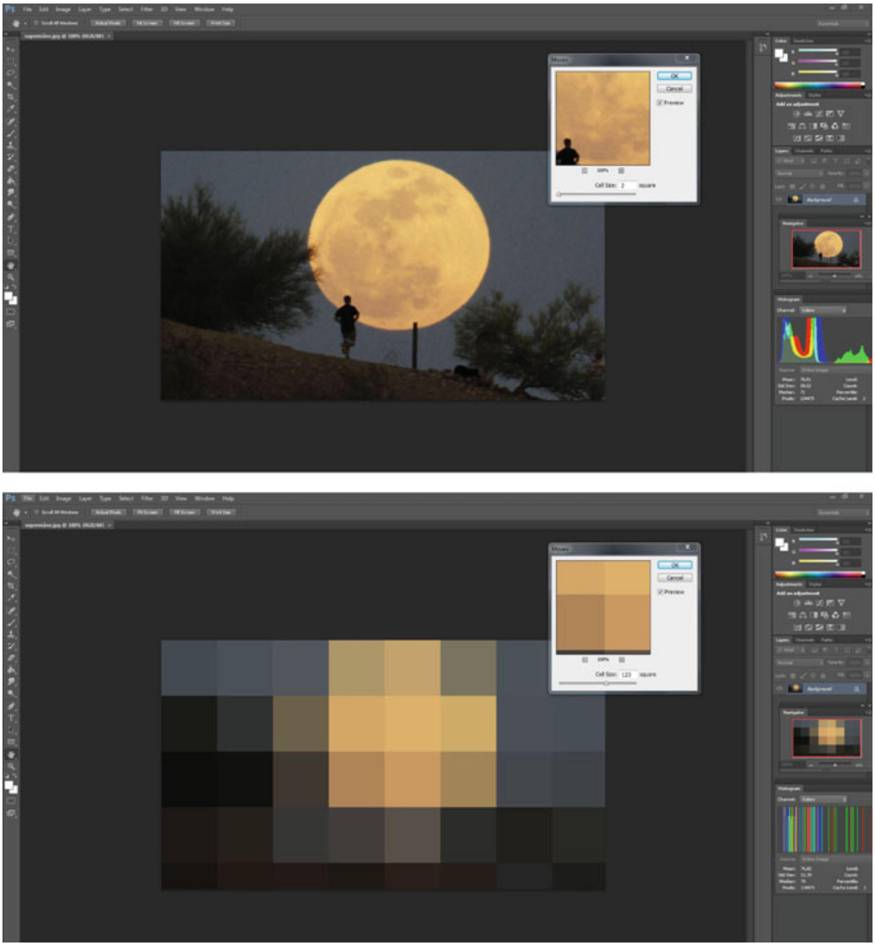


Fig. 11.6 An example of how to create a baseline image by using graphic software (in this case Adobe Photoshop: Filter → Pixelate → Mosaic) so that the average pixel brightness of the pixelated image remains the same (as revealed in the Histogram window) while keeping a roughly similar distribution of colors

is typically observed after the stimulus onset (e.g., Aboyoun & Dabbs, 1998; Young, Han & Wu, 1992). This downward change is typically followed by sustained dilations that are more directly related to the meaning of the stimulus or nature of the task. Remarkably, such pupillary responses can also be measured within the blind field of patients with primary visual cortex lesions and their pupil responses are similar to those for their sighted field, though attenuated in amplitude (Sahraiea, Trevethan, MacLeod, Urquharta, & Weiskrantz, 2013).

Box 1: The "Subtraction Method" in Pupillometry

A good experimental design of pupillometry experiments requires that before the presentation of any stimulus image a baseline measurement be taken by showing another image that is designed to control for and allow the removal of part of the variability in the pupillary response caused by light itself. When designing stimuli for an experiment, each baseline image can be made to preserve the same distribution and luminance of colors, by use of a 'mosaic' or pixelated version of its related stimulus image (as also shown in Fig. 11.7). In some cases it may be appropriate to use a baseline image that can simply consist of a grey neutral field (either "blank" or with a fixation cross) that is adjusted in its luminance to match closely its following stimulus. Standard image processing software, like Adobe Photoshop®, allows obtaining measurements of the mean pixels' brightness of each picture (given in RGB units, varying from the darkest level of 0 to the brightest of 255) and therefore any other image can be adjusted to be of the same luminance in RGB units by manually controlling the brightness levels. In addition, given that 'luminance contrast' in the image can also have measurable effects on pupil diameter (e.g., Wang & Munoz, 2014), this variable can also be controlled by adjusting the standard deviation of RGB luminance of the baseline image to approximate as nearly as possible that of the corresponding stimulus (Moulden, Kingdom, & Gatley, 1990). Implicit in the "subtraction method" is that the pupil response to a blank or pixelated image, which is devoid of content or meaning, will allow to measure—for a particular individual and at a particular time—the effect of light per se on the pupil. By subtracting the pupillary response to the baseline from that to the stimulus, the resulting "pupillary change" will provide a more direct measurement of changes due to internal mechanisms, like mental effort, without the contribution of changes of the pupil due to external factors like light conditions. In addition, subtractions in mm directly express changes in pupil size as 'dilations' or 'constrictions' of the pupil diameter, since these are clearly separated as positive and negative values respectively.

Although several classic studies (e.g., Hess & Polt, 1964) presented pupillary changes in terms of proportion of change (e.g., percentage of change in diameter from baseline), the subtraction method has the advantage of expressing measurement in term of actual physical and anatomical dimensions (e.g., in millimeters or, more typically, fractions of mm). Intuitively, expressing change in percentages would seem to be a corrective for individual differences in anatomical pupil size (e.g., adults' pupils are larger on average than children's). However, individual differences in baseline levels (either due to arousal levels or to anatomical differences) can be dealt effectively statistically by using repeated-measures analyses. Additionally, percentages might seem to offer the possibility of comparing meaningfully the effects obtained from qualitatively different experiments (i.e., different tasks, stimuli, samples of participants). However, according to several researchers, such a "correction" does not seem at all necessary (e.g., Beatty, 1982). In general, simple differences in pupil size provide the most reliable measure of pupil change in size, also considering non-linear changes in its variability with differing levels of illumination (Payne, Parry & Harasymiw, 1967, 1968). As Beatty and Lucero-Wagner (2000) remarked "Today, pupillary movements are conventionally reported as changes in diameter (in millimeters), since this measure appears to be independent of baseline pupillary diameter for a wide range of baseline values."

As the example in Fig. 11.8 shows, computing with the subtraction method actual changes in millimeters can yield greater effect sizes of the difference between conditions than when using the same data expressed as percentage values. According to t-tests, the percentage method may fail to reveal a significant statistical difference in the mean pupil size between the "green" condition and the "red" condition in the plots, $t(17) = 2.06$, $p = 0.056$, $d = 0.69$, whereas the subtraction (millimeter) would reveal it clearly and uncover a larger effect size, $t(17) = 2.46$, $p = 0.025$, $d = 0.82$.

Finally, some studies may compare to baseline the peak or maximal dilation that occurred during a specific time window (e.g., from onset of a stimulus and a behavioral response). However, this measure is more variable than computing the mean

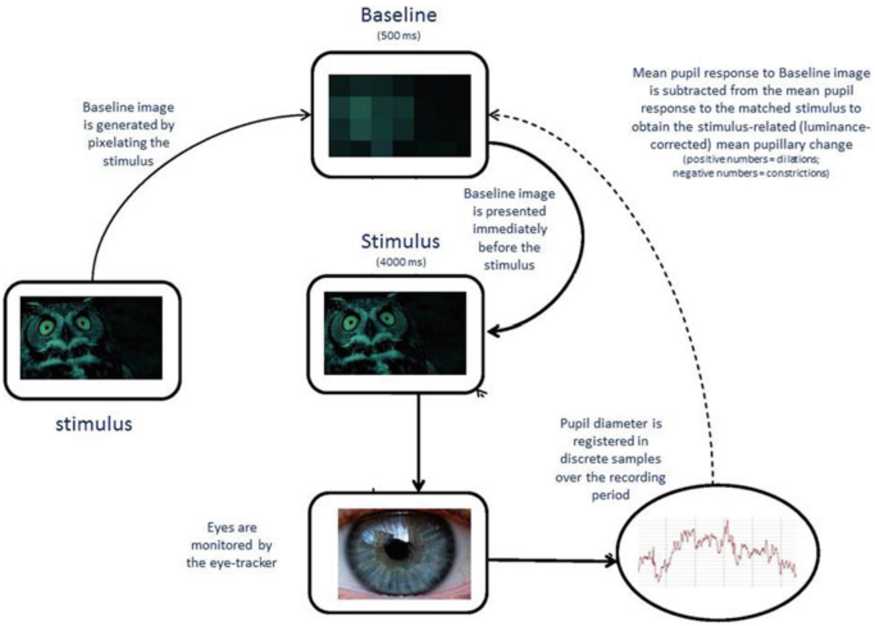


Fig. 11.7 An example of steps used in a typical pupillometry experiment with visual stimuli, where static pictures are presented for a few seconds after a luminance-matched baseline image is briefly shown

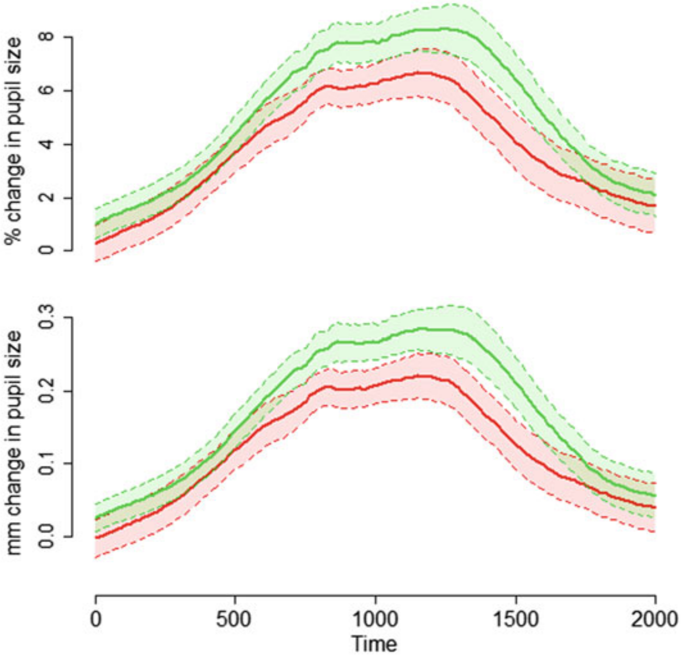


Fig. 11.8 An example from a pupillary experiment comparing two attentional set conditions (by Thomas Hagen, University of Oslo). As visible in the plots, the separation between the mean pupils' curves (solid red and green lines) and their 95% confidence intervals (colored shadow) is clearer when using difference in mm from baseline than percentage change

amplitude of the pupil during the same time window, since peak responses are based on a single value and therefore they are more susceptible to random variations and individuals' differences than the mean diameter (Beatty & Lucero-Wagner, 2000).

11.3.4 Advanced Analyses of the Pupil Signal

Some studies may be interested not only on how pupil size may change on average or over the time course of a trial but also in the dynamic parameters of change in pupillary movements. An example of this is the recent discovery that parameters like (maximum) acceleration in pupil change during the PLR (e.g., provoked by intermittingly spacing flashes of light while the observer is tested in a dark room) can signal abnormalities in the acetylcholine neurotransmitter system that is involved in signaling to the constrictor muscle of the iris (see Fig. 11.1). This measurement has revealed to be very sensitive in diagnosis the presence of neurodegenerative diseases affecting centrally the acetylcholine system, like Alzheimer's or Parkinson's (e.g., Fotiou et al., 2009; Prettyman, Bitsios, & Szabadi, 1997).

In recent years, a variety of new analysis methods of the pupillary response have been proposed based on advances in the analysis of either EEG or fMRI data. Several researchers have recommended the use of deconvolution methods used in neuroimaging (e.g., Hoeks & Levelt, 1993), which are particularly useful in disentangling pupillary responses to events that take place close in time to each other. Due to the slow build-up of the cognitive effects on the pupil, possibly reflecting the transmission and refractory periods of neuronal assemblies in the *Locus coeruleus*, the pupillary response may take from half a second to a second to peak. Wierda and colleagues (2012) applied successfully the "deconvolution" method, typically used in neuroimaging research (like fMRI), to the "attentional blink" paradigm, where targets can follow each other rapidly within a few hundreds of milliseconds and a repeated target is often not seen and missed. Alternative advanced analyses methods that could assist extracting cognitive-related pupillary events despite variations in luminance and fast-paced stimuli can be based on system identification frameworks (e.g., Zénon, 2017) or on principal/independent component analyses (Jainta & Baccino, 2010).

11.4 Pupillometry in Psychology

Pupillometry is the term that has become standard to indicate the measurement of pupil a diameter in psychology (e.g., Janisse, 1977). Today, the use of pupillometry is widespread within the various areas of experimental psychology: social (e.g., Berhnardt, Dabbs, & Riad, 1996); developmental (e.g., Fitzgerald et al., 1967; Gredebäck et al., 2013; Gredebäck & Melinder, 2010; Hepach, Vaish, & Tomasello, 2013; Hepach & Westermann, 2013; Jackson & Sirois, 2009), as well as the cognitive and affective neurosciences (e.g., Demos et al., 2008; Leknes et al., 2013; McRae et al., 2010; Siegle, Steinhauer, & Thase, 2004; Silvers, Weber, Wager, & Ochsner, 2014; Wu, Laeng, & Magnussen, 2011), including the study of animals like monkeys and dogs (e.g., Ebitz, Pearson & Platt, 2014; Iriki et al., 1996; Ogburn et al., 1998; Suzuki, Kunimatsu & Tanaka, 2016). Several clinical psychological studies have also used pupillometry in the study of patients with anxiety (e.g., Bakes, Bradshaw, & Szabadi, 1990; Janisse, 1976; White & Depue, 1999), depression (e.g., Laeng et al., 2013; Siegle, Steinhauer, & Thase, 2004), schizophrenia (Granholm & Verney, 2004; Hakerem, Sutton, & Zubin, 1963), and autism (Anderson, Colombo, & Shaddy, 2006; Blaser et al., 2014; Falck-Ytter, 2008; Nuske et al., 2014; Sepeta et al., 2012).

Part of the appeal in the monitoring of pupillary responses is that they are difficult to control voluntarily (Laeng & Sulutvedt, 2014) and they occur spontaneously (Loewenfeld, 1999; Weiskrantz, 1998). Because of their automatic expression and the non-invasiveness of the current technology, the measurement of the pupils seems particular relevant in psychology since they promise to offer a "window into the innermost mind" (Hess, 1975). Especially exciting is the possibility to tap into processes that occur without voluntary control or censorship (Berrien & Huntington, 1943; Lubow & Fein, 1996) or conscious awareness by the individual under study (Laeng, Sirois, & Gredebäck, 2012; Sahraiea et al., 2013; Tamietto et al., 2009; Yoshimoto et al., 2014).

11.4.1 High-Level Effects on Pupillary Adjustments

The term 'reflex' and the automaticity of the pupil responses and control by subcortical brain structures (Loewenfeld, 1999) do not necessarily imply that the pupil cannot be under the influence of top-down mechanisms, like mental processes that are ongoing in cortical areas supporting high-level perceptual processing. In fact, several studies have shown high-level effects on pupillary constrictions (e.g., Steinhauer, Condray, & Kasparek, 2000); despite the neural network controlling the PLR would seem to generally operate like a closed-loop servomechanism (e.g., similar to the photocell mechanism that controls the automatic opening and closing of a door in an elevator) and it is often described as independent from the (mainly cortical) parts of the visual system that interpret light energy information (cf. Barbur, 2004). The clinical observations of patients whose visual cortex was damaged and who lost conscious visual perception indicated that they still showed adjustment of the pupils to changes in illumination (e.g., Brindley, Gautier-Smith, & Lewin, 1969). It also suggested that the neural circuitry involved in the control of pupil size could act independently of cortical signals, although the pupillary responses of these patients were also greatly reduced.

A series of recent experiments confirm the existence of rapid top-down control on the pupil size by higher cognitive processes. For example, shifting attention covertly (i.e., without moving the eyes away from a central fixation point) towards a bright stimulus shown in the periphery evokes a pupillary constriction (Binda et al., 2013a; Mathôt et al., 2013, 2014; Naber et al., 2013). These responses can be interpreted as anticipations for a particular type of stimulus or cases of predictive processing or cognition. The pupil constricts more when viewing images of objects that are known to be bright, like pictures of the sun, when compared to equiluminant control images of less bright objects (e.g., the moon; see Binda et al., 2013b; Naber & Nakayama, 2013).

Moreover, Bruno Laeng and Tor Endestad (2013) showed images known as illusions of brightness or lightness; i.e., respectively, images where a surface may look like being a source of light or like reflecting more light than the background (Fig. 11.9 shows examples of the stimuli used in this study). When each pattern was seen in isolation on a computer screen, the two images quickly provoked different pupil responses, despite having the same luminance, with the "Asahi" pattern (on the left) constricting the pupil significantly more than "Evening Dusk" (on the right). The conclusion is that the pupillary responses are rapidly matched to the subjective perception of light intensity. In other words, the pupils reflect what we think we see. Such a pupillary response may seem "reflexive", given its speed, but it is the outcome of visual feedback information that is derived top-down (i.e., from the higher-level cor-

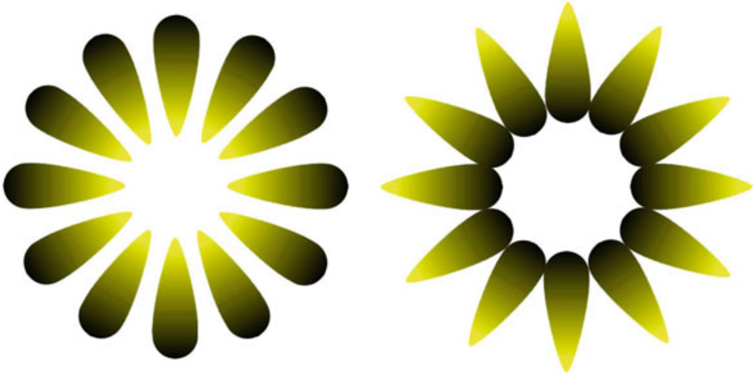


Fig. 11.9 The brightness and darkness illusions used in Laeng and Endestad (2013). On the left is a pattern created by the Japanese psychologist Akiyoshi Kitaoka called "Asahi" (meaning in Japanese: "sunrise"). The central part of the flowery pattern appears like a very bright white color, though the white is identical to that of the background. Moreover, the central part of Asahi has the same white of the pattern on the right by Bruno Laeng, called "Evening Dusk", which was created by simply rotating inside-out the same elements in the Asahi pattern. In this case, the reverse illusion is obtained, which is characterized by enhanced darkness in the center part though the color within the hole is still the same white as the background

tical areas to the low-level subcortical areas that control the pupillary light reflex). Using the same type of stimuli, and concurrently measuring pupil size and brain activity by EEG, Bombeke and colleagues (2016) demonstrated that the pupil diameter change was inversely correlated with the visual C1 component, a measure of the magnitude of feed-forward activity in the visual cortex. These effects of illusory lightness/brightness on pupillary responses further suggest that the low-level neural control of pupil size normally receives modulatory signals from higher-level areas that elaborate the percept.

In fact, merely thinking about stimuli with different luminance can induce pupillary constrictions when the stimulus is bright and dilations when the stimulus is dark (Laeng & Sulutvedt, 2014). Participants in Laeng and Sulutvedt's (2014) study were unable to voluntarily change the size of their eyes' pupils when explicitly asked to do so while looking at an empty gray background (e.g., "please try to enlarge/diminish your eye pupils while looking at the screen"). However, when they were shown shapes that varied in luminance or complexity, and they were subsequently asked to generate "mental images" of these on the same screen background, their pupils adjusted accordingly to the luminance of the imagined object. Participants were also requested to imagine familiar scenarios (e.g., think of a "sunny sky", a "night sky" or a "dark room") while looking at the same neutral screen and, again, participants' eye pupils dilated or constricted, respectively, in response to the dark or bright imagined scenarios. The reason pupillary adjustments occurred only in the imagery conditions and not after explicit commands is explained by the fact that (a) there is no voluntary control on pupil size, differently from eye movements and eye blinks that we can control at will; (b) mental imagery is accounted by processes based on brain states similar to those that arise during perception, which would trigger automatically the appropriate pupillary response whenever a particular scene is re-represented in the visual areas of the brain (cf. Ganis, Thompson, & Kosslyn, 2004). Taken together, the above findings strongly suggest that constrictions of the pupils cannot be just a reflexive response to the levels of light's physical energy that stimulates the eyes at a particular time, since the pupils appear to reflect what we think we see.

11.4.2 Emotion, Motivation, Attitude and the Pupil

11.4.2.1 Sexual Motivation and Feelings

It may not be surprising that the study that launched pupillometry within psychology was about reactions to sexually-relevant stimuli, i.e. nudes of women and men (Hess & Polt, 1960; Nunnally et al., 1967; Peavler & McLaughlin, 1967). Indeed, it was already common lore that viewing sexually-charged stimuli causes pupil dilations that are visible with the "naked eye" (no pun intended) if standing close to a person's face. Pupillary dilations to sexually relevant stimuli are not only large but they can be maintained for protracted periods of time, not only after visual stimulation but also for verbal or auditory stimuli (e.g., listening to erotic passages from a book; e.g., Dabbs, 1997; White & Maltzman, 1977). Just imagining erotic situations to the point of orgasm was able to evoke pupillary dilations equal to 30–40% (measured by from a rest condition or neutral baseline) in women; note that orgasms provoked by actual masturbation in the same female participants resulted in pupil dilations as large as 55% (Whipple et al., 1992). Hess and Polt (1960) had originally measured a 20% change when viewing pictures of nudes, as illustrated in Fig. 11.10. As one would expect, the eye pupils dilated most for nude pictures of the opposite sex (this is particularly clear in Fig. 11.7 for the male subjects), though women showed large dilations also for pictures that included a baby. Pictures of babies (alone or with the mother) were included by Hess and Polt in order to test sex-specific reactions, as well as landscapes as control or "neutral" non-arousing stimuli. Taken together, the above studies show that the sexually-relevant stimuli do not need to be either "visual" or physically present in order to observe dramatic pupil dilations.

Moreover, stimuli may not need to be labeled as "sexual" either in order to observe changes of the pupil related to changes in sexual motivation, as a study by Laeng and Falkenberg (2007) showed by measuring women's pupillary responses to facial portraits of their partners (i.e., sexually-significant others) during their hormonal cycle. Specifically, they found that pupil diameters increased during the follicular (ovulatory) phase of the cycle compared to the luteal or menstrual phases of the cycle and only in those women that were not using the contraceptive pill. Studies that have kept records of the frequency of female-initiated sexual activity during the hormonal cycle have found that this increases significantly during the follicular (or fertile) phase of the cycle (e.g., Adams, Gold, & Burt, 1978). To conclude, the pupillary responses may provide a non-invasive "window" on the implicit state of

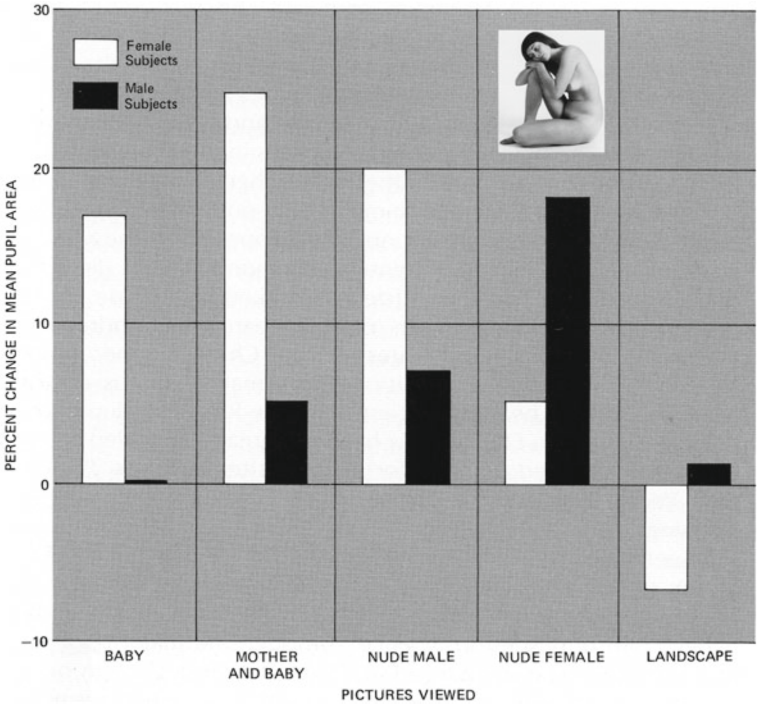


Fig. 11.10 Graph showing the results of the study by Hess and Polt (1960) reproduced in Eckhard Hess's (1975) book: *The tell-tale eye: How your eyes reveal hidden thoughts and emotions.* The inlet image shows one of the original pictures of a female nude that was used in the original study and reproduced in Eckhard Hess's (1975) book

sexual motivation even in clearly non-sexual contexts and without a participant's awareness and this method can be used as an alternative to invasive (e.g., intra-vaginal or penile plethysmography) sexual-response measurement tools (Rieger & Savin-Williams, 2012).

Given the crossover interaction in Fig. 11.10 of the pupil dilations across sexes of participants and nude models, it is likely that Hess and Polt's (1960) participants were heterosexuals. Indeed, later studies by Hess and his colleagues (e.g., Hess, Seltzer, & Shlien, 1965; Metalis & Hess, 1982) that included individuals of homosexual orientation, showed relatively larger pupillary dilations when viewing same-sex nudes. However, several researchers (e.g., Aboyoun & Dabbs, 1998; Garrett, Harrison, & Kelly, 1989) have cautioned against interpreting pupillary dilations to nudes or erotic materials as straightforward evidence for sexual interest, since it could just as well be provoked by stimuli that are novel or surprising to the viewers. In fact, Bernick, Kling and Borowitz (1971) found that any sexual content would provoke large dilations (i.e., also in heterosexual men viewing erotic movies with homosexual content). Hamel (1974) and Garrett et al. (1989) found that the amount of dilation was in general greatest to nude depictions of two males together, followed by images of heterosexual couples or with two females, the lowest dilations being provoked of nude images of either a male alone or a female alone. Rather than reflecting sexual interest, they argued that the pupillary responses could reflect generic arousal towards relatively novel material or to the typical frequency of exposure (cf. Janisse, 1974) to male vs. female nudity within the culture (being less frequent for male bodies than females). Nevertheless, recent studies with a large sample of participants (N = 100 or 325) by Rieger and Savin-Williams (2012) and Attard-Johnson Bindemann and Ó Ciardha (2016) have confirmed that self-reported sexual orientations (i.e., heterosexual, homosexual, bisexual) of both sexes correspond with systematic pupil dilations to erotic stimuli that differed between men or women. Specifically, among men, substantial dilation to both sexes was most common in bisexual-identified men but, among women, substantial dilation to both sexes was most common in heterosexual-identified women. In other words, bisexual men might have bisexual dilation patterns, while homosexual women show male-typical dilation patterns. Thus, it would seem that the pupillary changes could capture, to some extent, the sexual orientations of men and women. However, as Rieger and Savin-Williams (2012) emphasized, sexual-orientation differences in pupillary response can be significant on average but the correspondence is not perfect (i.e., not every participant's sexual orientation could be correctly classified, based on his or her pupil dilation to the same-sex or other-sex erotic stimuli).

Recent studies have also confirmed the Renaissance's common lore, mentioned earlier in Sect. 11.4, by showing that people of both sexes tend to be more attracted to the same face stimuli when shown with enlarged pupils than with either normal or small pupils (e.g., Grundl et al., 2012; Hess, 1965, 1975a; Tombs & Silverman, 2004). Looking at other people's eyes is clearly socially relevant and pupillary responses are typically greater to direct- than deviated-gaze faces, although the effect may be stronger in women than men (Caryl, Bean, Smallwood, Barron, Tully, & Allerhand, 2009; Porter, Hood, Troscianko, & Macrae, 2006), perhaps reflecting a greater interest of women in general than men in decoding facial expressions (McClure, 2000). Interestingly, studies that varied the size of others' pupils while keeping the rest of the face unaffected (see Fig. 11.11), not only showed that larger pupils enhance a face's attractiveness but that this can happen without awareness that the pupils differed in size.

Such an unconscious processing also modulates activity in reward areas of the brain, like the amygdala, as found by measuring brain activity with fMRI (e.g., Amemiya & Ohtomo, 2011; Demos, Kelley, Ryan, Davis, & Whalen, 2008). Interestingly, both humans and chimpanzees adjust the pupils when looking at pictures of co-specifics to mimic or "ape" others' pupil size, a phenomenon that got the name "pupillary contagion" and that may reflect an adaptive mechanism facilitating social communication (Harrison et al., 2006, 2009; Kret, Tomonaga, & Matsuzawa, 2014).

TWO PHOTOGRAPHS, almost identical, elicited very different responses from a group of male subjects. One in which a girl's eyes were retouched, as at left, to make the pupils large got a greater response than one in which the pupils were made small (*right*).

Fig. 11.11 Eckard Hess manipulated within the same photographs the size of the pupils and then asked participants to rate the attractiveness of the faces, finding that people who viewed the same face with large pupils liked better the person than people who viewed the same face but with small pupils

11.4.2.2 Hunger, Pain, and the Pupil

It was clear to Eckard Hess early in his research that the pupil was not simply a window into sexual interests but also to the other physiological drives or motivations which happened to arouse an animal or person. Hess (1975b), Metalis, Hess, and Beaver (1982) found that the pupils dilates when viewing pictures of food in a manner that seems to be consistent with the degree of caloric intake (e.g., the pupil dilated more for images of steak and potatoes than, for example, spaghetti) but only for hungry participants (Hess, 1965). In a study that largely predates the advent of today's "neuromarketing" studies, Hess and Polt (1966) also pioneered the use of pupillometry as a measure of taste differences by comparing presentations of water versus different types of orange drinks and measuring at the time of tasting the pupil size. They found that the pupil diameter was generally larger for the flavored drinks than water and that specific types of orange drinks yielded a significantly larger pupil than the others, possibly revealing taste preferences. Hess (1965) suggested that the pupil response could reveal preferences in some cases in which the actual taste differences are so slight that the subject cannot even articulate them; a possibility with clearly interesting implications for consumer and market research (Arch, 1979).

Sudden pain can causes a dramatic dilation of the pupil which peaks about a second after a person experiences pain (Chapman et al., 1999); size of the pupil dilation is also related to the intensity of the painful stimulus. Höfle et al. (2008)

showed that the response of the pupil to pain is measurable regardless of the level of ambient luminance. Ellermeier and Westphal (1995) showed that women not only rated the same stimuli as more painful than men did but they also showed larger pupil reactions; which they interpret as an objective index of perceived pain, since pupillary reactions are unlikely to reflect a response bias. Interestingly, Walter et al. (2006) measured pupillary responses to pain stimuli before and during analgesic-inducing hypnosis and found that the pain-related pupil dilation was significantly smaller during the hypnotic state than during the non-hypnotic state. Given that the pupillary response appear to be an automatic and objective signal of the actual pain experienced by an individual, it seem promising as a measure to be used in clinical settings, especially with children who may be unable to formulate verbally their experiences (e.g., Connelly et al., 2014).

11.4.2.3 Emotional Expressions, Feelings, and the Pupil

In his book "The expression of emotion in animals and man" (1872), Charles Darwin's aim was to explain why different muscles are brought into action under different emotions (see Fig. 11.12). He mentioned the observation of several scientists about 'fear' as being expressed with "enormously dilated" pupils (p. 242) and stating that "the pupils are always contracted in rage [...] but the movements of the iris under the influence of the different emotions is a very obscure subject (p. 242)." Today,

Fig. 11.12 Greatly dilated pupils and a sideways direction of gaze strongly communicate that the person we are observing is in a state of fear

much pupillometry research in psychology is related to social and affective responses and this line of work has greatly contributed to throw light on Darwin's "obscure subject."

Again Eckard Hess (1965, 1975a, b) should be credited for the initial studies on the central role of the eye pupil in human emotional expressions and communication. Hess (1965) commented that much common metaphorical language links the size of the pupils to emotional states (e.g., eyes like "saucers", "pinpoints of hate", beady-eyed", "bug-eyed", "hard-eyed", etc.). He also mentioned common lore about merchants or magicians guessing someone's intentions by watching their clients' pupillary reactions (e.g., when they are ready to pay a high-price for a specimen or when the "correct card" is turned up). Not only, as mentioned earlier, Eckard Hess found evidence of enlarged pupillary responses to pictures with "shock" contents (e.g., piles of corpses in a concentration camp; see Fig. 11.13) resulted in high galvanic responses measured on the hands as well as quick high pupil response. Stelmack and Mandelzys (1975) showed taboo words and emotional words to groups of introverts and extraverts, showing larger pupillary responses in the former than latter group.

However, there have been occasional observations of pupil constrictions with objects or events rated low in preference (e.g., Mudd, Conway, & Schindler, 1990)

Fig. 11.13 Morgue room next to the Dachau gas chamber, April 1945. Hess used similar pictures in his experiments and discovered that pupils would dilate also to unpleasant, negative, stimuli and not just to pleasant ones

could be due to decreasing attentional engagement with such non-preferred stimuli, which could over time sink below the level measured during baseline. In general, emotional stimuli are powerful in evoking pupillary dilations in proportion to their intensity though not their valence (Laeng, Suegami, & Aminihajibashi, 2016; Janisse, 1974; Partala & Surakka, 2003; Vanderhasselt et al., 2014), especially in individuals who are prone to emotional disorders like depression (e.g., Laeng et al., 2013; Steidtmann, Ingram, & Siegle, 2010). Consistently, pupillary responses to social and emotional stimuli can be reduced in autism (Anderson, Colombo, & Shaddy, 2006; Nuske et al., 2014; Sepeta et al., 2012).

Recent works have shown the usefulness of measuring the pupil as a predictor of empathic responses (e.g., Ellingsen et al., 2013; Harrison et al., 2007; Hepach, Vaish, & Tomasello, 2013; Kret et al., 2013, 2014) and social attitudes (e.g., prejudice; Dabbs & Milun, 1999; Wu, Laeng, & Magnussen, 2011) as well as individual differences in personality (e.g., Prehn et al. 2008). Pupillometry has also been implemented with some success to detect truth telling versus deception or malingering (Bradley & Janisse, 1970; Hacker et al., 2014; Heaver & Hutton, 2010; Lubow & Fein, 1996; Vendemia, 2003; Wang et al., 2010). Another promising application of pupillometry is the study of aesthetics responses to objects of art and design. A study by Carbon et al. (2006) suggested that innovative design of automobiles are cognitively more demanding and evoke more interest, this resulting in pupillary dilations (see also Blackburn & Schirillo, 2012; Kuchinke, Trapp Jacobs, & Leder, 2009; Johnson, Muday, & Schirillo, 2010; Powell & Schirillo, 2011; Yoshimoto et al., 2014).

11.5 Mental Effort and the Pupil

The “filtering” of information is often considered a crucial defining property of what attention is and does (e.g., Broadbent, 1958) but—perhaps—the most fundamental aspect of attention is the necessity of allocation of limited processing resources to task-relevant processes by the brain. As pointed out by Patrick Cavanagh (2004): “although we might all know what attention is (James, 1890), we don’t all agree on what it does, except that it only does a limited amount of it.” This “intensity” aspect of attention, as opposed to “selective”, underscores the fact that information processing by the brain should be understood as a form of biological “work” (Just, Carpenter, & Miyake, 2003). Kahneman (1973) coined the term “attentional effort”, using a phenomenological term to capture this latent intensity aspect of attention. According to Kahneman’s capacity theory, attentional resources are thought to be a limited commodity in the brain. So, when we experience an increase in task demands, and as we start moving towards the limit of the available processing resources, the intensity level at which our cognitive system operates increase. He metaphorically likened the expenditure of attentional resources to connecting an electrical appliance, e.g. a toaster, to the electricity or power grid: for the toaster do be able to do work, it expends electricity from the grid, resulting in a drop in voltage which then is compensated by an increase in energy production from the power plant to keep up with the current energy demands (Kahneman, 2011).

11.5.1 The Allocation of Mental Resources

As for cognition, when task demands increase, an increase in physiological arousal makes available the required processing resources to accomplish the task at hand (Libby, Lacey, & Lacey, 1973; Verney, Granholm, & Dionisio, 2001). Indeed, Otto Lowenstein (1920; p. 194; translation by Goldwater, 1972) had already claimed that dilation of the pupil can be observed "with every increase in attention by intellectual processes of every kind." As a journalist wrote in the magazine *The Scientist* (December 6, 2012): "What do an orgasm, a multiplication problem and a photo of a dead body have in common?" According to Kahneman's view, all of these can strongly grab our attention (despite their very different effects on our feelings) and all of them will require some degree of use of attentional resources.

What is most relevant for the present purposes is that Kahneman further suggested the task-evoked pupillary size could be taken as the most reliable measure of this "energetic" aspect of information processing or allocation of attentional resources (Kahneman et al., 1969). A clear advantage of pupillometry over other typical measures in psychology, like response times, is that it allows continuous monitoring of processing load over time (Bernhardt et al., 1996), thus providing a moment-by-moment estimate of changes in the intensity of processing (Beatty, 1982). Just and Carpenter (1992, 1993) described the intensity dimension of cognition as the consumption rate of limited capacity mental resources involved in maintenance and processing of information (see also Bradshaw, 1967, 1968). The intensity dimension of cognitive processing is thus proposed to be related to the degree of resource consumption relative to the total available processing resources. The term effort, as proposed by Kahneman (1973), is different from its common usage, in the sense that it is not intended to describe the degree to which an individual cares to perform or not or on changes in incentive value within the task (Kahneman & Peavler, 1969); in fact, resource allocation is really not under direct volitional control. However, as Sarter and colleagues (2006) point out, motivational factors like goals and intentions may determine whether or not an individual stays on task and/or return to it in response to challenges (see also Wykowska, Schubö, & Hommel, 2013). Therefore attentional effort is not exclusively a function of external task demands but the interplay of internal and external states. Kahneman (1973) suggested that a useful physiological measure of mental effort should be sensitive to levels of difficulty within tasks and be able to distinguish processing requirements between tasks domains, as well as be sensitive to between-subject differences in mental effort due to individual differences in processing capacity. Importantly, task-evoked changes in pupil diameter have been shown to meet all three of the above criteria (Beatty, 1982). As Kahneman also points out in his book *Thinking, Fast and Slow* (2011), "mental energy is more than a mere metaphor. The nervous system consumes more glucose than most other parts of the body, and effortful mental activity appears to be especially expensive in the currency of glucose. When you are actively involved in difficult cognitive reasoning or engaged in a task that requires self-control, your blood glucose level drops. The effect is analogous to a runner who draws down glucose stored in her muscles during a sprint." Interestingly, exerting physical effort does increase pupil diameter (Hayashi et al., 2010) in relation to the exercise intensity, probably because of the physiological arousal caused by the motor activity.

The existing literature supports the idea that task evoked pupil dilations constitute an important psychophysiological "reporter" variable for the intensity aspect of cognition (Beatty & Lucero-Wagoner, 2000). Importantly, as Just and Carpenter (1993) point out, effort is a characterization of an internal state of the cognitive system, which does not necessarily leave a trace on the output (behavior). While more demanding tasks typically increase error rates, the intensity of processing may very well be different for two participants performing a task at the same performance level as assessed by their behavioral responses. In some cases the pupil may even provide a window to study resource-demanding processes that never reach the threshold of conscious appraisal, or elicit an overt behavioral response (Laeng et al., 2011). The change in pupil diameter thus enable investigators to track online the mental effort, yielding information extending and supplementing the information contained in performance measures and reaction time data.

11.5.2 Cognitive Workload and the Pupil

The first psychological study that clearly related pupil size changes to cognitive processing demands was published by Hess and Polt (1964), were they photographed the participants pupil as they mentally solved two-digit multiplication problems. They reported a striking correlation between the difficulty in solving a task and pupillary size. A few years later, Kahneman and Beatty (1966) measured pupil dilations as participants performed a digit span task (i.e., holding in mind an increasing string of numbers, as when trying to remember a telephone number). They reported a gradual increase in pupil diameter as the digits were presented, before the pupil steadily decreased back to baseline diameter as the digits were reported and no longer had to be kept in working memory. As in Hess and Polt's (1964) study, the peak pupil dilations were strongly correlated with task difficulty, i.e. being largest for strings of seven digits, and smallest for the three-digit condition, demonstrating that the pupil is highly sensitive to within-task processing requirements. They further showed that pupil size measurements reliably indexed differences in processing demands between cognitive tasks, as larger pupillary dilations were observed when participants had to transform (e.g., by adding one digit to each one of the digit in the string), rather than just remembering the presented digit strings. In a subsequent article, also published in *Science*, Kahneman, Beatty, and Pollack (1967) used a dual-task paradigm in which a person performs two task simultaneously, although one (primary) task is given priority over the other (secondary) task; consider, for example, the common dual-task experience of driving the car (primary task) while being engaged in a conversation with a passenger (secondary task). Kahneman and colleagues predicted that intense mental activity should hinder perception, because fewer spared capacity can be allocated to external stimuli, and such a change in perceptual efficiency should vary with the pupillary response. They combined a digit transformation task (primary task) with a letter detection task (secondary task) and found that not only did the participants become “functionally blind” to the perception of the letters when deeply engrossed in thought but also that the degree of perceptual deficit in the dual task closely paralleled the pupillary response to the digit transformation (i.e., the physiological measure of mental effort).

Shiga and Ohkubo (1978) also used a dual-task paradigm to test Kahneman’s (1973) idea that the degree of spare capacity in the cognitive system can be revealed by a decrease in the amplitude of peak pupillary dilation to a probe as effort increases; participants had a primary task of “shadowing” (i.e., following a verbal auditory message) and a secondary task of pressing a key whenever a tone signal was heard. They found that while the pupil size increased with increasing cognitive load during the shadowing task, the time-locked pupillary response to the tones decreased accordingly; thus, supporting the idea that spare capacity decreases as the effort invested in the primary task increases and attention is withdrawn from the perceptual monitoring and concentrated on the main task. Karatekin (2004) has replicated these effects with dual-task with children (Karatekin et al., 2004, 2007).

In the field of psycholinguistics the pupil response has been used to investigate processing intensity or load by Just and Carpenter (1993), who compared pupil dilations in response to comprehension of simple versus more complex sentences and found larger pupil diameter changes with increased complexity. When analyzing the pupil’s time course they found that the largest diameter divergence, hence difference in processing demand, occurred at the position of the sentence where a syntactic complexity was first encountered. The pupil is also sensitive to differences in processing intensity with regard to word frequency effects in lexical decisions (Haro et al., 2016) or speech production (Papesh & Goldinger, 2012), the cognitive effort in translating from a language to another (Hyönä, Tommola, & Alaja, 1995; Kuipers & Thierry, 2013; Schmidtke, 2014) and in the comprehension of degraded speech (Zekveld, Kramer, & Festen, 2010) or prosody (Zellin et al., 2011). Hochmann and Papeo (2014) measured pupil sizes in infants and showed that they are sensitive to the distinction between frequent and infrequent acoustic stimuli and stop consonants coupled with different vowels; all this before they have yet to develop the ability to produce well-formed syllables.

A dilation of the pupil can also reveal the effects of increased load to counteract interference (Kahneman, 2012) during a task like in the Stroop interference effect (Laeng, Ørbo, Holmlund, & Miozzo, 2011), as in the Simon effect (van Steenbergen & Band, 2013), during the attentional blink (Zylberberg, Oliva, & Sigman, 2012), or as an absence of a dilation when the stimulus is “blinked” (Wierda et al., 2012), and during “inhibition or return” (Gabay, Pertzov & Henik, 2011). Overall increases in pupillary size during a task are also seen when a person is inattentive and the mind “wanders” (Franklin et al., 2013). Violations of expectations and predictions cause pupillary dilations (Raisig et al., 2007, 2010, 2012) as well as novel and surprising events (Friedman, Hakerem, Sutton, & Fleiss, 1973; Lavín et al., 2014; Preuschoff, t'Hart, & Einhäuser, 2011; Raisig et al., 2010), since all of these require a re-orienting of attention. Pupils dilate in relation to the degree of "load" on visual attentional resources (Alnæs et al., 2014; Porter, Troscianko, & Gilchrist, 2007). Thus, task-evoked pupil dilations signal processing intensity across, and irrespective of, cognitive domains. Note also that the pupillary dilations observed in these tasks cannot be reduced to the fact that the participant typically has to provide a verbal or manual response or prepare to it, which could per se cause in increment in autonomic activity (Campos & Johnson, 1966, 1967; Simpson, 1969; Simpson & Climan, 1971). In fact, pupil responses are typically greater for correct than incorrect responses (both of which are typically expressed in the same motor behavior) and they also occur when a target is presented despite no immediate response is required (e.g., Kahneman & Peavler, 1969; Paulsen & Laeng, 2006). It is more likely that a motor response will be related to the pupillary change when it is relevant and imposes demands on the participant to make a decision (Moresi et al., 2008, Simpson & Hale, 1969; Van der Molen et al., 1989).

The tight link between resource allocation and pupillary size changes is also evident from studies investigating the effect of processing overload: when load increases (e.g., the number of items to be held in mind) the pupil size reaches asymptote at the peak level of resource allocation. When demands of cognitive load surpass this upper limit, the result is a break-down of task performance, an event that is also reflected in the pupil size so that, instead of increasing, it rapidly declines and returns to baseline levels (Granholm, Asarnow, Sarkin, & Dykes, 1996; Peavler, 1974; Poock, 1973). Again, the power grid metaphor of Kahneman is very apt, since when loading electrical consumption in the household by turning on several appliances at once (e.g., oven, dishwasher, microwave, television, etc.), as a result the fuse will blow and the house will "black out."

11.5.3 Pupillometry in Ecological Conditions

There have been surprisingly few applications of the pupillometry method to the measuring of cognitive load in real-world situations, despite the fact that pupillometry appears to be a sensitive measure of attentional demands and the fact that many modern-time situations would seem to require high levels of attentional "management", which has high relevance for ergonomic performance and safety (e.g., human-machine interactions, surveillance). Driving is a paradigmatic example of a divided-attention type of task, where the strategic deployment of attentional resources plays a common but significant role in everyday life. As Schwalm et al. (2008) pointed out, one main reason why pupillometry as a measure for mental demand has been playing a minor role in applied research lies mainly in the fact that the size of the human pupil is dependent on a wide range of variables (such as light intensity, distance accommodation, etc.). In particular, changes in luminance are problematic in an ecological situation like driving, given the constant change within the visual field and of viewpoint, which makes seem impossible to disentangle the influence on the pupil of cognitive workload from the luminance-based factor. However, one possible solution is to use parameters of the pupillary data that are under the influence of cognitive load but not as strongly of luminance, like different frequencies in oscillations of the pupil diameter. This is basically what the so-called ICA (i.e., Index of Cognitive Activity; Marshall, Davis, & Knust, 2004) does, thus allowing one method for measuring mental workload during situation where ambient light cannot be controlled (see also Hampson, Opris, & Deadwyler, 2010). Schwalm et al. (2008) have applied ICA to pupillometry during a simulated driving task, where participants performed simultaneously the primary driving task that resembled a videogame, where the driver turns the wheel to change lanes, according to road signs appearing from time to time while performing a secondary visual search task on a separate side screen (this one simulating common activities during looking at and switching dials on a radio panel or with the GPS navigator). It was found that ICA increased each time the driver had to change lanes, signaling increased recruitment of attentional resources; moreover, a relative drop in the index occurred when a target in the visual search task appeared around the time a lane change was performed. Thus, these findings possibly indicated a reduction in the allocation of attentional resources (cf. Kahneman, Beatty, & Pollack, 1967) or overload (Granholm et al., 1996) when the requirement of splitting attention became too demanding.

11.5.4 Cognitive Abilities and the Pupil

Given that the amount of invested mental effort varies as learning proceeds (consider how a complex task like driving becomes effortless with practice), one would expect that the pupil diameter, being an index of the allocation of resources, should decrease as behavioral performance for a task improves with learning and training. Indeed, Ahern and Beatty (1979) compared high- and low-scoring individuals on the Scholastic Aptitude Test (SAT; i.e. a test of reasoning abilities) on a mental arithmetic task with different levels of difficulty and reported larger task-induced pupillary responses in the low-scoring compared to the high-scoring individuals. The higher scoring group thus operates at a lower intensity level at a given task difficulty compared to the low-scoring, indicative of more spare capacity or more efficient or automatic processing in the high than low group. Consistent with the idea that the ease of perceptual processing is reflected in pupil size, Yoshimoto et al. (2014) have explored the perceptual fluency phenomenon, where repeated presentations of the same stimuli (also subliminally, by backward masking, without awareness of the exposure by the participants) result not only in explicit preference for the "easy to process" stimuli but also in progressively smaller pupil sizes during exposure.

Using intelligence measurements with the Raven matrices, van der Meer and colleagues (2010) have compared participants with normal and high intelligence on both a choice reaction time task and a more complex geometric analogy task. For the choice reaction time task they reported higher accuracy levels and shorter reaction times for the high scoring group from the normal scoring group. However, on the geometric analogy task, higher performance on the task was accompanied with larger pupil dilations for the high intelligence compared to the normal intelligence group, which would seem at a first glance to run opposite to the above-mentioned findings by Ahern and Beatty (1979). As van der Meer and colleagues argue, in more challenging tasks than the simple math originally used by Ahern & Beatty, the group with higher fluid intelligence will be able to operate at a higher processing intensity (without incurring into overload and cognitive "black-out"), leading to larger pupil dilations on average in the task as well as lower error rates and RTs for the most difficult problems. Similarly, Takeuchi et al. (2011) trained participants on a complex conjunction search task, by practicing the task every day for 16 consecutive days. They found that despite response times became faster with training, this "expert" behavior was not accompanied by a decrease but an increase (see also Raisig et al., 2010) in the mental effort invested in the task (as also confirmed by the subjective reports of the participants). Those individuals who allocate resources to the processing of irrelevant information, as indexed by pupillary dilations to distractors, also perform more poorly on cognitive ability tests (Verney, Granholm, & Marshall, 2004) since they are displacing resources from the main task.

Not all studies on individual differences in cognitive abilities have been based on the fast changes in pupillary responses during a cognitive task and some have actually relied on more sustained differences of the pupil across individuals either at rest, before a task has actually been started, or during repeated baseline measurements in the course of the task. In some studies, a pre-test baseline measured before a specific cognitive task yielded individual differences that were positively correlated to performance in working memory tasks (Unsworth & Robison, 2017) as well as "fluid intelligence" (Tsukahara, Harrison, & Engle, 2016). These authors have suggested that such pre-test baseline measurements reveal capacity aspects of attentional functioning that are important for maintaining moment-to-moment optimal arousal levels and for controlling attention in a goal-directed manner. Although cognitive abilities and pupil size at rest may show a significant relationship, as these studies suggest, pupillary differences before a testing session could reflect motivations, interests and plans and therefore the level engagement of underlying cognitive/neural resources. In such a case, pupil size may predict the scores of subsequent cognitive tests without it reflecting stable differences in cognitive capacity (e.g., if I do not run now, it does not necessarily imply that I am unable to do that).

Based on the above evidence, it seems that pupillometry can reliably index individual differences in mental effort within and between task domains (as originally proposed by Kahneman, 1973; Beatty, 1982). The above results are similar when comparing effects of age on performance efficiency and pupillary responses. By adjusting the task-evoked pupillary response for age-related differences in pupillary reactivity, Piquado (2010) reported larger pupillary responses in a digit span task for older adults compared to younger adults. Interestingly, the performance level did not significantly differ between the groups, so that the increased dilations in the older adults suggested that they needed to allocate more resources than the younger participants needed to do in order to perform the task at the same level of performance. Thus, pupillary differences among groups might signal the presence of compensation mechanisms.

11.5.5 Consciousness and the Pupil

Changes in pupil diameter are sensitive even to small changes in effort or load on attentional capacity. Simply detecting the appearance of a target stimulus leads to a pupillary dilation (Privitera et al., 2010) whereas the presentation of an irrelevant distractor typically results in no pupillary change; whether the stimuli are visual or auditory (Beatty, 1982; Klinger, Tversky, and Hanrahan, 2011) or a near-threshold visual stimulus (Hakerem & Sutton, 1966). Recent studies have basically replicated the original results of Harms (1937) with binocular rivalry, by showing that the pupil size can reveal which of a variety of concomitant stimuli dominates awareness (e.g., Fahle, Stemmler, & Spang, 2011; Naber, Frassle, & Einhauser, 2011). In a study with bistable stimuli like the Necker cube (Einhäuser et al., 2010) found that the pupil began to dilate slightly earlier (about 500 ms) than the time the participants signaled by a key press that a switch had occurred in their perception and this dilation peaked a few hundred milliseconds after such a response. Thus the pupillary dilations were related to likelihood that a particular item would enter consciousness, in other words revealing the ability of pupillometry to accomplish "mind reading" (Einhäuser et al., 2010). Indeed, the pupillary response has also been successfully used to aid the communication with locked-in syndrome patients (Stoll et al., 2013) and it holds the promise of establishing itself as a relatively simple or inexpensive (compared to fMRI; e.g., Owen and Coelman 2006) but reliable method for obtaining responses from patients who are unable to verbalize or express their mental states.

The above evidence strongly suggests that pupillary responses can provide an easily observable signal of the moment in which one event becomes relevant to awareness while alternative events are consigned to a negligible status. Stimuli presented at the visual threshold (i.e., barely perceptible) tended to provoke significant pupillary dilations more when they were detected than when they went undetected (e.g., Hakerem & Sutton, 1966). However, more recent studies have clearly revealed that also unreported targets can trigger pupillary dilations (Privitera et al., 2010) and that subliminal 'reward cues' evoke changes proportional to the cues' value as well as the level of demands of a cost/benefits decision-making task (Bijleveld, Custers, & Aarts, 2009; Yoshimoto et al., 2014). Thus, pupillometry can effectively index changes in mental state that would have led, normally, to their conscious acknowledgement but that remain under the threshold of awareness. A clear advantage of monitoring pupillary responses over behavioral measurements (like key presses) is that the pupillary changes do not depend on an overt and controlled response that, despite allowing the evaluation of the correctness and efficiency of processing, it is limited to representing the final state or behavior of the system. Hence, pupillometry can capture processes "in the making" that may still be in a "preconscious" state

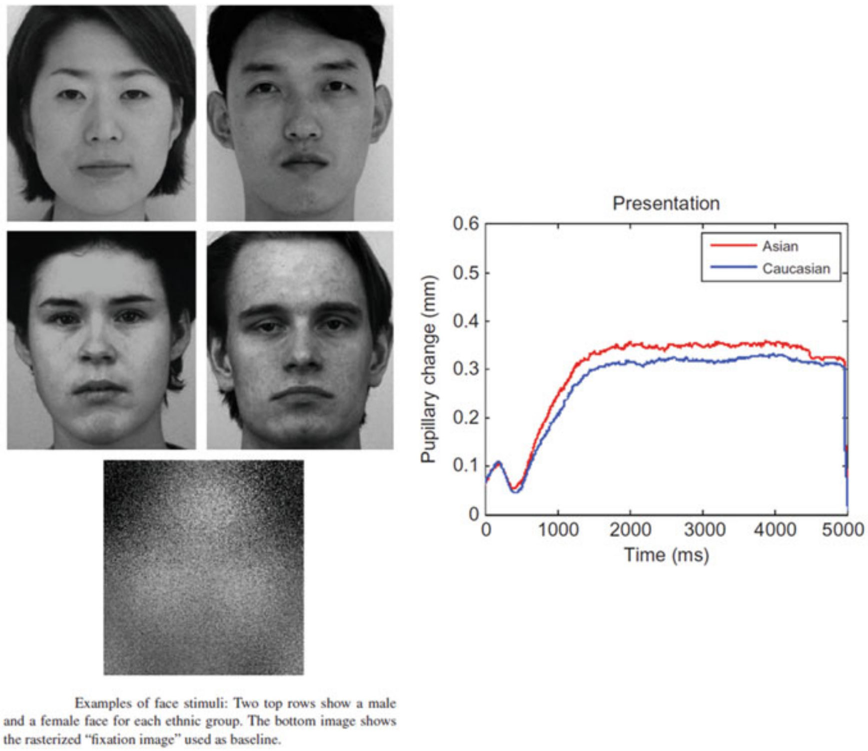
(Fahle, Stemmler, & Spang, 2011; Laeng, Sirois, & Gredebäck, 2012) as well as when they are not able to reach a conscious state, as a consequence of brain damage interrupting the flow of information between various regions of the brain (Laeng et al., 2007; Tamietto et al., 2011; Weiskrantz, 1998; Weiskrantz et al., 1998, 1999).

11.5.6 Memory and the Pupil

Kahneman and Beatty (1966) measured pupil dilations during a verbal memory task (i.e., the digit span task by holding in mind an increasing string of numbers) and found that the pupil size was proportional to load on (working) memory. It is now clear that load on spatial memory can also cause the pupil to dilate proportionally to the load for item locations (Porter et al., 2007; Unsworth & Robison, 2014). However, with aging, the pupillary responses to load on memory decrease in size (Van Gerven et al., 2004).

Pupils are larger when participants recall information than when they recognize the same information (Goldinger & Papesh, 2012; Stanners, Headly, & Clark, 1972), suggesting that rehearsal or maintaining information on-line while retrieving it from long-term memory may require more mental effort. In general, the repetition of stimuli leads to adaptation of the pupil response and stimuli that are unexpected, surprising, or low in probability typically result in dilations of the pupil (Beatty, 1982; Friedman et al., 1973; Lavìn et al., 2014; Preuschoff, t'Hart, & Einhäuser, 2011; Qiyuan et al., 1985; Raisig et al., 2010). Consistently, Laeng and colleagues (2007) showed that amnesic patients, after damage to the hippocampus of the brain, showed pupillary dilations to pictures of new items compared to old items, despite the latter were not explicitly recognized by the patients. However, in healthy participants, the pupils dilate to a greater extent when participants view old compared to new items during recognition memory tests (Gardner, Mo, & Borrego, 1974; Heaver & Hutton, 2011; Otero, Weekes, & Hutton, 2011; Papesh, Goldinger, & Hout, 2012; Võ et al., 2008). Hence in normal recognition tests, the pupil response may be driven by the conscious acknowledgment of the recollected target and the "dwelling" of the mind on the memory.

Recognition memory strength can be predicted by pupillary responses at encoding (Kafkas & Montaldi, 2011; Otero, Weekes, & Hutton, 2011). In an fMRI study, Sterpenich and colleagues (2006) found that the *Locus coeruleus* was engaged during memory retrieval in proportion to the pupillary dilation shown at encoding for the successfully retrieved items, suggesting a role of the LC–NE system in the consolidation of memories and confirming the use of the pupillary response as an index of learning. Learning and remembering faces of other races is also effortful and require cognitive resources; thus, Wu, Laeng, and Magnussen (2011) found that pupils dilated more when learning faces of another race (see Fig. 11.14). Goldinger, He, and Papesh (2009) found the degree of pupil dilation at encoding predicted the memory performance with face of different races.



Examples of face stimuli: Two top rows show a male and a female face for each ethnic group. The bottom image shows the rasterized "fixation image" used as baseline.

Fig. 11.14 Example of a pupillometry study of memory retrieval within the social and affective neurosciences by Wu, Laeng, and Magnussen (2011). The study investigated the so-called other-race bias or the difficulty in recognition memory for faces belonging to another ethnicity than that of the observer. The pupillary curves show that pupil size was larger when making seen/unseen decisions with Asian faces (all participants were Caucasians), which can be interpreted as an indication of greater effort in retrieving past visual information about faces of other than the own race

Studies of learning have revealed that the pupil size reflects the degree of uncertainty about stimuli in a task or the consequences of events (e.g., Einhäuser, Stout, Koch, & Carter, 2008; Friedman, Hakerem, Sutton, & Fleiss, 1973; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011; Preuschoff, t'Hart, & Einhäuser, 2011; Raisig et al., 2010; Richer & Beatty, 1987). In particular, pupillary changes appear to index a specific form of uncertainty; that is, after change points in information about an event signal the need for rapid learning on the basis of the most recent information (Nassar et al., 2012). Thus, task uncertainty is related to the amount of mental effort exerted at any given time and this is reflected in the pupillary changes (cf. Aston-Jones & Cohen, 2005a, b).

11.6 Neuro-pupillometry: The Pupil as a "Window" to Brain Function

During the last decades our knowledge of the brain, and how its structure and function enables mental operations has, and continuous to, increase at a fast rate, owing much to techniques such as functional magnetic resonance imaging (fMRI) and other brain imaging methods allowing us to probe the brain in action. While Kahneman's toaster-metaphor is still a useful one, recent advances in the field of cognitive neuroscience may allow us to translate the metaphor in terms of specific brain mechanisms or networks that support effortful processing and resource allocation.

11.6.1 Pupil Dilations and the Brain Stem's LC-NE System

Pupil dilations related to cognitive processes are thought to result principally from an inhibitory effect on the parasympathetic oculomotor complex by release of norepinephrine (NE) from the *Locus coeruleus* (LC) (Wilhelm et al., 1999). A few studies have combined single cell recordings of LC-neurons with concurrent measurement of pupil diameter in monkeys, have reported a tight correlation between LC-activity and pupil dilations (Joshi, Li, Kalwani, & Gold, 2016; Rajkowski, Kubiak, & Aston-Jones, 1993), suggesting that online pupillary size changes may be used as a proxy, or indicator, of the neuromodulatory activity by the LC-NE system. Joshi et al. (2016) found that microstimulating neurons within the LC led to pupillary dilations similar to those shown by the same animal to startle sounds. Moreover the pupillary dilations caused by the electrical microstimulation of the LC occurred earlier than those after stimulation of two other subcortical regions, the inferior colliculus and the superior colliculus, that have both been linked to pupillary changes that are not mediated by luminance. Moreover, pharmacologically induced increases in LC-NE-activity also result in increased pupillary dilations in humans (Hou, Freeman, Langley, Szabadi, & Bradshaw, 2005; Koss, 1986; Phillips, Szabadi, & Bradshaw, 2000). Interestingly, lesion studies of the LC in animals show that following bilateral destruction of the *Locus coeruleus* complex (with 85–95% depletion of cortical norepinephrine) fails to induce any change in EEG desynchronization. This suggests that the effects of *Locus coeruleus* stimulation are not consistent with a generalized tonic activating system or "arousal" system, but with more specific components of arousal. In fact, the most relevant operation of such components might be to increase signal-to-noise ratio or gain (Aston-Jones & Cohen, 2005a, b; Foote & Bloom, 1979) as well as in switching emphasis from one set of inputs to another (Einhäuser et al., 2008).

These studies thus suggest a coupling between LC-activity and the task evoked pupil response, and given the relationship between the pupil response and mental effort, a role of the LC in processing resource allocation and cognition has been hypothesized. Using fMRI, recent studies have probed the activity of the LC in order to investigate this hypothesis. Identifying LC-activity using fMRI is challenging due to the limited spatial resolution of fMRI and the small size of the LC nuclei: it requires careful steps for aligning the participants' brainstems to a common template, steps which are usually not performed for fMRI studies which primarily investigates cortical activations. Two recent studies employing such steps suggest that a relationship between task-evoked pupil diameter changes and LC neural activity does indeed exist in humans (see Box 2; Alnæs et al., 2014; Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014).

11.6.2 The Role of the LC-NE System in Perception and Attention

In general, the LC-NE system is suggested to play a role in the regulation of cortical attention and cognitive processes (Berridge & Waterhouse, 2003). It accomplishes this by interacting with cortical top-down modulatory systems, and regulating the interaction between brain networks responsible for goal driven, focused attention and those responsible for initiating attentional shifts (Aston-Jones & Cohen, 2005a; Aston-Jones, Rajkowski, & Cohen, 2000; Bressler, Tang, Sylvester, Shulman, & Corbetta, 2008). Goal-driven attention depends on a cortical network encompassing dorsal fronto-parietal areas and it is thought to both mediate stimulus selection and mapping of appropriate responses to task relevant stimuli. Stimulus driven attention on the other hand depends on a ventral fronto-parietal network, which is involved in interrupting and reorienting attention to behaviorally relevant stimuli outside the current focus (Corbetta, Patel, & Shulman, 2008). Recent theorizing about the role of the LC-NE system based on animal models relates different modes of LC-activity to specific behavioral states and modes of attention (adaptive gain theory; Aston-Jones & Cohen, 2005b): in what is termed the 'phasic mode', the LC exhibits an intermediate tonic firing rate, and responds with bursts of activity to behaviorally relevant stimuli. This mode of LC-activity facilitates goal-directed behavior in contexts where reward-contingencies are known and are being exploited. The phasic activity is suggested to optimize or consolidate task related decision processes, i.e. the process of mapping task-relevant stimuli to appropriate responses within a task context. In the tonic-mode the LC exhibits high baseline activity and a reduction of phasic activity, behaviorally reflected in an increased distractibility, and hence at the same time decreases task-performance and instead facilitates exploration (i.e., distractibility) of alternative sources of reward. In fact, tonic pupil diameter shows an inverted U-shape to the magnitude of P3-amplitudes during the auditory oddball experiment (Murphy, Robertson, Balsters, & O'Connell, 2011). The P3 is an event-related potential (i.e. time-locked to a stimulus) measured using EEG, and is evoked by presenting a less frequent (oddball) stimulus among a stream of repeated identical stimuli. The magnitude of both task-evoked (phasic) pupil dilations and the P3 decreases when tonic pupil diameter deviates from the intermediate level associated with the phasic LC-mode. Thus, both measures track changes in attentional states, as predicted by the adaptive gain theory of the LC-NE system. In fact, the P3 has been hypothesized to have a noradrenergic origin, signaling the outcome of neural decision making processes in the cortex (Nieuwenhuis, Aston-Jones, & Cohen, 2005).

Interestingly, several studies do suggest a role of the LC-NE system in internal cognitive (Einhäuser et al., 2010) and perceptual (Einhäuser, Stout, Koch, & Carter, 2008) decision processes, suggesting a general role of the LC-NE system in facilitating reconfigurations of neural networks (Bouret & Sara, 2005). By regulating the interaction between cortical networks, the LC-NE system would seem to play a pivotal role in both maintaining attentional states associated with goal driven behavior and in 'interrupting' and 'resetting' of neural networks facilitating attentional, cognitive or perceptual shifts (Bouret & Sara, 2005; Einhäuser et al., 2008; Sara & Bouret, 2012).

11.6.3 The LC-NE System's Role in Mental Effort

In the context of resource models of attention, the pupil reliably indexes the moment-to-moment changes in mental effort, reflecting allocation of attentional resources mediated by regulation of cortical arousal states by the LC-NE system (Sara, 2009). The pupil may also signal NE-mediated cortical network plasticity enabling adaptive behavior in response to changes in environmental or task imperatives. Indeed, Sarter et al. (2006) suggested that the term attentional effort should be understood in terms of the brain's top-down control of attentional processes, a conceptualization that fits well with current views on the role of LC-NE system in cognition. The adaptive gain theory has also gained supported from recent pupillometry studies with human participants (Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010).

So far only a few studies have combined pupillometry with fMRI measurement in order to investigate and validate the implication of specific neural substrates for mental effort. Modeling individual differences in mental effort, Siegle, Steinhauer, Stenger, Konecky, and Carter (2003) employed concurrent pupillometry and fMRI measurements while participants performed a digit sorting task, in order to explore whether modeling of individual differences in effort could further inform the analysis of task-fMRI. They compared models using the grand mean pupil dilations for each of the conditions as a measure of task difficulty, as well as modeling individual differences using mean condition and trial-to-trial pupil dilations, to a standard approach using task-based explanatory variables only. They reported that models that included pupil dilations increased the explained variance of the model, especially when modeling individual mean pupil responses and trial-to-trial differences (see also Nassar et al., 2012), and revealed activation in prefrontal and parietal areas previously reported to be involved in executive and working memory processes. Importantly, they did not find any condition-related behavioral effects, and all participants performed the task at low error rates. Using task-evoked pupil dilations they were able to model individual differences in mental effort, which was not captured by variables modeling the levels of task-difficulty. This study extends the usefulness of the pupil size as a reporter variable of effort to neuroimaging studies, enabling researchers to model individual differences, which may not leave a trace on the behavioral performance.

Recent fMRI studies have specifically addressed not only the question of a correlation between LC-activity and pupil diameter, but also the role of the LC-NE system in resource allocation when task demands vary. A study by Raizada and Poldrack (2008) reported activity in a brain stem area encompassing the anatomical location of the LC related to sudden and unpredictable changes in task difficulty (see Fig. 11.15). Also, functional connectivity between the LC and the right frontal cortex increased with task difficulty, and the LC could signal to frontal executive areas the need for resource allocation in response to changes in task demands. During multiple object tracking (MOT; Cavanagh & Alvarez, 2005; Scholl, 2009) pupil diameter increases as a function of the number of visual objects (see Fig. 11.16) that has to be simultaneously tracked by the attentional system (Drew, Horowitz, & Vogel, 2013), and the

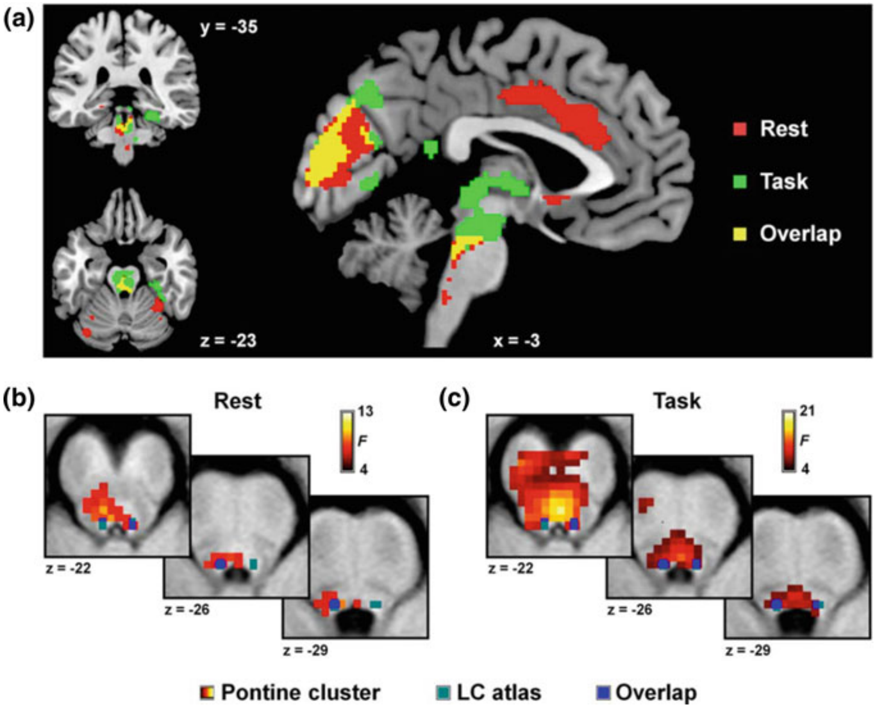


Fig. 11.15 Brain areas which show activations that correlate with concurrently measured pupil diameter changes during resting state fMRI and during performance of an oddball experiment is shown in panel **a**. LC-activations during the same conditions are shown in panels **b** and **c**. During an attention demanding task, the LC-pupil correlations increase in strength and encompasses a larger cluster of voxels compared to rest

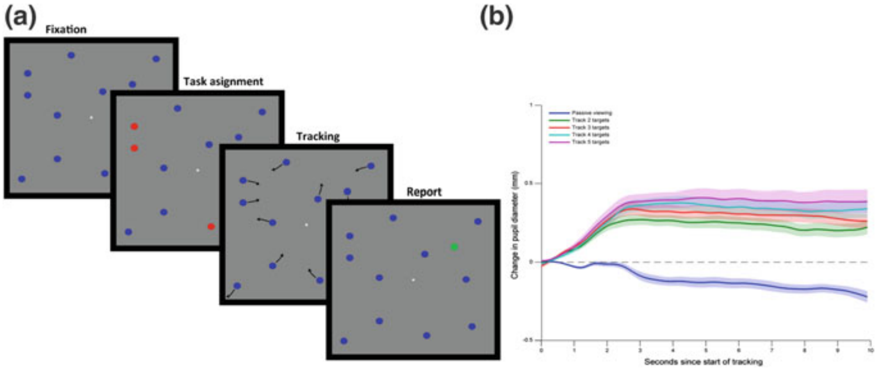


Fig. 11.16 **a** The multiple object tracking task. The task starts with a display presenting the objects (i.e., blue colored disks). Then a subset of the disks turns red (in tracking trials). All colors return to blue before they starting moving randomly inside the tracking area (gray). At the end of the tracking period, the participants respond to a probe (i.e., one of the disks changes to green), where the participant answers "yes" or "no" on whether the probed object was one of the original target objects (fMRI session). **b** Group averages of the pupil time series during Multiple Object Tracking. The x-axis represents the tracking period in seconds, y-axis the pupil size in mm. The pupil is clearly sensitive to changes in cognitive load

magnitude of these pupil dilations correlate with activity LC (Alnæs et al., 2014). In a non-visual task of mental effort, Landgraf, van der Meer and Krueger (2010) investigated math-specific cognition measuring pupil dilations and fMRI activations in two different sessions of a task in which participants solved multiplication problems of different levels of difficulty. They reported activation in a frontoparietal circuit, which showed monotonically increasing activation with problem difficulty. The pupil was also sensitive to these variations in computational intensity and when correlating the peak pupil diameter change with BOLD signal change for the regions identified in the fMRI session, a significant association was found with the intraparietal sulcus (IPS), also supporting a role for this region in math-specific cognition. In both studies of Landgraf et al. (2010) and Alnæs et al. (2014), the average pupil dilation within each level of task load was computed for each participant in a separate session outside the fMRI environment, thus serving as an index or 'trait' of each participant's cognitive system (like IQ scores) rather than mapping moment-to-moment changes in pupil diameter and BOLD-signal (see Fig. 11.17).

Concurrently measuring pupil dilations and BOLD-activity (as in Murphy et al., 2014) allows a direct comparison with pupil-related activations. This approach was used by Zekveld, Heslenfeld, Johnsrude, Versfeld, and Kramer (2014) for measuring task-evoked pupil dilations simultaneously to BOLD signal amplitude while participants listened to sentences with different degrees of degradations. Larger pupil dilations were observed with increasing speech intelligibility. When modeling brain activation using individual trial-by trial peak dilations, Zekveld and colleagues reported pupil related activity both in areas implicated in speech perception and in frontal executive areas, thus revealing individual differences in areas assumed to be sources

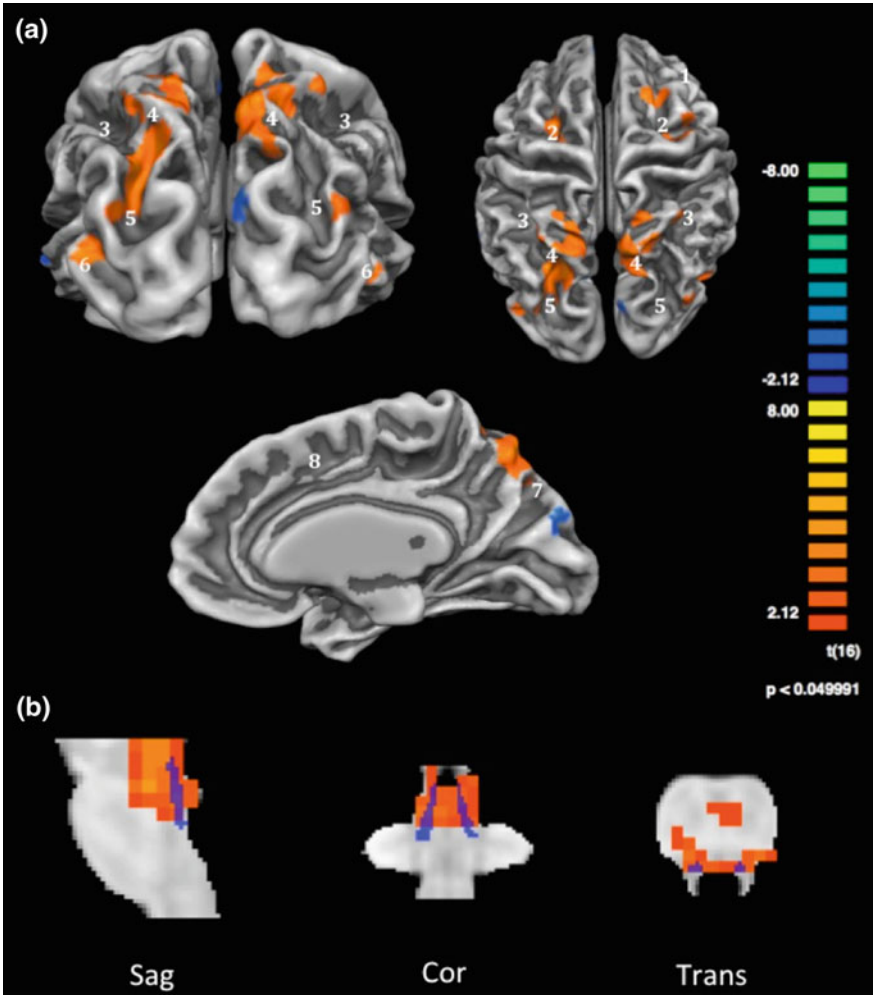


Fig. 11.17 Brain activation related to pupil diameter changes, where pupillometry and fMRI data were collected for the same participants in two different sessions of multiple object tracking. In panel **a** the activation map is projected on the cortical surface, showing pupil related activity in a dorsal fronto-parietal network which is critical for goal driven attention. In panel **b** the brain stem activation is shown. Here the location of the *Locus coeruleus* is shown in blue. The color bar shows the statistical values (t) for the cluster level significance test

and targets of top-down modulatory influences related to effortful processing of speech.

Attention is often thought to reflect a common, general purpose processing resource. However, a caveat with capacity models of attention and cognition is that they refer to a limited commodity labeled capacity or processing resources, without explicitly specifying what these resources are in the brain (Just & Carpenter, 1992; Navon, 1984). Alvarez and Cavanagh (2005) proposed that capacity limits, instead of stemming from a central resource (e.g., metabolic supply, as in Kahneman's previous statements), should be understood in terms of interactions between several limited-capacity processing stages. A recent proposal (Franconeri, Alvarez, & Cavanagh, 2013) suggest that these stage-specific resource limitations may also be related to a physical or structural, rather than a process (or metaphorical) commodity in the brain, which they term "cortical real estate." Their account specifically relates capacity limitations to the representational space in cortical two-dimensional representational maps, in which representations of objects or events compete for representational space. Just, Carpenter and Miyake (2003) have suggested that pupil dilations reflect an aggregate measure or summed index of brain activity across separate pools of resources, some of which are limited by the lower-level architecture of processing modules, i.e. for spatial and language processing, but also with regard to the processing capacity of higher order systems involved in executive control. Hence, the concept of processing resources is likely to refer to multiple areas of the brain, to their structural organization, and to how they interact with one another.

Box 2: Pupil Dilations and the LC-NE System

Studies using single-cell recording and microstimulation of LC neurons in monkeys have documented a robust relationship between pupil dilations and activity of LC neurons (Aston-Jones & Cohen, 2005; Joshi, Li, Kalwani, & Gold, 2016). Importantly, such a relationship suggests that we can measure pupil diameter changes and use them as a proxy measure of noradrenergic modulation by the LC. In humans, however, such a relationship has to be investigated using more indirect, non-invasive methods of measuring brain activity, such as fMRI. Two recent fMRI studies have produced evidence for such a relationship also in humans. Murphy and colleagues (2014) performed a time series correlation of concurrently measured BOLD activity and pupil diameter during both a resting state run and during a run where participants performed a visual oddball task. In both runs they reported pupil-related activation in a cluster encompassing the LC, localized using neuromelanin-sensitive structural imaging (Keren, Lozar, Harris, Morgan, & Eckert, 2009), with a stronger and more spatially extended cluster during the task-engaged run (Fig. 11.15). This relationship was robust when probing the effect of spatial smoothing and persisted when correcting fMRI time series for physiological noise. This study provide evidence that moment-to-moment changes in pupil diameter can be used as a proxy measure of LC-NE activity in humans. In another study, Alnæs, Sneve, Espeseth, Endestad, van de Pavert and Laeng (2014) measured pupil dilations and BOLD activity in the same participants across two sessions using the Multiple Object Tracking (MOT; see Fig. 11.5) task. As previously shown (Beatty, 1982; Siegle, Ichikawa, & Steinhauer, 2008; Steinhauer et al., 2004), pupil dilations were clearly sensitive to sustained processing and increased parametrically with the number of tracked objects. In Alnæs et al's study, each individual's mean pupil diameter during different levels of attentional load was calculated and used as an index of individual differences in attentional effort, to model brain activation during a separate fMRI-session of the MOT. Consistent with the proposed role of the LC-NE system in resource allocation during top-down attention, the individuals' pupillary changes explained variance beyond the tracking load, in areas belonging to both the top-down dorsal attention network and the brain stem including the LC (see Fig. 11.17). Importantly, using each individual's pupil dilations to model brain activity revealed clearly related LC activity to mental effort and more extensively than using the levels of load, as operationalized in the task by the number of targets to be tracked. Together, these two studies support the notion that pupil dilations can be used as a proxy of LC-NE activity in humans, and implicates the LC-NE system in allocating attentional resources during effortful cognitive processing.

11.6.4 The LC-NE System and Brain Network Interactions

According to Bouret and Sara (2005), the LC-NE system plays a cortical role in regulating the dynamic interactions between neural systems. According to their network reset theory, phasic LC-activity acts to interrupt activity in neural networks and facilitate their reorganization into new functional configurations. One example where such a reorganization might be necessary is when an important stimulus or event occurs outside our current focus of attention. Experiments probing such attentional shifts have revealed that unattended but task-relevant stimuli evoke both pupillary dilations as well as the P3. Further, imaging studies have shown that attentional shifts engage a right lateralized ventral attention network, acting to interrupt the ongoing activity of a bilateral fronto-parietal network involved in goal-driven, focused, attention (Corbetta, Patel, & Shulman, 2008). A possible mechanism in which the LC-phasic response may act as a network reset signal in the human brain may thus be through signaling to the ventral attention system to send a circuit-breaking signal to the dorsal attention system in order to disengage from the current task and engage in a new (Sara & Bouret, 2012). In fact, pharmacological blocking of cortical NE have been shown to attenuate such functional reorganization in humans: by exposing participants to fear-related stressors, Herman and colleagues (2011) reported large scale reorganization of cortical networks, in which the ventral attention system increased connectivity with sensory visual networks as well as the amygdala and the thalamus. Blockade of NE-receptors damped this reconfiguration by reducing the resulting connectivity change, an effect, which seemed specific to NE since the inhibition of cortisol did not have the same effect. In another study, Eldar, Cohen, and Niv (2013) found that pupillary measurements were strongly correlated with global changes in the strength and clustering of functional connectivity and suggested that brain-wide fluctuations of gain modulate the breadth of attention and the degree to which processing is focused on aspects of the environment to which one is predisposed to focus attention.

11.7 Suggested Readings

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the *Locus coeruleus*. *Journal of Vision, 14*(4), 1–20. https://doi.org/10.1167/14.4.1

- *A study combining fMRI and pupillometry measurements of cognitive load in the same individuals. It revealed that individual differences in pupil responses to cognitive load were closely related to activity within the Locus coeruleus of the brain.*

Aston-Jones, G., & Cohen, J. D. (2005). Adaptive gain and the role of the *Locus coeruleus*-norepinephrine system in optimal performance. *Journal of Comparative Neurology, 493*, 99–110. https://doi.org/10.1002/cne.20723

- *A comprehensive review of the neurophysiological evidence of the role of the locus coeruleus in attention and learning and for "adaptive gain theory".*

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276–292.

- *A comprehensive review of the relation between pupillary responses and cognitive resources.*

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proc Natl Acad Sci U S A, 105*, 1704–1709. https://doi.org/10.1073/pnas.0707727105

- *One of the first studies that revealed how pupillary dilations can signal changes in conscious perception.*

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 140*, 1190–1192. https://doi.org/10.1126/science.143.3611.1190

- *One of the earliest and most influential studies on the relation between levels of cognitive load or task difficulty and the amplitudes of pupillary dilations.*

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron, 89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028

- *A neurophysiological study with monkeys presenting evidence that pupillary dilations are controlled by the locus coeruleus.*

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Strauss, Giroux.

- *A book on many aspects of decision making by a Nobel-winner scientist. In chapter 2, the author presents a synthesis of his model of attentional effort and the usefulness of the pupillary response as a measure of effort.*

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583–1585.

- *One of the earliest and most influential studies on the relation between levels of cognitive load or task difficulty and the amplitudes of pupillary dilations.*

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A Window to the Preconscious? *Perspectives on Psychological Science*, 7(1), 18–27. https://doi.org/10.1177/1745691611427305

- *A comprehensive and recent review of the relation between pupillary responses and cognitive and affective processes.*

Mathôt, S., & Van der Stigchel (2015). New light on the mind's eye: The pupillary light response as active vision. *Current Directions in Psychological Science*, 24(5), 374–378.

- *A recent review of the relation between pupillary responses and visual processing.*

11.8 Questions Students Should Be Able to Answer

What is the main function of the changes in the opening of the eye pupil?
What parts of the brain control adjustments of the pupil?
How would you set up with an eye-tracker a controlled experiment to measure, for example, people's responses to emotional stimuli?
If we observe that someone's pupil dilates, although nothing in the ambient illumination has changed, what could we conclude that such a person is doing?
Describe what happens to the pupil when someone imagines something.
What is the relationship between cognitive workload ("mental effort") and pupillary changes?
What is the effect of the activity of the norepinephrine system on the brain and on the eye pupil?

Bibliography

Aboyoun, D. C., & Dabbs, J. M. (1998). The Hess pupil dilation findings: Sex or novelty? *Social Behavior and Personality, 26,* 415–420.

Adams, D. B., Gold, A. R., & Burt, A. D. (1978). Rise in female-initiated sexual activity at ovulation and its suppression by oral contraceptives. *The New England Journal of Medicine, 21,* 1145–1150.

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science, 205,* 1289–1292. https://doi.org/10.1126/science.472746.

Alnæs, D., Sneve, M. H., Espeseth, T., Endestad, T., van de Pavert, S. H. P., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the *Locus coeruleus*. *Journal of Vision, 14*(4), 1–20. https://doi.org/10.1167/14.4.1.

Alvarez, G. A., & Cavanagh, P. (2005). Independent resources for attentional tracking in the left and right visual hemifields. *Psychological Science, 16,* 637–643. https://doi.org/10.1111/j.1467-9280.2005.01587.x.

Arch, D. C. (1979). Pupil dilation measures in consumer research: Applications and limitations. In William L. Wilkie (Ed.), *NA - Advances in consumer research* (Vol. 06, pp. 166–168). Ann Abor, MI : Association for Consumer Research.

Aston-Jones, G., & Cohen, J. D. (2005a). Adaptive gain and the role of the locus coeruleus-norepinephrine system in optimal performance. *Journal of Comparative Neurology, 493,* 99–110. https://doi.org/10.1002/cne.20723.

Aston-Jones, G., & Cohen, J. D. (2005b). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience, 28,* 403–450. https://doi.org/10.1146/annurev.neuro.28.061604.135709.

Aston-Jones, G., Rajkowski, J., & Cohen, J. (2000). *Locus coeruleus* and regulation of behavioral flexibility and attention. *Progress in Brain Research, 126,* 165–182. https://doi.org/10.1016/s0079-6123(00)26013-5.

Attard-Johnson, J., Bindemann, M., & Ó Ciardha, C. (2016). Heterosexual, Homosexual, and bisexual men's pupillary responses to persons at different stages of sexual development. The Journal of Sex Research, 00(00), 1–12.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91,* 276–292.

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5,* 371–372.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo, L. G. Tassinary, & G. G. Berntson (Eds.), (2nd ed., pp. 142–162). New York: Cambridge University Press.

Bernhardt, P. C., Dabbs, J. M., & Riad, J. K. (1996). Pupillometry system for use in social psychology. *Behavior Research Methods, Instruments & Computers, 28*(1), 61–66.

Berridge, C. W., & Waterhouse, B. D. (2003). The *Locus coeruleus*-noradrenergic system: Modulation of behavioral state and state-dependent cognitive processes. *Brain Research. Brain Research Reviews, 42,* 33–84.

Binda, P., Pereverzeva, M., & Murray, S. (2013). Pupil constrictions to photographs of the sun. *Journal of Vision, 13*(6), 1–9.

Blackwell, R., Miniard, P., & Engel, J. (2001). *Consumer Behavior* (9th ed.). Orlando: Harcourt.

Bombeke, K., Duthoo, W., Mueller, S. C., Hopf, J., & Boehler, C. N. (2016). Pupil size directly modulates the feedforward response in human primary visual cortex independently of attention. *NeuroImage, 127,* 67–73. https://doi.org/10.1016/j.neuroimage.2015.11.072.

Bouret, S., Duvel, A., Onat, S., & Sara, S. J. (2003). Phasic activation of locus ceruleus neurons by the central nucleus of the amygdala. *The Journal of Neuroscience, 23,* 3491–3497.

Bouret, S., & Sara, S. J. (2005). Network reset: A simplified overarching theory of locus coeruleus noradrenaline function. *Trends in Neurosciences, 28*(11), 574–582. https://doi.org/10.1016/j.tins.2005.09.002.

Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., & Corbetta, M. (2008). Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 28,* 10056–10061. https://doi.org/10.1523/jneurosci.1776-08.2008.

Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods, 45*, 1322–1331. https://doi.org/10.3758/s13428-013-0327-0.

Campbell, F. W., & Gregory, A. H. (1960). Effect of size of pupil on visual acuity. *Nature, 4743,* 1121–1123.

Carbon, C. C., Hutzler, F., & Minge, M. (2006). Innovativeness in design investigated by eye movements and pupillometry. *Psychology Science, 48*(2), 173–186.

Cavanagh, P., & Alvarez, G. A. (2005). Tracking multiple targets with multifocal attention. *Trends Cognitive Science, 9,* 349–354. https://doi.org/10.1016/j.tics.2005.05.009.

Corbetta, M., Patel, G., & Shulman, G. L. (2008). The reorienting system of the human brain: From environment to theory of mind. *Neuron, 58,* 306–324. https://doi.org/10.1016/j.neuron.2008.04.017.

Cornsweet, T. N. (1970). *Visual perception*. New York: Academic Press.

Damsma, A., & van Rijn, H. (2017). Pupillary response indexes the metrical hierarchy of unattended rhythmic violations. *Brain and Cognition, 111,* 95–103.

Denton, E. J. (1956). The responses of the pupil of Gekko gekko to external light stimulus. *J. Gen. Physiol., 40,* 201–2016.

Denton, E. J., & Pirenne, M. H. (1952). On the functional stability of the retina. *The Journal of General Physiology, 117,* 55P.

Drew, T., Horowitz, T. S., & Vogel, E. K. (2013). Swapping or dropping? Electrophysiological measures of difficulty during multiple object tracking. *Cognition, 126*, 213–223.

Einhäuser, W., Koch, C., & Carter, O. L. (2010). Pupil dilation betrays the timing of decisions. *Frontiers in Human Neuroscience, 4,* 18. https://doi.org/10.3389/fnhum.2010.00018.

Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proceedings of National Academy Science United States America, 105,* 1704–1709. https://doi.org/10.1073/pnas.0707727105.

Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience, 16*(8), 1146–1153. https://doi.org/10.1038/nn.3428.

Fan, J., Gu, X., Guise, K.G., Liu, X., Fossella, J., Wang, H., & Posner, M. I. (2009). Testing the behavioral interaction and integration of attentional networks. *Brain and Cognition, 70*(2), 209–220. https://doi.org/10.1016/j.bandc.2009.02.002.

Feinberg, R., & Podolak, E. (1965). Latency of pupillary reflex to light stimulation and its relationship to aging. In A. T. Welford (Ed.), *Behavior, Aging, and the Nervous System*. Springfield, IL: C.C. Thomas Publisher.

Foote, S. L., & Morrison, J. H. (1987). Extrathalamic modulation of cortical function. *Annual Review of Neuroscience, 10,* 67–95. https://doi.org/10.1146/annurev.ne.10.030187.000435.

Franconeri, S. L., Alvarez, G. A., & Cavanagh, P. (2013). Flexible cognitive resources: Competitive content maps for attention and memory. *Trends Cognitive Science, 17,* 134–141. https://doi.org/10.1016/j.tics.2013.01.010.

Gagl, B., Hawelka, S., & Huzler, F. (2011). Systematic influence of gaze position on pupil size measurement: Analysis and correction. *Behavioral Research, 43,* 1171–1181.

Gamlin, P. D. R., McDougal, D. H., Pokorny, J., Smith, V., Yau, K.-W., & Dacey, D. M. (2007). Human and macaque pupil responses driven by melanopsin-containing retinal ganglion cells. *Vision Research, 47,* 946–954.

Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive Affected Behaviour Neuroscience, 10,* 252–269. https://doi.org/10.3758/cabn.10.2.252.

Gingras, B., Marin, M. M., Puig-Waldmüller, E., & Fitch, W. T. (2015). The eye is listening: Music-induced arousal and individual differences predict pupillary responses. *Frontiers in Human Neuroscience, 9,* 619. https://doi.org/10.3389/fnhum.2015.00619.

Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin, 77,* 340–355.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology, 33*(4), 457–461. https://doi.org/10.1111/j.1469-8986.1996.tb01071.x.

Haro, J., Guasch, M., Vallès, B., et al. (2016). Is pupillary response a reliable index of word recognition? Evidence from a delayed lexical decision task. *Behavioral Research*. https://doi.org/10.3758/s13428-016-0835-9.

Hermans, E. J., van Marle, H. J. F., Ossewaarde, L., Henckens, M. J. a. G., Qin, S., van Kesteren, M. T. R., … Fernández, G. (2011). Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. *Science, 334*, 1151–1153. https://doi.org/10.1126/science.1209603.

Hess, E. H. (1965). *Attitude and pupil size. Scient. American, 212,* 46–54.

Hess, E. H. (1975). The role of pupil size in communication. *Scientific American, 233*(5), 110–112, 116–119.

Hess, E. H., & Polt, J. M. (1960). Pupil size as related to interest value of visual stimuli. *Science, 132,* 349–350.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 140,* 1190–1192. https://doi.org/10.1126/science.143.3611.1190.

Hou, R. H., Freeman, C., Langley, R. W., Szabadi, E., & Bradshaw, C. M. (2005). Does modafinil activate the locus coeruleus in man? Comparison of modafinil and clonidine on arousal and autonomic functions in human volunteers. *Psychopharmacology (Berl), 181*(3), 537–549. https://doi.org/10.1007/s00213-005-0013-8.

Jainta, S., & Baccino, T. (2010). Analyzing the pupil response due to increased cognitive demand: An independent component analysis study. *International Journal of Psychophysiology, 77,* 1–7.

Janisse, P. (1974). *Pupillary dynamics and behavior*. New York: Plenum.

Janisse, P. (1977). *Pupillometry: The psychology of the pupillary response*. Hemisphere Publishing, Wiley.

Joshi, S., Li, Y., Kalwani, R. M., & Gold, J. I. (2016). Relationships between pupil diameter and neuronal activity in the *Locus coeruleus*, Colliculi, and Cingulate Cortex. *Neuron, 89*(1), 221–234. https://doi.org/10.1016/j.neuron.2015.11.028.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99*(1), 122–149. https://doi.org/10.1037/0033-295x.99.1.122.

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology, 47,* 310–339.

Just, M. A., Carpenter, P. A., & Miyake, A. (2003). Neuroindices of cognitive workload: Neuroimaging, pupillometric and event-related potential studies of brain work. *Theoretical Issues in Ergonomics Science, 4,* 56–88. https://doi.org/10.1080/14639220210159735.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, New Jersey: Prentice Hall.

Kahneman, D. (2011). *Thinking, fast and slow*. New York: Farrar, Strauss, Giroux.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154,* 1583–1585.

Kahneman, D., & Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Perception and Psychophysics, 2,* 101–105.

Kahneman, D., Beatty, J., & Pollack, I. (1967). Perceptual deficit during a mental task. *Science, 157,* 218–219.

Kahneman, D., & Peavler, W. S. (1969). Incentive effects and pupillary changes in association learning. *Journal of Experimental Psychology, 79,* 312–318.

Kahneman, D., Tursky, B., Shapiro, D., & Crider, A. (1969). Pupillary, heart rate, and skin resistance changes during a mental task. *Journal of Experimental Psychology, 79*(1), 164–167.

Kawasaki, A., & Kardon, R. H. (2007). Intrinsically photosensitive retinal ganglion cells. *Journal of Neuro-Ophthalmology, 27*(3), 195–204.

King, A. S. (1972). Pupil size, eye direction, and message appeal: Some preliminary findings. *Journal of Marketing, 36*(3), 55–58.

Koss, M. C. (1986). Pupillary dilation as an index of central nervous system alpha 2-adrenoceptor activation. *Journal of Pharmacological Methods, 15,* 1–19.

Krugman, H. E. (1965). The impact of television advertising: Learning without involvement. *Public Opinion Quarterly, 29*(3), 349–356.

Laeng, B., & Endestad, T. (2013). Bright illusions reduce the eye's pupil. *PNAS, 109,* 2162–2167.

Laeng, B., Eidet, L. M., Sulutvedt, U., & Panksepp, J. (2016). Music chills: The eye pupil as a mirror to music's soul. *Consciousness and Cognition, 44*, s 161–178. https://doi.org/10.1016/j.concog.2016.07.009. ISSN 1053-8100.

Laeng, B., Sirois, S., & Gredeback, G. (2012). Pupillometry: A window to the preconscious? *Perspectives on Psychological Science, 7*(1), 18–27. https://doi.org/10.1177/1745691611427305.

Laeng, B., Suegami, T., & Aminihajibashi, S. (2016). Wine labels: An eye-tracking and pupillometry study. *International Journal of Wine Business Research, 28*(4), 327–348.

Laeng, B., & Sulutvedt, U. (2014). The eye pupil adjusts to imaginary light. *Psychological Science, 25*(1), 188–197. https://doi.org/10.1177/0956797613503556.

Laeng, B., Waterloo, K., Johnsen, S. H., Bakke, S. J., Låg, T., Simonsen, S. S., et al. (2007). The eyes remember it: Oculography and pupillometry during recollection in three amnesic patients. *Journal of Cognitive Neuroscience, 19*(11), 1888–1904. https://doi.org/10.1162/jocn.2007.19.11.1888.

Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary stroop effects. *Cognitive Processing, 12*(1), 13–21. https://doi.org/10.1007/s10339-010-0370-z.

Landgraf, S., Meer, E., & Krueger, F. (2010). Cognitive resource allocation for neural activity underlying mathematical cognition: A multi-method study. *ZDM Mathematics Education, 42*(6), 579–590. https://doi.org/10.1007/s11858-010-0264-7.

Lavie, P. (1979). Ultradian rhythms in alertness: A pupillometric study. *Biological Psychology, 9,* 49–62.

Lee, D. H., Susskind, J. M., & Anderson, A. K. (2013). Social transmission of the sensory benefits of eye widening in fear expressions. *Psychological Science, 24*(6), 957–965.

Libby, W. L., Lacey, B. C., & Lacey, J. I. (1973). Pupillary and cardiac activity during visual attention. *Psychophysiology, 10*(3), 270–294.

Loewenfeld. (1958). Mechanisms of reflex dilatation of the pupil. Historical review and experimental analysis. *Documenta Ophthalmologica, 12*, 185.

Loewenfeld, I. E. (1999). *The pupil: Anatomy, physiology, and clinical applications*. Boston (MA): Butterworth Heinemann.

Lowenstein, O., Feinberg, R., & Loewenfeld, I. E. (1963). Pupillary movements during acute and chronic fatigue. *Investigative Ophthalmology & Visual Science, 2,* 138–157.

Lowenstein, O., & Loewenfeld, I. E. (1958). Electronic pupillography. A new instrument and some clinical applications. *AMA Archives of Ophthalmology, 59*, 352.

MacLachlan, C., & Howland, H. C. (2002). Normal values and standard deviations for pupil diameter and interpupillary distance in subjects aged 1 month to 19 years. *Ophthalmic and Physiological Optics, 22,* 175–182.

Martinez-Conde, S. (2006). Fixational eye movements in normal and pathological vision. In S. Martinez-Conde, S. Macknik, L. M. Martinez, L. M. Alonso, & P. Tse (Eds.), *Progress in brain research* (Vol. 154). Elsevier B.V.

McConkie, G. W., Kerr, P. W., Reddix, M. D., & Zola, D. (1988). Eye movement control during reading: I. The location of initial eye fixations on words. *Vision Research, 28,* 1107–1118.

McDougal, D. H. M., & Gamlin, P. D. R. (2008). Pupillary control pathways. In R. H. Masland, & T. Albright (Eds.), *The senses: A comprehensive reference* (pp. 521–536). San Diego, California: Academic Press.

Moulden, B., Kingdom, F., & Gatley, L. F. (1990). The standard deviation of luminance as a metric for contrast in random-dot images. *Perception, 19*(1), 79–101.

Mouton, P. R., Pakkenberg, B., Gundersen, H. J., & Price, D. L. (1994). Absolute number and size of pigmented locus coeruleus neurons in young and aged individuals. *Journal of Chemical Neuroanatomy, 7,* 185–190.

Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping*, n/a–n/a. https://doi.org/10.1002/hbm.22466.

Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology, 48,* 1532–1543. https://doi.org/10.1111/j.1469-8986.2011.01226.x.

Murray, R. B., Adler, M. W., & Korczyn, A. D. (1983). The pupillary effects of opioids. *Life Sciences, 33,* 495–509.

Navon, D. (1984). Resources—A theoretical soup stone? *Psychological Review, 91*(2), 216–234.

Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin, 131,* 510–532. https://doi.org/10.1037/0033-2909.131.4.510.

Owen, A. M., & Coelman, M. R. (2006). Functional neuroimaging of the vegetative state. *Nature Reviews Neuroscience, 9,* 235–243. https://doi.org/10.1038/nrn2330.

Papesh, M. H., & Goldinger, S. D. (2012). Pupil-BLAH-metry: Cognitive effort in speech planning reflected by pupil dilation. *Attention, Perception, & Psychophysics, 74*(4), 754–765. https://doi.org/10.3758/s13414-011-0263-y.

Payne, D. T., Parry, M. E., & Harasymiw, S. J. (1967). A comparison of percentage of pupillary dilation with other measures of difficulty of mental multiplication items. *Research Bulletin* (67-42). Princeton, NJ: Educational Testing Service. http://onlinelibrary.wiley.com/doi/10.1002/j.2333-8504.1967.tb00551.x/pdf.

Payne, D. T., Parry, M. E., & Harasymiw, S. J. (1968). Percentage pupillary dilation as a measure of item difficulty. *Perception and Psychophysics, 4,* 139–143.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology, 11,* 559–566. https://doi.org/10.1111/j.1469-8986.1974.tb01114.x.

Phillips, M. a., Szabadi, E., & Bradshaw, C. M. (2000). Comparison of the effects of clonidine and yohimbine on spontaneous pupillary fluctuations in healthy human volunteers. *Psychopharmacology, 150*, 85–89.

Poock, G. K. (1973). Information processing vs pupil diameter. *Perceptual and Motor Skills, 37*(3), 1000–1002.

Porter, G., & Troscianko, T. (2003). Pupillary response to grating stimuli. *Perception, 32,* 156.

Porter, G., Troscianko, T., & Gilchrist, I. D. (2007). Effort during visual search and counting: Insights from pupillometry. *Quarterly Journal of Experimental Psychology, 60*(2), 211–229. https://doi.org/10.1080/17470210600673818.

Raizada, R. D. S., & Poldrack, R. a. (2008). Challenge-driven attention: Interacting frontal and brainstem systems. *Frontiers in Human Neuroscience, 1*, 3. https://doi.org/10.3389/neuro.09.003.2007.

Rajkowski, K., Kubiak, P., & Aston-Jones, G. (1993). Correlations between *Locus coeruleus* (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society of Neuroscience Abstracts*, 974.

Richman, J. E., Golden McAndrew, K., Decker., & Mullaney, S. (2004). An evaluation of pupil size standards used by police officers for detecting drug impairment. *Optometry, 75*, 2–8.

Samuels, E. R., & Szabadi, E. (2008). Functional neuroanatomy of the noradrenergic *Locus coeruleus*: Its roles in the regulation of arousal and autonomic function. *Current Neuropharmacology, 6,* 235–253.

Sara, S. J. (2009). The locus coeruleus and noradrenergic modulation of cognition. *Nature Reviews Neuroscience, 10,* 211–223. https://doi.org/10.1038/nrn2573.

Sara, S. J., & Bouret, S. (2012). Orienting and reorienting: The *Locus coeruleus* mediates cognition through arousal. *Neuron, 76,* 130–141. https://doi.org/10.1016/j.neuron.2012.09.011.

Sarter, M., Gehring, W. J., & Kozak, R. (2006). More attention must be paid: The neurobiology of attentional effort. *Brain Research Reviews, 51,* 145–160. https://doi.org/10.1016/j.brainresrev.2005.11.002.

Scholl, B. J. (2009). What have we learned about attention from multiple-object tracking (and vice versa)? In D. Dedrick & L. Trick (Eds.), *Computation, cognition, and Pylyshyn* (pp. 49–77). Cambridge, MA: MIT Press.

Schwalm, M., Keinath, A., & Zimmer, H. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. In D. de Waard, F. O. Flemisch, B. Lorenz, H. Oberheid, & K. A. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 1–13). Maastricht, The Netherlands: Shaker Publishing.

Siegle, G. J., Steinhauer, S. R., Stenger, V. A., Konecky, R., & Carter, C. S. (2003). Use of concurrent pupil dilation assessment to inform interpretation and analysis of fMRI data. *Neuroimage, 20,* 114–124. https://doi.org/10.1016/s1053-8119(03)00298-2.

Silverman, J., Mooney, J. M., & Shepherd, F. D. (1992, March). Infrared video cameras. *Scientific American*, 78–83.

Simms, T. M. (1967). Pupillary response of male and female subjects to pupillary difference in male and female picture stimuli. *Perception & Psychophysics, 2*(11), 533–555.

Steinhauer, S. R., Siegle, G. J., Condray, R., & Pless, M. (2004). Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. *International Journal of Psychophysiology, 52*, 77–86.

Susskind, J. M., Lee, D. H., Cusi, A., Feiman, R., Grabski, W., & Anderson, A. K. (2008). Expressing fear enhances sensory acquisition. *Nature Neuroscience, 11,* 843–850. https://doi.org/10.1038/nn.2138.

Suzuki, T. W., Kunimatsu, J., & Tanaka, M. (2016). Correlation between pupil size and subjective passage of time in non-human primates. *The Journal of Neuroscience, 36*(44), 11331–11337.

Svensson, T. H. (1987). Peripheral, autonomic regulation of locus coeruleus noradrenergic neurons in brain: Putative implications for psychiatry and psychopharmacology. *Psychopharmacology, 92,* 1–7.

Szabadi, E. (2012). Modulation of physiological reflexes by pain: role of the locus coeruleus. *Frontiers in Integrative Neuroscience, 6,* 94. https://doi.org/10.3389/fnint.2012.00094.

Tsukahara, J. S., Harrison, T. L., & Engle, R. W. (2016). The relationship between baseline pupil size and intelligence. *Cognitive Psychology, 91*, 109–123.

Ueda, T., Nawa, Y., Yukawa, E., Taketani, F., & Hara, Y. (2006). Change in dynamic visual acuity (DVA) by pupil dilation. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 48,* 651–655.

Unsworth, N., & Robison, M. K. (2017). A *Locus coeruleus*-norepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin and Review*.

van der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., … Wartenburger, I. (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology, 47*, 158–169. https://doi.org/10.1111/j.1469-8986.2009.00884.x.

Wang, C.-A., & Munoz, D. P. (2014). Modulation of stimulus contrast on the human pupil orienting response. *European Journal of Neuroscience, 40,* 2822–2832.

Warga, M., Lüdtke, H., Wilhelm, H., & Wilhelm, B. J. (2009). How do spontaneous pupillary oscillations in light relate to light intensity. *Vision Research, 49,* 295–300.

Watson, A. B., & Yellott, J. I. (2012). A unified formula for light-adapted pupil size. *Journal of Vision, 12*(10), 1–16.

Whipple, B., Ogden, G., & Komisaruk, B. R. (1992). Physiological Correlates of Imagery: Induced Orgasm in Women. *Archives of Sexual Behavior, 21*(2), 121–133.

Wilhelm, B., Wilhelm, H., & Lüdtke, H. (1999). Pupillography: Principles and applications in basic and clinical research. In J. Kuhlmann & M. Böttcher (Eds.), (pp. 1–10). München, Germany: Zuckscjwerdt Verlag.

Wilhelm, B. J., Wilhelm, H., Moro, S., & Barbur, J. L. (2002). Pupil response components: Studies in patients with Parinaud's syndrome. *Brain, 125,* 2296–2307.

Winn, B., Whitaker, D., Elliott, D. B., & Phillips, N. J. (1994). Factors affecting light-adapted pupil size in normal human subjects. *Investigative Ophthalmology & Visual Science, 35*(3), 1132–1137.

Wu, E. X. W., Laeng, B., & Magnussen, S. (2011). Through the eyes of the own-race bias: Eye-tracking and pupillometry during face recognition. *Social Neuroscience*. https://doi.org/10.1080/17470919.2011.596946.

Zekveld, A. a., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., & Kramer, S. E. (2014). The eye as a window to the listening brain: Neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage*. https://doi.org/10.1016/j.neuroimage.2014.06.069.

Zekveld, A. a., Kramer, S. E., & Festen, J. M. (2010). Pupil response as an indication of effortful listening: The influence of sentence intelligibility. *Ear and hearing, 31*(4), 480–490. https://doi.org/10.1097/aud.0b013e3181d4f251.

Zénon, A. (2017). Time-domain analysis for extracting fast-paced pupil responses. *Scientific Reports, 7*, 41484. https://doi.org/10.1038/srep41484.

Zhu, Y., Iba, M., Rajkowski, J., & Aston-Jones, G. (2004). Projection from the orbitofrontal cortex to the locus coeruleus in monkeys revealed by anterograde tracing. *Society of Neuroscience Abstracts, 30,* 211–213.

Zylberberg, A., Oliva, M., & Sigman, M., (2012). Pupil dilation: A fingerprint of temporal selection during the Attentional Blink. *Frontiers in Psychology*, *3*(316). https://doi.org/10.3389/fpsyg.2012.00316.

Chapter 12
Functional Magnetic Resonance Imaging of Eye Movements: Introduction to Methods and Basic Phenomena

Sharna D. Jamadar and Beth Johnson

Contents

S. D. Jamadar (✉) · B. Johnson
Monash Institute for Cognitive and Clinical Neuroscience, School of Psychological Sciences, Monash University, Melbourne, Australia
e-mail: sharna.jamadar@monash.edu

S. D. Jamadar
Monash Biomedical Imaging, Monash University, Melbourne, Australia

Australian Research Council Centre of Excellence for Integrative Brain Function, Melbourne, Australia

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_12

Abstract The advent of functional magnetic resonance imaging (fMRI) in the early 1990s led to a rapid increase in the study of the neural bases of cognition. fMRI has made it possible to non-invasively study the spatial distribution of the neural processing of eye movements in humans, on a scale that was only previously achieved using invasive methods in animals and non-human primates. With increasing accessibility and affordability of fMRI, the field of functional neuroimaging has grown in usage, sophistication, impact and range of uses (Bandettini, 2012). Here, we present a didactic introduction to the fMRI method, with a specific focus on eye movement research. We introduce the principles of magnetic resonance signal generation and the physiological basis of the fMRI signal; how to set up an fMRI experiment for eye movement research, and fundamental principles of fMRI data preparation and analysis. We then discuss the basic phenomena of the neural bases of eye movements that have been studied using fMRI.

12.1 Introduction and Learning Objectives

The study of eye movements occupies an important and unique place in experimental psychology and cognitive neuroscience. Eye movements have a long history of intense research in humans, non-human primates and other animal models. As a result of these studies, the ocular motor system is one of the best-mapped systems in the brain (see Pierrot-Deseilligny, Milea, & Müri, 2004; McDowell, Dyckman, Austin, & Clementz, 2008; Enderle, 2002 for reviews). Studying the convergence of the ocular motor system with another well-mapped neural system, vision, has provided many important insights into the way the brain performs and integrates sensory perception and action planning with other higher-order cognitive processes such as cognitive control, attention, and working memory. Thus, the study of the neural bases of eye movements occupies a somewhat unique place in experimental psychology: unlike many other neural systems, we have a high degree of understanding into which parts of the brain contribute to multiple aspects of ocular motility.

In this Chapter, we will present a general introduction to the magnetic resonance imaging (MRI) and functional MRI (fMRI) techniques, with a specific focus on challenges and issues in eye movement research. The learning objectives for this Chapter are to:

(a) understand the fundamentals of MR signal generation and the physiological basis of fMRI
(b) learn how to set up an fMRI experiment in the field of eye movement research, including setting up lab equipment, how to address technological challenges in image acquisition, and how to design an eye movement study
(c) understand the role of fMRI data preparation (preprocessing) approaches, single-subject and group-level analyses

(d) gain an overview of how fMRI has shed light on the basic phenomena of the neural bases of eye movements in the human brain.

In this chapter, we will often refer to the antisaccade paradigm as an example to illustrate important points. The antisaccade paradigm is explored in depth in the chapter by Pierce et al. (this volume), and compares *antisaccade* trials to *prosaccade* trials. In a prosaccade task, participants are instructed to make a saccade toward a target, whereas the antisaccade task requires participants to inhibit the prepotent response to look at the target, and instead to make a volitional saccade to the mirror-opposite target location (see chapter by Pierce et al. in this volume). The antisaccade paradigm is often used in fMRI studies to explore response inhibition, attention, working memory, and decision-making.

12.2 Historical Annotations

Historically, much of what we know about the ocular motor system in the brain has come from invasive electrophysiological and lesion studies of animal behavior, most notably non-human primates, owing to the ocular motor network of non-human primates being highly analogous to humans (Hikosaka, Takikawa, & Kawagoe, 2000; Johnston & Everling, 2008; Voogd & Barmack, 2006). The swathe of electrophysiological and lesion studies in non-human primates throughout the 1980s and ’90s not only mapped the highly distributed network of brain regions associated with ocular motor control, and the connections between them, but also informed our understanding the pharmacological properties of neurons in these regions, and their precise timing and activity. For example, the seminal work of Okihide Hikosaka in non-human primates was central to our current understanding of the role of the basal ganglia in volitional saccades, such as inhibition of unwanted saccades and making saccades to remembered locations (e.g., Hikosaka et al., 2000). Microinjections of pharmacological agonists and antagonists to regions within the basal ganglia were used alongside electrophysiological recordings to explore the relationship between different regions within the basal ganglia and their effects on saccadic gating, saccade inhibition and spatial working memory. The ability to lesion (both pharmacologically and permanently) and directly record neuronal activity within subregions of the brainstem and cerebellum in non-human primates, much of which was pioneered by Charles A. Scudder, Albert F. Fuchs and David A. Robinson (e.g., Becker & Fuchs, 1985; Fuchs & Robinson, 1966; Fuchs, 1967; Lisberger, Evinger, Johanson, & Fuchs, 1981; Lund, Lund, Hendrickson, Bunt, & Fuchs, 1975; Robinson, 1970, 1995; Robinson & Fuchs, 2001; Robinson, Fuchs, & Noto, 2002; Scudder, 1988; Scudder & Fuchs, 1992; Scudder, Kaneko, & Fuchs, 2002; Scudder et al., 2002), gave rise to our current models of how oculomotor coordinates are encoded (*where* an eye movement is to be made to), the release of eye movements (*when* the eye movement is made), and velocity characteristics (*how* eye movements are made), takes place. The rapid and precise timing of ocular motor signals, and the position of key ocular motor brain regions being in the midbrain, brainstem and cerebellum, mean that these highly invasive methods undertaken in animal models have been critical to mapping the ocular motor network.

In humans, many studies have used behavioural measures like reaction time, error rates and ocular motor characteristics, such as saccade accuracy and velocity, to understand ocular motor control. However these measures represent the cumulative endpoint of multiple sensory, cognitive and motor processes that contribute to an overt response, and cannot provide direct evidence of the *neural* bases of oculomotor control. Furthermore, these measures cannot directly measure covert processes in the absence of a response (e.g. no-go or stop-signal trials), nor the covert processes leading up to a response including anticipatory, expectancy or interference processes (Karayanidis & Jamadar, 2014). So, over the past two decades, researchers have increasingly used a range of functional neuroimaging tools to bridge the gap between animal and human models and to non-invasively characterise the ocular motor system of humans. The most common functional imaging tools include electroencephalography/event-related potentials (EEG/ERPs) and functional magnetic resonance imaging (fMRI).[1]

The EEG is a continuous measure of the spontaneous electrical activity of the brain. EEG, sometimes colloquially referred to as 'brain waves', represents the summation of synchronous activity of thousands or millions of neurons that are radially or tangentially oriented towards the scalp. ERPs are extracted from the EEG and are variations in electrical activity time-locked to an event (Rugg & Coles, 1994). The temporal resolution of EEG/ERPs is in the order of milliseconds, and is unmatched by other non-invasive human functional imaging methods. Thus, EEG and ERPs are particularly useful for studying the electrophysiological and temporal dynamics of ocular motility (see e.g. Greenlee, 2000; Jagla, Jergelova, & Riecanský, 2007 for reviews).

By comparison, fMRI has superior spatial resolution compared to EEG/ERPs, and has largely been used to non-invasively map the neural regions involved in cognition in the intact human brain. fMRI has been widely used to map the neural ocular motor network in the human brain, and to study changes in the relative contribution of neural regions to ocular motor behaviour with experimental manipulation.

12.3 General Introduction to fMRI Methods

We present a brief introduction to magnetic resonance (MR) signal generation, image formation and blood oxygenation level-dependent (BOLD) fMRI that will give the reader a sufficient foundation to understand the use of the technology in ocular motor

---

[1] A small number of studies have used positron emission tomography (PET), which is another neuroimaging technique, to map the ocular motor system in the brain (e.g. Paus et al., 1993 or Sweeney et al., 1996). While PET possesses a number of unique advantages over fMRI such as relative insensitivity to participant motion, it is a more invasive technique with reduced temporal resolution. It has therefore largely been superseded by fMRI in recent years.

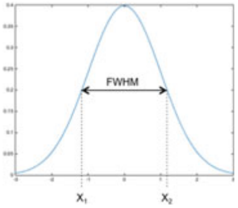
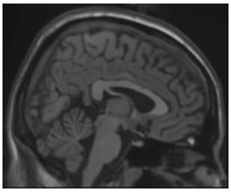
research. Table 12.1 summarises important parameters in fMRI acquisition. For a more complete introduction to the topic, the reader is referred to the comprehensive introduction presented by Huettel, Song, and McCarthy (2008).

12.3.1 A Brief Introduction to MR Signal Generation

A magnetic resonance image (MRI) is a three dimensional spatial matrix of numbers that corresponds to the spatial distribution of some quantity (e.g. T1 relaxation time, haemodynamic response). Figure 12.1 summarises the MR signal generation process. Briefly, in the form of MRI typically used in experimental psychology, the signal that is usually measured is a change in the hydrogen atoms (i.e. protons) in the brain tissue. Each proton *spins* around an axis and when placed in a magnetic field *precesses* (i.e. moves in a gyroscopic fashion) around an axis parallel to the magnetic field (Fig. 12.1a). Under normal conditions, the spins are aligned randomly and there is little to no net magnetisation M. When body tissue is brought into the strong external magnetic field in the MR scanner bore, the spins align with the external magnetic field B either parallel or antiparallel to the magnetic field (Fig. 12.1b). The parallel state is a slightly more stable, lower energy state, whereas the antiparallel state is a slightly unstable, higher energy level. At equilibrium, a small majority of the spins align parallel to the magnetic field, resulting in a small net magnetisation M parallel to B. In order to detect an MR signal, it is necessary to perturb the equilibrium state of the spins (Fig. 12.1c). Radiofrequency (RF) coils in the MR scanner bombard the spins with photons, and some spins in the low-energy parallel state absorb this energy and shift to the less stable high-energy state in a process known as *excitation*. After the energy source is removed, some spins return to the low-energy state, emitting electromagnetic energy that is detected by the RF coil; measurement of this emitted energy provides the MR signal depicted in MR images.

The MR signal following an excitation pulse decays over time in a process known as *relaxation*. Figure 12.2 summarises the MR relaxation process. The relaxation time is determined by two primary mechanisms: *longitudinal relaxation*, and *transverse relaxation*. Before excitation, the net magnetisation M is aligned with the z-axis in the longitudinal plane (Fig. 12.2a). When the spin system is then excited by the RF pulse the net magnetisation flips 90° into the x-y transverse plane (Fig. 12.2b). The time taken for the net magnetisation axis to recover parallel to the external magnetic field B (Fig. 12.2c) is determined by time T_1. When the net magnetisation is tipped into the x-y plane, the spins are initially coherent; that is, they begin precessing around the main magnetic field vector z at the same phase. Over time, the coherence between the spins is lost and they begin to precess out of phase; this is transverse relaxation (Fig. 12.2d). There are two main causes of transverse relaxation: an intrinsic and extrinsic cause. The intrinsic cause is due to *spin-spin interaction*: when many spins are excited at once, interactions between spins causes some spins to precess faster and others to precess slower, ultimately resulting in the dephasing of the spins. T_2 is the time it takes for the decay of the transverse component of the net magnetisation

Table 12.1 Important parameters in MRI

Parameter	Explanation	Example
T1	The time taken for net magnetisation to recover into the longitudinal plane (Fig. 12.2). A T1-weighted scan is used as a ‘structural’ or ‘anatomical’ image in fMRI, as it provides good contrast between grey and white matter	
T2	The time taken for relaxation or loss of spin coherence in the transverse plane due to spin-spin interaction (Fig. 12.2). T2—weighted images are provide good contrast in fluid-filled regions such as CSF and lesions	
T3	The time taken for relaxation or loss of spin coherence in the transverse plane due to spin-spin interaction and field inhomogeneities (Fig. 12.2). T2* images are sensitive to the amount of deoxygenated haemoglobin in a sample and form the basis of fMRI	
TR	Repetition time. The time between excitation pulses; effectively the time to acquire an image	
TE	Echo time. The time between an excitation pulse and the peak of the signal	
FWHM	Full width at half maximum. Describes the width of a curve (e.g. Gaussian curve) at the two points (x1, x2) that are equal to half of the maximum height of the curve	
Stereotaxic space	Standardised space for images that allows coordinate systems to be used to define positions in an image. Talairach space (Talairach & Szikla, 1967; Talairach & Tournoux, 1988) was traditionally used in brain imaging and was based on the postmortem of a single human brain. Montreal Neurological Institute (MNI) space is more commonly used today and is based on a large number of brain images (e.g., 305 individuals, MNI305). One image, the ‘colin27’ image is commonly used to visualise fMRI results, and is normalized to the MNI305 template	

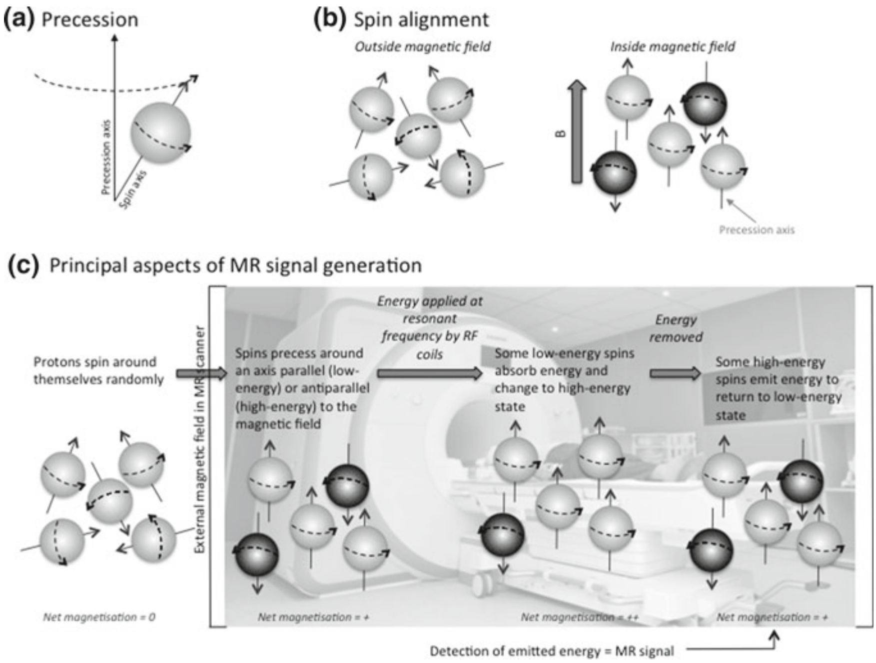


Fig. 12.1 Magnetic resonance (MR) signal generation. **a** When placed in an external magnetic field, protons change their orientation, initiating a gyroscopic motion that aligns with the magnetic field, known as precession. **b** The precession axis aligns with the magnetic field B. **c** Summary of the cascade of events necessary to detect an MR signal. Abbreviation: RF, radiofrequency. MRI scanner image (Siemens Skyra) courtesy of Monash Biomedical Imaging

due to spin-spin interactions. The extrinsic cause of transverse relaxation is inhomogeneities in the external magnetic field. Since the precession frequency of each spin is proportional to the local field strength of the magnetic field, local variations of the field strength cause spins to precess at different frequencies, speeding up the loss of spin coherence in the transverse plane. The combined effects of spin-spin interaction and field inhomogeneities result in faster signal decay, given by T_2*.

Different tissues have different T_1, T_2 and T_2* properties (Table 12.1). For example, at 3T the human brain shows a range of T1 values, from approximately 850 ms for white matter, 1300 ms for grey matter and 4500 ms for cerebrospinal fluid (CSF; Wansapura, Holland, Dun, & Ball, 1999), as such T_1-weighted images are typically used for visualising and quantifying brain morphology (grey/white matter volumetry, thickness, curvature etc.) In contrast, the range of T_2 values is much narrower for white and grey matter, around 80 and 110 ms respectively, providing little contrast between these tissues (Wansapura et al., 1999). T_2 scans are often used to quantify lesions, as the T_2 value for CSF is much longer (approximately 2000 ms), providing stronger contrast for CSF and fluid-filled regions such as lesions. T_2* scans are

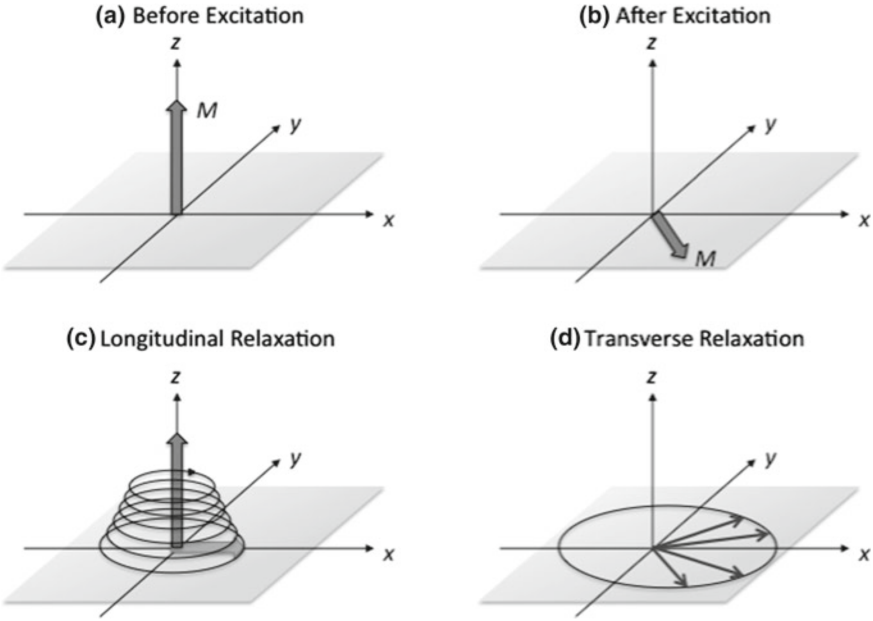


Fig. 12.2 Magnetic Resonance (MR) relaxation. **a** Before excitation, the net magnetisation M is aligned with the main magnetic field vector z. **b** Excitation tips the net magnetisation vector into the x-y plane. **c** Longitudinal relaxation occurs when the spins return to the parallel state, causing the net magnetisation vector to return to the longitudinal plane. **d** Transverse relaxation occurs when the spins dephase in the x-y plane

particularly relevant to experimental psychology, as it is sensitive to the amount of deoxygenated haemoglobin present in the sample, and serves as the primary signal in fMRI.

12.3.2 BOLD fMRI

The most common form of fMRI in use today is BOLD fMRI, which is used to infer neuronal activity *indirectly* from changes in blood oxygenation (Ogawa, Lee, Kay, & Tank, 1990). Eye movement studies usually examine BOLD fMRI responses to stimuli or responses (eye movements), which are often referred to as 'events', in the literature. The BOLD fMRI signal arises from the differences in the magnetic properties of oxygenated and deoxygenated haemoglobin in the blood. Hb is diamagnetic, with no unpaired electrons and zero magnetic moment, and so is weakly repelled from a magnetic field. dHb is paramagnetic, with unpaired electrons and greater than zero magnetic moment and so is attracted to a magnetic field. Thus, dHb but not Hb affects spin dephasing in the transverse plane and so can be imaged using T_2*-weighted scans.

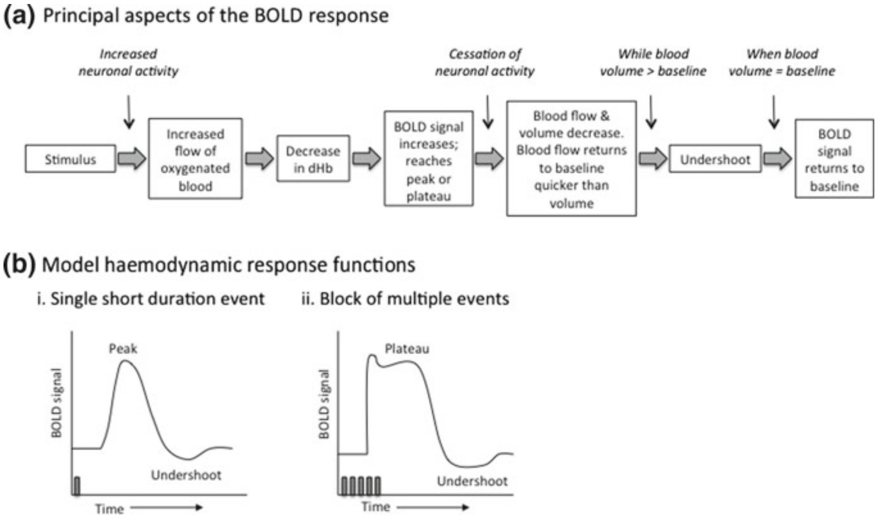


Fig. 12.3 **a** Cascade of events that leads to the BOLD response. **b** Model haemodynamic response functions for (i) a single short duration event and (ii) a block of multiple events. Vertical bars symbolise stimuli/events

Figure 12.3 summarises the cascade of events that leads to the BOLD response [as summarised by the balloon model; for discussion of alternate approaches to modeling the BOLD response, see Buxton (2012)]. Following the presentation of a stimulus, the metabolic demands of increased neuronal activity results in an increase in the flow of oxygenated blood, and hence the displacement of dHB in the sample. The resulting increase in the BOLD signal reaches its peak approximately 5 s after a single short-duration event, or reaches a plateau in case of multiple presentations of the same stimulus. Following cessation of the event, blood flow decreases more rapidly than blood volume, resulting in a relative increase in the amount of dHb compared to Hb in the sample, and the BOLD signal decreases below baseline (‘undershoot’). The BOLD signal returns to baseline as the blood volume of the tissue returns to normal.

12.3.3 fMRI in Experimental Psychology

MRI and fMRI possesses a number of unique advantages and disadvantages in the study of cognition and motor behaviour as compared to other techniques such as behavioural ocular motor recordings and EEG/ERPs. MRI possesses better spatial resolution, and is arguably the best way to visualise the human brain non-invasively. At high field strength, microstructures of the brain are detectable, with reports of fMRI mapping at the resolution of the microvasculature (e.g. Duong et al., 2002) and orientation columns (Yacoub, Harel, & Ugurbil, 2008). In practice, current fMRI resolution typically used in experimental psychology is 1 × 1 × 1 mm or 3 × 3 × 3 mm voxels. This spatial resolution is far superior to scalp-recorded ERPs, which largely represent the summated postsynaptic activity of multiple populations of pyramidal cells in distributed regions of the brain. These are difficult to localise (due to the ill-posed nature of the problem, Helmholtz, 1853; for an introduction see Luck, 2005), however there have been recent advances in this field (e.g. BESA, Scherg et al., 1989; LORETA, Pascual-Marqui, Michel, & Lehmann, 1994; among others).

By comparison, given that the BOLD response does not peak until approximately 5 s after the stimulus, the temporal resolution of fMRI is inferior to that of ERPs. The latter has a temporal resolution in the order of milliseconds, although advances in experimental design and MR sequences can substantially improve fMRI temporal resolution (see Feinberg & Setsompop, 2013, for a review). In addition, fMRI represents a much more indirect measure of neural activity that relies on neurometabolic and neurovascular coupling (Sect. 1.2), the basis of which is still under active investigation (Logothetis, 2007). By comparison, ocular motor measures such as eye tracking directly measure the behaviour of interest, and EEG measures the summated activity of populations of pyramidal cells.

12.3.4 Image Processing: Preparation for Analysis

The analysis of fMRI data is complex and relies upon techniques from signal processing and advanced statistics. Several steps are necessary to prepare the data for statistical modeling and inference. This “preprocessing” procedure has largely been standardised in recent years, although ongoing work continues to improve and refine this process.

Figure 12.4a summarises a standard preprocessing procedure (Poldrack, Mumford, & Nichols, 2011). Note that the order of preprocessing steps varies between studies, and most steps can be considered optional (particularly slice time correction) with one possible exception: motion correction. Several software packages have been developed to preprocess and analyse fMRI data, e.g. Statistical Parametric Mapping (SPM), Functional MRI of the brain Software Library (FSL), Analysis of Functional NeuroImages (AFNI), and BrainVoyager,[2] each of which emphasise different methods and approaches at each stage. At all stages, visual inspection of the data is imperative to ensure quality control.

---

[2] SPM: www.fil.ion.ucl.ac.uk/spm.
FSL: fsl.fmrib.ox.ac.uk/fsl/fslwiki.
AFNI: afni.nimh.nih.gov/afni.
BrainVoyager: www.brainvoyager.com.

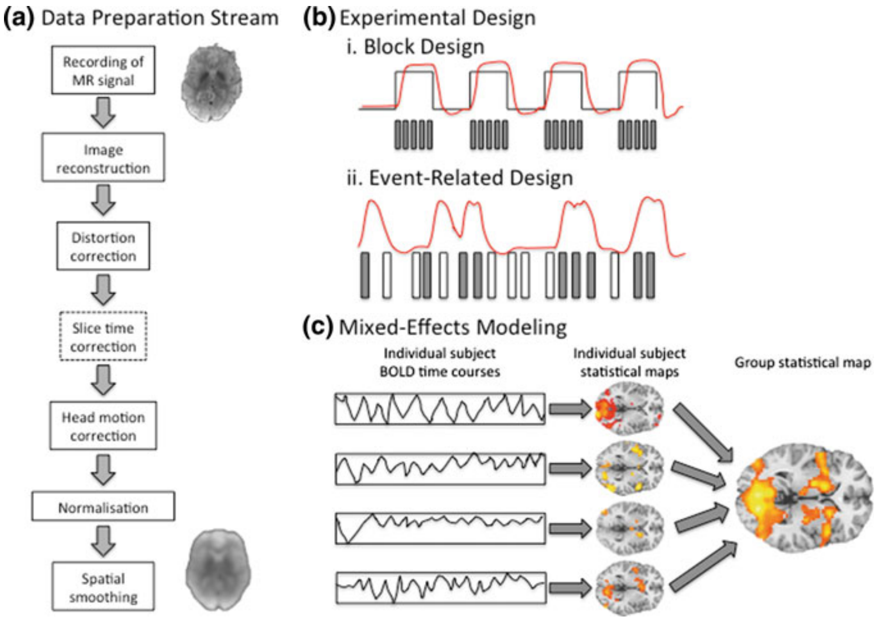


Fig. 12.4 **a** Example data preparation stream. Most stages can be considered optimal, and variation exists in the order that each stage is conducted. Slice time correction in particular may or may not be conducted. Images on the right show an example image prior to preprocessing (top) and after preprocessing (bottom). **b** fMRI experimental design. (i) Block designs present multiple trials (represented by grey bars) within a block alternating with either null trials or control condition in an ‘on/off design. A model haemodynamic response function is shown in red. Note that the response to each block saturates so that responses to individual events within a block cannot be determined. (ii) Event-related designs randomise presentation of different trial types (represented by grey and white bars) and jitter intertrial intervals to estimate the haemodynamic response for each individual trial. A stylised estimated haemodynamic response function is shown in red, note that when trials are presented closely together with little intertrial jittering, the response saturates and becomes similar to that seen in the block design. **c** In mixed-effects (or random-effects) modeling, first-level analyses involves the construction of a general linear model (GLM) for each subject and calculating single-subject statistical maps. These individual statistical maps are entered into a second- level analysis to test for significance at the group level

12.3.4.1 Distortion Correction

Following image reconstruction, which is largely completed by the image acquisition software, correction of image distortion is performed to account for artefacts related to the fast acquisition of T_2* fMR images. These artefacts include signal dropout and distortion near air-tissue interfaces such as the sinuses and ear canals (see for example the anterior portion of the frontal cortex in the unprocessed image in Fig. 12.4). *Distortion correction* uses a field map, which characterises the main magnetic field *B* in an attempt to correct for the effects of artefactual field inhomogeneities in these regions (Jezzard & Balaban, 1995).

12.3.4.2 Slice Time Correction

During image acquisition, full brain volumes are acquired in *slices*, in ascending (1, 2, 3, …), descending (46, 45, 44, …) or interleaved (1, 3, 5, … 2, 4, 6) sequences. Thus, the last slice acquired in the volume is taken some time after the first slice (determined by the repetition time, *TR*, of the acquisition sequence). As most analysis procedures assume that the full brain volume has been acquired at the same point in time, slice time correction uses temporal interpolation to correct for the difference in acquisition time between slices in a volume (Henson, Buchel, Josephs, & Friston, 1999). Recently, there has been a shift away from the use of slice time correction, as it has been determined that inferential statistics are relatively robust to slice time differences, and the procedure tends to introduce or exacerbate image artefacts (Poldrack et al., 2011).

12.3.4.3 Head Motion Correction

Participant motion represents a significant issue for fMRI data, and is always an issue in living, breathing subjects. Head motion presents at least four problems for fMRI:

- First, head motion results in a shift in the sampling target of each voxel in the image; that is, a voxel that indexes a location in one image will index a different location in subsequent images. The effects of this on the data can be subtle (i.e. from indexing one region in the visual cortex to another) to severe (e.g. from indexing a region in the visual cortex to a region in the cerebellum).
- The second problem introduced by participant motion is that movements that occur during the acquisition of a volume distorts the MR signal itself, distorting the image intensities between slices and creating a striping effect across the image volume. This form of motion artefact is much more difficult to correct and is not typically possible using standard software, although it can potentially be corrected using sophisticated algorithms (e.g. Friston, Williams, Howard, Frackowiak, & Turner, 1996b).
- A third problem is stimulus- or event-related motion, where head motion is correlated with an event of interest (Hajnal et al., 1994). This type of motion can occur if participants are required to make a verbal response (moving the jaw), engage in self-talk during the task (e.g. verbalising task rules), or make a body movement simultaneous with stimulus or response onset (e.g. startle response). It is also possible to obtain modest event-correlated motion by chance. Event-correlated motion results in larger signal changes in voxels at the boundaries of tissue types, leading to the erroneous inference that these changes have occurred in response to the event (false positives). In addition, motion correction routines may erroneously remove event-related signal variance, leading to a loss of statistical power to detect true event-related activity (false negatives). The best way to deal with this type of motion artefact is to use preventative measures including

experimental designs that do not encourage event-related motion, and instructing participants to minimise verbalisation during the task.

- The fourth problem that head motion can cause for fMRI data occurs in studies of non-healthy populations. Many patient populations will show increased head motion in the fMRI scanner compared to young healthy adults, which can lead to false positives in between-groups comparisons (Bullmore et al., 1999; Seto et al., 2001).

Realignment algorithms typically use a rigid body transformation with 6 parameters (x, y, z, pitch, roll, yaw) to estimate and correct for head motion. These routines essentially shift each image to match with a reference image (e.g. the first image in the series or a mean of the time series), and tend to deal best with modest movements (~½ voxel size). To account for residual effects of motion, including between-group differences, it is standard to include the *realignment parameters* describing volume-by-volume motion in the first-level model (see below). Note however that recent evidence suggests that these standard approaches to motion correction are not sufficient for all studies (Power et al., 2014; Van Dijk, Sabuncu, & Buckner, 2012) and may require further action such as including higher-order derivatives of the realignment parameters in the first-level model (e.g. Power et al., 2014) or independent component analysis (ICA)-based removal of motion artefact (Satterthwaite et al., 2012).

The best method of dealing with motion in fMRI data remains prevention: clear instructions that are to the point but also stresses the importance of staying still. (In our lab, participants are told at multiple times by the researcher and radiographer of the importance of staying still: “Remember that it is very important for you to stay as still as possible during the scan, because even motion of one centimeter will mean that we cannot use your data”).

Unfortunately, data loss due to motion remains a common problem in fMRI studies. While researchers try to avoid exclusion of data due to motion, it is often unavoidable especially in patient data. Our experience is around 5% data loss (i.e. excluded participants) in young healthy control samples and 10–15% data loss in patient populations. Researchers should closely examine the realignment profiles of every individual participant to screen for excessive motion.

12.3.4.4 Normalisation

Normalisation is typically only necessary in multi-subject studies that average across individual data to make a statistical inference on the neural response of the sample or population. In its simplest conceptualisation, normalisation ensures that a voxel in a certain location in one subject indexes the same location in all subjects in the sample. There are a number of approaches to normalisation, the most common of which are the volume-based (affine linear and non-linear transformation) and surface-based (FreeSurfer) approaches. In multi-subject designs, subject images are typically normalised to an atlas template (e.g. MNI) to facilitate interpretation of task activations.

Note that software packages differ substantially in the algorithms used for spatial normalisation and the stage during the processing stream it is conducted (e.g. SPM: during preprocessing; FSL: after statistical analysis).

12.3.4.5 Spatial Smoothing

Spatial smoothing is performed in multi-subject fMRI designs for three main reasons. Firstly, task-based fMRI signals of interest tend to extend across many voxels, so removing small-scale high-frequency changes in the image increases the ability to detect signals of interest (signal-to-noise ratio). Secondly, it is recognised that normalisation algorithms are imperfect, so spatial smoothing further reduces the impact of anatomical mismatch between subjects. Thirdly, a number of statistical approaches require a certain level of spatial smoothness in the image (e.g. Gaussian random field theory). Spatial smoothing is usually conducted by applying a 3D Gaussian smoothing filter with full-width at half-maximum (FWHM) set to approximately twice the acquired voxel size.

12.4 fMRI Experimental Design

The properties of the haemodynamic response influence the experimental design and thus the types of experimental hypotheses that can be tested in fMRI studies of ocular motor control. Like EEG/ERPs, fMRI designs used in experimental psychology and studies of ocular motor control typically rely on signal averaging and mental chronometry techniques. These include the Donders (1868) subtractive and Sternberg (1969) additive factors logics, which have recognised limitations (Friston et al., 1996a). Like ERPs, fMRI is also best suited to examining responses with well-defined onsets and offsets, such as stimulus-triggered task-based saccadic responses. In contrast, responses with long durations and poorly defined onset and offsets (such as memory guided saccade tasks, self-paced saccades or saccade adaptation tasks) are more difficult to test using fMRI.

Here, we review the two main approaches used in fMRI studies of ocular motor control, block and event-related designs by using the *antisaccade paradigm* (see chapter by Pierce et al. in this volume) as a model task. The popularity of these approaches reflects the traditional *localist* focus of the field (Friston, 1997), which tries to localise cognitive processes to single brain regions or a network of brain regions. The choice of experimental design depends on a number of considerations, including whether the goal is to *detect* an effect or *estimate* the response to individual trials, and whether the goal is to study *sustained* or *transient* responses to a task. For reviews to other approaches used less often in ocular motor studies, see Haynes and Rees (2006), Calhoun, Liu, and Adali (2009), and Friston (2011).

12.4.1 Block Design

A block design fMRI experiment is most suited for studies where the goal is to *detect* a signal, e.g. '*which regions are activated by task X?*'. This simple design compares the BOLD signal from an experimental condition (e.g. antisaccade trials) to a control condition (e.g. fixation trials or prosaccade trials; Fig. 12.4b). Trials are blocked according to trial type: for example 12 antisaccade trials alternating with 12 null trials (fixation). The estimated haemodynamic response is therefore very large and sustained across the duration of the block, and thus block designs are very good at detecting voxels with significant activity. However, because the haemodynamic response is *sustained* across the entire block, these designs are insensitive to activity to individual trials within a block—that is, the ability to *estimate* the haemodynamic response to single events is rather poor. Thus, if participants make a performance error within a block, it is usually not possible to extract a response to that error trial separate to the correctly performed trials within the block.

12.4.2 Event-Related Design

Event-related fMRI designs are most suited for studies where the goal is to *estimate* the shape of the haemodynamic response associated with each trial. Because block designs estimate sustained responses across multiple consecutive trials, they are relatively insensitive to experimental design considerations such as stimulus onset asynchrony (SOA), intertrial intervals, and so on. By comparison, event-related designs must carefully consider timing intervals, SOAs, and trial order to ensure robust estimation of the shape of the haemodynamic response; thus event-related fMRI designs are often more complex than block designs and have reduced detection power. In the simplest of event-related designs, trials (events) are presented slowly, around the resolution of the haemodynamic response (10–15 s) to ensure that the signal has fully resolved prior to the onset of the subsequent event. In fast event-related designs, trial presentation is randomised (Fig. 12.4b) to effectively increase the time between similar events; so sometimes a trial type will occur several times in a row, whereas at other times there will be a long interval between successive presentations. This *jittering* of the effective interval between trials of the same type allows for much shorter intertrial intervals than would otherwise be possible in fMRI. Additional jitter is often added to a design by randomising the intertrial interval between all trials regardless of trial type (Fig. 12.4b). Without jittering, the BOLD signal will saturate, becoming equivalent to a block of trials, reducing the estimation efficiency of the design. Event-related designs are best suited to paradigms that require randomised presentation of trial types and that aim to examine *transient* responses to individual trials—so, separate responses can be estimated for trials where participants make a performance error, for example. Although challenging, it is possible to mathematically optimise the trial randomisation and intertrial interval jittering to ensure the best possible estimation of the haemodynamic response in a design, and there are several methods to do so including M-sequences (Liu & Frank, 2004) and genetic algorithms (Wager & Nichols, 2003). Note however that procedures for optimizing task design is still a matter of active study.

Event-related designs allow for fMRI paradigms to more closely resemble tasks used in behavioural and ERP research, where randomised trial presentation, short within-trial SOA and variable intertrial intervals are often used. An important caveat is that unlike ERPs, the response to a trial in an event-related design is the response to the entire trial—it is often not possible to extract responses to individual events within a trial, as it is possible in ERPs. For example, in a cue-stimulus-response trial design, separate ERPs can be extracted locked to the cue, stimulus and response; in fMRI the estimated response is averaged across all events in a trial. It is possible (and common) to introduce different trial types to target events within a trial, such as cue-only, stimulus-only and no-response trials, although it is important to fully test that such designs do not change the effect of interest, such as introducing no-go effects and expectancy violation (Ruge, Jamadar, Zimmerman, & Karayanidis, 2013).

12.5 fMRI Analysis

Following preprocessing, fMR images are usually analysed using a *mixed-effects* analysis by first submitting them to single-subject general linear model (GLM) analysis, followed by group analysis.

12.5.1 Single-Subject Analysis

Typically, the BOLD signal is modeled for every voxel in the brain, in a *voxel-wise mass-univariate* approach. While mass-univariate approaches must deal with the substantial problem of multiple comparisons correction (see Sect. 3.2), this approach is powerful in that it allows for inferences about regionally specific events. The GLM of the time series of an individual voxel is expressed as

$$y = [X]\beta + \varepsilon$$

where y is the BOLD (haemodynamic) response, β are the parameter weights to be estimated, ε is the residual error and $[X]$ is the design matrix constructed during study design and containing the timing parameters (e.g. stimulus onsets) of the effects of interest. Each column in the design matrix represents a prediction on how the haemodynamic response should change in response to that factor. So, in a block design this would resemble the on/off shape shown in Fig. 12.4b (i, black line) and in an event-related design this is usually modeled as a short-duration canonical haemodynamic response function *convolved* (mathematically blended) with the event onsets (for a discussion of the choice of haemodynamic response function and convolution see Poldrack et al., 2011 or Henson, 2004).

For example, in an event-related design with two trial types, antisaccade and prosaccade trials, the GLM can be conceptualised as

$$observed\ BOLD\ response = antisaccade\ variance + prosaccade\ variance \\ + residual\ error$$

Following estimation of the GLM, summary statistics can be calculated to examine effects of interest. In our example this summary statistic may be the $antisaccade - prosaccade$ contrast, which—consistent with the mass-univariate approach—is calculated for every voxel in the brain and can be summarised in a *contrast image*.

12.5.2 Group Analysis

Following estimation of the single-subject GLM, summary statistic images and their associated within-subject variance are entered into a second-level mixed-effects model where a mean is estimated and tested for significance (Fig. 12.4c). In our simple example, a one-sample t-test may be calculated to test for differences between antisaccade and prosaccade trials, but a variety of tests are possible depending on the experimental design, including t-tests, full factorial ANOVAs, ANCOVA, multiple regression and so on. Note that this method of mixed-effects analysis allows statistical inferences to be generalised to the population of interest as it accounts for both within- and between-subject variance. The alternate approach, *fixed-effects* analysis, does not account for between-subject variance so inferences only apply to the sample under investigation.

In a *whole-brain* analysis, the statistical test (here, one-sample t-test) is carried out for each voxel in the brain, and is visualised by creating a *statistical parameter map*, where the statistical map is thresholded (to remove non-significant voxels), colour-coded to represent the statistic (here, T-values) and overlaid on a representative (average of group, template or atlas) high-resolution scan of the brain. This whole-brain voxel-wise approach can be powerful in that effects of interest can be detected throughout the brain, including in regions not included in original hypotheses or in regions the size of a small cluster of voxels.

However, the mass-univariate whole-brain voxel-wise approach is particularly susceptible to the *multiple comparisons* problem—that the probability of a false positive (type I error) increases with the number of statistical tests. So, at an $\alpha = 0.05$ level, if a test is repeated 100 times, then 5% of results will be false positives. In a whole-brain voxel-wise analysis of images that can contain 100,000 voxels, at an $\alpha = 0.05$ level, 5000 voxels will be false positives. Thus, *multiple comparisons* *correction* is a critical issue for fMRI, and a number of approaches are used to account for the problem.

- *Family-wise error* (*FWE*) *rate correction* controls the chance of one or more false positives anywhere in the image. So, at $\alpha_{FWE} = 0.05$, there is a 95% confidence that there are no false positive voxels. Examples of family-wise error corrections include Bonferroni correction and correction based on virtual voxels (Resolution Elements or RESELs) derived from random field theory (Nichols & Hayasaka 2003). In practice, family-wise error correction tends to be conservative and many experiments will produce no positive results, particularly with small sample sizes.
- *False-discovery rate correction* controls the number of false positives in the image so that at $\alpha_{FDR} = 0.05$, 5% of the detected voxels will be false positives. So, false discovery rate correction procedures tend to be more sensitive to effects of interest than family wise error correction.
- *Voxel level versus cluster level inference*: When applied at the voxel level, family-wise error and false-discovery rate correction is focused on controlling the false positive rate of individual voxels in the image. However, voxel level correction makes no use of the spatial information in the image. fMRI data in neighboring voxels tends to be spatially correlated due to spatial smoothing applied during preprocessing and due to the fact that anatomic structures are likely to span multiple voxels. In other words, effects of interest are unlikely to be constrained to a single voxel; rather, they are likely to occur in multiple voxels clustered together. Cluster level correction takes advantage of the spatial structure of fMRI data to make inferences on *clusters* of voxels, rather than individual voxels. The significance test is determined by the size of the cluster. In practice, cluster level correction is more sensitive than voxel level correction, and so is a popular choice for multiple comparisons correction. The greater sensitivity of cluster level correction is offset by its reduced spatial specificity, that is, it is not possible to identify a single voxel within a cluster and definitively say that the signal does (or does not) originate from that voxel. All that can be inferred is that the signal arises from one or more voxels within the cluster.

An alternative to whole-brain voxel-wise analysis is *region-of-interest (ROI) analysis* (Poldrack, 2007). Whole-brain analyses are suited to studies where it is important not to constrain the search for effects of interest within a predefined set of regions, such as in exploratory studies. ROI analyses take advantage of the fact that the researcher often has an a priori defined set of regions where they expect the effects of interest to occur. This is particularly the case in studies of eye movements, where the rich body of animal research has yielded a well-defined network of regions known to be involved in the behaviour. ROI analyses can therefore diminish the burden of multiple comparisons correction by constraining the search volume for effects of interest, and increase the sensitivity of the test by averaging across voxels within a predefined ROI. Like cluster level inference, ROI analyses are less spatially specific than voxel-wise inference, and challenges exist in creating ROI definitions that map on to the same anatomical region across subjects. Furthermore, ROI analyses can introduce the additional problem of creating definitions of regions that are defined functionally, rather than anatomically, such as the frontal eye fields. Lastly, constraining the search volume to a priori defined ROIs means that other, unexpected effects of interest in other regions will not be identified. Thus, in practice, many researchers use a combination of whole-brain and ROI analyses for inference.

12.5.3 Baselines in fMRI Research

The concept of a 'baseline' in fMRI research has been a contentious issue in the literature. The BOLD signal has no inherent baseline (Gusnard & Raichle, 2001), and so a number of methods have been used to estimate a baseline response.

Many studies use rest periods, where the participant is instructed to lie quietly with eyes open or closed for a period of time. These trials, sometimes called 'null' trials, can be interspersed with activation blocks in a block design (where the null block may last 10–30 s) or interspersed with task trials in an event-related design (where the null trial has a similar duration to task trials). This method assumes that null trials involve no activity within the task-related network, which the large body of resting-state fMRI has shown to be false (e.g. Biswal, Zerrin Yetkin, Haughton, & Hyde, 1995; Fox & Raichle, 2007). 'On' versus 'off' periods may be useful for some sensory neural regions like V1 (e.g. light vs. dark), but for other cortical regions it is usually unclear what participants are doing and what the brain activity represents during these null trials. Further, depending on their design null trials can become true events when they violate expectancy and task contingencies (see Ruge, et al., 2013 for a discussion).

An alternate to null trials is to use cognitive subtraction logic (e.g. Donders method, 1868/1969) to create an active baseline in the design. In this case, the 'baseline' condition is assumed to differ from the 'experimental' condition only in a single cognitive process of interest. Although this assumption of 'pure insertion' underlies much research in the fields of cognitive psychology and neuroscience, it is known to be invalid (e.g. Kulpe, 1909; Friston et al., 1996a). For example, in the antisaccade paradigm, a researcher may assume that the antisaccade-prosaccade contrast isolates the process of inhibitory control that is required in antisaccade but not prosaccade trials. However, antisaccade and prosaccade trials also differ in the amount of foveal stimulation, which is greater in pro versus antisaccade trials; as well as the requirement to hold task rules in working memory (look away from the target) in antisaccade versus prosaccade trials. Antisaccades are also more effortful than prosaccades, and so may differentially affect motivation and task engagement. An alternate method of simple cognitive subtraction is to use a factorial designs, which test for interactions between conditions and avoids the assumption of pure insertion (Friston et al., 1996a).

In an event-related design, many researchers compare their activation conditions to 'implicit' baseline, which is basically anything not entered into the model—i.e., all unmodelled variance in the general linear model. When planning to compare a condition against implicit baseline, it is important to ensure that the model is as accurate as possible. In the antisaccade example, what happens if the participant makes an antisaccade error? Are the errors collapsed into the antisaccade term, or are the errors unmodelled? In the first case, error-related activity is included in the antisaccade parameter, but in the second case error-related activity is included in the implicit baseline. Thus when using an implicit baseline it is crucial to ensure that the model is accurate and accounts for all brain activity relevant to the study.

In sum, the problem of an ambiguous baseline in fMRI should be considered when designing and critically evaluating fMRI research. The BOLD response has no intrinsic baseline, and the brain is never truly ‘off’—it never stops using oxygen, firing neurons etc.—while the participant is still living and breathing. An inappropriate baseline can reduce, eliminate, or even reverse the sign of the effects of interest (Gusnard & Raichle, 2001). Factorial designs can avoid the baseline issue by examining interactions between conditions, but many studies use null trials, subtractive logic (active baseline) or implicit baselines, acknowledging the limitations of their method. Importantly, while the issue of baseline has been an area of active investigation in the fMRI literature, the philosophical issue of how to define a baseline in a continuously active brain is common across many aspects of cognitive psychology/neuroscience, including event-related potential research (see e.g. Coles, Smid, Scheffers, & Otten, 1994).

12.6 Practical Issues for fMRI Studies of Eye Movements

12.6.1 Recording Eye Movements

Recording eye movement data in the constrained space of the fMRI scanner represents a challenge for the ocular motor researcher. These challenges are not inconsequential, and are likely to explain why a small number of studies have not recorded eye movement data in the scanner at all. As an example case, we discuss here the challenges encountered with setting up the ocular motor fMRI equipment at Monash Biomedical Imaging, Melbourne, Australia, and the solutions we devised to overcome them (Table 12.2). Given that we have found that even very experienced operators of the non-MR eye tracking system need a considerable amount of training and practice in operating the system in the MR environment, our aim here is to assist others in obtaining a reliable eye tracking-fMRI recording system in their own laboratories.

We acquire our data with an MR-compatible SR Research (Ontario, Canada) EyeLink 1000 infrared video-based eye tracker and Siemens (Erlangen, Germany) 3T Skyra MRI scanner. During fMRI scanning, participants lie supine in the scanner bore with their head in a radiofrequency head coil (Fig. 12.5a). A mirror is placed on the head coil and is positioned to capture the stimulus display that is projected on a screen positioned behind the scanner bore. The eye tracking camera is positioned between the projector screen and the back of the scanner bore, with the camera and

Table 12.2 Common problems and solutions encountered during eye-tracking in MRI scanner

Problem	Solution
Calibration difficulties or poor eye tracking	• Check camera and infrared light have not been moved • Turn light in the bore as high as it can go, turn scanner room lights on • With the guidance of a radiographer, ensure the participant is positioned quite high in the bed to minimise shadowing from the head coil • Moving the coil mirror can change the position of shadowing on the eye • Some people with small heads may sit a little low in the head coil, try boosting the head with foam pillows or slightly tilting the head, with the guidance of a radiographer • A slight head tilt can encourage participants to open eyes wider during the task • Ensure mascara and eyeliner is removed • Make sure that there is something on the screen for the participant to focus on during calibration • Try using a lower acquisition rate. In our study, 500 Hz provided a more stable and less noisy eye trace than 1000 Hz
Inconsistency in capturing eye between participants	• Design a semi-permanent installation that is not moved between participants and has been calibrated in height for a range of participant head sizes. Projector screen mounting is often not sturdy enough • Use a head coil that does not cast a shadow over the eye (Fig. 12.5)
Participant fatigue	• Ensure fan is off in the scanner bore • Provide regular rest breaks where participants can close their eyes for a minute or two

infrared illuminator directed to capture the image from the mirror attached to the head coil. As the camera receives the image at an acute angle, the eye can quickly fall outside the field of view because of small movements in camera position (such as those associated with vibrations of the scanner during image acquisition) and variation in participant head size. Given the space constraints in the area behind the scanner bore, it is often suggested to mount the eye tracking equipment to the projector screen; however we found that a sturdy non-ferromagnetic stand for the eye tracker was required to ensure that the camera mount is secure and not prone to small movements and vibrations from the scanner and to ensure the projected image is not occluded.

Head coils come in multiple designs (Fig. 12.5b), which can influence the signal quality of the acquired ocular motor data. For example, some head coils have a mask that partially occludes the eye (Fig. 12.5b, left). While it is sometimes possible to adjust the camera or mirror angle to project under the mask, this is problematic as the mask itself casts a shadow over the eye, and adjustments need to be made on a subject-by-subject basis which can be cumbersome and is often not feasible given time constraints during data acquisition. In addition, we found that even when the eye was in view of the tracker, there was significant signal drop-out, noisy eye traces, and the pupil tracking software regularly “lost” the eye even when it was in the field of view. We found that a head coil with an open-mask design (Fig. 12.5b, right) offers the best view of the eye in a majority of subjects. In our case, the open-mask head

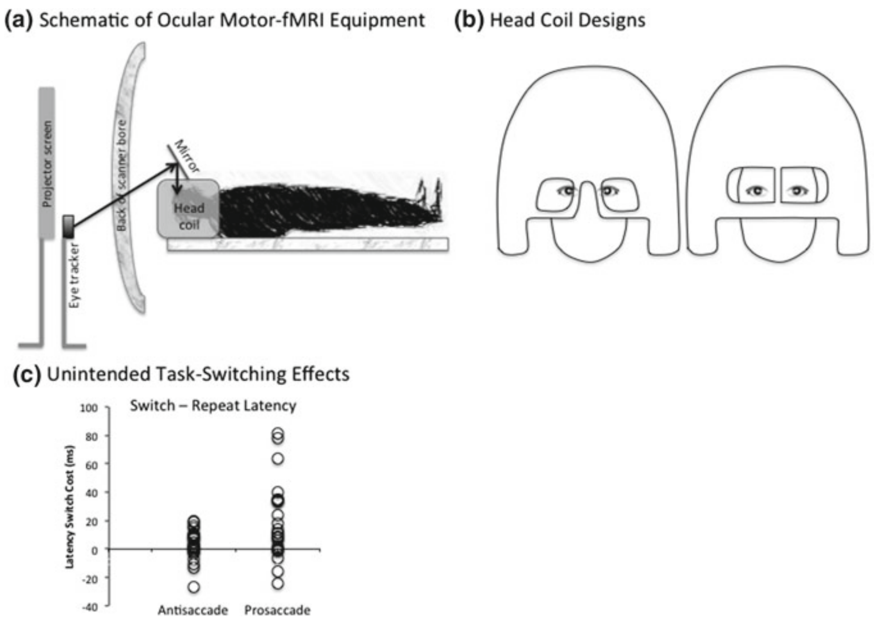


Fig. 12.5 Practical issues for ocular motor-fMRI data acquisition. **a** Schematic of ocular motor-fMRI equipment. Note the acute angle at which the eye tracker camera and infrared light acquires the eye data via the mirror mounted on the head coil. **b** Two head-coil designs: the left mask occludes and casts a shadow over the eye, making eye tracking difficult; the right (open) mask does not occlude the eye and shadowing is minimal, making eye tracking more feasible. **c** Randomising trial order can induce task-switching effects. In unpublished data from our lab, subjects (n = 23) switched randomly between antisaccade and prosaccade trials in an event-related design. Figure shows latency switch costs (switch-repeat) for antisaccade and prosaccade trials for each individual (represented by circles). Some individuals showed substantial positive and negative switch costs, with greater switch effects for prosaccade than antisaccade trials (consistent with the asymmetric switch cost, Allport et al., 1994)

coil is a 20-channel coil, whereas the head coil that occludes the eye and often casts a shadow over the eye is a 32-channel coil; thus a compromise was required between MRI signal to noise ratio and the ability to easily and consistently track the eye. For our purposes, this compromise is acceptable as the loss in MRI signal-to-noise ratio between the 32- and 20-channel coils does not significantly impact our research studies.

Additional considerations that we have found that optimise the quality of the ocular motor data acquisition includes turning the light on in the scanner bore to its maximum luminance, switching the fan in the bore low or off to prevent eye fatigue, positioning the subject quite high in the coil (i.e. shifted towards the back of the coil) to prevent shadowing from the lower bar of the head coil (Fig. 12.5b right; this should be done in consultation with the radiographer as it may affect the signal-to-noise ratio of the MR data), subjects with smaller head sizes should be boosted with extra foam to sit higher off the bed to prevent shadowing from the head coil, and sometimes tilting the head slightly backwards with foam under the base of the head/neck to optimise the ability of the infrared light to illuminate the eye.

12.6.2 Imaging Deep Brain Structures with fMRI

Functional imaging of the deep midbrain structures known to be important for ocular motor control represents a significant challenge for fMRI studies, for a number of reasons. Midbrain nuclei are difficult to delineate on standard whole-brain anatomical (T_1) scans by either automated or manual methods, as there is little variation in image intensity between the structures (Fischl et al., 2002). fMRI studies examining midbrain function are often optimised by restricting the field of view of the image over the midbrain only, restricting the ability to examine cortico-subcortical interactions known to be important for ocular motor control. The reduced field of view of midbrain fMRI presents an additional problem with registering the functional T_2* images to the anatomical T_1 images and normalising the data to stereotaxic space, due to the decreased spatial information in the midbrain T_2* image. In addition, the problem of registration and normalisation is exacerbated by traditional registration and normalisation algorithms that are optimised for whole-brain registration, and perform poorly on the small and tightly packed midbrain nuclei (Limbrick-Oldfield et al., 2012). The move towards higher dimension arrays (>12-channel) in head coils tends to result in reduced signal intensities in central brain structures; algorithms such as prescan normalise can improve the signal in central structures, but at a loss to signal in cortical structures (Kaza, Klose, & Lotze, 2011). Lastly, imaging of the midbrain nuclei is particularly susceptible to pulsatile and respiratory artefacts, due to the close proximity to large vasculature (Poncelet, Wedeen, Weisskoff, & Cohen, 1992), creating artefactual changes in image intensities and blurring of BOLD signal to nearby voxels (Limbrick-Oldfield et al., 2012).

Despite these challenges, it is possible to optimise the fMRI acquisition and analysis protocols to detect BOLD changes in deep structures (Box 1). While this remains an active field of research, algorithms have been developed that improve the registration and normalisation of reduced field of view MRI of the midbrain to anatomical and stereotaxic space (e.g. Limbrick-Oldfield et al., 2012), although any study acquiring data with a reduced field of view will be limited in its ability to examine cortico-subcortical interactions. Methods such as cardiac gating (Guimaraes et al., 1998), inclusion of cardiac and respiratory signals as covariates in first-level subject GLM (Wall, Walker, & Smith, 2009) and image correction algorithms (e.g. Retrospective Image Correction [RETROICOR], Glover, Li, & Ress, 2000; Physiological Noise Model, Limbrick-Oldfield et al., 2012) have shown success in reducing physiological noise artefacts in the fMRI data. However note that any method that relies upon measurement of physiological data using electrocardiogram, pneumatic belt for respiratory data, and pulse oximetry considerably increases the setup time and invasiveness of the scanning session for the subject.

In sum, the difficulty in imaging deep brain structures is likely to account for the unreliable detection of fMRI effects in the midbrain regions in studies of ocular motor control (Jamadar, Fielding, & Egan, 2013). While challenging, it is possible to reliably detect changes in fMRI signal in these regions, and given their known importance in ocular motor control, fMRI studies in this field should endeavor to include techniques to increase the ability to detect effects in midbrain structures.

Box 1: Imaging Deep Brain Structures Using fMRI

fMRI of deep brain structures is of particular interest in oculomotor research, but can be challenging (see text, Sect. 4.2). A detailed discussion of the challenges of midbrain fMRI is presented in Duzel, Guitart-Masip, Maass, et al. (2015).

Improving registration of reduced field-of-view images of the midbrain

The midbrain nuclei are difficult to image as they are small, tightly packed structures in an area of the brain susceptible to signal dropout or interference. To improve localization of activity to specific nuclei, many studies use a limited field-of-view to image this region. This presents a challenge for registration and normalization of these images (Sect. 1.4) as there is less structural information in these images to register with the structural and atlas images is difficult and error-prone. Multi-step registration can significantly improve the registration of limited field-of-view images to anatomical images. For example, Limbrick-Oldfield et al. (2012) demonstrated significantly improved registration using four-step registration compared to three- or two-step registration. In four-step registration, the midbrain T2*-image was (1) registered to a whole-brain T2*-image that was acquired immediately after the midbrain images using the same parameters. This registration was then optimised for the midbrain using a hand-drawn mask of the midbrain to weight the accuracy of registration towards these regions. The midbrain image was then (2) registered to a T2-weighted image, again weighting the accuracy towards the midbrain mask. The image was then (3) registered to the T1-weighted image and the (4) MNI atlas image (Fig. 12.6).

Reducing cardiorespiratory artifacts in midbrain fMRI

Cardiorespiratory noise is a significant source of artifact in midbrain fMRI. The simplest way to control for this uses cardiac or respiratory gating, which involves recording the cardiac or respiratory response simultaneously with the MRI data. In cardiac gating, cardiac contraction is detected using EKG or a peripheral pulse transducer. In respiratory gating, a pneumatic respiratory belt is used. The MR data is then corrected using the phase of the cardiorespiratory cycle (Guimaraes et al., 1998). Cardiac and respiratory measures can then also be entered into the GLM (Sect. 4.1) as regressors of no interest to further control for these effects (e.g. Wall et al., 2009).

RETROICOR (Glover et al., 2000) is a widely-used image-based correction method to reduce respiratory and cardiac noise. A Fourier series is fit

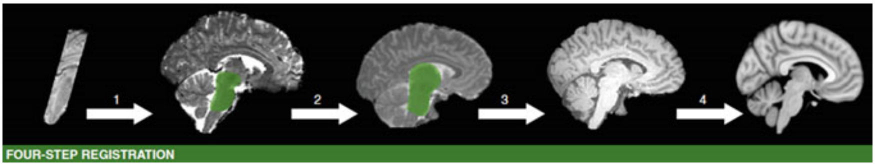


Fig. 12.6 Four-step registration can improve registration of limited field of view images of the midbrain. Figure adapted from Limbrick-Oldfield et al. (2012)

to the image-based time-series data based on the phase of the respiratory or cardiac cycle during the acquisition of each image. Like gating methods, the RETROICOR method relies upon measurement of cardiac and respiratory signals to inform the method.

Another method involves correlating the mean timecourse of voxels in a region-of-interest (ROI) in the midbrain image to either a resting-state scan (de Zwart, van Gelderen, Fukunaga, & Duyn, 2008) or a region of no interest (Wall et al., 2009). Common fluctuations between the rest image or region-of-no-interest is assumed to represent noise; the resulting ‘noise’ timecourse is then entered into the GLM as a regressor of no interest. Wall et al. (2009) showed significant improvement of statistical detection power in the superior colliculus using this method; with the region-of-no-interest method performing better than the de Zwart resting-state method.

12.6.3 Imaging the Cerebellum with FMRI

On the basis of the large body of animal and human lesion studies (see Sect. 4; Scudder et al., 2002; Stoodley, Valera, & Schmahmann, 2012), it is known that the cerebellum plays a central role in the control of eye movements. However, this region can be difficult to image using fMRI (Box 2). In the field of fMRI as a whole (including but not limited to eye movements) significantly less studies report activation of the cerebellar nuclei (Habas, 2010). This is due to anatomical and methodological constraints. The size of the cerebellar nuclei are typically small, vary considerably in position and size between individuals (Kozlova, 1984) and large nuclei like the dentate nucleus are composed of a number of functional subdivisions (Dum & Strick, 2003); each of these issues can introduce errors of activation mis-location. Unlike MRI atlases of the cortex, MRI atlases of the cerebellum that define structures and are used for stereotaxic normalisation are still under development (e.g. Schmahmann et al., 1999; Diedrichsen, 2006) and tend to be limited due to small sample sizes, normalisation problems, and difficulty in delineating individual nuclei on the scans (see Kuper, Thurling, Maderwald, Ladd, & Timmann, 2012 for a discussion). Problems with normalisation of the cerebellum tend to be exacerbated by normalisation algorithms, which can distort and elongate the cerebellum (e.g. standard SPM nonlinear normalisation; Diedrichsen, 2006). At the group statistics level, normalisation problems and inter-individual variability in the location of cerebellar nuclei tends to smear activation across cerebellar nuclei and sometimes into non-cerebellar regions leading to mis-localisation and reduced statistical significance.

The physiological basis of the BOLD response in the cerebellum differs from that of the cortex. In the cortex, a linear relationship exists between the increase in the cerebral metabolic rate of oxygen consumption ($CMRO_2$) induced by neuronal activity, the change in local concentration of deoxygenated haemoglobin and changes in cerebral blood flow (Sect. 1.2; Buxton, Wong, & Frank, 1998; Hoge et al., 1999). In the cerebellum, the relationship between $CMRO_2$, cerebral blood flow and neuronal activity is less well defined and is known to differ from cortical dynamics (see Attwell & Iadecola, 2002 for a review), resulting in smaller BOLD signal changes (Thomsen, Piilgaard, Gjedde, Bonvento, & Lauritzen, 2009). Like the brainstem and subcortical regions, the cerebellar BOLD signal also shows large physiological artefacts related to cardiac pulsatile and respiratory factors, with up to 40% of the variance in cerebellar BOLD signal attributable to these artefacts (Diedrichsen, Verstynen, Schlerf, & Wiestler, 2010).

As a consequence of these considerations, the presence or absence of cerebellar activations in an fMRI study should be interpreted with caution. Given this region's importance for eye movement control, studies specifically interested in cerebellar activation should use current best-practice methods, including use of separate normalisation templates for the cerebellum (e.g. SUIT, Diedrichsen, 2006) and normalisation methods that are optimised for the cerebellum and probabilistic atlases (Diedrichsen, Balsters, Flavell, Cussans, & Ramnani, 2009) for activation localisation. Physiological artefacts are best dealt with by recording physiological data and including the noise regressors into the first-level subject GLM models (Diedrichsen et al., 2010).

Box 2: Normalising the Cerebellum in fMRI

fMRI has significantly expanded our understanding of the neural bases of cognition and eye movement control, particularly in cortical and subcortical networks. While more work is required, we currently have a satisfactory understanding of the physiological basis of the BOLD fMRI response in the cortex, and how it relates to the underlying neuronal activity. However by comparison, progress in understanding the cerebellar contributions to cognition generally, and eye movements more specifically, using fMRI has been much slower. This is linked to a poorer understanding of the physiological contributors of the BOLD response in the cerebellum, and the substantial differences in neuronal architecture and firing rates in the cerebellum (see Diedrichsen et al., 2010 for a review).

Another major issue in imaging the cerebellum using fMRI is the anatomical architecture of the region - the cerebellum consists of a number of small, tightly packed lobules and sub-lobules with considerable inter-individual variability in size, shape and location (Diedrichsen, 2006; Schmahmann, Doyon, et al., 2000). Thus, standard whole-brain normalisation of images across individuals tends to show a high level of error and low reliability in the cerebellum. Furthermore, during image acquisition, it is not uncommon for researchers to preferentially image the cortex over the cerebellum. This leads to high rates of partial sampling of the cerebellum, particularly in individuals with head size greater than the study field-of-view. Researchers conducting eye movement studies should ensure they use a field-of-view that is adequate to accommodate both cortex and cerebellum in all participants.

Diedrichsen and colleagues have made significant advances in improving the normalisation of the cerebellum, using cerebellum-specific, non-linear normalisation (Diedrichsen, 2006). Figure 12.7 describes the modified procedure for preprocessing and analysis required by SUIT. Briefly, slice time correction and realignment are conducted in the typical manner. The anatomical scan is then centred at the anterior commissure, and the functional images are then registered to the anatomical scan. The single subject GLM is then conducted, and the subsequent analyses are conducted on the parameter estimate images. The cerebellum and brainstem are then isolated from the rest of the brain on the anatomical scan. The resulting image is then corrected manually, and the automated cropped image and manually-corrected image are then normalised to the SUIT template image. The resulting deformation field is then applied to the functional parameter estimate scans, and these images are then entered into the group-level random effects analysis.

SUIT improves normalisation of the cerebellum up to 8% over whole-brain methods (Fig. 12.8). Used together with their probabilistic atlas of the cerebellum (Diedrichsen et al., 2009), SUIT leads to substantial improvements in accuracy in the identification of anatomical regions of activity in functional analyses.

12.6.4 *Influence of Experimental Design on Effects of Interest*

As outlined in beginning of the chapter, fMRI studies require unique experimental design considerations to maximise the ability to detect a signal. These experimental design constraints can impact the behavioural effect, and hence cognitive processes of interest—as such, researchers should exercise caution when changing the parameters of a behavioural task to make it suitable for fMRI.

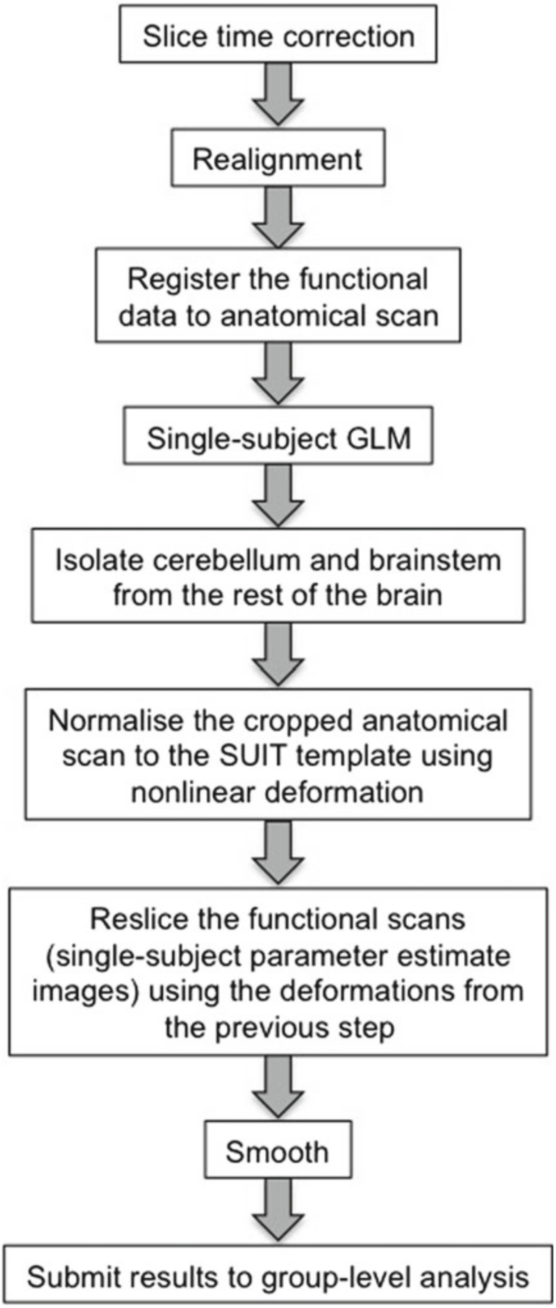


Fig. 12.7 Normalisation of the cerebellum using SUIT (http://www.diedrichsenlab.org/imaging/suit_fMRI.htm)

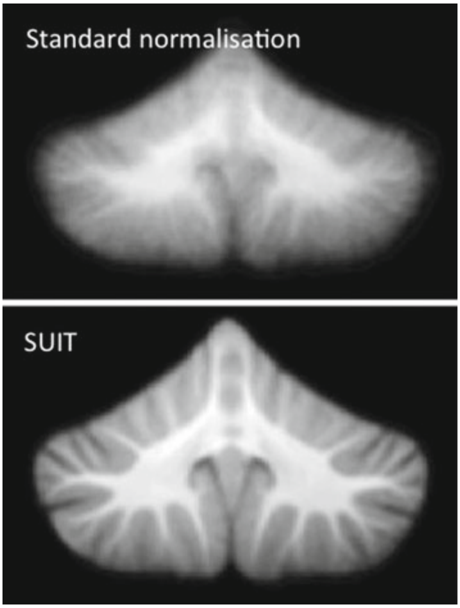


Fig. 12.8 Normalisation of the cerebellum using the SUIT procedure shows significant improvement over standard methods. Top: affine whole-brain alignment to ICBM152 template. Bottom: non-linear normalisation using SUIT. Figure adapted from Diedrichsen et al. (2006)

The sluggish haemodynamic response necessitates the use of slower fMRI paradigms than those used in the behavioural domain or in EEG/ERP. For example, in two studies examining the effects of behavioural training on the antisaccade task, Dyckman, McDowell and colleagues required participants to hold fixation on a peripheral stimulus for 600 ms in the behavioural study (Dyckman & McDowell, 2005) and for 1000 ms in the fMRI study (Lee et al., 2012). Other studies of the antisaccade task have used long inter-trial intervals (e.g. jittered 3, 4, 5 s, Brown, Vilis, & Everling, 2007; 13.3 s, Ford, Goltz, Brown, & Everling, 2005), long trial durations (e.g. 12 s, De Weijer et al., 2010; 12 s, Ettinger et al., 2008) or required subjects to fixate on the peripheral antisaccade or prosaccade target for long periods (e.g. 12 s, DeSouza et al., 2003) to account for the slow haemodynamic response. These timing parameters are much slower than those used in behavioural studies, and researchers should consider how these long durations affect task performance and subject fatigue and engagement with the task.

If the researcher is interested in examining the neural response to performance error trials separate to correctly performed trials, an event-related design must be used. This necessitates randomisation of trial types which can have unintended consequences for the task. For example, in an unpublished study, we have seen that randomising presentation of antisaccade and prosaccade trials results in substantial *switch costs* [difference in latency between trials where the trial type changes between trials (i.e. antisaccade → prosaccade or prosaccade → antisaccade) vs. trials where the trial type repeats between trials (i.e. antisaccade → antisaccade or prosaccade → prosaccade)] in many individuals, particularly for prosaccade trials (Fig. 12.5c). This switch cost, particularly the larger switch cost for prosaccade trials, is consistent with the task-switching and asymmetric task switching effects (Allport, Styles, & Hsieh, 1994) and suggests that a paradigm that randomises antisaccade and prosaccade tasks for the sake of fMRI experimental design may no longer be a pure index (free of confounds) of the antisaccade and prosaccade behaviours (Dyckman, Camchong, Clementz, & McDowell, 2007).

12.6.5 Strengths of Ocular Motor fMRI Research

In this section, we have focused on difficulties that fMRI studies of eye movements may encounter and outlined potential solutions. We have tried to highlight some practical and methodological constraints to the field so that the researcher can conduct and evaluate fMRI studies in an informed manner. Indeed, the issues discussed in Sects. 4.2–4.4 are shared by the fMRI neuroimaging field as a whole, and are not specific to the ocular motor field. Despite the caveats that must be held in mind when conducting and evaluating fMRI research, ocular motor fMRI research possesses one particular strength: the large body of animal electrophysiology, and animal and human lesion evidence of the neural bases of ocular motor control (see also chapters by Johnson & Everling as well as Müri et al. in this volume). The ocular motor system represents one of the few systems in the brain with a rich a priori framework for study with fMRI; in other words, the animal and human lesion data give us a good idea of the components of the network, the dynamics between regions within the network, and the neurons and their firing patterns involved in generating eye movements. Few other research fields in experimental psychology enjoy this advantage. The a priori knowledge of the ocular motor system offers a number of distinct advantages.

Extensive knowledge of the ocular motor system can be used to validate the results of fMRI studies. If a particular region is not significantly activated by an ocular motor task, it can be concluded with reasonable assurance that the absence of an effect is due to a methodological consideration such as those outlined in this section. For example, in a recent meta-analysis of fMRI studies of the antisaccade task, we did not find consistent activation of a number of regions known to be critical for eye movement control such as the superior colliculus and ocular motor cerebellum (Jamadar et al., 2013). Consistent findings of no effect in critical components of the ocular motor system across 18 studies of the antisaccade task suggests that it is likely to be due to methodological constraints of imaging these regions with fMRI. In other fields where the neural network is less well defined, it remains unclear if the absence of an effect in a region indicates that it is not involved in that task or is due to methodological factors.

Extensive knowledge of the ocular motor system allows fMRI activity to be interpreted as "involved but not necessary" for task performance. For example, our meta-analysis of the antisaccade task found consistent activity in the insula, precuneus, lingual gyrus and cuneus (Jamadar et al., 2013). While the consistent activation of these regions suggest they are likely to perform important processes during antisaccade task performance, we conclude that these processes are probably not strictly necessary, as they do not play a circumscribed role in animal and human lesion models of the task. In other research fields it usually remains unclear on the basis of fMRI findings alone which regions are necessary and which regions play supportive but non-critical roles in task performance.

fMRI studies of ocular motor control have strong priors for definition of regions of interest. As discussed in Sect. 3.2, ROIs can ameliorate the problem of multiple comparisons and increase the effective power of a study by restricting the search volume to a predefined set of regions.

Ocular motor fMRI research can quickly move on from defining the network. While some refinement of the ocular motor network in humans has been made on the basis of fMRI research (Sweeney, Luna, Keedy, McDowell, & Clementz, 2007), the detailed knowledge of the network allows secondary research aims to be explored. These include the understanding of the dynamics of connectivity and interactions between regions of the network (e.g. De Weijers et al., 2010), parcellating functional subdivisions of the network (e.g. Neggers et al., 2012) and developmental trajectories (e.g. Velanova, Wheeler, & Luna, 2009), and its disruption in psychiatric and neurological conditions (e.g. Raemaekers et al., 2002).

12.7 fMRI of Eye Movements: Basic Phenomena

fMRI has allowed us to further characterise the role of the key human ocular motor brain regions used in a variety of contexts. The properties of these key cortical, midbrain and cerebellar regions in controlling ocular motor functions have been elucidated using tasks such as antisaccade, gap/overlap, delayed response (memory guided), and smooth pursuit tasks. By using the relationship between the BOLD signal and key eye movement variables, such as saccade latency and accuracy, fMRI methods have been used to inform us about which brain regions are involved in the various aspects of ocular motor control. These include, amongst others, saccade preparation and initiation, precision of timing of eye movements (reaction time variability), and accuracy.

12.7.1 Prosaccade and Antisaccade Tasks

The most commonly used tasks in fMRI studies of the ocular motor network are the prosaccade and antisaccade tasks. Pro- and antisaccades are frequently compared in fMRI studies as a means of exploring several cognitive processes including response inhibition, attention, working memory, and decision making.

Prosaccades reliably activate the supplementary eye fields, frontal eye fields, and intraparietal sulcus compared to fixation or implicit baseline (Fig. 12.9; Jamadar et al., 2013; see also Aichert, Williams, Moller, Kumari, & Ettinger, 2012; Brown, Goltz, Vilis, Ford, & Everling, 2006; Ettinger et al., 2009; Kimmig et al., 2001;

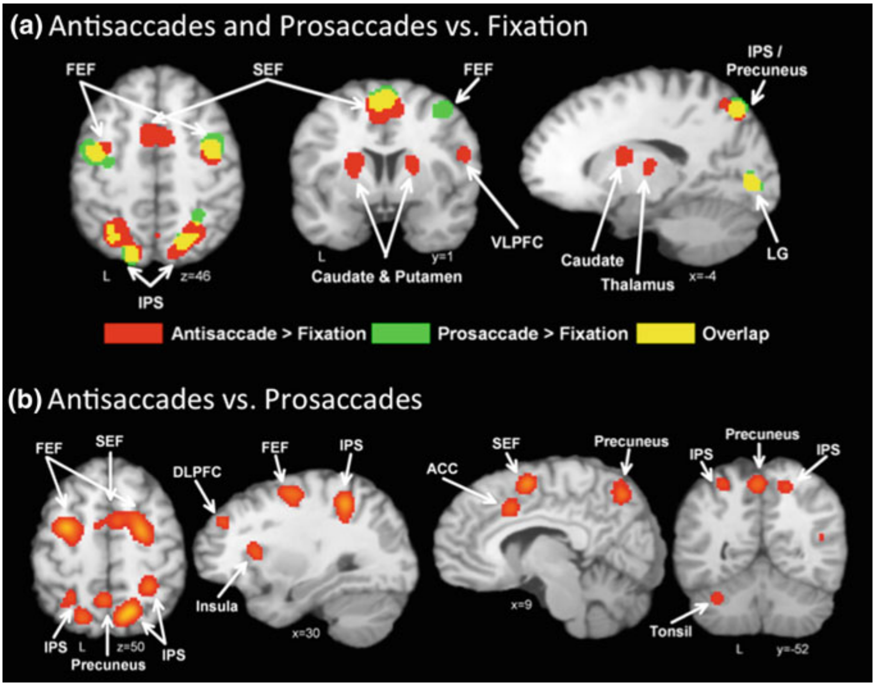


Fig. 12.9 Meta-analysis results of fMRI and PET studies of antisaccades and prosaccades (Figure adapted from Jamadar et al., 2013). **a** Activation likelihood estimation maps for antisaccade > fixation and prosaccade > fixation. **b** Activation likelihood estimation maps for antisaccade > prosaccade. Abbreviations: L, left; FEF, frontal eye field; SEF, supplementary eye field; SPL, superior parietal lobule; IPS, intraparietal sulcus; IFG, inferior frontal gyrus; LG, lingual gyrus; DLPFC, dorsolateral prefrontal cortex; ACC, anterior cingulate cortex

Matsuda et al., 2004). Antisaccades activate the ventrolateral and dorsolateral prefrontal cortices, frontal eye fields, supplementary eye fields, intraparietal sulcus, thalamus, caudate, putamen, supramarginal gyrus, and visual areas relative to fixation or implicit baseline (Jamadar et al., 2013; see also Aichert et al., 2012; Brown et al., 2006, 2007; Chikazoe, Konishi, Asari, Jimura, & Miyashita, 2007; Ettinger et al., 2008, 2009; Ford et al., 2005; Manoach et al., 2007). The additional activity in the dorsolateral and ventrolateral prefrontal cortices highlights the importance of increased cognitive control and inhibitory control required in antisaccades relative to prosaccade trials. fMRI-behavioural correlations have further characterised the role that individual regions in the ocular motor network play in ocular motor control. Activation of the frontal eye fields, but not the supplementary eye fields or intraparietal sulcus, is significantly associated with saccade reaction time, while both the frontal eye fields and supplementary eye fields are significantly associated with target preparation in both pro- and antisaccades (Connolly, Goodale, Goltz, & Munoz, 2005).

The design of prosaccade and antisaccade trials make them well-suited to study with fMRI, as they are characterised by a quantifiable short duration response with distinct onset and offset, and are amenable to both block and event-related designs. Additionally, prosaccades provide an excellent active control condition in subtractive designs, as they require many of the processes invoked on antisaccade trials, except the additional processes that are usually of interest to the investigator—i.e., cognitive control, inhibitory control and vector inversion. However, evidence does suggest that randomising task presentation for event-related designs may change the neural process under investigation. Dyckman et al. (2007) showed that while behaviour was similar between single pro- and antisaccade task blocks and mixed pro/antisaccade blocks, antisaccade versus prosaccade fMRI activity across the ocular motor network was attenuated in the mixed task block compared to the single task blocks, particularly in the supplementary and frontal eye fields. This is consistent with an increased requirement for preparation of the saccade response in prosaccades presented within a mixed task block compared to a single task block.

fMRI results show that individuals with difficulty in performing antisaccades show widespread disruption of the ocular motor network. Generally, disruption tends to occur in the regions important for cognitive and inhibitory control (dorsolateral prefrontal and ventrolateral prefrontal cortices, respectively) and regions involved in the preparation of the saccade response (frontal and supplementary eye fields). Individuals with schizophrenia show higher rates of directional errors and correctional saccades compared to healthy controls, and this is associated with reduced fMRI activity in dorsolateral and ventrolateral prefrontal cortices and thalamus (Fukumoto et al., 2009; Tu et al., 2006), and frontal and supplementary eye fields (Camchong, Dyckman, Austin, Clementz, & McDowell, 2008; see also Dyckman et al., 2011) during antisaccades. This is compatible with an overall impairment in cognitive and inhibitory control, particularly with preparation of the difficult antisaccade response. Individuals with Parkinson's disease show impaired activity of the frontal and supplementary eye fields, with attenuated activity in the frontal eye fields related to Parkinsonian symptoms (Cameron et al., 2012). Adults with autism show reduced activity and connectivity of the frontal eye fields and anterior cingulate, with this reduced activity associated with higher error rate and more severe restricted repetitive behaviour (Agam, Joseph, Barton, & Manoach, 2010). Finally, during ageing, older and middle aged adults show an anterior-posterior shift in activity during prosaccades (but not antisaccades), with elevated activity in frontal areas (frontal and supplementary eye fields) and reduced activity in posterior occipital and parietal areas relative to young adults (Raemaekers, Vink, van den Heuvel, Kahn, & Ramsey, 2006).

12.7.2 Gap-Overlap Task

The gap task requires individuals to focus on a central fixation point, which then disappears, and after a brief delay a peripheral target appears. In the overlap task, the peripheral target appears while the central fixation point is still present (there is an overlap in appearance between the two targets). The gap task elicits short latency saccades. During the gap period, there is no sensory or motoric activity, and therefore this task can be used to elucidate regions associated with saccadic preparation. The overlap task produces longer latency saccades; as both the central and peripheral targets are simultaneously present, there is greater difficulty "disengaging" from the central target. The overlap task is therefore useful in assessing task shifting and disengagement of attention (see also the chapter by Pierce et al. in this volume).

Although there have been a considerable number of electrophysiological studies exploring gap/overlap paradigms, there are relatively few fMRI studies that have utilised the gap task, and even fewer that have explored the overlap task. Connolly and colleagues (Connolly, Goodale, Cant, & Munoz, 2007; Connolly, Goodale, Menon, & Munoz, 2002; Connolly et al., 2005) have used the gap task in fMRI studies to systematically demonstrate the association of the frontal eye fields with motoric preparation of saccades (consistent with evidence from the antisaccade task), and importantly, the linear correlation of the frontal eye field activation with the duration of the gap (Curtis & Connolly, 2008). A gap/overlap study conducted by Ozyurt and Greenlee (2011) was unable to replicate the frontal eye fields BOLD activity–saccade reaction time relationship during the gap task, but found that saccade reaction times during the overlap task were significantly positively correlated with activation in a cluster in the right ventrolateral prefrontal cortex, a region well established for its role in response inhibition and fixation. Greater ventrolateral prefrontal activation during the overlap task is likely to help overcome inhibitory processes that maintain fixation, and thereby foster disengagement to initiate a saccade to an alternative, competing location (Fig. 12.10).

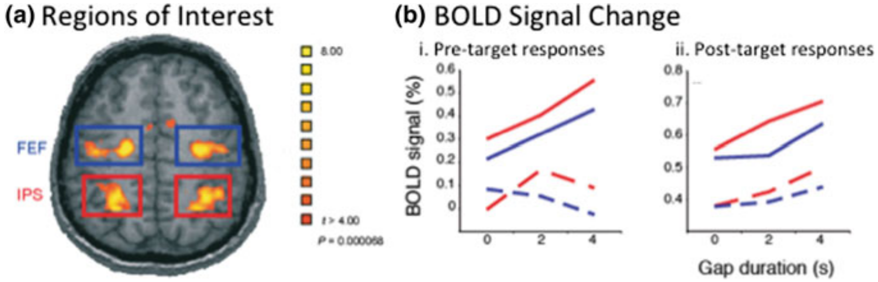


Fig. 12.10 fMRI results from a gap task (figure adapted from Connolly et al., 2002). **a** Frontal eye field and intraparietal sulcus regions of interest were identified from a prosaccade localiser scan. **b** Percent signal change as a function of gap duration (solid lines = frontal eye field, dashed lines = intraparietal sulcus) for (i) pre-target and (ii) post-target responses

12.7.3 *Memory Guided (Delayed Response) Task*

The memory guided, or delayed response task (Funahashi, Bruce, & Goldman-Rakic, 1989), begins with the presentation of a central fixation target. A peripheral cue briefly flashes, which the participants are instructed to remember the location of but not initiate a saccade to; when the central target disappears, participants are asked to initiate a saccade to the peripheral location (see chapter by Pierce et al., in this volume). Manipulating the position of the peripheral cue and the timing of when to initiate the saccade to the peripheral location in event-related fMRI studies can help to inform about which regions are essential to maintaining visuo-spatial coordinates, motor coordinates, and initiation of saccades. This task is also useful for exploring response inhibition, because the prepotent response is to look towards the briefly presented peripheral target. One difficulty with this task is that it can be difficult to design an active control task for. In the antisaccade and gap/overlap tasks, it is relatively straightforward to design a prosaccade comparison task with similar timing parameters as the experimental condition. Unlike these tasks, it can be difficult to design a prosaccade comparison task with similar timing parameters as the memory-guided task.

fMRI studies by Curtis, Sun, Miller, and D'Esposito (2005) found that maintaining the position of a target in working memory was underpinned by coherent activations across a small network involving the frontal eye fields, supplementary eye fields and dorsal anterior cingulate. Curtis et al. proposed that, when there were a small number of motor representations to be stored in working memory, this network maintained the locations of these potential saccade goals (Curtis et al., 2005). By contrast, storage of spatial information, which underpins the spatial accuracy of memory guided saccades, was maintained by a small network of cortical regions. Bilateral activation of the frontal eye fields, right intraparietal sulcus and right superior frontal sulcus was found to predict saccade accuracy, depending on the period of delay (Curtis, Rao, & D'Esposito, 2004). Broadly, however, fMRI studies using memory guided paradigms have elaborated on our understanding of the role of parietal and frontal regions in ocular motor control, namely that that coding of visual space is underpinned by the posterior parietal cortex, while coding of saccadic targets is more prominent in the frontal eye fields (Curtis et al., 2004, 2005).

In fMRI investigations of individuals with first-episode schizophrenia, performance on the memory guided task was associated with reduced task-related activation the frontal eye fields, supplementary eye fields, parietal and cingulate cortex, and dorsolateral prefrontal cortex, compared to matched control subjects. This suggests a dysregulation of sensorimotor control regions associated with ocular motor control and spatial working memory (Keedy, Ebens, Keshavan, & Sweeney, 2006). fMRI investigations by Luna et al. (2002) in individuals with high-functioning autism also found impairments in spatial working memory using the memory-guided task: task related activation was found to be significantly reduced in the dorsolateral prefrontal cortex and posterior cingulate cortex, but not sensorimotor regions such as

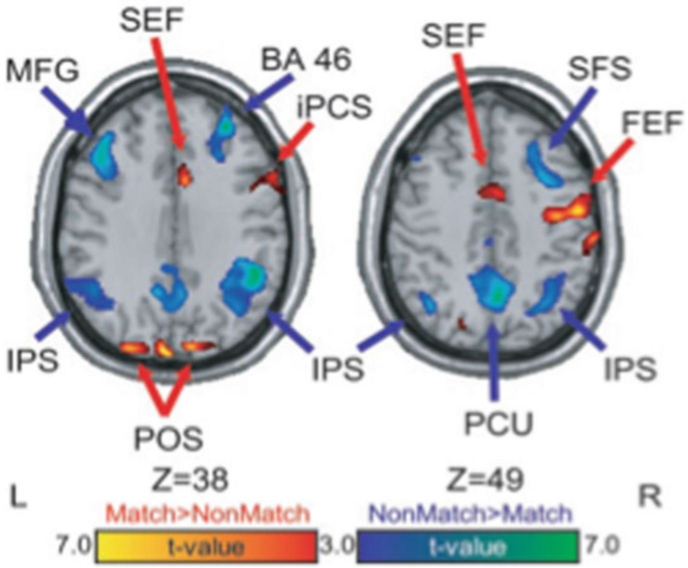


Fig. 12.11 fMRI activity for memory-guided saccades (figure adapted from Curtis et al., 2004). fMRI contrasts for memory-guided saccades during the delayed match-to-sample versus nonmatch-to-sample tasks. Warm colours represent regions with greater saccade period activity on matching versus non-matching tasks, and cool colours represent regions with greater saccade activity on nonmatching versus matching trials. Abbreviations: L, left; R, right; FEF, frontal eye fields, IPS, intraparietal sulcus, MFG, Middle frontal gyrus; SFS, superior frontal sulcus; iPCS, inferior precentral sulcus; POS, parietal-occipital sulcus; PCU, precuneus; BA, Brodmann area

the frontal or supplementary eye fields, anterior cingulate, midbrain regions or cerebellum (Fig. 12.11).

12.7.4 *Smooth Pursuit Tasks*

Pursuit eye movements (see also chapter by Lencer et al., in this volume) are often assessed by examining the response (latency and accuracy of tracking) to constant velocity target motion, typically referred to as a ramp stimulus. Often, a step-ramp task is used in order to eliminate catch-up saccades: the target motion starts just to the side of midline, and then crosses the midline in a ramp at a constant velocity. Alternatively, pursuit eye movements may be assessed using a sinusoidal stimulus, in which a stimulus moves back and forth along a horizontal or vertical plane.

fMRI investigations of pursuit eye movements show regions of activation that parallel saccadic eye movements. These include the frontal eye fields, supplementary eye fields, parietal eye fields, anterior cingulate, striatum, areas MT/MST (middle temporal and middle superior temporal visual areas), and cerebellum (Tanabe, Tregellas, Miller, Ross, & Freedman, 2002; Petit & Haxby, 1999). fMRI is thus used to decompose visual detection of moving objects, sensorimotor transformations, and the initiation of pursuit eye movements. The visual component of pursuit eye movements, such as detection of moving objects, was associated with activation in area MT, which is well established for its role in motion detection, and also the posterior parietal cortex, which is classically associated with sensorimotor transformations. Initiation of pursuit eye movements was associated with greater activation in ocular motor control regions, namely the frontal and supplementary eye fields (Kimmig et al., 2008). Pursuit eye movements often rely on predictive mechanisms to compensate for delays in detecting a moving object and also initiating pursuit; these predictive mechanisms can include the estimation of a target's speed and trajectory based on prior experience (Burke & Barnes, 2008). fMRI was used to investigate these elements in pursuit eye movements, and it was reported that the frontal and supplementary eye fields, as well as the supramarginal gyrus were involved in timing and maintaining eye velocity, while activation in the superior parietal lobe was associated with sensorimotor integration. Finally, activation in the dorsolateral prefrontal cortex was related to selection of appropriate target velocities from working memory (Burke & Barnes, 2008).

Slow pursuit eye movements and difficulty sustaining pursuit eye movements, has been reported in both autism and schizophrenia. fMRI investigations of these eye movement phenomena in autism revealed reduced activation in the superior temporal sulcus (STS), bilateral FEF, left superior frontal gyrus, and V5 during pursuit. The authors reported that this was suggestive of an impaired top–down modulation of visual sensory processing (Takarae, Luna, Minshew, & Sweeney, 2014). Additionally, the authors suggested network-wide abnormalities to be associated with visual motion processing in autism (Takarae et al., 2014). By contrast, in schizophrenia, reduced pursuit eye movement velocity was related to reduced BOLD activation of the right ventral premotor cortex, but an opposing increased activation of the left dorsolateral prefrontal cortex, right thalamus, and left cerebellar Crus II (Nagel et al., 2007). Tregellas et al. (2004) also reported the unusual finding of greater hippocampal activation during pursuit eye movements in schizophrenia patients, with further region-of-interest analysis indicating lower activation in traditional ocular motor regions, including the frontal eye fields, cingulate gyrus, and occipital region (Fig. 12.12).

12.8 Conclusions

fMRI research has grown to be one of the most widely used methods in human neuroscience. Ongoing research and development of the field has led to a highly sophisticated method which has resulted in many high-impact discoveries about the human brain. In this Chapter, we have introduced the concepts of MR signal generation, and presented an overview of the physiology of the BOLD fMRI response. We have discussed standard fMRI experimental designs and data preparation and anal-

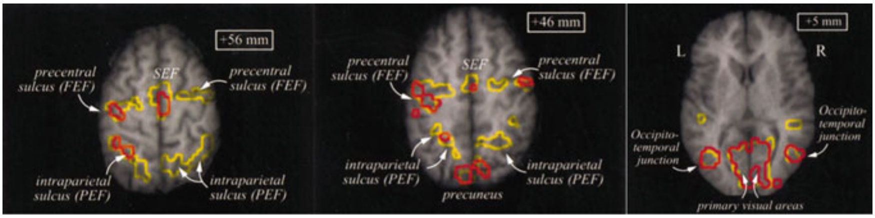


Fig. 12.12 fMRI activity for a single subject for the pursuit (red) and saccade-related (yellow) activity overlaid on the individual's normalised T1 image (figure adapted from Petit & Haxby, 1999). Abbreviations: L, left; R, right

ysis procedures. These 'bread and butter' methods have led to a large improvement in our understanding of the neural bases of eye movements in the human brain, as described in Sect. 12.7, fMRI of eye movements: basic phenomena.

As the field continues to develop, the sophistication of methods continues to develop. Optimisation of data preparation and preprocessing continues, and alternate methods of inference and modeling of data are becoming more widespread (e.g. Bayesian inference). Advances in multiple comparisons correction continue to improve the statistical robustness of the field (e.g., Ecklund, Nichols, & Knutsson, 2016), and advances in motion correction (e.g. Power et al., 2014) and normalisation (e.g. Diedrichsen et al., 2010) improve the quality of the data entered into analysis. Multivariate analysis techniques, such as independent component analysis (Calhoun et al., 2009) and multi-voxel classification algorithms, are powerful ways to examine the latent spatiotemporal structure present in the data. Furthermore, the rapidly growing field of connectomics (Fornito, Zalesky, & Breakspear, 2015) is throwing off the shackles of the traditional localist approach to understand the inherent interconnectedness of the continuously active brain.

fMRI research into the neural bases of eye movements is going to continue to develop in the number of publications, methodological complexity and scientific impact for many years to come. With a good understanding of the fundamentals of fMRI presented as in this Chapter, the reader can critically evaluate existing studies and develop their own research in this vibrant field.

12.9 Suggested Readings

Books:

Huettel, S.A., Song, A.W., & McCarthy, G. (2008). *Functional Magnetic Resonance Imaging*. Sinauer Associates, Inc: Sunderland, Massachusetts.

– *Thorough and accessible introduction to fMRI methods*

Poldrack, R.A., Mumford, J.A., Nichols, T.E. (2011). *Handbook of Functional MRI Data Analysis*. Cambridge University Press: New York.

– *Accessible introduction to fMRI data analysis*

Journal Articles:

Pooley, R.A. (2005). Fundamental physics of MR imaging. *Radiographics, 25,* 1087–1099.

– *Introduction to MR physics, including relaxation time and basic pulse sequences.*

Logothetis, N.K. (2008). What we can do and what we cannot do with fMRI. *Nature, 453*, 869–878.

– *Discussion of the physiology of the fMRI signal.*

Diedrichsen, J., Verstynen, T., Schlerf, J, Wiestler, T. (2010). Advances in functional imaging of the human cerebellum. *Current Opinion in Neurology, 23,* 382–387.

Limbrick-Oldfield et al. (2012). Identification and characterization of midbrain nuclei using optimised functional magnetic resonance imaging. *Neuroimage, 59*, 1230–1238.

– *Discusses challenges and solutions for midbrain imaging.*

Jamadar, S.D., Fielding, J, Egan, G.F. (2013). Quantitative meta-analysis of fMRI and PET studies reveals consistent activation in fronto-striatal-parietal regions and cerebellum during antisaccades and prosaccades. *Frontiers in Psychology*, https://doi.org/10.3389/fpsyg.2013.00749.

– *Recent meta-analysis of imaging studies of antisaccade task.*

12.10 Questions to Students

1. Discuss the difference between T1, T2 and T2* relaxation times. How does this affect what the MR image will show?
2. Outline the physiology of the BOLD-fMRI response. What implications does the BOLD physiology have on the kinds of questions we can use fMRI to study? How does this affect how we use fMRI in eye movement research?
3. Design two simple example fMRI experiments using a

   i. block design and an
   ii. event-related design

   for an eye-movement task. This doesn't have to be a long experiment, just a few (say 10–15) trials of a task. Consider trial durations, trial transitions, stimulus-response maps etc. Trial transitions in particular will differ quite a lot between the two experimental designs. You may also want to consider the practical aspects of how you would set up such an experiment on your MRI scanner.

(a) Contrast the advantages and disadvantages of each task—consider psychological aspects (expectancy, task contingencies, etc.) and technical aspects (signal to noise, detection power).
(b) In what situations would the block design be better suited? When would the event-related design be more suitable?
(c) What will your general linear model look like? What factors will you model? How will you deal with instances where things don't go to plan—what will you do if participants make an error? How will you deal with trials where the eye tracking signal drops out and you have no eye movement measure?
(d) What will you use as a baseline in your experiment? What are the advantages/disadvantages of your choice?
(e) Consider which regions of interest you would like to target in your experiment. Will you need to consider specialised acquisition/analysis strategies to adequately image these regions?

Acknowledgements S. Jamadar is supported by an Australian Research Council Discovery Early Career Researcher Award (ARC DECRA) DE150100406. B. P. Johnson is supported by a National Health and Medical Research (NHMRC) Peter Doherty Biomedical Early Career Research Fellowship (APP1112348).

References

Agam, Y., Joseph, R. M., Barton, J. J. S., & Manoach, D. S. (2010). Reduced cognitive control of response inhibition by the anterior cingulate cortex in autism spectrum disorders. *NeuroImage, 52,* 336–347.

Aichert, D. S., Williams, S. C. R., Moller, H.-J., Kumari, V., & Ettinger, U. (2012). Functional neural correlates of psychometric schizotypy: An fMRI study of antisaccades. *Psychophys, 49,* 345–356.

Allport, A., Styles, E.A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing*. MIT Press.

Atwell, D., & Iadecola, C. (2002). The neural basis of functional brain imaging signals. *Trends in Neurosciences, 25,* 621–625.

Bandettini, P. A. (2012). Twenty years of functional MRI: The science and the stories. *Neuroimage, 62*, 575–588.

Becker, W., & Fuchs, A. F. (1985). Prediction in the oculomotor system: Smooth pursuit during transient disappearance of a visual target. *Experimental Brain Research, 57*(3), 562–575.

Biswal, B., Zerrin Yetkin, F., Haughton, V. M., & Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magnetic Resonance in Medicine, 34*(4), 537–541.

Brown, M. R. G., Goltz, H. C., Vilis, T., Ford, K. A., & Everling, S. (2006). Inhibition and generation of saccades: Rapid event-related fMRI of prosaccades, antisaccades and no-go trials. *Neuroimage, 33,* 644–659.

Brown, M. R. G., Vilis, T., & Everling, S. (2007). Frontoparietal activation with preparation for antisaccades. *Journal of Neurophysiology, 98,* 1751–1762.

Bullmore, E. T., Brammer, M. J., Rabe-Hesketh, S., Curtis, V. A., Morris, R. G., Williams, S. C. R., et al. (1999). Methods for diagnosis and treatment of stimulus-correlated motion in generic brain activation studies using fMRI. *Human Brain Mapping, 7*(1), 38–48.

Burke, M. R., & Barnes, G. R. (2008). Brain and behavior: A task-dependent eye movement study. *Cerebral Cortex, 18*(1), 126–135.

Buxton, R. B. (2012). Dynamic models of BOLD contrast. *Neuroimage, 62,* 953–961.

Buxton, R. B., Wong, E. C., & Frank, L. R. (1998). Dynamics of blood flow and oxygenation changes during brain activation: The balloon model. *Magnetic Resonance in Medicine, 39,* 855–864.

Calhoun, V. D., Liu, J., & Adali, T. (2009). A review of group ICA for fMRI data and ICA for joint inference of imaging, genetic and ERP data. *Neuroimage, 45,* S163–S172.

Camchong, J., Dyckman, K. A., Austin, B. P., Clementz, B. A., & McDowell, J. E. (2008). Common neural circuitry supporting volitional saccades and its disruption in schizophrenia patients and relatives. *Biological Psychiatry, 64,* 1042–1050.

Cameron, I. G. M., Pari, G., Alahyane, N., Brien, D. C., Coe, B. C., Stroman, P. W., et al. (2012). Impaired executive function signals in motor brain regions in Parkinson's disease. *Neuroimage, 60,* 1156–1170.

Chikazoe, J., Konishi, S., Asari, T., Jimura, K., & Miyashita, Y. (2007). Activation of right inferior frontal gyrus during response inhibition across modalities. *Journal of Cognitive Neuroscience, 19,* 69–80.

Coles, M. G. H., Smid, H. G. O. M., Scheffers, M. K., & Otten, L. J. (1994). Mental chronometry and the study of human information processing. In M. D. Rugg & M. G. H. Coles (Eds.), *Electrophysiology of mind: Event-related brain potentials and cognition* (pp. 86–131). USA: Oxford University Press.

Connolly, J. D., Goodale, M. A., Cant, J. S., & Munoz, D. P. (2007). Effector-specific fields for motor preparation in the human frontal cortex. *Neuroimage, 34,* 1209–1219.

Connolly, J. D., Goodale, M. A., Goltz, H. C., & Munoz, D. P. (2005). fMRI activation in the human frontal eye field is correlated with saccadic reaction time. *Journal of Neurophysiology, 94,* 605–611.

Connolly, J. D., Goodale, M. A., Menon, R. S., & Munoz, D. P. (2002). Human fMRI evidence for the neural correlates of preparatory set. *Nature Neuroscience, 5,* 1345–1352.

Curtis, C. E., & Connolly, J. D. (2008). Saccade preparation signals in the human frontal and parietal cortices. *Journal of Neurophysiology, 99,* 133–145.

Curtis, C. E., Rao, V. Y., & D'Esposito, M. (2004). Maintenance of spatial and motor codes during oculomotor delayed response tasks. *Journal of Neuroscience, 24,* 3944–3952.

Curtis, C. E., Sun, F. T., Miller, L. M., & D'Esposito, M. (2005). Coherence between fMRI time-series distinguishes two spatial working memory networks. *Neuroimage, 26,* 177–183.

de Zwart, J. A., van Gelderen, O., Fukunaga, M., Duyn, J. H. (2008). Reducing correlated noise in fMRI data. *Magnetic Resonance Medicine, 59*, 939–945.

De Weijer, A. D., Mandl, R. C. W., Sommer, I. E. C., Vink, M., Kahn, R. S., & Neggers, S. F. W. (2010). Human fronto-tectal and fronto-striatal-tectal pathways activate differently during anti-saccades. *Frontiers in Human Neuroscience, 4,* 41.

DeSouza, J. F., Menon, R. S., & Everling, S. (2003). Preparatory set associated with pro-saccades and anti-saccades in humans investigated with event-related fMRI. *Journal of Neurophysiology, 89,* 1016–1023.

Diedrichsen, J. (2006). A spatially unbiased atlas template of the human cerebellum. *Neuroimage, 33,* 127–138.

Diedrichsen, J., Balsters, J. H., Flavell, J., Cussans, E., & Ramnani, N. (2009). A probabilistic MR atlas of the human cerebellum. *Neuroimage, 46,* 39–49.

Diedrichsen, J., Verstynen, T., Schlerf, J., & Wiestler, T. (2010). Advances in functional imaging of the human cerebellum. *Current Opinion in Neurology, 23,* 382–387.

Donders, F. C. (1868). On the speed of mental processes. In W. G. Koster (Ed.), *Acta Psychologica, 30: Attention and Performance II* (1969), 412–431.

Dum, R. P., & Strick, P. L. (2003). An unfolded map of the cerebellar dentate nucleus and its projections to the cerebral cortex. *Journal of Neurophysiology, 89,* 634–639.

Duong, T. Q., Yacoub, E., Adriany, G., Hu, X., Ugurbil, K., et al. (2002). High-resolution, spin-echo BOLD and CBF fMRI at 4 and 7 T. *Magnetic Resonance Medicine, 48,* 589–593.

Duzel, E., Guitart-Masip, M., Maass, A., et al. (2015). Midbrain fMRI: applications, limitations and challenges. In K. Uludag et al. (Eds.). *fMRI: From Nuclear Spins to Brain Functions*. Biological Magnetic Resonance 30. New York: Springer.

Dyckman, K. A., Camchong, J., Clementz, B. A., & McDowell, J. E. (2007). An effect of context on saccade-related behaviour and brain activity. *Neuroimage, 36,* 774–784.

Dyckman, K. A., Lee, A. K. C., Agam, A., Vangel, M., Goff, D. C., Barton, J. J. S., et al. (2011). Abnormally persistent fMRI activation during antisaccades in schizophrenia: A neural correlate of perseveration? *Schizophrenia Research, 132,* 62–68.

Dyckman, K. A., & McDowell, J. E. (2005). Behavioral plasticity of antisaccade performance following daily practice. *Experimental Brain Research, 162,* 63–69.

Ecklund, A., Nichols, T. E., & Knutsson, H. (2016). Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. *PNAS, 113,* 7900–7905.

Enderle, J. D. (2002). Neural control of saccades. In J. Hyona, D. P. Munoz, W. Heide, & R. Radach (Eds.), *Progress in brain research* (Vol. 140, pp. 21–49).

Ettinger, U., Ffytche, D., Kumari, V., Kathmann, N., Reuter, B., Zelaya, F., et al. (2008). Decomposing the neural correlates of antisaccade eye movements using event-related fMRI. *Cerebral Cortex, 18,* 1148–1159.

Ettinger, U., Williams, S. C. R., Patel, D., Michel, T. M., Nwaigwe, A., Caceres, A., et al. (2009). Effects of acute nicotine on brain function in healthy smokers and non-smokers: Estimation of inter-individual response heterogeneity. *Neuroimage, 45,* 549–561.

Feinberg, D. A., & Setsompop, K. (2013). Ultra-fast MRI of the human brain with simultaneous multi-slice imaging. *Journal of Magnetic Resonance, 229,* 90–100.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., et al. (2002). Whole-brain segmentation: Automated labeling of neuroanatomical structures in the human brain. *Neuron, 31,* 341–355.

Ford, K. A., Goltz, H. C., Brown, M. R. G., & Everling, S. (2005). Neural processes associated with antisaccade task performance investigated with event-related fMR. *Neurophysiology, 94,* 429–440.

Fornito, A., Zalesky, A., & Breakspear, M. (2015). The connectomics of brain disorders. *Nature Neuroscience, 16,* 159–172.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience, 8*(9), 700–711.

Friston, K. J. (1997). Introduction: Experimental design and statistical parametric mapping. In R. S. J. Frackowiak, K. J. Friston, C. Frith, R. Dolan, & J. C. Mazziotta (Eds.), *Human brain function*. USA: Academic Press.

Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connectivity, 1,* 13–36.

Friston, K. J., Price, C. J., Fletcher, P., Moore, C., Frackowiak, R. S. J., & Dolan, R. J. (1996a). The trouble with cognitive subtraction. *Neuroimage, 4,* 97–104.

Friston, K. J., Williams, S., Howard, R., Frackowiak, R. S., & Turner, R. (1996b). Movement-related effects in fMRI time-series. *Magnetic Resonance Medicine, 35,* 346–355.

Fuchs, A. F. (1967). Saccadic and smooth pursuit eye movements in the monkey. *The Journal of Physiology, 191*(3), 609–631.

Fuchs, A. F., & Robinson, D. A. (1966). A method for measuring horizontal and vertical eye movement chronically in the monkey. *Journal of Applied Physiology, 21*(3), 1068–1070.

Fukumoto-Motoshita, N., Matsuura, M., Ohkubo, T., Ohkubo, H., Kanaka, N., Matsushima, E., et al. (2009). Hyperfrontality in patients with schizophrenia during saccade and antisaccade tasks: A study with fMRI. *Psychiatry and Clinical Neurosciences, 63,* 209–217.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331–349.

Glover, G. H., Li, T. Q., & Ress, D. (2000). Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magnetic Resonance in Medicine, 44,* 162–167.

Greenlee, M. W. (2000). Human cortical areas underlying the perception of optic flow: Brain imaging studies. *International Review of Neurobiology, 44,* 269–292.

Guimaraes, A. R., Melcher, J. R., Talavage, T. M., Baker, J. R., Ledden, P., Rosen, B. R., et al. (1998). Imaging subcortical auditory activity in humans. *Human Brain Mapping, 6,* 33–41.

Gusnard, D. A., & Raichle, M. E. (2001). Searching for a baseline: Functional imaging and the resting human brain. *Nature Reviews Neuroscience, 2*(10), 685–694.

Habas, C. (2010). Functional imaging of the deep cerebellar nuclei: A review. *Cerebellum, 9,* 22–28.

Hajnal, J. V., Myers, R., Oatridge, A., Schwieso, J. E., Young, I. R., & Bydder, G. M. (1994). Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magnetic Resonance in Medicine, 31*(3), 283–291.

Haynes, J.-D., & Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience, 7,* 523–534.

Helmholtz, H. V. (1853). Ueber einige Gesetze der Vertheilung elektrischer Ströme in körperlichen Leitern mit Anwendung auf die thierisch-elektrischen Versuche. *Annalen der Physik, 165*(6), 211–233.

Henson, R. N. A. (2004). Analysis of fMRI time series. In R. S. J. Frackowiak, K. J. Friston, C. D. Frith, R. J. Dolan, C. J. Price, S. Zeki, J. T. Ashburner, & W. D. Penny (Eds.), *Human brain function* (2nd ed.). USA: Academic Press.

Henson, R. N. A., Buchel, C., Josephs, O., & Friston, K. J. (1999). The slice-timing problem in event-related fMRI. *Neuroimage, 9,* 125.

Hikosaka, O., Takikawa, Y., & Kawagoe, R. (2000). Role of the basal ganglia in the control of purposive saccadic eye movements. *Physiological Reviews, 80*(3), 953–978.

Hoge, R. D., Atkinson, J., Gill, B., Crelier, G. R., Marrett, S., & Pike, G. B. (1999). Linear coupling between cerebral blood flow and oxygen consumption in activated human cortex. *Proceedings of the National Academy of Sciences, 96,* 9403–9408.

Huettel, S. A., Song, A. W., & McCarthy, G. (2008). *Functional magnetic resonance imaging*. Sunderland, Massachusetts: Sinauer Associates Inc.

Jagla, F., Jergelova, M., & Riecanský, I. (2007). Saccadic eye movement related potentials. *Physiological Research, 56*(6), 707.

Jamadar, S. D., Fielding, J., & Egan, G. F. (2013). Quantitative meta-analysis of fMRI and PET studies reveals consistent activation in fronto-striatal-parietal regions and cerebellum during antisaccades and prosaccades. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2013.00749.

Jezzard, P., & Balaban, R. S. (1995). Correction for genoetric distortion in echo planar images from B0 field variations. *Magnetic Resonance Medicine, 34,* 65–73.

Johnston, K., & Everling, S. (2008). Neurophysiology and neuroanatomy of reflexive and voluntary saccades in non-human primates. *Brain and Cognition, 68*(3), 271–283.

Karayanidis, F., & Jamadar, S. (2014). ERP measures of proactive and reactive control processes in task-switching paradigms. *Task Switching and Cognitive Control*, 200–236.

Kaza, E., Klose, U., & Lotze, M. (2011). Comparison of a 32-channel with a 12-channel head coil: Are the relevant improvements for functional imaging? *Journal of Magnetic Resonance Imaging, 34,* 173–183.

Keedy, S. K., Ebens, C. L., Keshavan, M. S., & Sweeney, J. A. (2006). Functional magnetic resonance imaging studies of eye movements in first episode schizophrenia: Smooth pursuit, visually guided saccades and the oculomotor delayed response task. *Psychiatry Research: Neuroimaging, 146*(3), 199–211.

Kimmig, H., Greenlee, M. W., Gondan, M., Schira, M., Kassubek, J., & Mergner, T. (2001). Relationship between saccadic eye movements and cortical activity as measured by fMRI: Quantitative and qualitative aspects. *Experimental Brain Research, 141,* 184–194.

Kimmig, H., Ohlendorf, S., Speck, O., Sprenger, A., Rutschmann, R. M., Haller, S., & Greenlee, M. W. (2008). fMRI evidence for sensorimotor transformations in human cortex during smooth pursuit eye movements. *Neuropsychologia, 46*(8), 2203–2213.

Kozlova, G. P. (1984). Individual anatomic variations in cerebellar nuclei. *Neuroscience and Behavioural Physiology, 14,* 63–67.

Kulpe, O. (1909). *Outlines of psychology: Based upon the results of experimental investigation* (3rd ed.) (translation of original work published in 1893). New York: Macmillan.

Kuper, M., Thurling, M., Maderwald, S., Ladd, M. E., & Timmann, D. (2012). Structural and functional magnetic resonance imaging of the human cerebellar nuclei. *Cerebellum, 11,* 314–324.

Lee, J., Park, C., Dyckman, K. A., Lazar, N. A., Austin, B. P., Li, Q., et al. (2012). Practice-related changes in neural activation patterns investigated via wavelet-based clustering analysis. *Human Brain Mapping, 34,* 2276–2291.

Limbrick-Oldfield, E. H., Brooks, J. C. W., Wise, R. J. S., Padormo, F., Hajnal, J. V., Beckmann, C. F., Ungless, M. A. (2012). Identification and characterization of midbrain nuclei using optimised functional magnetic resonance imaging. *Neuroimage, 59,* 1230–1238.

Lisberger, S. G., Evinger, C., Johanson, G. W., & Fuchs, A. F. (1981). Relationship between eye acceleration and retinal image velocity during foveal smooth pursuit in man and monkey. *Journal of Neurophysiology, 46*(2), 229–249.

Liu, T., & Frank, L. R. (2004). Efficiency, power and entropy in event-related fMRI with multiple trial types. Part 1: Theory. *Neuroimage, 21,* 387–400.

Logothetis, N. K. (2007). The ins and outs of fMRI signals. *Nature Neuroscience, 10,* 1230–1232.

Luck, S. J. (2005). *An Introduction to the event-related potential technique*. Cambridge: MIT Press.

Luna, B., Minshew, N. J., Garver, K. E., Lazar, N. A., Thulborn, K. R., Eddy, W. F., & Sweeney, J. A. (2002). Neocortical system abnormalities in autism An fMRI study of spatial working memory. *Neurology, 59*(6), 834–840.

Lund, J. S., Lund, R. D., Hendrickson, A. E., Bunt, A. H., & Fuchs, A. F. (1975). The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by retrograde transport of horseradish peroxidase. *Journal of Comparative Neurology, 164*(3), 287–303.

Manoach, D. S., Thakkar, K. N., Cain, M. S., Polli, F. E., Edelman, J. A., Fischl, B., Barton, J. J. S. (2007). Neural activity is modulated by trial history: A functional magnetic resonance imaging study of the effects of a previous antisaccade. *The Journal of Neuroscience, 27,* 1791–1798.

Matsuda, T., Matsuura, M., Ohkubo, T., Ohkubo, H., Matsushima, E., Inoue, K., et al. (2004). Functional MRI mapping of brain activation during visually guided saccades and antisaccades: Cortical and subcortical networks. *Psychiatry Research: Neuroimaging, 131,* 147–155.

McDowell, J. E., Dyckman, K. A., Austin, B. P., & Clementz, B. A. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades: Evidence from studies of humans. *Brain and Cognition, 68*(3), 255–270.

Nagel, M., Sprenger, A., Nitschke, M., Zapf, S., Heide, W., Binkofski, F., et al. (2007). Different extraretinal neuronal mechanisms of smooth pursuit eye movements in schizophrenia: An fMRI study. *Neuroimage, 34*(1), 300–309.

Neggers, S. F. W., van Diepen, R. M., Zandbelt, B. B., Vink, M., Mandl, R. C. W., & Gutteling, T. P. (2012). A functional and structural investigation of the human fronto-basal volitional saccade network. *PLoS ONE, 7.*

Nichols, T. E., & Hayasaka, S. (2003). Controlling the family-wise error rate in functional neuroimaging: A comparative review. *Statistical Methods in Medical Research, 12,* 419–446.

Ogawa, S., Lee, T. M., Kay, A. R., & Tank, D. W. (1990). Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proceedings of the National Academy of Science, 87,* 9868–9872.

Ozyurt, J., & Greenlee, M. W. (2011). Neural correlates of inter- and intra-individual saccadic reaction time differences in the gap/overlap paradigm. *Journal of Neurophysiology, 105,* 2438–2447.

Pascual-Marqui, R. D., Michel, C. M., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology, 18*(1), 49–65.

Paus, T., Petrides, M., Evans, A. C., & Meyer, E. (1993). Role of the human anterior cingulate cortex in the control of oculomotor, manual, and speech responses: A positron emission tomography study. *Journal of Neurophysiology, 70*(2), 453–469.

Petit, L., & Haxby, J. V. (1999). Functional anatomy of pursuit eye movements in humans as revealed by fMRI. *Journal of Neurophysiology, 82*(1), 463–471.

Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current Opinion in Neurology, 17*(1), 17–25.

Poldrack, R. A. (2007). Region of interest analysis for fMRI. *Social Cognitive and Affective Neuroscience, 2*(1), 67–70.

Poldrack, R. A., Mumford, J. A., & Nichols, T. E. (2011). *Handbook of functional MRI data analysis*. New York: Cambridge University Press.

Poncelet, B. P., Wedeen, V. J., Weisskoff, R. M., & Cohen, M. S. (1992). Brain parenchyma motion: Measurement with cine echo-planar MR imaging. *Radiology, 185,* 645–651.

Power, J. D., Mitra, A., Laumann, T. O., Snyder, A. Z., Schlaggar, B. L., & Petersen, S. E. (2014). Methods to detect, characterise, and remove motion artifact in resting state fMRI. *Neuroimage, 84,* 320–341.

Raemaekers, M., Jansma, J. M., Cahn, W., Van der Geest, J. N., van der Linden, J. A., Kahn, R. S., & Ramsey, N. F. (2002). Neuronal substrate of the saccadic inhibition deficit in schizophrenia investigated with 3-dimensional event-related functional magnetic resonance imaging. *Archives of General Psychiatry, 59*, 313-320.

Raemaekers, M., Vink, M., van den Heuvel, M. P., Kahn, R. S., & Ramsey, N. J. (2006). Effects of aging on BOLD fMRI during prosaccades and antisaccades. *Journal of Cognitive Neuroscience, 18,* 594–603.

Robinson, D. A. (1970). Oculomotor unit behavior in the monkey. *Journal of Neurophysiology, 33*(3), 393–403.

Robinson, F. R. (1995). Role of the cerebellum in movement control and adaptation. *Current Opinion in Neurobiology, 5*(6), 755–762.

Robinson, F. R., & Fuchs, A. F. (2001). The role of the cerebellum in voluntary eye movements. *Annual Review of Neuroscience, 24*(1), 981–1004.

Robinson, F. R., Fuchs, A. F., & Noto, C. T. (2002). Cerebellar influences on saccade plasticity. *Annals of the New York Academy of Sciences, 956*(1), 155–163.

Rugg, M. D., & Coles, M. G. H. (Eds.). (1995). *Electrophysiology of mind: Event-related brain potentials and cognition*. Oxford psychology series, No. 25. New York, NY, US: Oxford University Press.

Ruge, H., Jamadar, S., Zimmerman, U., & Karayanidis, F. (2013). The many faces of preparatory control in task switching: Reviewing a decade of fMRI research. *Human Brain Mapping, 34,* 12–35.

Satterthwaite, T. D., Wolf, D. H., Loughead, J., Ruparel, K., Elliott, M. A., Hakonarson, H., et al. (2012). Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. *Neuroimage, 60*(1), 623–632.

Scherg, M., Vajsar, J., & Picton, T. W. (1989). A source analysis of the late human auditory evoked potentials. *Journal of Cognitive Neuroscience, 1*(4), 336–355.

Schmahmann, J. D., Doyon, J., Toga, A. et al. (2000). *MRI Atlas of the Human Cerebellum*. San Diego: Academic Press.

Schmahmann, J. D., Doyon, J., McDonald, D., Holmes, C., Lavoie, K., et al. (1999). Three-dimensional MRI atlas of the human cerebellum in proportional stereotaxic space. *Neuroimage, 10,* 233–260.

Scudder, C. A. (1988). A new local feedback model of the saccadic burst generator. *Journal of Neurophysiology, 59*(5), 1455–1475.

Scudder, C. A. (2002). Role of the fastigial nucleus in controlling horizontal saccades during adaptation. *Annals of the New York Academy of Sciences, 978*(1), 63–78.

Scudder, C. A., & Fuchs, A. F. (1992). Physiological and behavioral identification of vestibular nucleus neurons mediating the horizontal vestibuloocular reflex in trained rhesus monkeys. *Journal of Neurophysiology, 68*(1), 244–264.

Scudder, C. A., Kaneko, C. R., & Fuchs, A. F. (2002). The brainstem burst generator for saccadic eye movements. *Experimental Brain Research, 142*(4), 439–462.

Seto, E., Sela, G., McIlroy, W. E., Black, S. E., Staines, W. R., Bronskill, M. J., et al. (2001). Quantifying head motion associated with motor tasks used in fMRI. *Neuroimage, 14*(2), 284–297.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30,* 276–315.

Stoodley, C. J., Valera, E. M., & Schmahmann, J. D. (2012). Functional topography of the cerebellum for motor and cognitive tasks: An fMRI study. *Neuroimage, 59*, 1560–1570.

Sweeney, J. A., Luna, B., Keedy, S. K., McDowell, J. E., & Clementz, B. A. (2007). fMRI studies of eye movement control: Investigating the interaction of cognitive and sensorimotor brain systems. *Neuroimage, 36,* T54–T60.

Sweeney, J. A., Mintun, M. A., Kwee, S., Wiseman, M. B., Brown, D. L., Rosenberg, D. R., & Carl, J. R. (1996). Positron emission tomography study of voluntary saccadic eye movements and spatial working memory. *Journal of Neurophysiology*, *75*(1), 454–468.

Takarae, Y., Luna, B., Minshew, N. J., & Sweeney, J. A. (2014). Visual motion processing and visual sensorimotor control in autism. *Journal of the International Neuropsychological Society, 20*(01), 113–122.

Talairach, J., & Szikla, G. (1967). *Atlas of stereotactic concepts to the surgery of epilepsy.*

Talairach, J., & Tournoux, P. (1988). *Co-planar stereotaxic atlas of the human brain*. Thieme Medical Pub.

Tanabe, J., Tregellas, J., Miller, D., Ross, R. G., & Freedman, R. (2002). Brain activation during smooth-pursuit eye movements. *Neuroimage, 17,* 1315–1324.

Thomsen, K., Piilgaard, H., Gjedde, A., Bonvento, G., & Lauritzen, M. (2009). Principal cell spiking, postsynaptic excitation, and oxygen consumption in the rat cerebellar cortex. *Journal of Neurophysiology, 102,* 1503–1512.

Tregellas, J. R., Tanabe, J. L., Miller, D. E., Ross, R. G., Olincy, A., & Freedman, R. (2004). Neurobiology of smooth pursuit eye movement deficits in schizophrenia: An fMRI study. *American Journal of Psychiatry, 161*(2), 315–321.

Tu, P. C., Yang, T. H., Kuo, W. J., Hsieh, J. C., & Su, T. P. (2006). Neural correlates of antisaccade deficits in schizophrenia, an fMRI study. *Journal of Psychiatric Research, 40,* 606–612.

Van Dijk, K. R., Sabuncu, M. R., & Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage, 59*(1), 431–438.

Velanova, K., Wheeler, M. E., & Luna, B. (2009). The maturation of task set-related activation supports late developmental improvements in inhibitory control. *The Journal of Neuroscience, 29*(40), 12558–12567.

Voogd, J., & Barmack, N. H. (2006). Oculomotor cerebellum. *Progress in Brain Research, 151,* 231–268.

Wager, T. D., & Nichols, T. E. (2003). Optimisation of experimental design in fMRI: A general framework using a genetic algorithm. *Neuroimage, 18,* 293–309.

Wall, M. B., Walker, R., & Smith, A. T. (2009). Functional imaging of the human superior colliculus: An optimised approach. *Neuroimage, 47,* 1620–1627.

Wansapura, J. P., Holland, S. K., Dun, R. S., & Ball, W. S., Jr. (1999). NMR relaxation times in the human brain at 3.0 tesla. *Journal of Magnetic Resonance Imaging, 9,* 531–538.

Yacoub, E., Harel, N., & Ugurbil, K. (2008). High-field fMRI unveils orientation columns in humans. *Proceedings of the National Academy of Science, 105,* 10607–10612.

Chapter 13
Eye Movement Recordings in Natural Settings

Benjamin W. Tatler, Dan Witzner Hansen and Jeff B. Pelz

Contents

Abstract In this chapter we consider why it is important to study eye movements in natural, real world settings and the practical considerations that need to be taken into account when collecting eye movement data in these situations. Conducting a study in a real world environment poses very different challenges than those present in

---

B. W. Tatler (✉)
School of Psychology, University of Aberdeen, Aberdeen, UK
e-mail: b.w.tatler@abdn.ac.uk

D. W. Hansen
IT University Copenhagen, Copenhagen, Denmark

J. B. Pelz
Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, USA

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_13

laboratory-based settings. Variables are hard to control and this restricts the kinds of paradigms that can be employed. Thus, careful consideration of the research question and whether it is appropriate for real world study is a necessary first step. Mobile eye trackers are often the most obvious choice for studying eye movements in real world settings. There exists a range of mobile eye tracking systems available to buy, and researchers also build their own mobile eye trackers to solve particular issues. Selecting the right system is an important consideration for mobile eye tracking and in this chapter we will highlight some of the key choices that the researcher should consider when buying or building a mobile eye tracker. Care must be taken to ensure that the data collected are sufficiently reliable to address the research question. While the principles of eye tracking - how we detect features of the eye and use these to estimate where someone is looking - are very similar for all eye trackers, the challenges faced when collecting data in natural settings are far greater than those that face eye trackers in a controlled laboratory setting. We will consider the key threats that real world settings pose for collecting reliable eye tracking data in the hope that we not only raise awareness of these issues but also offer some suggestions for how these can be minimised or at least identified in the data. For those considering whether to run an eye tracking study in the real world, we hope that this chapter will help you as you consider your research questions and choices of mobile eye tracker. For those already committed to studying eye movement behaviour in a real world setting, we hope that our discussion of the problems and possible solutions to the challenges faced in such settings will help minimise data loss from the common challenges that we face when conducting mobile eye tracking studies.

13.1 Introduction

Running experiments in the laboratory has many attractions, especially when running eye tracking experiments. In the laboratory, it is possible to precisely control variables and bulky eye tracking equipment can be accommodated without interfering with the experimental protocol. The latter consideration certainly constrained early eye tracking studies to the laboratory given the bulky nature of early photographic eye trackers (Fig. 13.1).

However, if our goal is to understand human behaviour then there is a potential risk in exclusively running reductionist paradigms in the laboratory. The perceptual and cognitive processes that we aim to understand are not isolated systems but are part of a larger organism that interacts with the world around it. Moreover, perceptual and cognitive processes are influenced by these interactions: our ability to detect and recognise objects, and to make perceptual and cognitive judgements about them are changed in the presence of action (Barsalou, 2008; Bekkering & Neggers, 2002; Fagioli, Hommel, & Schubotz, 2007). Gaze is no different: how we deploy our gaze differs in the presence of action, compared to situations in which we do no interact with the world (Epelboim et al., 1995). Indeed, if we simplify any aspect of the task - the responses that we require participants to make, or the stimuli that we ask

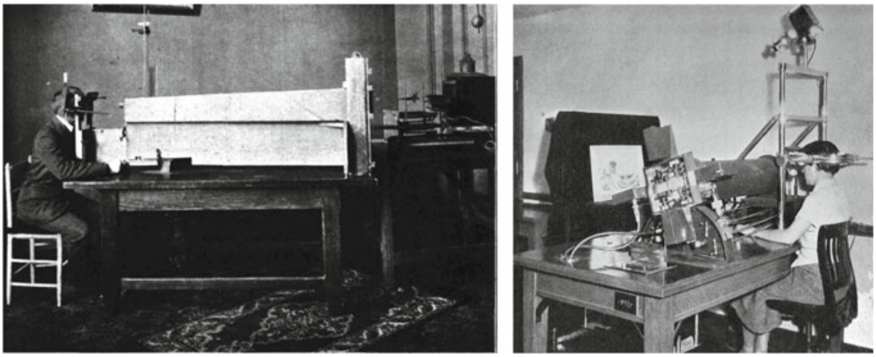


Fig. 13.1 Left, the tracker developed by Raymond Dodge and used in Diefendorf and Dodge (1908). Right, Guy Buswell's eye tracker (Buswell 1935)

them to respond to - we run the risk of disrupting the behaviour that we set out to investigate. A clear example of this can be found in how goal keepers respond to a free kick in football. Dicks, Button, and Davids (2010) showed that simplifying either the response required (from an actual save, to a manual indication of direction or a verbal response) or the stimulus (from an actual individual kicking the ball toward the goal keeper to a life-sized video of the kick) changed the goal keeper's gaze behaviour. This study demonstrates that any attempt to simplify a behaviour to make it easier to control and study experimentally can disrupt the normal use of gaze in that behaviour.

How well laboratory-based studies of cognition and behaviour generalise to real world behaviours, how these two settings differ and why we should study cognition and behaviour in the real world have become topics of growing interest and prominence in research. For more in-depth discussions of these issues and comparisons between findings from the laboratory and real world we recommend that the reader look at Foulsham, Walker, and Kingstone (2011), Kingstone, Laidlaw, Nasiopoulos, and Risko (2017), Kingstone, Smilek, and Eastwood (2008), Tatler and Land (2016) and the 2017 special issue of the *Canadian Journal of Experimental Psychology* on this topic (Risko & Kingstone, 2017). Unfortunately, we cannot simply translate experimental approaches and practices from the laboratory to real life. Natural environments remove the ability to precisely control the environment and conditions in which an experiment is conducted. A cornerstone of experimental psychology in the laboratory is the precise control over experimental and extraneous variables. This is not always possible in natural settings. For example, imagine that we wish to conduct an experiment about the influence of mobile phone distraction on driving in an urban environment. We simply cannot control environmental conditions like lighting and weather; nor can we control the amount of traffic on the road, the behaviour of other drivers or the behaviour of pedestrians. Even controlling the experimental variable under test - in this case whether the driver is using a mobile phone or not - can be very hard to achieve in real world environments. Indeed, in this example, the experiment itself would be rather hard get off the ground: placing drivers in a busy urban environment and then distracting them might not impress many ethics boards. But knowing the effects of distraction during driving are obviously important for public safety and can offer considerable theoretical insights into our understanding of vision, attention and cognition, so how do we study this? It is not surprising that most researchers choose to tackle issues like this one in more controlled settings, using virtual reality, driving simulators or reductionist paradigms typical of experimental psychology. In such situations variables can be controlled precisely. But these laboratory-based approaches make a critical sacrifice: the participant knows that the setting is not real and so the consequences of their actions are not the same as in the natural setting. Driving into another car in virtual reality might be rather embarrassing but it has very different consequences than doing so in the real world. Removing the real consequences of actions by designing more controlled settings in the laboratory might mean that we are no longer really studying the behaviour that we set out to understand (Kingstone et al., 2008).

This imagined example illustrates the challenges and opportunities that face those interested in studying aspects of natural human behaviour. A natural setting sacrifices experimental control, but more than this it constrains the types of research question that can feasibly be asked. But in the end it is clear that if we avoid studying the behaviour in its natural context then we risk failing to study the true behaviour.

Of course, if our aim is to study eye movements in natural behaviour then it is important to remember that this aim may be impossible to achieve! Until mobile eye trackers are small enough to be deployed covertly (but with careful ethical consideration) it will remain that participants know that their eyes are being tracked during the study. Simply knowing that your eyes are being tracked is enough to alter eye movement behaviour (Risko & Kingstone, 2011). How much of a problem this is will likely depend on the research question - awareness of eye movement recording will perhaps be more likely to impact behaviour during social interactions than in tasks such as driving or making tea. Other threats to truly natural behaviour come from any compromises that the researcher makes when trying to overcome the challenges of collecting reliable real world data (see Sect. 13.6). For example, even asking participants to wear a peaked cap during the study could be considered a manipulation and thus a step away from the natural behaviour. Without some compromise - and the necessity of wearing a mobile eye tracker - study of eye movements in real world settings is not possible. It is up to the researcher to be aware of these issues and to decide what compromises are acceptable for their research aims and how these could impact the study.

A practical limit on the kinds of questions that can be addressed in natural settings comes from the current limitations in eye tracking hardware (e.g., temporal sampling capabilities of eye trackers) and robustness of gaze estimation (e.g., the need to deal with greater challenges from environmental conditions than in the laboratory). Thus a key part of our discussion will be about how to ensure good data collection, and we will relate this to how eye trackers work and where threats to good data are particularly important and problematic for real world settings. Toward the end of the chapter we discuss the next steps in data analysis once the study has been collected and illustrate why even the most basic first step that is taken in laboratory-based eye tracking analyses - parsing data into saccades and fixations - can be very challenging in mobile eye tracking data.

13.2 Learning Objectives

By the end of this chapter, you should understand the following key issues for conducting mobile eye tracking research:

- The different hardware choices for mobile eye trackers
- That the task and environment are key factors in deciding what eye tracker you need
- That not all research questions are feasible in real world settings
- The challenges that are faced for getting reliable mobile eye tracking data, and how to mitigate against them
- That detecting saccades and fixations is much harder for mobile eye tracking data than for eye tracking data collected in the lab
- The importance of having access to the image of the eye recorded by the eye tracker and (if possible) the fit of the gaze estimation model to this image.

Terminology used in this chapter. In this chapter we use the term *gaze* to refer to the *point of regard* (the point in 3D space where a person is looking), and *gaze direction* to refer to the direction in which a person is looking. A change in gaze direction (a *gaze shift*) in real world settings is achieved by the coordinated combination of eye, head and body turns. We refer to such changes of gaze direction as *saccades*, *eye movements* or *gaze shifts* in the sections that follow. *Gaze behaviour* is used to refer to the manner in which gaze is deployed to support our activities and to gather information from the world around us. In the case of laboratory-based eye movement data, the raw samples from the eye tracker can be classified into saccade and fixation events. For mobile eye tracking data, the estimates of gaze calculated by the eye tracker (which we refer to as *gaze samples*) can be classified into similar events, which we refer to as *gaze events*.

13.3 A Brief History of Mobile Eye Tracking

The need to take eye movement research out of the laboratory and into more natural settings has been recognised for some time.

Hartridge and Thomson (1948) discussed a number of methods for tracking eye movements. They considered the advantages and disadvantages of tracking the cornea, iris, sclera, retina, and a contact lens firmly attached to the cornea, weighing the relative benefits of using photographic methods and photoelectric measurements. The paper is impressive in its scope and creativity, including suggestions to tattoo the sclera to provide trackable landmarks; projecting a bright light onto the blind spot; and attaching a scleral mirror to magnify rotations of the eye. They proposed an ingenious optical system that allowed eye movements to be measured with great precision without requiring the head to be fixed in position. This was accomplished by suspending the optical apparatus seen in Fig. 13.2a in front of an observer, who

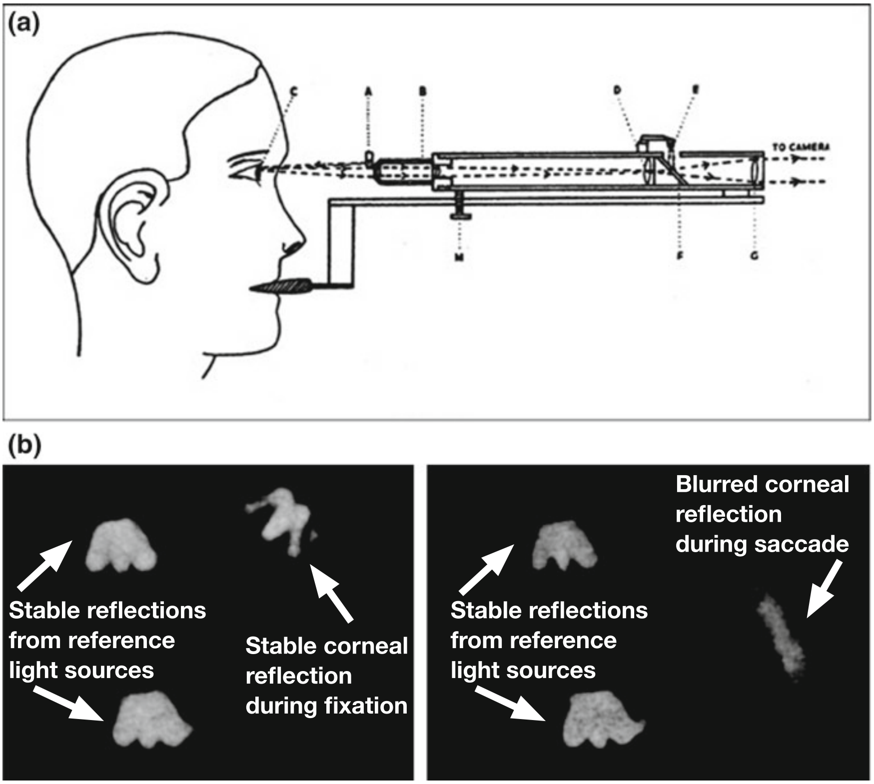


Fig. 13.2 **a** The apparatus proposed by Hartridge and Thomson (1948). **b** Two reference sources and corneal reflection from Hartridge and Thomson (1948), Figs. 4 and 5. Left, a frame captured during a fixation, showing a stable image of the corneal reflection of the light source. Right, a frame captured during a saccade; the corneal reflection image is blurred because it is moving during this frame

stabilised the system with a mouth plate. The optical system provided a light source that produced a corneal reflection, two additional reference light sources, and a collimator that allowed all three images to be photographed with a cine camera running at up to 64 frames per second at high magnification. Figure 13.2b shows two frames from the film. On the left is a frame captured during a fixation, showing the stable corneal reflection to the right of the reference sources. On the right is a frame captured during a saccade, showing the blurred corneal reflection. The system offered three important advantages over a previous design in which the microscope was suspended from a plaster hat fitted to the observer's head: the system did not inhibit head movements, those head movements could be recorded by measuring the offset of the reflected reference sources, and artefacts due to heart beats that were caused by movement of the microscope in the older system were avoided.

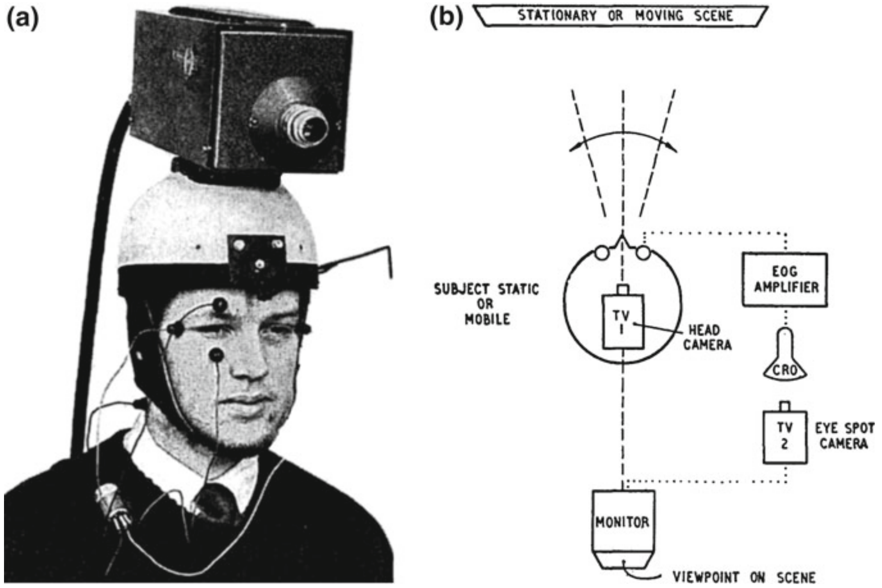


Fig. 13.3 **a** The head-mounted camera and EOG-electrodes from Shackel (1960). **b** Schematic layout of the eye tracker shown in (**a**)

Shackel (1960) built a hybrid system that combined video recording with electro-oculography (EOG). In EOG, electrodes placed on the skin, near the eyes can be used to detect rotations of the electrostatic dipole and these can be used to infer eye position information. Shackel's "mobile eye viewpoint recording" system monitored gaze via electro-oculography, but superimposed the resultant 'viewpoint' over a video recording of the scene by mixing the video from a head-mounted video camera. Figure 13.3a shows the apparatus and Fig. 13.3b shows the manner in which the EOG signals were converted to a visual signal and mixed with the scene video.

While Shackel implemented the hybrid video/EOG system shown in Fig. 13.3, he also proposed an all-optical system that monitored gaze by utilising the corneal reflection and mixing it optically before capture by the video camera. Mackworth and Thomas (1962) implemented the all-optical system that Shackel (1960) proposed, mounting an 8 mm motion-picture camera on the observer's head for a system they termed a "Head-Mounted Eye-Marker Camera". Figure 13.4a shows the system in schematic form and attached to a helmet mounted on an observer's head. Because the system worked by overlaying the image of the illuminator reflected from the eye (the corneal reflection, CR) onto the scene image, calibration consisted of adjusting the magnification and offset of the CR image so that the overlay image was in the correct position. They reported a calibrated accuracy of approximately 2° horizontally and 4° vertically near the centre of the camera's field of view. They suggested a literal

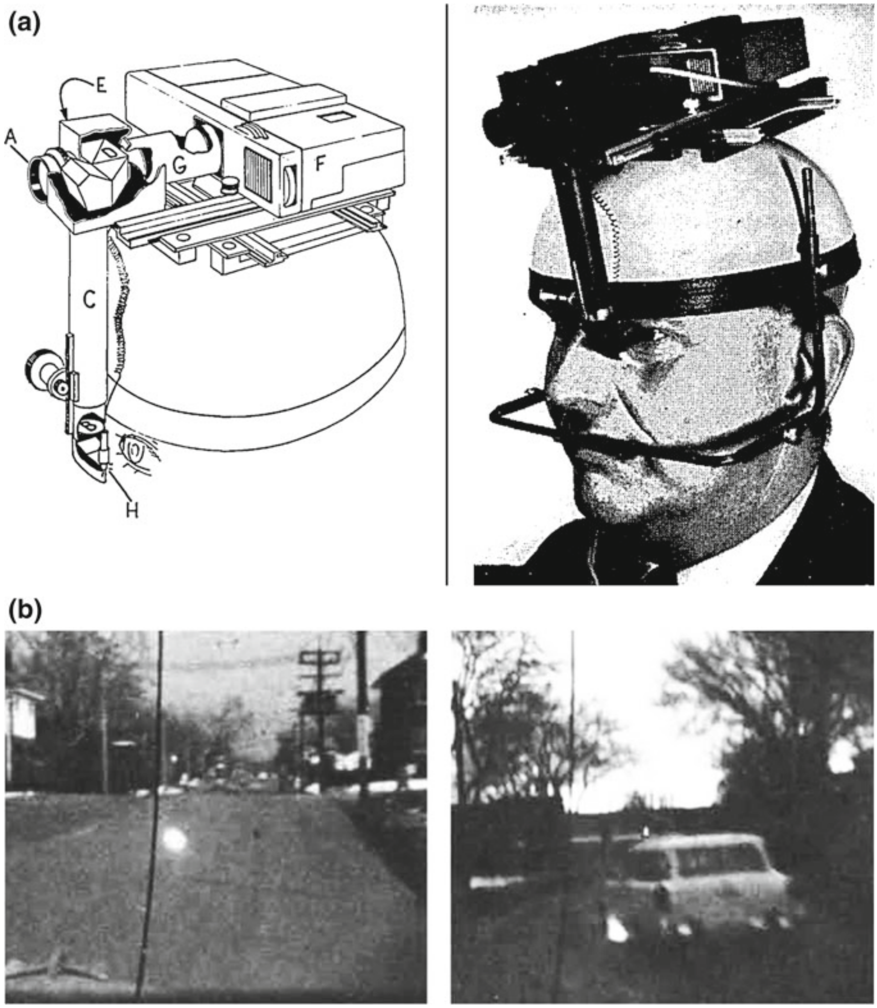


Fig. 13.4 **a** The "Head-mounted Eye-Marker Camera" system used by Mackworth and Thomas (1962). **b** Frames from 8 mm film indicating drivers' gaze as superimposed corneal reflection

rule of thumb to describe this performance, "the spot will usually fall within plus or minus a thumb's breadth of the true gaze position." (Mackworth & Thomas, 1962, p. 716). Thomas (1968) used this system to study the eye movements of drivers in traffic. Rather than report numerical angular gaze positions, Thomas' results were displayed as motion-picture frames with a light point representing the overlaid CR indicating gaze in the frame, as shown in (Fig. 13.4b).

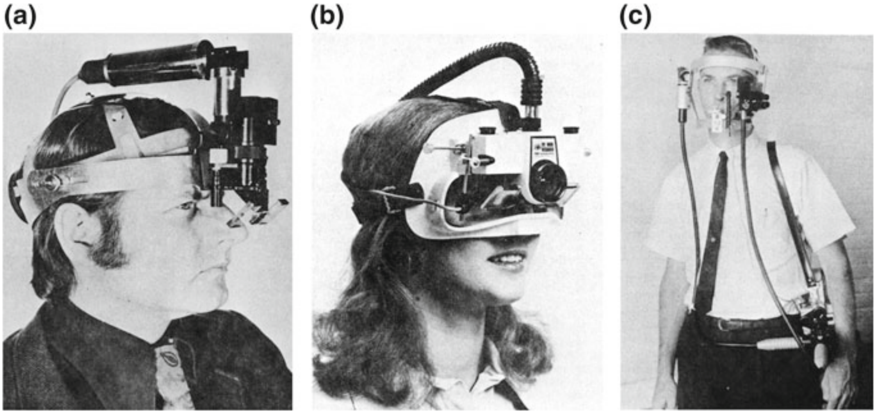


Fig. 13.5 **a** Wearable corneal-reflection tracking system by NAC Visual Systems (Young & Sheena, 1975). **b** Fiber-optics coupled corneal-reflection tracking system by NAC Visual Systems (Young & Sheena, 1975). **c** Fiber-optics coupled corneal-reflection tracking system by Polymetric Company (Young & Sheena, 1975)

These systems were made smaller and lighter by miniaturizing the optical components and/or moving some of the recording apparatus off the head (Young & Sheena, 1975). Figure 13.5a shows one such system, made by NAC Visual Systems in the 1970s. Its weight was reduced to under 4 lbs and the system was stabilized with a multi-point headband. Figure 13.5b is a system from NAC that moved the recording cameras off the head, connected by a coherent fiber-optics bundle, seen above the headgear, reducing the weight of the headgear to less than 2 lbs. Figure 13.5c shows another system from the same era, made by the Polymetric Company that also used fiber-optics to reduce the weight of the instrumentation on the head. The Polymetric tracker weighed less than 2 lbs, and was stabilized with both a headband and a bite-bar to minimize movement of the imaging optics with respect to the head.

All of these systems required that relative motion between the observer and the imaging system used to monitor the eye be minimized, because if only the corneal reflection (CR) is tracked, it is impossible to distinguish between image motion due to a real eye movement and image motion due to head movement, and motion as small as 1/8 mm results in errors of 1° in the CR systems described above (Young & Sheena, 1975). Merchant (1966) made a very important contribution by first proposing a system that tracked both the pupil and the corneal reflection, allowing the disambiguation of rotational motion (due to actual eye movements) and lateral motion (due to relative motion of the observer and eye camera). In the pupil/CR system originally proposed by Merchant, the CR was created by a small bright spot on a small cathode ray tube (CRT) rather than a discrete light source so that its position could track eye movements and be maintained in the centre of the pupil. The eye, imaged onto the face of a photocathode, was tracked with a scanning photomultiplier tube that provided the error signal to keep the CR centred on the pupil and the pupil location (each at 120 Hz). The curvature of the cornea and the position of the iris within the anterior chamber are such that when imaged from in front of an observer, the image of the pupil viewed through the cornea, and the virtual image of a distant source reflected off the cornea (the corneal reflection), are nearly coincident inside the eye, so they can be focused at the same time. This coincidence allowed Merchant to use a single optical system to scan both the pupil and CR with his active CRT/scanning photomultiplier system, and allows contemporary video-based eye trackers to use a single camera/lens to track the eye.

Merchant, Morrissette, and Porterfield (1974) built on the original active CRT/scanning photomultiplier system, proposing the now wide-spread "pupil minus CR" technique for monitoring gaze angle using a stationary light source and a standard video camera. Rather than actively moving the CR to maintain it in the centre of the pupil, pupil-CR systems calibrate the vector difference between the pupil and CR to gaze position. Merchant et al. (1974) reported a typical accuracy of 1° over the useful field. This technique is still the basis of many modern approaches to estimating gaze.

Pupil-CR trackers became the standard for laboratory based work, but the next landmark in the development of of mobile eye tracking systems came some years later when Land adapted a system originally designed to measure eye movements in the mantis shrimp (Land, Marshall, Brownless, & Cronin, 1990) in order to measure human eye movements (e.g., Land, 1993). This system used a single video camera mounted on the head and pointing downwards toward mirrors that were used to record the image of the eye and scene simultaneously (Fig. 13.6). This system recorded a video comprising images of the eye and scene, and gaze estimation involved offline fitting of a computer-generated model of the eyeball, by manually aligning an ellipse

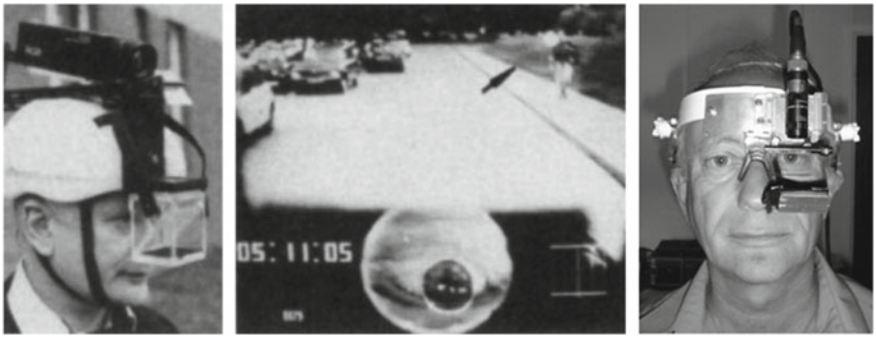


Fig. 13.6 Left, an early helmet-mounted mobile system developed by Land. Middle, a sample frame from the processed video data collected using the eye tracker shown in the left panel. The bright spot (highlighted by the black arrow in this example frame) indicates the estimate of gaze direction, derived from offline fitting of a computer-generated model of the eye. Right, a descendant of the eye tracker shown on the left that was used in a variety of studies of natural behaviour in the late 1990s (including Land et al., 1999) and throughout the 2000s

Fig. 13.7 Top left, The EyeSeeCam (Kohlbecher et al., 2010) with two scene cameras, one of which moves with the eyes to keep it aligned with the current direction of gaze. Top right, a Tobii mobile eye tracker with two cameras for each eye and a scene camera. Bottom row, Research-based eye trackers used for a skeet shooting and whitewater kayaking experiments, designed and built by researchers at IT University of Copenhagen for these purposes

with the border of the iris. This eye tracker and its immediate descendants were used in studies of a wide range of real world activities including driving (Land & Lee, 1994), table tennis (Land & Furneaux, 1997), tea making (Land, Mennie, & Rusted, 1999), musical sight reading (Furneaux & Land, 1999) and cricket (Land & McLeod, 2000).

Commercial mobile eye trackers reached the market only relatively recently. A selection of commercial and research-based eye trackers are shown in Fig. 13.7. Current eye trackers are fundamentally similar to early systems but benefit from recent advances in camera technology, materials and computing capabilities. The general improvement in consumer camera technology has resulted in eye trackers that are lighter and provide images with much better quality and higher spatial and temporal resolution. Cameras can be integrated nicely into a given design due to their size and at low production cost. Most systems have one or more cameras monitoring the eyes and typically one scene camera. Some systems deviate a bit from the standard eye tracker by allowing the scene cameras to be reoriented and aligned with gaze: for example the EyeSeeCam (Kohlbecher, Bartl, Bardins, & Schneider, 2010). This will ensure that the point of regard is located in the central region of the scene image. Mirrors and hot mirror glasses are sometimes still used to avoid the cameras obstructing the view of the subject (e.g., the EyeSeeCam eye trackers on the top left of Fig. 13.7). The combination of improved hardware and better understanding of the mathematical underpinnings of eye tracking has resulted in improved eye tracking results (Hansen & Ji, 2010).

13.4 Buying or Building a Mobile Eye Tracker

Whether you choose to buy an off-the-shelf mobile eye tracker or follow in the footsteps of the pioneers in this field and build your own, it is important to think about the hardware configuration of the eye tracker headset. Commercial systems vary in terms of how they are supported on the head, where the scene camera is positioned, whether they track one eye or both eyes and where the eye cameras are positioned. All of these hardware choices have consequences and should be considered relative to the tasks and situations that the experimenter is interested in using the eye tracker to study. Box 1 summarises the hardware choices and consequences for mobile eye trackers. This list in itself is not a basis on which to decide what system the researcher needs to buy or build. It is intended to highlight the main hardware options and the relevance of each of these options for studies using mobile eye trackers. However, when deciding what system is needed, a key question for the researcher is what task or behaviour they want to study. This question can define the key hardware choices and threats to data collection. These issues will be outlined in Box 2 and Sect. 13.6.

BOX 1: QUESTIONS TO ASK WHEN BUYING OR BUILDING A MOBILE EYE TRACKER

What kind of headset is the eye tracker built around?

Typical options: Mobile eye trackers can be mounted on a helmet, headband or built into glasses or goggles

Practical considerations: A headband- or helmet-mounted system may help overcome problems of combining the eye tracker with a participant's glasses (a problem for eye trackers that are glasses or goggles). However, helmet and headband systems are susceptible to slippage when participants talk or smile, whereas glasses systems which use the bridge of the nose for support do not suffer as much from this problem. In both cases, the mount must be rigid enough to avoid the eye tracking cameras from moving relative to the head as a result of bodily movements by the participant or vibrations (e.g., when driving). Any slippage of the eye or scene camera that occur during testing will negate the calibration and thus make the estimated gaze unreliable.

Are one or both eyes recorded?

Typical options: Many commercial systems now record both eyes, but some still record only one eye.

Practical considerations: A system that records only one eye will suffer from problems in the accuracy of the gaze estimation over changes in depth (see Sect. 13.6.7), whereas systems that record both eyes can use the vergence of the two eyes to overcome this. For a single-eye system, the positioning of the scene camera relative to the eye is an important consideration (see below).

Where is the scene camera positioned?

Typical options: Above one eye or between the two eyes.

Practical considerations: For single-eye recording systems, the further the camera is from the tracked eye, the greater the error will be in gaze estimation over changes in depth (see Sect. 13.6.7). Thus for such systems the scene camera should be as close to the tracked eye as possible.

What is the scene camera's field of view?

Typical options: Variable, but typically in the range of 60–100° horizontally

Practical considerations: Beyond the obvious point that a wider field of view captures more of the environment, a wider field of view will also decrease the risk of data loss for large amplitude eye movements. Saccades in natural settings can be quite large and if they target a location outside the scene-camera's field of view it can be hard to work out what is being fixated. However, wide angle lenses that distort the scene image (like fisheye lenses) may result in inaccurate gaze estimates. Furthermore, it can be harder for the experimenter to code some variables in later analysis - for example it can be harder to measure saccade amplitudes and distances to particular locations in the scene for such lenses.

Can the scene camera be adjusted?

Typical options: Increasingly, commercial systems use fixed, unmovable scene cameras, but in some the camera can be rotated

Practical considerations: For very wide angle lenses there is no need to re-position the scene camera, but for narrower field of view cameras, it may be necessary to position the scene camera differently depending on the task and individual in order to capture the behaviour. For example, in driving and walking gaze is typically directed to distant locations whereas in tasks that involve manipulating objects close to the observer (such as making tea) gaze is typically directed more downward relative to the head. Rotating the camera upward for the former and downward for the latter will increase the chances of capturing the majority if the gaze behaviour in each task. Any rotatable scene camera must be held in a mount that is rigid enough to avoid slippage during the experiment as the participant moves around.

Can the eye camera(s) be adjusted?

Typical options: Many systems have fixed position eye cameras, but some are mounted in moveable housings

Practical considerations: A rigid housing for the eye camera(s) minimises the risk that the cameras will slip as a result of movements of the participant or environmental factors and so offers a more stable recording of the eye. However, a rigid eye camera must have a sufficiently wide field of view to capture the eye of different individuals despite anatomical differences between individuals. A moveable eye camera on the other hand can have a narrower field of view, devoting more pixels in the image to the features that will be tracked (e.g., the pupil and corneal reflection) and thus potentially offering more precise tracking. Flexibility of course comes at the potential cost of rigidity and stability during recording and increases the setup time for each participant at the start of recording.

What are the frame rates of the scene and eye cameras?

Typical options: 30, 60 or 120 Hz

Practical considerations: A faster camera will allow finer differentiation of when saccades start and end - an entire saccade can begin and end within a single camera frame in a 30 Hz system. A faster camera also increases the ability to employ velocity-based algorithms for detecting saccades (as is common in fast, laboratory-based eye tackers). However, more frames also means more data to code, and most of the coding procedure for mobile eye trackers remains manual (see Sect. 13.7). Thus, whether a high frame rate is needed depends on the research question: a high frame rate is needed for questions where precise timings of saccades and fixation are central, but for many real world studies where it is only the spatial allocation of gaze that is of interest this is not the case and lower frame rates will suffice.

13.5 How the Eye Tracker Tracks the Eye

Whether an eye tracker is wearable, remote or mounted on a tower they are typically based on very similar hardware and software constructs (Hansen & Ji, 2010). The common steps are shown in Fig. 13.8.

Most modern wearable eye trackers resemble the early systems by using one or more cameras (image/video data) to observe the users' eyes and one or more cameras directed towards the viewed scene. Where the scene camera is located relative to the eye may influence the accuracy of the eye tracker (details given in Sect. 13.6.7). Most eye trackers use infrared (IR) lights and IR sensitive cameras to obtain images of the user's eye(s). The eye images are subsequently analysed and eye features extracted using various algorithms. For example, the pupil is typically a recognisable dark circular region, the sclera is white, and reflections from the surface of the cornea can generate an identifiable pattern of bright spots. For more information about how algorithms identify and track features on the eye see Hansen and Ji (2010). The location and shape of the image features are used to estimate where the user is looking. This process is called gaze estimation in Fig. 13.8. An important difference between mobile eye trackers and stationary eye trackers (e.g., tower-mounted or remote) is the surface on which gaze is estimated. For stationary eye trackers, the point of regard is typically estimated on a planar surface like a screen on which images are displayed to the participant. For mobile eye trackers, gaze is estimated on the 2D scene-image captured by a camera pointed at the scene being viewed (often called the scene camera). However, this 2D surface is a projection of the viewed 3D scene. As such the estimation will suffer from limitations and ambiguities in depth when the experimenter attempts to map the gaze estimates in the collected data onto objects in the 3D environment.

In the sections that follow we will use the terms "eye detection" and "gaze estimation" to differentiate the two critical steps in eye tracking, where eye detection represents eye localization in the image and gaze estimation means estimating where the user is looking.

13.5.1 *Eye Detection*

Most eye trackers use the pupil and/or reflections of the illuminating light source(s) on the corneal surface (a.k.a. glints). These prominent features can be detected reliably

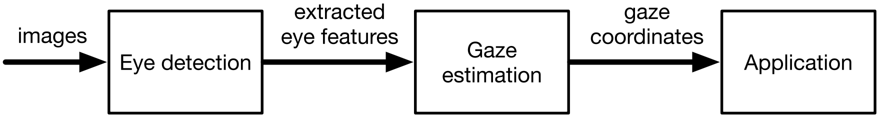


Fig. 13.8 Components of video-based eye detection and gaze estimation

in the images under relatively stable light conditions. However, to be robust the eye trackers need strong models to filter and remove outliers and spurious features in the eye images. Even small errors in pupil centre estimation can have a significant effect on the measured gaze.

Under infrared illumination the pupil will appear relatively bright in the image if the light source is placed near to the optical axis of the camera and relatively dark if the source is positioned away from the optical axis of the eye. Bright pupil imaging techniques are rarely used in mobile eye trackers since the bright pupil may become weak or disappear when the ambient light is too strong as is often the case in outdoor settings. Eye trackers often represent the pupil either as the centroid of the pupil pixels or via a model (e.g., circle or ellipse) fitted to the pupil boundary. Some eye trackers allow the user to change which pupil representation to use, but as the two representations do not necessarily give the same pupil centre estimate, this can lead to different gaze estimation results (Hansen & Ji, 2010; Mulligan, 2006).

The corneal reflection (a.k.a. 'glint', or first Purkinje image) is the reflection of a light source on the corneal surface (see Fig. 13.9). Glints appear as small clusters of high intensity pixels in the eye image. The location, shape and intensity of the glints are governed by the relative orientation/normals of the corneal surface and the camera and light source positions.

Glints move with eye movements but much less than when the head moves relative to the light source. Thus, glints are essential for the robustness of current remote eye trackers but are technically not needed for eye trackers where the head is fixed relative to the camera and light source (i.e. head mounted and mounted in towers (see Hansen & Li, 2010). Since the eye tracker uses IR light to illuminate the eye a glint is likely to appear on the corneal surface regardless of whether the glints are actually used for gaze estimation. In fact using both glints and pupil for gaze estimation in head mounted eye trackers may in fact introduce more noise on the gaze estimation results. For mobile eye trackers, slippage of the head-mount is possible during a recording session and the use of glints may partially compensate for slippage.

13.5.2 Gaze Estimation

To get a high resolution, clear view of an object, the eye is rotated such that light from the object falls on the fovea (the region of the retina with highest resolution vision; Fig. 13.9). Gaze direction is consequently assumed to be the line connecting a point on the fovea and the back surface of the lens (a.k.a. visual axis). This means that humans do not look directly through the pupil centre but at an angular offset (see Fig. 13.9) called *angle kappa*. The angle *kappa* differs significantly between individuals - for example, Gharaee et al. (2015) suggest that in normal vision, *kappa* could vary by 4° between individuals. Angle *kappa* is furthermore not directly observable from images of the eye.

On the other hand, the *optical axis* is the line connecting the back surface of the lens and the pupil centre. The pupil centre is measurable from images of the eye. The

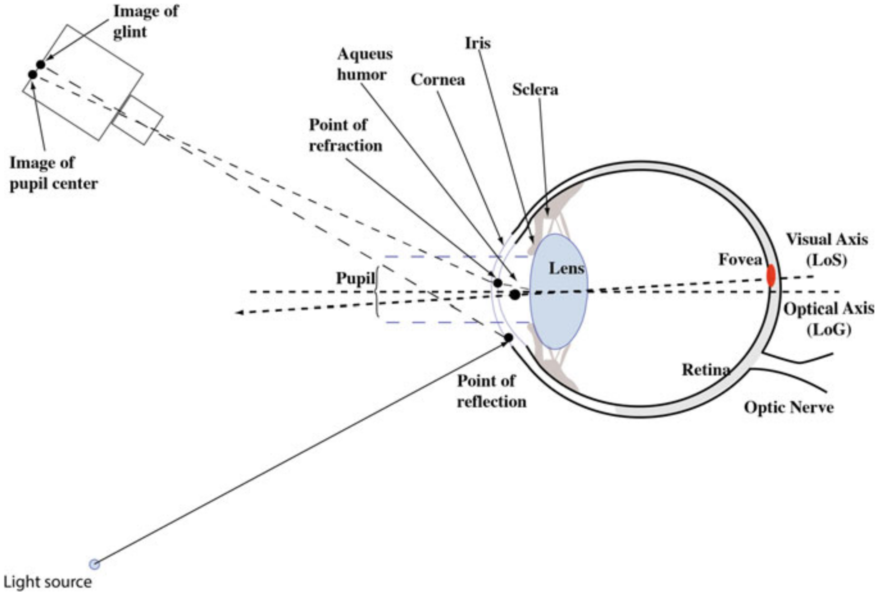


Fig. 13.9 General setup of eye trackers, with camera, light sources and the eye. Note that our depiction of the corneal reflection is an over-simplification for the purpose of illustrating the principle image formation for corneal reflections. More accurately, the light source will produce a distribution of light over the surface of the cornea, which will be imaged by the camera as if arising from a point behind the cornea, on the surface of the lens (this is a virtual corneal reflection)

angle *kappa* can be estimated by explicitly having the user look at a set of targets and through this obtain an estimate of the visual axis and how it relates to the measurable optical axis. Therefore, how eye features such as glints and the pupil centre map onto where an individual is directing his or her fovea toward depends on these physiological properties of the individuals' eye. Furthermore, the configuration of the eye tracking hardware can influence the mapping between detected eye features and gaze direction, particularly if the relative placements of the illuminator, eye camera and scene camera can be varied between individuals and due to factors such as the eye tracker slipping on the wearer's head during recording sessions.

Detecting features on the eye image, therefore, is not in itself sufficient to determine where an individual is looking. Thus all eye trackers require that the participant takes part in a calibration procedure once the eye tracker has been fitted in order to be sufficiently precise. This process is needed to generate a set of observations of the eye image for which the ground truth points of regard are known. In this way, image data of the eye can be mathematically linked to fixations of particular locations in the scene image or the screen. Gaze modelling consequently focuses on the (mathematical) relations between the image data and the point of regard[1] (that is, the

---

[1] Keep in mind that gaze information does not necessarily mean that the person is in an attentive state.

location being looked at). As with any other model, gaze models are approximations of an unknown underlying function.

Under specific assumptions only a single calibration point is needed (Hansen & Li, 2010). However, more calibration points are typically needed to compensate for measurement noise and other a priori unknown system and person specific parameters. Most current eye trackers solely rely on image features (glints and pupil) but other sources such as head pose or external position devices may, depending on application and setup, also be included in the calculations (Hansen, 2014). Current models obtain gaze estimation accuracies of less than 1° of visual angle. However, there is still room for improvements when compared to dual Purkinje (DP) eye trackers. These old DP eye tracking techniques obtain accuracies of around 0.1° and are considered the gold-standard for eye trackers. These techniques cannot be applied directly to mobile eye tracking data, but demonstrate that it is possible to achieve very high degrees of accuracy for estimating gaze with the right techniques; thus the challenge for mobile eye tracking is to find suitable techniques to improve the accuracy of gaze estimation beyond that which is currently possible.

13.6 Practical Considerations for Mobile Eye Tracking

Methods for detecting eye features in images and gaze estimation models face considerable challenges in mobile eye tracking data. These challenges come from environmental factors such as light conditions and from physiological factors such as the fact that the range of rotations of the eye within its socket is much larger in natural behaviour than when viewing computer screens or other spatially restricted stimuli. In this section we discuss the key challenges that threaten data quality in mobile eye tracking studies. All examples of problematic feature detection or gaze estimation in the sections that follow are taken from data collected by the first author. Primarily these are derived from videos collected for Tatler, Macdonald, Hamling, and Richardson (2016), which involved data collection in an indoor environment that was brightly lit by a combination of sunlight from two windows on one wall of the room and by bright spotlights positioned on stands within the room itself.

Data quality can be adversely affected for entire recording sessions by factors that make eye feature detection difficult such as poor eye tracker setup, or environmental lighting, or by factors that interfere with gaze estimation such as poor calibration. However, these systematic errors are not the main threat faced for the researcher as they are generally easy to detect and compensate for. More problematic for analysis and interpretation of the data are situations in which data quality is transiently disrupted. It is these factors that we will discuss in this section of the chapter because they are quite common issues in mobile eye tracking data and ones that the experimenter must be aware of and carefully check for in their data.

Transient disruptions to eye feature detection or gaze estimation are particularly problematic if they result in estimates of gaze that look plausible; if maintained for a few consecutive frames, such errors can look like an eye movement and fixation that did not occur. Figure 13.12 shows an example of data in which the bright light (and subsequent shadows on the eye) from the window gives incorrect inputs to the gaze estimation algorithm. Here the participant is fixating the books on the table, but appears from the gaze cross hairs to make a saccade away to the wall, where fixation is maintained for 17 frames of the video (about 566 ms) before saccading back to the books. These 17 frames are an artefact of the algorithm mis-identifying the shadows from the upper eyelid as the pupil (as can be seen from the cross hairs and green marker in the eye image). In fact, the participant maintained fixation on the books throughout this entire sequence of frames. The tracking error persisted stably for 17 frames because the eye was stable during this time and the constant lighting during this period resulted in a stable mis-identification over the 17-frame sequence. Such an error in the gaze estimation can have a clear influence on the interpretation of the data if the researcher does not notice the mis-identification in the eye image. For example, if we were concerned with the number of fixations made on the books or the time spent fixating the books, both of these measures would be contaminated by the transient tracking error: we would record two fixations of the books instead of one and would fail to record 17 frames of fixation time on the books.

In the sections that follow, we outline some of the key challenges that face experimenters when conducting mobile eye tracking studies. However, it is important to remember that there will inevitably be some data loss and noisy, inaccurate data. Being aware of this is important for two reasons. First, it means that steps can be taken to reduce the impact of this during data collection. Second, it means that the experimenter can look for, identify and exclude (or if possible, correct) sections of erroneous gaze estimation that might otherwise add noise to the analysis and lead to misinterpretations.

While we outlined some key available hardware options for mobile eye trackers and their relative strengths in Box 1, the researcher should try to answer the following three questions before deciding what eye tracker is required: (1) What task(s) will be studied? (2) In what environment(s) will testing take place? (3) What research questions and measures will be tested? Answering these three interrelated questions will identify the most likely threats to obtaining reliable mobile eye tracking data and will help not only to select the most suitable type of mobile eye tracker, but also to identify any strategies that can be employed for minimising data loss during testing. Box 2 highlights the key threats that can be identified by asking these questions and the implications for choosing an eye tracker.

BOX 2: THE RIGHT TOOL FOR THE JOB: CHOOSING AN EYE TRACKER TO SUIT YOUR STUDY

Key questions about your task and environment for mobile eye tracking

1. Does the task require the participant to move around the environment?

If so, consider the need for portable light-weight recording equipment and a headset that is rigid enough to withstand the participant's activities without slipping or too much camera vibration.

2. Does the task itself impose restrictions on the eye tracker hardware?

Tasks that require specialist clothing or equipment may interfere with the placement of the eye tracking hardware. This may prevent off-the-shelf eye trackers being used and require that you build your own eye tracker. For example, if a task requires a specific safety helmet (such as that of a racing driver) then off-the-shelf helmet- or glasses-based eye trackers may not be usable. Similarly, where the recording equipment is worn can depend on constraints such as protective clothing and the behaviour of the participant during the task. For example, tasks that require the participant to be seated for part or all of the task are problematic for backpack-housed recording equipment.

3. Is the task conducted indoors, outdoors or both?

Consider the lighting conditions that the eye tracker must work in (Sect. 13.6.3). Sunlight can be problematic because reflections or large amounts of light will make it difficult for the eye tracker to localise the eye in the camera image.

4. Does the task require the participant to look over large changes in depth?

Calibration is most accurate at the depth at which the calibration procedure was conducted (see Sect. 13.6.7). When participants will be looking at objects over a large depth range, consider strategies for minimising this issue such as calibrations at multiple depths (if this option is available for your eye tracker).

5. Expect imperfect data

Because of the issues highlighted above, and for the reasons we explain in Sect. 13.6, it is inevitable that there will be periods of data loss or incorrectly estimated gaze. Being aware of this allows you to work to reduce data loss during experiments and avoid including erroneous data in subsequent analyses.

13.6.1 What Can Be Studied in the Real World?

When planning a real world study it is important to consider what is realistic and feasible to investigate. As we explained in the introduction to this chapter, precise control over extraneous and even experimental variables is often unrealistic when studying natural behaviour in real world environments. For example, a study of natural urban driving cannot possibly control the density of other traffic (beyond considerations of likely traffic levels at different times of the day and places in the urban environment), yet traffic density has a profound influence on gaze behaviour in driving. Even the difference between driving slowly in traffic and being stationary in traffic is considerable (Land & Tatler, 2009), with not only differences in what is looked at but also in how long each gaze event lasts (Fig. 13.10). In this way the level of congestion during a recording session will strongly influence the data collected and must be considered carefully in any interpretation of the collected data.

That factors like traffic density might influence gaze behaviour is not surprising and can often be dealt with during coding and analysis of the collected data. For example, in their classic study of tea making, Land et al. (1999) removed periods where the participant was waiting for the kettle to boil as these periods were qualitatively different in terms of the requirements for visuomotor coordination. Throughout the rest of the task, the participant is engaged in searching for objects, relocating previously-found objects, and co-ordinating motor actions. In contrast, none of these processes occur when waiting for the kettle to boil. Not only were these periods qualitatively different in terms of task demands, but also key variables of interest during this period such as eye-hand latency, fixation duration and the tendency to look to targets of future actions all differed considerably when waiting for the kettle to boil and these would have skewed the reported characteristics of visuomotor control while making tea.

Of course, it is not only extraneous variables that cannot be controlled in natural settings, but also the experimental variables under investigation. Without the ability to intervene or interrupt, it is necessarily impossible to precisely limit exposure times, the number of fixations that an object will receive or even whether or not an object is looked at. Similarly, individual differences in how tasks are performed abound and so even with a simple task like tea making the order in which objects are used may differ considerably between individuals and thus the task requirements will also differ between sub-tasks for different individuals.

In addition to the lack of control available in natural settings, what can feasibly be measured also differs. The sudden onsets of a peripheral target is a commonly-used method in the lab to consider time courses of responses (Carpenter, 1981; Hallett, 1978), the factors that might attract and hold attention (Theeuwes, 1991), and whether or not a participant is attending to a location (MacLeod, Mathews, & Tata, 1986). Such events are not really feasible in natural, real world settings. Indeed, if we are interested in measuring covert attention then paradigms that have been developed in the lab (e.g., Posner, 1980; Wolfe, 1998) are almost impossible to implement in a natural setting.

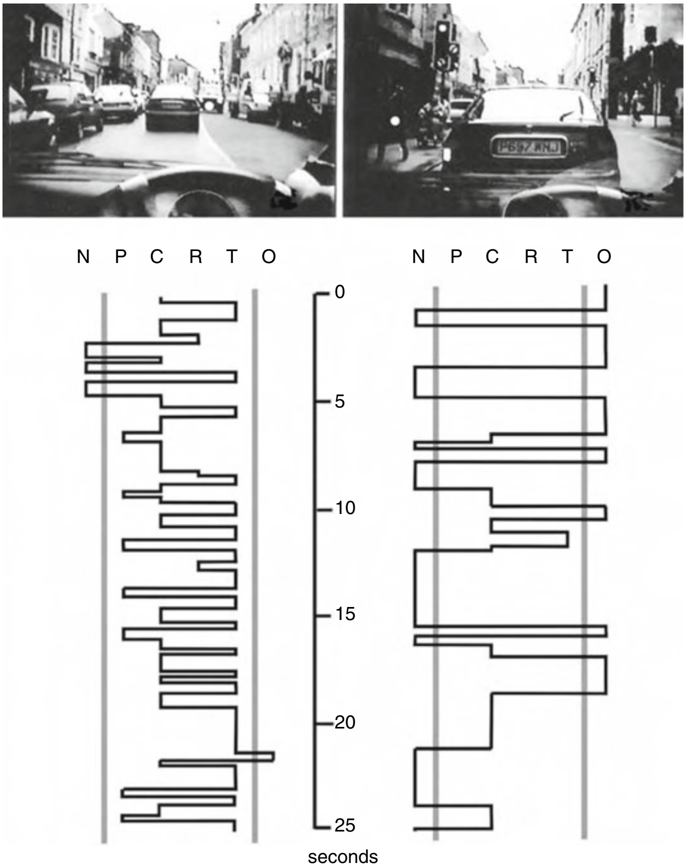


Fig. 13.10 Allocation of gaze when driving slowly in traffic (Left) and when stationary (Right) to (N) nearside, off-road locations, (P) parked vehicles on near-side, (C) car in front, (R) on or above open roadway, (T) oncoming traffic, (O) off-side, off-road locations

These limitations on what can be controlled and what can be studied in natural settings should not overshadow the central advantage afforded by studying eye movements in real world settings: the possibility to study gaze behaviour in the context of natural behaviour. If our goal is to understand how the eyes are used to gather information from our surroundings in order to support ongoing behaviour then it is clear that approaches must be employed that disrupt natural behaviour as little as possible in order to study it.

13.6.2 Calibrating Eye Trackers in Natural Settings

In the laboratory, calibration normally involves asking the participant to fixate in sequence an array of markers, displayed on the surface on which the experimental stimuli will be presented. This procedure is conducted prior to the experiment and is typically used to estimate the parameters of the gaze model. In real world settings the same approach for displaying the calibration pattern is often employed - arranging a set of markers on a surface in the environment (such as a wall or door). However, this method can be problematic in real settings for at least two reasons (explained below) and we therefore suggest an alternative approach to conducting calibrations in natural settings. Furthermore, some mobile systems have the option of either fitting the gaze estimation online, prior to the start of the recording session or offline, using the recorded videos of the eye and scene during as the participant looked at the array of markers to fit the model to the observed data after the experiment. The advantage of offline model fitting is that more time can be taken to ensure a good fit of the model without encroaching on data collection time.

In practice, the first problem is to arrange markers on a calibration surface in a real environment so that these markers cover as much of the view of the scene camera as possible. This is because the gaze model typically works best within the area of the scene camera enclosed by the outermost points of the calibration pattern. Finding a surface that allows markers to be placed such that the full extent of the scene camera image is encompassed by the calibration array can be difficult. A second problem with calibrations in the real world is ensuring that the participant's head does not move during calibration. When we move our eyes between locations in the environment we tend to move eyes and head together and this is particularly the case when the gaze shift is large (Land, 2004). If the participant's head moves toward each fixated marker during calibration, each external marker will no longer be located in the periphery of the scene camera image when fixated but rather will be displaced toward the centre of the image, thus producing a calibration over a restricted portion of the scene camera image.

One useful way of minimising the two problems discussed above is to project the calibration targets onto the scene from the eye tracker headset itself. For example, using a laser pointer fixed to the headset, with a diffraction grating to split the beam into a pattern of calibration points (Babcock & Pelz, 2004); this calibration system is used by the EyeSeeCam (Kohlbecher et al., 2010). This approach has the advantage that no physical markers need be placed in the environment and also that the calibration pattern will move with any head movements, so as long as the participant fixates the markers, head movements do not intrude on the extent of space within the scene camera that the calibration pattern covers.

However, an alternative method for collecting calibration data is to generate a calibration pattern by asking the participant to fixate a single point in the scene and then move their head systematically so that the single point covers the required extent of the scene camera image, effectively using the vestibular ocular reflex (VOR). For example, if the participant is asked to fixate a central target when the head is in a comfortable resting position and then to move their head such that they effectively trace the shape of a + with their nose before returning their head to resting position and then trace the shape of a x with their nose, while maintaining fixation on the target throughout, then the calibration point will pass through the scene camera image along these two paths (Evans, Jacobs, Tarduno, & Pelz, 2012). For online model fitting, calibration points can be obtained by asking the participant to pause during the process such that the calibration target falls on the desired position in the scene camera image. For offline recording the calibration pattern can be generated by pausing the videos at frames where the calibration target falls on the desired positions that the experiment wishes to use to form the calibration pattern (see Fig. 13.11). The advantage of this technique when used for offline model fitting is that the number of points selected to form the calibration pattern can be decided by the experimenter - the more frequently the video is paused during the process then more points the model fit can be estimated over, which may help remove outliers.

Mobile eye trackers are unsurprisingly at risk from deterioration in calibration accuracy over the course of a data collection session: natural movement may result in slippage of the headset, and participants may inadvertently displace the eye tracker slightly if they touch their face during the data collection session. For this reason it is important to repeat the calibration procedure whenever possible, balancing the

Fig. 13.11 A 9-point calibration grid formed by asking the participant to fixate a single point (the experimenter's hand) while tracing a + followed by a x with their nose

need for repeated calibration procedures with the need to interrupt natural behaviour as little as possible. At the least, a calibration procedure should be conducted at the end of the recording session. These repeated calibration procedures serve either as checks for the spatial accuracy of the calibrated gaze model (i.e. the model fitted during calibration to relate the detected eye features to where the eyes are directed in the scene video, that produces the estimates of gaze direction in the scene video) or to provide calibration patterns to re-fit the gaze model when necessary.

A further issue with calibration in natural settings is the relationship between the depth plane at which the calibration is conducted and the depths at which fixated locations occur during subsequent recording sessions. In most real world eye tracking systems the calibration accuracy decreases with distance in depth from the plane at which the calibration was conducted. This systematic error through depth is known as parallax error and will be discussed separately in Sect. 13.6.7.

In the end, how accurate can a mobile eye tracker be? The spatial accuracy of gaze estimation will depend not only on the hardware and software of the mobile eye tracker, but also on the environmental conditions and how they impact on the reliability of the data (see Sect. 13.6). It is important to use the recording of the calibration processes at the start and end of the testing session, and any recordings of calibrations conducted during testing, to estimate the spatial accuracy of the calibration that has been achieved. This determines the accuracy that can be claimed in the analysed data, and also should inform criteria for deciding whether an object is fixated or not (if the gaze estimation model shows fixation to within 1° of an object when the achieved accuracy during calibration is 1° or more, this should probably be counted as a fixation on the object). See Table 13.1 for some suggestions to bear in mind when calibrating a mobile eye tracker.

13.6.3 How Light Conditions Influence Eye Tracking Data

Environmental lighting poses a number of problems for eye trackers that use the pupil or corneal reflections. The problems associated with outdoor recordings are well known: the high levels of infrared light are problematic for infrared-based eye cameras and this interferes with eye feature detection. However, this problem can be reduced by shading the eye tracker from direct sunlight, for example by asking the participant to wear a peaked cap (but this will not help reflections coming from the side or below e.g., reflections from water). Furthermore, eye trackers may control for some light changes by adjusting the camera parameters (such as shutter time, aperture, white balance etc.) in order to overcome such problems.

However, even for indoor environments, changes in light within an environment may lead to overexposed eye images and can therefore pose transient problems for eye feature detection from the images. Light from the side (say a window) may result in strong shadows or highlights on the eye that create difficulties for the gaze estimation algorithm. Eye glasses and contact lenses can seriously disrupt eye feature detection by either generating spurious glints or reflections from external light sources on the surface of the glasses that occlude the pupil.

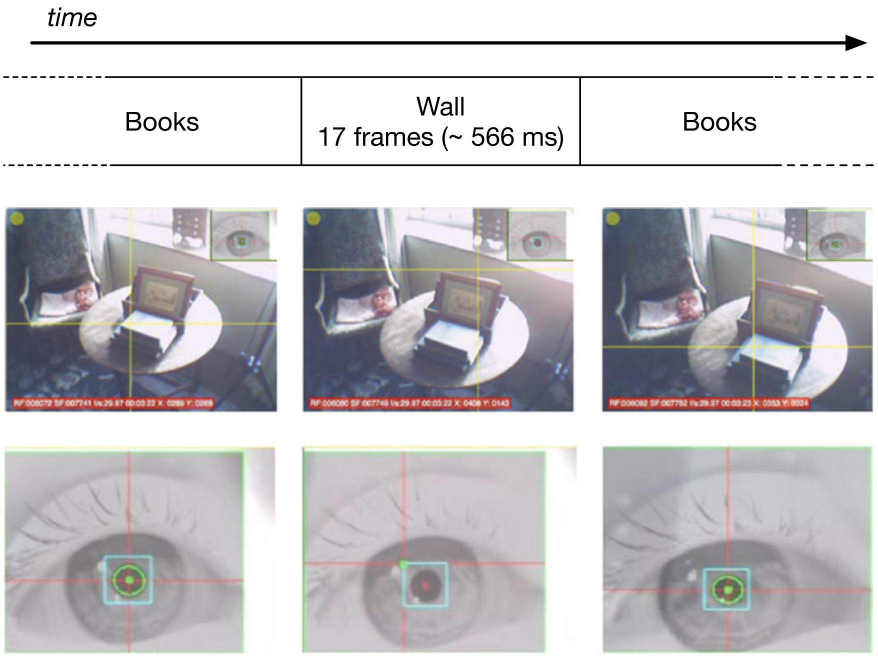


Fig. 13.12 What looks like a fixation of the wall for 17 frames from the gaze direction crosshairs is in fact an error in the gaze detection algorithm arising from mis-tracking of shadows from the upper eyelid. This mis-track arises from shadows cast by the strong light from the window. This 17-frame mistaken tracking appears as a plausible fixation if working only from the output of the gaze detection algorithm (that is working only from the cross-hairs in the scene video)

Table 13.1 Key points to consider when calibrating a mobile eye tracker

Tips for calibrating in the wild
Make sure calibrations points cover as much of the scene camera image as possible
If using external markers for calibration, make sure participants don't move their heads during calibration. If they move their head to point it at each marker, then you will end up with all markers in the middle of the scene camera image - so only the centre of the scene camera image will be calibrated
Calibrate before and after testing in order to check for any slippage of the eye tracker during testing
Calibrate at a depth appropriate for the study. For monocular systems, only the depth at which calibration is conducted will be properly calibrated. (For more on this see Sect. 13.6.7)
If your eye tracker requires you to fit the gaze model during testing, check that you can later re-fit the model if necessary to improve the fit

Figure 13.12 shows an example of a transient mistake in the gaze estimation as a result of lighting issues and Fig. 13.13 shows several examples in which indoor environmental lighting conditions can result in challenging eye images that can cause problems for techniques that use either boundary information, pixel intensities or both. In some cases it may even be difficult for a human observer to precisely identify the pupil and corneal reflections coming from the system light sources (Table 13.2).

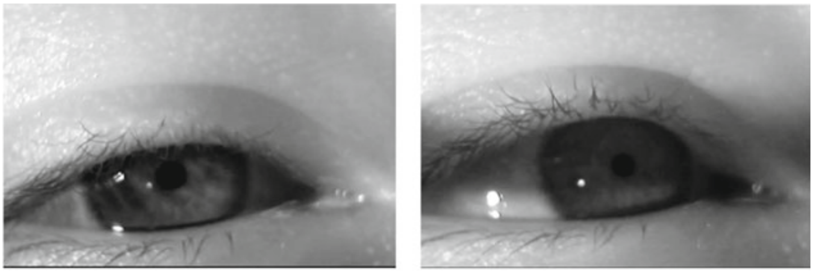
Bright light - either as a result of sunlight or strong artificial lighting - can result in periods in which the eye image is disrupted by shadows (Fig. 13.13a, b), or overexposed (Fig. 13.13c, d). Both overexposed and underexposed regions make it difficult to ensure robust eye feature detection. Shadows may come from the eye itself or from external sources. In the example shown in Fig. 13.13a shadows from the upper eyelid obscures the pupil and in Fig. 13.13b a shadow from an object in the environment falls across the eye and obscures the pupil. In both cases the shadows result in reduced contrast between the pupil and surrounding eye regions and result in areas within the eye region that have about the same intensity as the pupil. Artificial light sources can result in discrete, bright reflections on the cornea that are hard to distinguish from the corneal reflection arising from the system light sources. In the examples shown in Fig. 13.13e, f, the extra corneal reflections provided by spotlights in the environment make identification of the correct corneal reflection problematic. Even dirt on the camera sensor or lens can alter the light conditions: the dust particles evident in Fig. 13.13d may not only change the overall intensity in the image, but also distract the feature detection methods with spurious dark and bright spots.

13.6.4 When the Pupil Is Not Fully Visible

Blinks interrupt eye tracking data because the eyelids occlude the pupil and iris. However, more problematic are situations in which partial occlusions of the eye by the eyelids result in inaccurate tracking of the eye features. Partial occlusions can happen because of extreme positions of the eyes in the orbits, are common in older participants and can be found in some pathological conditions. Partial occlusion may also happen as a consequence of the camera capturing the eye in the middle of a blink. Figure 13.14 shows an example where partial occlusion of the eye results in an apparent downward movement of gaze as shown by the displacement of the cross-hairs in the scene image; however, there is no corresponding movement of the eye during this period.

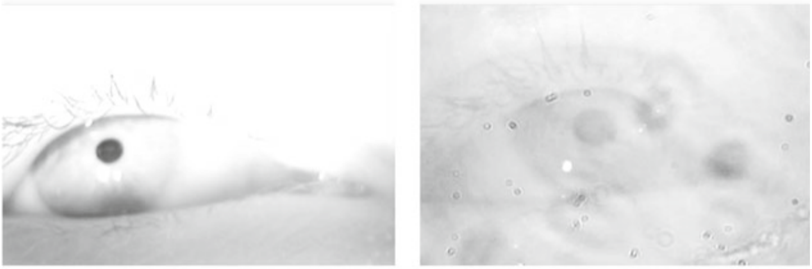
13.6.5 Estimating Gaze During Extreme Eye Rotations

Extreme rotations of the eye in its socket present challenges for detection algorithms based on pupil or corneal reflection tracking even in the absence of any partial occlusion of these features by the eyelids. The pupil changes shape from circular to



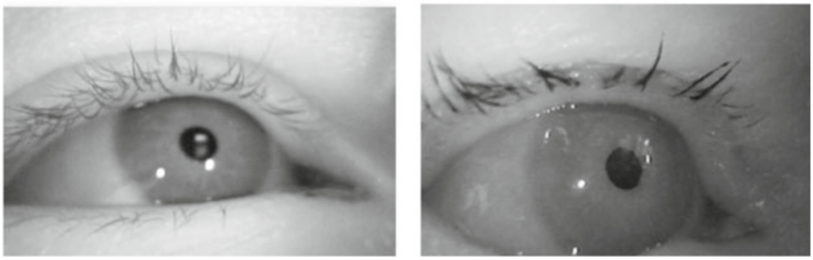
(A) Shadowing from eyelid obscures pupil

(B) Shadow from external object falls across pupil



(C) Bright daylight floods scene camera and obscures corneal reflection of IRLED

(D) Bright daylight results in artefacts in image from dust on scene camera lens, obscuring pupil



(E) A second corneal reflection arising from a strong light source (spot light) in environment

(F) Multiple corneal reflections arising from a combination of a wet eye and strong light sources (spot lights) in environment

Fig. 13.13 Examples of eye images in which eye features are transiently obscured by lighting conditions in the environment

Table 13.2 Key points about lighting conditions to consider when conducting studies with a mobile eye tracker

Tips for minimising the problems of light conditions
For outdoor environments, a peaked cap can shade the eye camera from light coming from above, providing this does not impede the behaviour under test
For indoor environments consider whether testing can be conducted in a part of the environment that is away from windows and strong artificial light sources. Can curtains/blinds be drawn without impeding the behaviour being studied?
Calibrate in light conditions as close to those in the location where testing takes place as possible, so that the pupil is of a similar size during calibration and during testing

elliptical as the eye rotates away from the optical axis of the eye camera (Fig. 13.15a–c). Factors such as refraction and partial occlusion have a significant influence on observations of the pupil as the eye moves away from the camera center. Refraction occurs because the pupil is located behind the liquid filled aqueous humor. The effects of refraction are observable in the eye images in particular for head mounted systems as the camera is close to the eyes. Refraction can have a considerable impact on the gaze estimation data, with larger deviations in gaze estimates as the participant looks further away from the centre of their orbit. These errors can be exaggerated by the fact that the pupil becomes more elliptical at extreme rotations, so small errors in estimating the pupil centre will have increasingly large effects on gaze estimation.

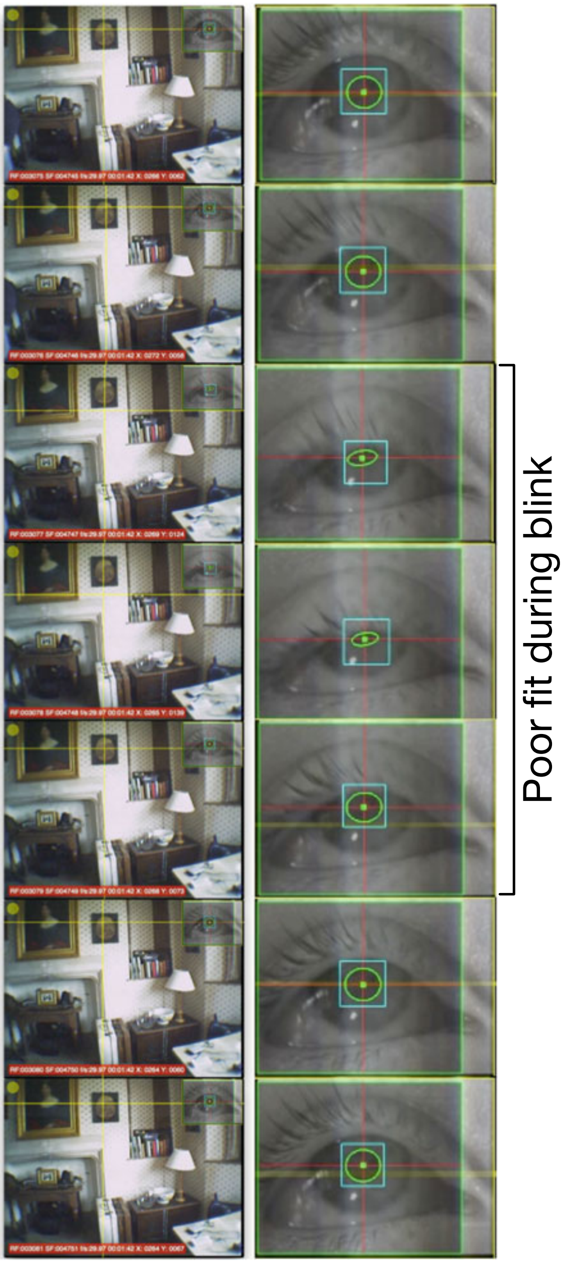
This is a particular problem for mobile eye tracking because the range of rotations of the eye within the socket is much larger in natural behaviour than is typically the case in the lab where stimuli are spatially restricted (often presented on computer screens). The gaze estimation model used for mobile eye tracking must therefore be able to model eye direction even for very extreme rotations of the eye, for which the image of the pupil will be distorted by viewing angle, partial occlusion and refraction.

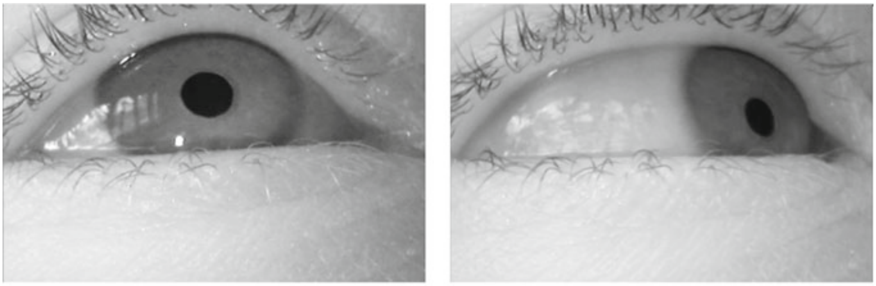
Corneal reflections also suffer with large rotations of the eye and may be lost at large rotations (Fig. 13.15b, c). Indeed corneal reflections are best modelled when they fall upon the spherical part of the cornea. Gaze estimation methods often do not model the occurrence of reflections on the non-spherical parts or on the sclera and hence the gaze results may be less accurate. Reflections on these areas may furthermore be non-linearly deformed and can even break-up in several parts due to the unevenness of the surfaces making them even harder to use (Fig. 13.15d; Table 13.3).

13.6.6 The Problems of Changes in Pupil Size and Shape

As we have discussed above, the shape of the pupil in the eye camera image can have a significant influence on the gaze estimation results, but additionally and perhaps surprisingly, gaze estimation accuracy is affected by pupil dilation and constriction. The influences of pupil size and shape on gaze estimation increase as the participant

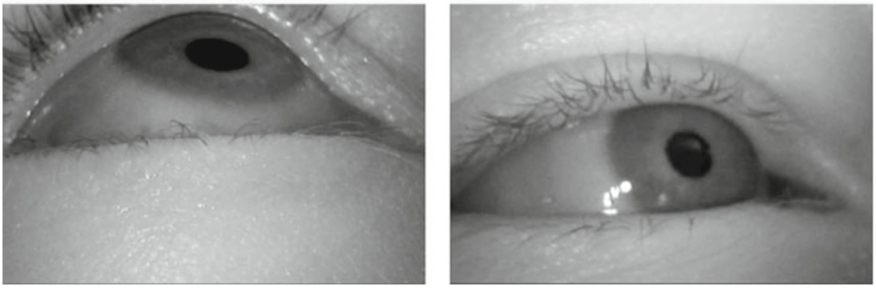
Fig. 13.14 A sequence of consecutive frames showing a partial occlusion of the eye results in misfitting of the pupil. Here the centre of the visible portion of the pupil is misclassified as the centre of the pupil, and the ellipse is fitted only to the visible portion of the pupil. The result is an apparent downward deflection of gaze for a few frames indicated by the crosshairs in the scene image, but there is no corresponding downward movement of the eye





(A) Eye centred in orbit showing circular pupil and clearly identifiable corneal reflection of IRLED on iris

(B) Lateral rotation of the eye showing horizontal shortening of pupil and loss of corneal reflection of IRLED



(C) Upward rotation of the eye showing vertical shortening of pupil and loss of corneal reflection of IRLED

(D) Disruption of corneal reflection as it passes from corneal bulge to sclera

Fig. 13.15 **a** A view of the eye relatively centred in its orbit, showing circular pupil and clearly identifiable corneal reflection of the system light sources. **b**, **c** Show large rotations of the eye in which the pupil appears elliptical in the resultant eye image, and there is a loss of corneal reflection system light sources. **d** Here the corneal reflection is fractured by falling at the border between corneal bulge and sclera

looks to the side (Ahmed, Mardanbegi, & Hansen, 2016; Narcizo & Hansen, 2015) meaning that these effects contribute to the problems of extreme rotations of the eye discussed above. These issues for gaze estimation arising from changes in pupil size and shape mean that if the pupil was calibrated in one light condition (say dark) and the actual experiment is conducted in a different environment then this may lead to errors in gaze estimation that might influence the apparent results of the

Table 13.3 Key points for dealing with issues of pupil detection in mobile eye tracking data

Tips for minimising the problems of extreme eye rotations, pupil occlusion and pupil shape
Realistically, these issues cannot be avoided entirely in real world environments: people blink, make large saccades and undergo considerable pupil size variation in natural settings
Extreme eye rotations can be less of an issue if the objects required from the task all fall in a small region of the environment, but this is not the case for many real behaviours
A wide angle scene camera reduces the chance that gaze will fall on objects outside the scene camera's field of view, but does not remove the issue of poorer gaze estimation for pupils at extreme rotation
Be aware of these potential problems in gaze estimation and check your data for these instances. Treat these sections of data with more caution during analysis

study. Similarly, transient changes to pupil size due to transient changes in lighting conditions can impact the reliability of gaze estimation.

13.6.7 *Parallax Error: The Challenge of Tracking Across Changes in Depth*

A common problem with wearable eye trackers is that they become less accurate when the distance between the eye tracker to fixated objects is different from the eye tracker to the (e.g., virtual) plane of calibration targets during calibration. These errors are the so-called parallax error. An illustration of the parallax error is shown geometrically in Fig. 13.16.

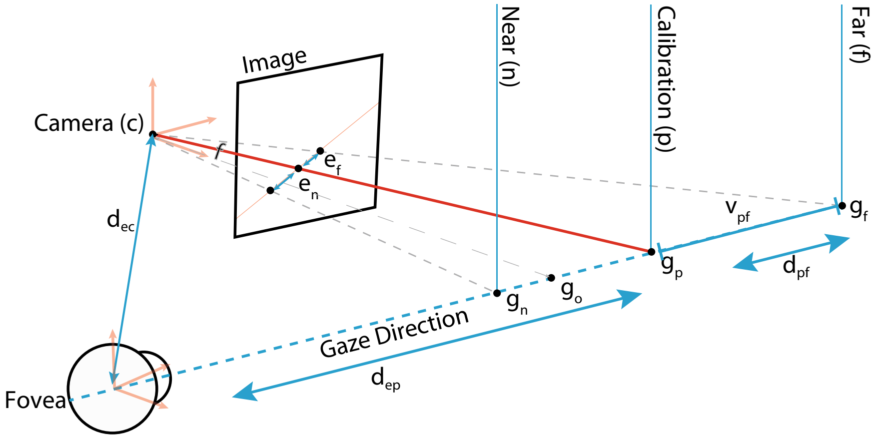


Fig. 13.16 The parallax error occurs when the camera and the eye are not co-axial. The figure is exaggerated for illustrative purposes

In Fig. 13.16 the eye (e) of the user is dislocated from the camera (c) by a distance d_{ec}. In wearable eye trackers the gaze is typically estimated in scene camera coordinates (e.g., as a cross-hair in the video) Suppose the user gets calibrated with targets located on the "calibration plane" p at a distance d_{ep} from the eye. After calibration the eye tracker should (ideally) not yield any errors when gazing at a point, g_p, on the calibration plane. The effect of the parallax error unfolds when considering the user is looking towards g_p but may in reality be fixating on any given point along the direction of gaze (e.g., g_n,g_p and g_f). The eye tracker cannot easily detect where the user is focusing so the projection of the gaze direction line (e.g., the uncertainty of at which distance the user is fixating) onto the image yields the line containing all points the user may potentially be fixating for the given gaze direction. If the user is actually gazing at g_n this yields an error e_n in the scene image and equivalently when the user is fixating on the point g_f this will yield an error e_f in the scene image. The direction of the parallax error (e_f) is directly related to the relative direction of the gaze vector and camera-eye vector and can be approximated by the projection of vector v_{pf} to the image. Figure 13.16 shows that the parallax error e_f can be reduced by minimizing (1) the eye-to-camera distance ($d_e c$) and (2) distance between the calibration and fixation planes, d_{pf}

13.6.7.1 Parallax Error Correction Methods

Parallax error can be corrected either by having the stimulus presented only on the same plane as the calibration plane (e.g., a computer screen), by calibrating the user at multiple depth planes, through eye tracker construction or using prior information about objects in the scene (Holmqvist et al., 2011; Mardanbegi & Hansen, 2012; Narcizo & Hansen, 2015). Constructing eye trackers with a half-silvered mirror in the observer's optical path (e.g., Land et al. 1999) allows the scene camera and the eye image to be co-axial and so avoids the parallax error. However the mirror reduces the quality of both scene video and the observer's view. This option is not employed on any current commercial systems.

The extent of ocular convergence in binocular eye trackers can be used to track the distance of focused objects, but the correction is only practical for short distances. For example, an observer fixated on an object 2 m away, tracked with binocular trackers with angular accuracies of $\pm 1°$, provides an uncertainty in fixated depth between 1 m and infinity. Having calibration targets and stimuli presented on the same plane can be difficult in mobile eye-tracking outside the laboratory, where fixated objects may appear at a variety of distances and participants can move closer to or farther from these objects during the recording. Table 13.4 highlights important issues to consider for minimising parallax error in mobile eye tracking data.

13.7 Analysing Mobile Eye Tracking Data

After collecting data and carefully calibrating the gaze model, the result from any contemporary mobile eye tracker is a video of the scene image with the gaze estimates superimposed on each image in the video. For most systems this is accompanied by a file that contains the actual parameters of the eye features in the eye image and gaze coordinates in the scene image together with other data available from the model fitting procedures. There exist some software packages that parse mobile eye tracking data into saccades and fixations, but as we will see below, these are often of limited use unless accurate data is available about head movements. As a result, mobile eye-tracking analysis often involves laborious manual frame-by-frame coding of the scene video with gaze data. Furthermore, if we wish to relate the gaze estimate in the scene video to objects in the environment, there is usually no option but to go through the video data and manually code this.

13.7.1 Identifying Problematic Sections of Data

Gaze estimation errors can be misleading if they are not identified by the researcher and corrected before subsequent analysis. Most eye tracker manufacturers do no reveal details of their techniques. The results shown in the scene camera are therefore the output of a black-box. Gaze estimation is highly dependent on accurate detection of eye features and how outliers and noise are handled in both the eye tracker and gaze mapping function. Greater transparency of how feature detection and gaze estimation are handled in eye trackers would be an important step toward identifying and understanding problematic sections of data in mobile systems. It may also provide the eye tracker companies with valuable information to improve their trackers even more.

Table 13.4 Key points for dealing with the need to calibrate over changes in depth in the environment

Tips for minimising parallax error in your data
A binocular system will reduce the problem of calibration through depth
For a monocular system, a scene camera as close as possible to the tracked eye will reduce parallax error
For a monocular system, calibrate at multiple depths - over a range appropriate for the task being studied - if this is feasible
Be aware of these problems in gaze estimation when analysing your data and be cautious about the gaze position marker for objects far from the depth at which calibration was conducted

However, the above examples of problems in reliably detecting features in the eye image and the danger not of lost data but of misleading estimates of gaze direction illustrate perhaps the most important requirement for obtaining reliable data from mobile eye trackers: the need to be able to access and evaluate the image of the eye, with corresponding model fit, for each frame of the recorded data. Without this image, the above problems cannot be identified easily. These problems are common in data collected in real world settings, even in indoor settings with controlled lighting. As a result, if they are not detected and dealt with during analysis, they can influence key measures of eye movement behaviour.

At present not all commercial mobile eye tracking systems provide access to an image of the eye with corresponding model fit. The reasons for this seem to be twofold. First, there is insufficient recognition of how vital this information is for researchers in order to obtain reliable data. Even if the manufacturer and researcher are confident that the gaze detection algorithm is good, identifying transient mistakes such as those outlined above is essential for usable mobile eye tracking data and however good the underlying algorithm these issues will inevitably be found in studies conducted in real world settings. Second, manufacturers are understandably wary about providing access to any details that may reveal proprietary information, hardware specific details and limitations of their algorithms.

A record of the image of the eye alone for each frame is certainly helpful for identifying some problems and some manufacturers share this as a compromise. This image can be used to easily identify blinks, pupil occlusions and losses of corneal reflection, but cannot be used to identify any transient tracks of incorrect features in images. For example, the mis-localisation of the pupil in Fig. 13.12 would be hard to detect from the eye image alone, although any apparent deflection of the gaze marker in the scene video that is not accompanied by a corresponding movement of the eye in the eye image would indicate this particular mistake.

We strongly encourage the requirement for access to not only the image recorded by the eye camera corresponding to each frame of the scene video, but also that this eye image includes the eye feature estimates and the fit of the gaze direction model on each frame, as illustrated in the examples provided above. If this model fit is included then it is clear to the researcher when the position of the crosshair in the scene video is reliable and when it is not. The researcher can then make judgements about which frames to include in data coded for subsequent analysis and whether it is possible from the surrounding frames to manually adjust the coded data. For example, for the sequence show in Fig. 13.12 the 17 frames where the crosshair is on the wall can be reassigned as fixation of the books. One helpful option in this respect is the ability to refit the model during identified problematic sections of data, or even to manually override these samples, and manually indicate the centre of the pupil in the eye image (e.g., the Dikablis eye tracker offers these software options).

BOX 3: THE ONE KEY FEATURE TO ASK FOR IN AN EYE TRACKER

Access to the eye camera image, preferably with the eye model fit shown (e.g., the detected pupil and glint positions and shapes), is an essential requirement for mobile eye tracking.

Seeing this image allows for instances of poor tracking due to the reasons outlined in Sect. 13.6 to be identified and corrected (if possible), treated with caution or excluded from analysis.

Without access to the eye image, it is hard to judge the reliability of the gaze marker that the gaze estimation model creates on the scene image.

13.7.2 Eye Event Detection in Mobile Data

The first step when analysing eye movement data collected in the laboratory is typically to categorise the stream of samples into saccades and fixations (a.k.a. event detection). For lab-based setups where the participant's head is stabilised or in remote eye trackers, algorithms tend to either use methods based around the sample-to-sample dispersion of gaze co-ordinates, counting any that fall within a defined area as belonging to the same fixation, or use sample-to-sample velocity estimates to detect the onset and offset of saccades, and these work very well. But translating these algorithms for use with mobile eye tracking data is not straight forward.

Velocity-based methods are limited by the low sampling rates of most mobile systems and the velocity of eye rotations produced by the vestibular ocular reflex during fixations that accompany fast rotations of the head can be greater than the velocity of eye rotations produced by small amplitude saccades. Dispersion methods suffer from the fact that VOR during head movements can result in large variations in eye-in-head position (and thus the position of gaze in the scene camera image) during steady fixation of an object while the head is moving (Fig. 13.17a); indeed, changes in eye-in-head position can be greater within fixations than within saccades when the head is free to move (Fig. 13.17b, c). The consequence of this is that even with relatively modest head movements, it is not possible to find a single threshold for dispersion of eye-in-head samples that would successfully classify the fixations in the data shown in Fig. 13.17b, c. For example, a dispersion threshold high enough to classify the fixation illustrated in Fig. 13.17b as a single fixation without any eye movement, would be too high to identify the saccade illustrated in Fig. 13.17c. This is because the change in position during the saccade highlighted in Fig. 13.17c (c. 8°) is less than the change in position during the fixation highlighted in 13.17b. As a result, the two fixations either side of the illustrated saccade in 13.17c would be classified incorrectly as belonging to the same fixation event. In contrast, a threshold small enough to correctly identify the distinction between the gaze events lying either

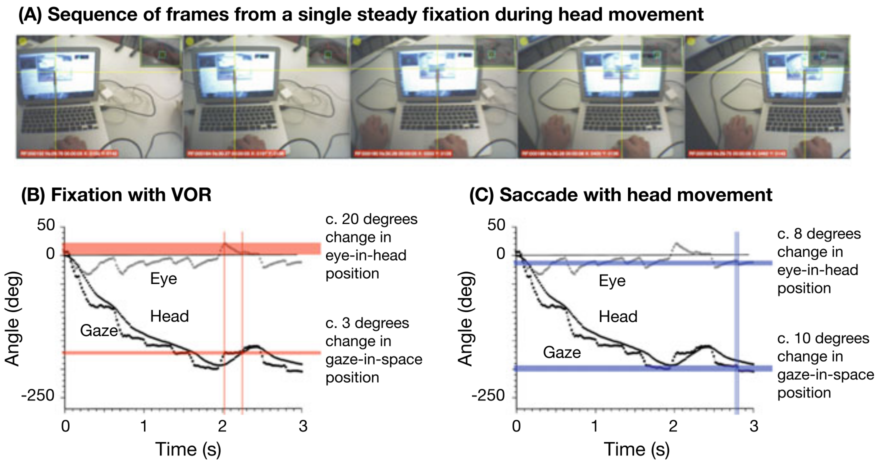


Fig. 13.17 **a** Sequence of frames from a gaze-fitted scene camera video during steady fixation accompanied by head movement. The eye remains fixated on the tip of the pen - as shown by the yellow crosshairs - while the head is rotated from left to right. As the head moves, the position of the pen in the head-centred scene video moves in the opposite direction. Counter-rotation of the eye with respect to the head occurs as a result of the vestibular ocular reflex and can be seen by the change in eye orientation in the eye image shown in the top right corner of each frame. The lower two panels **b**, **c** show a short sequence of data from a recording of Mike Land making tea and include the change in position of the eye within the head (open circles), the head within space (filled circles) and gaze within space (filled squares). Counter-rotation of the eye during head movement can be seen clearly by the saw-tooth pattern in the eye-in-head data. Here we highlight a single fixation (**b**) and saccade (**c**) that illustrate the problems of identifying fixations using dispersion methods. **b** In the highlighted fixation, in order to achieve a relatively stable gaze fixation while the head is moving, the position of the eye-in-head changes by about 20° over the course of the fixation. **c** For this highlighted saccade, the shift in gaze is about 10°, but as it is accompanied by head movement in the same direction, the position of the eye in the head changes by only about 8°. A dispersion method could not be used to set a single threshold that would classify (**b**) as a fixation without failing to classify the two fixations either side of (**c**) as separate fixation events

side of the saccade illustrated in Fig. 13.17c would classify the samples within the fixation illustrated in Fig. 13.17b as belonging to several distinct fixation events.

Of course, if the motion of the scene camera can be accurately tracked and synchronised to the eye recordings then the co-ordinates of head-in-space (derived from the movement of the scene camera) can be used to correct those for eye-in-head (from the gaze estimates) in order to derive gaze-in-space co-ordinates; indeed this is shown in the data plotted in Fig. 13.17b, c. Head position can be monitored using commercial head trackers (e.g., EyeSeeCam includes head motion tracking), and many eye tracking systems integrate with available head trackers. While these provide precise data, many head trackers constrain the space within which a study can be conducted, requiring to be within an electrical field or within sight of an optical tracking system. The alternative approach for monitoring head position is to estimate this from the image in the scene camera video. Because the scene camera is essentially fixed to the wearer's head, the image of a static target in the environment will move in the scene camera as the head moves although how this external object moves will depend on how far it is in depth from the participant. Thus, the movement of the head can be calculated from the movement of static objects in the scene camera video, provided we know about (or can infer from the image) the structure of the environment and the properties of the camera. Some commercial eye trackers now include this feature in their software for coding mobile eye tracking data, but the manner in which these calculations are made differs between manufacturers. At present at least three different solutions to this approach have been taken: (1) tracking external infrared markers placed in the environment within which the study is conducted, (2) tracking visible marker patterns placed in the environment within which the study is conducted, (3) estimating movement from frame-to-frame changes in the video video image, without any externally-placed markers. The first two typically allow more precise estimates than the third method, but constrain the space within which the study can be conducted and introduce additional visible elements in the environment that may interfere with the participant's natural behaviour. The data plotted in Fig. 13.17b, c for gaze position were calculated using a manual version of the third method, where head position was tracked and compensated for by measuring the displacement of background objects in the scene camera video. Kinsman, Evans, Sweeney, Keane, and Pelz (2012) proposed a method for compensating for head rotations in event detection from the scene video and this 'head-motion compensation' method was used by Evans et al. (2012).

At present, all of these methods for calculating head movements and using these in combination with the eye tracker's record of eye position in scene camera coordinates are somewhat limited in their ability to produce data that will reliably be classifiable into saccade and fixation events by software. This is in part due to the fact that sample-to-sample estimates of eye direction in mobile eye tracking data can be noisy for reasons such as those outlined in Sect. 13.6. It remains, therefore, important to manually check the recorded data to ensure that saccades and fixations are correctly identified: when the eye moves is evident in the image of the eye recorded by the eye tracker, so if this image is available, it is possible (and important) to manually check when saccades occur.

13.7.3 Measures of Eye Movement Behaviour in Mobile Eye Tracking Data

In principle, all of the measures of eye movement behaviour (fixations, saccades, scan paths, etc.) that are used in the lab, and the insights that these provide (Holmqvist et al., 2011) can be measured and analysed in mobile eye tracking data. However, in practice, some of these measures are currently hard to estimate with sufficient accuracy. For example, estimating the peak velocity within a saccade or even estimating saccade duration accurately is limited by camera frame rate, which is typically relatively low in mobile eye trackers.

Typically in mobile eye tracking studies the key measures are whether an object was fixated, and the number and durations of entries to a particular region of interest or object. These measures are relatively easy to extract from mobile eye tracking data. It should be remembered, however, that the limited capabilities of fixation detection algorithms in mobile eye tracking data (see Sect. 13.7.2) mean that manual coding of these variables from the eye tracking videos may be required or at least manual checking of any software-derived estimations of these variables should be conducted. Whatever the method for identifying fixations and saccades, these periods must be linked to what is being fixated in the environment. This typically requires manual coding by the experimenter, but has the potential for automated identification if the 3D space is mapped (e.g., using external markers in the scene) or with advances in computer vision and machine learning automated identification and tracking of objects in the scene camera's view.

More precise measures such as the durations of individual fixations and the amplitudes of saccades are also feasible in mobile eye tracking data but are less commonly used because they require more laborious coding or checking of the data. For fixation duration it is important to derive accurate time estimates of the beginning and end of each fixation, which typically requires manual checking of the eye image (this measure is typically more laborious to code than the duration of entries to an object or region of interest because multiple fixations may be made within the region of interest). For saccade amplitude, what we are typically interested in is the distance between two consecutive fixated points in space. The coordinates of the gaze estimate within the scene camera view is insufficient to calculate this distance if the head is also moving (as it usually will be when shifting gaze in a real world setting). The problem is that the gaze shift between two points will be achieved by a combination of eye-in-head movement (which is calculable from the change in gaze estimation within the scene camera view) and head-in-space movement. This typically requires an estimate of the change head position from a head position sensor or by estimating it from the camera image (Sect. 13.7.2). Alternatively, making use of advances in computer vision and machine learning this could be calculated from the gaze estimate and scene image if the experimenter has a model of the 3D world and camera pose. As a result of these difficulties, measures of saccade amplitude and individual fixation durations are less commonly reported in mobile eye tracking studies.

In laboratory-settings, the value of analysing sequences of eye movements (or scan paths) is increasingly being recognised: there is information and insight available in the order in which locations are looked at. Consequently, methods for analysing such sequences of eye movements are available (e.g., Jarodzka, Holmqvist, & Nyström, 2010). In real world settings, the value of considering the order of looking is the same, but analysis of this is more difficult because the locations looked at are distributed in 3D space. Thus methods for scan path analysis used in the laboratory are not easily applied to mobile eye tracking data. Information about the sequence of eye movements in real world tasks is reported in some studies, but analysis does not benefit from the sophisticated and multi-dimensional methods for quantifying different aspects of the sequence that are employed in the laboratory.

13.8 Emerging and Future Directions in Mobile Eye Tracking

There is a growing recognition of the importance of conducting studies of vision, attention and cognition in natural settings (e.g., Kingstone et al., 2008), with eye tracking a key tool in this field. With this recognition comes greater pressure to move mobile eye tracking forward in order to make data more reliable and analyses less laborious. As we have discussed we are at present some way from ideal solutions to these problems.

In terms of hardware, cameras are likely to continue becoming faster (more frames per second), smaller and with better resolution. Smaller cameras allow for less obtrusive headgear and this will be an important step toward minimising the intrusions that eye trackers can have on the wearer. It is certainly the case that when a participant knows that he or she is being tracked, it can influence the eye movement behaviour that they display (Risko & Kingstone, 2011), but this effect can diminish over the course of a recording session, unless the participant is reminded that they are being tracked (Nasiopoulos, Risko, Foulsham, & Kingstone, 2015). As such, smaller and lighter mobile eye trackers reduce the chance that the equipment itself reminds the participant that they are being tracked. However, a lighter eye tracker is typically more vulnerable to movement during the experiment. The same developments in hardware will help with reducing any interfering effects of the equipment worn by participants on other individuals that they might be interacting with in a social setting. Commercial trackers increasingly look more like glasses and less like eye trackers, but there is still some way to go before they become unnoticeable (and the fixed geometry of the glasses may not be suitable for all eye tracking situations). It also seems likely that the current move toward binocular rather than monocular systems will continue as cameras get smaller and better and this will help with problems of calibration in depth.

Faster cameras, thus higher frame rates, will aid velocity based event detection algorithms and higher resolution cameras may achieve some improvements in gaze estimation accuracy since feature detection may become more accurate. However, while higher spatial and temporal resolution of the cameras offers the opportunity for more detailed analysis, it may actually reveal that the detailed eye tracking signal is convoluted with other signals that may or may not be of any use. Higher spatial and temporal resolution adds to the computational requirements and it remains unknown to what degree the added amount of data is beneficial to the experiment. A higher temporal resolution requires more light but with the current trends in more sensitive cameras this may not become a problematic issue.

It seems likely that incorporation of head tracking systems into eye tracking systems will also become more widespread as the technology improves. The obvious assistance that knowing head movements provides for event detection in mobile eye tracking data makes it likely that mobile eye trackers will include head tracking systems by default in the future (indeed commercial systems are starting to incorporate head tracking already, and the EyeSeeCam has done so for some time).

Software for eye tracking is of course likely to get faster and better. However it is not obvious that this will provide a significant change due to the limitations of the eye and gaze detection techniques. Perhaps we need to think of how to extract eye information in a different way and even use different hardware setups. Current trends from computer vision and machine learning offer multiple opportunities to make eye tracking and eye tracking analysis more automatic, but may require a different approach for solving analytical problems: At present eye tracking hardware and software are designed to be general purpose - to obtain data in as wide variety of tasks and environments as possible. However, if we instead produce solutions that are targeted at specific tasks we can use knowledge of the task as priors for more targeted analytical solutions that capitalise on new techniques from machine learning and big data. Thus a move away from general purpose solutions to task-specific solutions may help in the development of mobile eye tracking.

13.9 Questions the Reader Should Be Able to Answer

When planning a mobile eye tracking study, deciding about an eye tracker to buy or build, setting up a study or analysing mobile eye tracing data, you should be able to answer the following key questions:

- Why is it important to study eye movement behaviour in the context of a natural task and environment?
- What hardware choices are best for my task, environment and research question?
- What are the main threats to collecting reliable mobile eye tracking data for my study?
- Why is it important to have access to the image of the eye in the recorded data?

BOX 4: KEY TESTS WHEN EVALUATING A MOBILE EYE TRACKER

Calibrate and then verify that the system is sufficiently accurate in determining gaze on the calibration points.

Then test the eye tracker on points (same plane) that were not used during calibration.

Test the eye tracker on points at different depths to that calibrated to get an idea of the parallax error.

Rotate the head while fixating on a target (do this several times for some points used and some unused during calibration) and measure how much the gaze deviates. This tests how well the system deals with extreme rotations of the eye and distorted views of the pupil. Ideally the gaze cursor should not deviate at all during this test.

Try wearing the eye tracker over glasses to see how comfortable and stable it is.

Test stability by calibrating, wearing the eye tracker for a while then testing the accuracy for the calibration points.

Test any event detection software by (1) maintaining fixation on a single point while rotating the head and (2) maintaining fixation on a smoothly moving target (for example, the tip of your own or someone else's finger as they move it). See whether the software returns these tests correctly as a single fixation event or incorrectly as a series of fixations and saccades.

13.10 Suggested Readings

Duchowski, A. T. (2007). *Eye tracking methodology: theory and practice.* Springer.
Fawcett, J., Risko, E., & Kingstone, A. (2015). *The handbook of attention.* MIT Press.
Hansen, D. W. & Pece, A. E. (2005). Eye tracking in the wild. *Computer Vision and Image Understanding*, 98 (1), 155–181.
Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: a comprehensive guide to methods and measures.* OUP Oxford.
Land, M. F. & Tatler, B. W. (2009). *Looking and Acting: Vision and Eye Movements in Natural Behaviour.* Oxford: Oxford University Press.
Hansen, D. W. & Ji, Q. (2010). In the eye of the beholder: a survey of models for eyes and gaze. *IEEE Transactions on pattern analysis and machine intelligence*, 32 (3), 478–500.

References

Ahmed, Z., Mardanbegi, D., & Hansen, D. W. (2016). Pupil center as a function of pupil diameter. In Proceedings of the ninth biennial acm sym- posium on eye tracking research & applications (pp. 283–286). ETRA '16. Charleston, South Carolina: ACM. https://doi.org/10.1145/2857491.2857536

Babcock, J. S. & Pelz, J. B. (2004). Building a lightweight eyetracking headgear. In Proceedings of the 2004 symposium on eye tracking research & applications (pp. 109–114). ACM.

Barsalou, L. W. (2008). Grounded cognition. *Annu. Rev. Psychol.*, *59*, 617–645.

Bekkering, H., & Neggers, S. F. (2002). Visual search is modulated by action intentions. *Psychological science*, *13*(4), 370–374.

Buswell, G. T. (1935). *How people look at pictures: a study of the psychology and perception in art*. University of Chicago Press.

Carpenter, R. H. S. (1981). Oculomotor procrastination. In D. F. Fisher, R. A. Monty, & J. W. Senders (Eds.), Eye movements (pp. 237–246). Hillsdale: Lawrence Erlbaum: Eye Movements: Cognition and Visual Perception.

Dicks, M., Button, C., & Davids, K. (2010). Examination of gaze behaviors under in situ and video simulation task constraints reveals differences in information pickup for perception and action. *Attention Perception & Psychophysics*, *72*(3), 706–720.

Diefendorf, A. R., & Dodge, R. (1908). An experimental study of the ocular reactions of the insane from photographic records. *Brain*, *31*(3), 451–489.

Epelboim, J. L., Steinman, R. M., Kowler, E., Edwards, M., Pizlo, Z., Erkelens, C., et al. (1995). The function of visual search and memory in sequential looking tasks. *Vision Research*, *35*, 3401–3422.

Evans, K. M., Jacobs, R. A., Tarduno, J. A., & Pelz, J. B. (2012). Collecting and analyzing eye tracking data in outdoor environments. *Journal of Eye Movement Research*, *5*(2), 6.

Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: action planning primes action-related stimulus dimensions. *Psychological research*, *71*(1), 22–29.

Foulsham, T., Walker, E., & Kingstone, A. (2011). The where, what and when of gaze allocation in the lab and the natural environment. *Vision research*, *51*(17), 1920–1931.

Furneaux, S., & Land, M. F. (1999). The effects of skill on the eye-hand span during musical sight-reading. *Proceedings of the Royal Society of London B*, *266*, 2453–2440.

Gharaee, H., Shafiee, M., Hoseini, R., Abrishami, M., Abrishami, Y., & Abrishami, M. (2015). Angle kappa measurements: normal values in healthy iranian population obtained with the orbscan ii. *Iranian Red Crescent Medical Journal*, *17*(1),

Hallett, P. (1978). Primary and secondary saccades to goals defined by instructions. *Vision research*, *18*(10), 1279–1296.

Hansen, D. W. (2014). Proceedings of the symposium on eye tracking research and applications 2014. In Conference etra' 14 eye tracking research and applications. Association for Computing Machinery.

Hansen, D. W., & Ji, Q. (2010). In the eye of the beholder: a survey of models for eyes and gaze. *IEEE Transactions on pattern analysis and machine intelligence*, *32*(3), 478–500.

Hartridge, H., & Thomson, L. (1948). Methods of investigating eye movements. *The British Journal of Ophthalmology*, *32*(9), 581–591.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). Eye tracking: a comprehensive guide to methods and measures. OUP Oxford.

Jarodzka, H., Holmqvist, K., & Nyström, M. (2010). A vector-based, Multidimensional scanpath similarity measure. In Proceedings of the 2010 sym- posium on eye-tracking research & applications (pp. 211–218). ACM.

Kingstone, A., Laidlaw, K., Nasiopoulos, E., & Risko, E. (2017). Cognitive ethology and social attention. In On human nature (pp. 365–382). Elsevier.

Kingstone, A., Smilek, D., & Eastwood, J. D. (2008). Cognitive Ethology: A new approach for studying human cognition. *British Journal of Psy-chology*, *99*(3), 317–340.

Kinsman, T., Evans, K., Sweeney, G., Keane, T., & Pelz, J. (2012). Egomotion compensation improves fixation detection in wearable eye tracking. In Proceedings of the symposium on eye tracking research and ap- plications (pp. 221–224). ACM.

Kohlbecher, S., Bartl, K., Bardins, S., & Schneider, E. (2010). Low-latency combined eye and head tracking system for teleoperating a robotic head in real-time. In Proceedings of the 2010 symposium on eye-tracking research & applications (pp. 117–120). ACM.

Land, M., Marshall, J., Brownless, D., & Cronin, T. (1990). The eye-movements of the mantis shrimp odontodactylus scyllarus (crustacea: stomatopoda). *Journal of Comparative Physiology A*, *167*(2), 155–166.

Land, M. F. (1993). *Eye-head coordination during driving*. IEEE Systems, Man and Cybernetics: Le Touquet.

Land, M. F. (2004). The coordination of rotations of the eyes, head and trunk in saccadic turns produced in natural situations. *Experimental brain research*, *159*(2), 151–160.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. Philosophical Transactions of the Royal Society of London. *Series B: Biological Sciences*, *352*, 1231–1239.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature*, *369*(6483), 742–744.

Land, M. F., & McLeod, P. (2000). From eye movements to actions: how batsmen hit the ball. *Nature Neuroscience*, *3*, 1340–1345.

Land, M. F., Mennie, N., & Rusted, J. (1999). The roles of vision and eye movements in the control of activities of daily living. *Perception*, *28*(11), 1311–1328.

Land, M. F., & Tatler, B. W. (2009). *Looking and Acting: Vision and Eye Movements in Natural Behaviour*. Oxford: Oxford University Press.

Mackworth, N. H., & Thomas, E. L. (1962). *Head-mounted eye-marker camera. JOSA*, *52*(6), 713–716.

MacLeod, C., Mathews, A., & Tata, P. (1986). Attentional bias in emotional disorders. *Journal of Abnormal Psychology*, *95*(1), 15–20.

Mardanbegi, D. & Hansen, D. W. (2012). Parallax error in the monocular head-mounted eye trackers. In Proceedings of the 2012 acm conference on ubiquitous computing (pp. 689–694). ACM.

Merchant, J. (1966). Oculometry. *Optical Engineering*, *4*(2), 040258–040258.

Merchant, J., Morrissette, R., & Porterfield, J. L. (1974). Remote measurement of eye direction allowing subject motion over one cubic foot of space. *Biomedical Engineering, IEEE Transactions on*, *4*, 309–317.

Mulligan, J. B. (2006). Optical eye models for gaze tracking. In Proceedings of the 2006 symposium on eye tracking research & applications (pp. 51–51). ACM.

Narcizo, F. B. & Hansen, D. W. (2015, December). Depth compensation model for gaze estimation in sport analysis. In Proceedings of the 2015 ieee international conference on computer vision workshop (pp. 788–795). ICCVW '15. Washington, DC, USA: IEEE Computer Society. https://doi.org/10.1109/ICCVW.2015.107

Nasiopoulos, E., Risko, E., Foulsham, T., & Kingstone, A. (2015). The social presence effect of wearing an eye tracker: now you see it, now you don't. *PERCEPTION*, *44*, 79–80.

Posner, M. I. (1980). Orienting of attention. *Quarterly Journal of Experi- mental Psychology*, *32*(1), 3–25.

Risko, E. F., & Kingstone, A. (2011). Eyes wide shut: implied social presence, eye tracking and attention. *Attention, Perception, & Psychophysics*, *73*(2), 291–296.

Risko, E. F., & Kingstone, A. (2017). Everyday attention. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie experimentale*, *71*(2), 89–92.

Shackel, B. (1960). Note on mobile eye viewpoint recording. *JOSA*, *50*(8), 763–768.

Tatler, B. W., & Land, M. F. (2016). Everyday visual attention. *Handbook of attention*, 391–422.

Tatler, B. W., Macdonald, R. G., Hamling, T., & Richardson, C. (2016). Looking at domestic textiles: an eye-tracking experiment analysing in- uences on viewing behaviour at owlpen manor. *Textile History*, *47*(1), 94–118.

Theeuwes, J. (1991). Exogenous and endogenous control of attention: the effect of visual onsets and offsets. *Perception & psychophysics*, *49*(1), 83–90.

Thomas, E. L. (1968). Movements of the eye. *Scientific American*, *219*(2), 88–95.

Wolfe, J. M. (1998). Visual Search. In H. Pashler (Ed.), *Attention* (pp. 13–73). London: University College London Press.

Young, L. R., & Sheena, D. (1975). Survey of eye movement recording methods. *Behavior research methods & instrumentation*, *7*(5), 397–429.

Chapter 14
Eye Movements in Primates—An Experimental Approach

Kevin Johnston and Stefan Everling

Contents

K. Johnston · S. Everling (✉)
Brain and Mind Institute, University of Western Ontario, London, Canada
e-mail: severlin@uwo.ca

Department of Physiology and Pharmacology, University of Western Ontario,
London, Canada

Department of Psychology, University of Western Ontario, London, Canada

S. Everling
Robarts Research Institute, London, Canada

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_14

Abstract The system controlling movements of the eyes provides an outstanding window into not only motor control, but also the neural basis of cognitive control. To derive a thorough understanding of this system, studies of neural mechanisms are necessary, and these require invasive investigations. Such investigations require an animal model. This chapter will discuss a systems neuroscience approach to understanding cognitive and oculomotor control in our best available model—the macaque monkey. We will outline the scientific rationale for the use of this model, and describe physiological methods for studying neural circuits, including neural recordings, electrophysiological identification of projection neurons, and reversible deactivation of network components. We then illustrate the use of this model and these techniques by discussing our own approach to investigating a long-standing theory of the role of the prefrontal cortex in oculomotor control—inhibitory control.

14.1 Introduction and Learning Objectives

We humans are visual animals. To convince yourself of this, imagine you are looking for a friend in a crowded café. Most likely the first strategy you will adopt is moving your eyes quickly around the scene in an attempt to locate them. This allows you to analyze the scene in greater detail by orienting the fovea, the high-resolution portion of your retina, toward objects of greatest interest or relevance. Your eye movements might be guided by a variety of factors. Objects that are conspicuous because they are brighter than the surroundings, like a shiny teapot, for instance, might attract your gaze in a bottom-up fashion. Alternatively, your gaze may be guided by prior knowledge about your friend. If you know they are likely to be wearing a red shirt, or that they have dark hair, you will most likely preferentially look toward red objects, and dark-haired people. Thus, your eye movements in this scenario may also be guided not only by the simple sensory aspects of the scene, but by more cognitive, top-down factors. With a little luck you'll find your friend.

It should be obvious from the above example that although eye movements may seem very simple, they are critical for our ability to actively gather visual information about the world around us, enormously flexible in the manner in which they can be deployed, and influenced by a number of cognitive factors. As such, they provide an intriguing window into cognitive processes. Often, changes in eye movements offer valuable insights into the brain in cases of acute neurological impairments such as stroke, neurodegenerative disorders such as Parkinson's disease, and psychiatric illnesses such as schizophrenia. Indeed, as early as 1908, eye movements in schizophrenic patients were being investigated in the pioneering work of Diefendorf and Dodge (Diefendorf & Dodge, 1908).

Such a critical and ubiquitous system obviously attracts our attention as scientists, and understanding it fully requires a variety of approaches. This chapter will discuss an approach that aims to understand directly the neural basis of the cognitive factors influencing the control of eye movements at the systems level. It is a truism that if one wishes to develop an understanding of a neural system, it is important to investigate the mechanisms by which that system operates. Ultimately, we believe that this requires knowledge of how the fundamental units of information processing in the brain, single neurons, interact to instantiate cognitive processes, and control the movements themselves. This approach is one that necessitates invasive studies in an animal model. Our goal here is to discuss the techniques of this approach in one such animal model, the behaving non-human primate, and more specifically, the macaque monkey. As with any research program using animals, it is important to understand and justify scientifically the validity of the model being used. In the case of non-human primates this is especially important due to the ethical-sensitivity of their use.

By the end of this chapter, you should be familiar with:

(1) The scientific justification for use of the non-human primate model in studies of cognitive and oculomotor control.
(2) The basic concepts of primate training and single neuron recordings in behaving animals.
(3) The neurophysiological techniques which can be used to investigate circuit properties of the network controlling eye movements, including identification of projection neurons and temporary deactivation of network components.
(4) How those techniques have been applied to evaluate a model of saccade control by the prefrontal cortex.

14.2 Historical Annotations

As we will discuss in this chapter, the system controlling eye movements is arguably the best understood sensorimotor system of the brain. This is particularly remarkable considering that the bulk of our knowledge of the neural circuits and physiological details of this system has been developed in only the last 50 years. As we have discussed here, this has been the result of the synthesis of behavioural and physiological techniques. To be sure, we have been aided by the facts that eye movements are relatively easy to measure in exquisite detail, and are readily amenable to behavioural training (Fuchs, 1967). Following the initial development of techniques for recording in awake, unanesthetized animals by Hubel (1958) and Jasper, Ricci, and Doane (1960), the leap to combining performance of a trained task with single neuron recordings was developed by Evarts (1966), to investigate the dynamics of force and position coding in the motor cortex in the rhesus macaque model. In the oculomotor field, the combination of eye movement monitoring with neural recordings in alert animals commenced at around this time, though many early studies, such as those carried out by Bizzi in the FEF (Bizzi, 1968; Bizzi & Schiller, 1970), and Fuchs, Robinson, and others, in the brainstem (Luschei & Fuchs, 1972; Robinson, 1970), correlated neural activity with spontaneous, but untrained eye movements.

The pioneering work of Wurtz (1969) extended this to include trained eye movements in investigations of response properties of striate cortex, and was soon extended to the midbrain SC (Wurtz & Goldberg, 1972). This was followed by the proliferation of studies investigating the response properties of neurons in cortical areas in the late 1970s to early 1990s. The visual and saccade-related properties of neurons in FEF (Bruce & Goldberg, 1985), DLPFC (Boch & Goldberg, 1989; Funahashi, Bruce, & Goldman-Rakic, 1989; Kojima, 1980), SEF (Schlag & Schlag-Rey, 1987; Schall, 1991), and PPC including area LIP (Andersen, Essick, & Siegel, 1987; Lynch, Mountcastle, Talbot, & Yin, 1977) were all established at this time. The oculomotor responses of basal ganglia neurons were also first investigated (Hikosaka & Wurtz, 1983). These studies laid the foundation for the many current studies investigating the neural basis of cognition using single neuron recording in the oculomotor system.

14.3 What Makes an Appropriate Animal Model?

Before discussing the non-human primate in detail, it is important to think about what we mean when using the term “animal model”. Intuitively, most of us have an understanding of what an animal model is, but it is worthwhile to think about this in some detail. Researchers in different fields employ different models such as rodents and pigeons for various reasons, such as tradition in their field and experimental convenience. Results from different labs are more easily interpreted if the same species are used. Animals of course need to be housed and cared for, and there is a large body of existing knowledge regarding how to best do this in common laboratory species. These are not inconsiderable factors, but we can go a little further. The neurobiologist Todd Preuss has written extensively on the use of animal models in neuroscientific research, and has suggested that what is truly critical is to identify an animal species in which the system of interest is similar to that of humans in terms of function and structure (Preuss, 2000). Having done so, we may then record from or modify activity in the system to determine its principles of operation, and make inferences about how the system might function in humans during health and disease. We contend that the non-human primate is the model *par excellence* for systems-level investigations of oculomotor and cognitive control, so we will look at how primates fit the functional and structural criteria in the following sections.

14.3.1 Oculomotor Behavior Is Functionally Similar in Macaques and Humans

Let us start with the functional criterion. We might ask specifically whether non-human primates exhibit eye movements similar to those of humans. With respect to the most commonly used non-human primate in oculomotor research, the macaque, this is certainly the case. As with humans, macaque monkeys are foveate animals, and the overt oculomotor behaviour of humans and macaques is strikingly similar. Both species produce saccadic and smooth pursuit eye movements with only minor quantitative differences (Fuchs, 1967). In this chapter, we will focus on control of saccades, for the reason that most cognitive studies employing oculomotor responses have used variations of saccade tasks. There is substantial overlap between the saccade and smooth pursuit systems, and it is our hope that this chapter will provide an entry point for the interested reader. There are a number of excellent reviews covering the smooth pursuit system (see Krauzlis, 2005 in suggested readings).

We can, with some effort, train monkeys to perform sophisticated oculomotor tasks, and their performance on most of these tasks is again only quantitatively different than that of human participants. One example is provided by a paradigm initially introduced for human participants by Hallett in (1978), and which we have employed extensively with rhesus macaques to investigate the neural basis of saccade control—the antisaccade task (see also Chap. 2 by Pierce et al., in this volume). In this task, participants are required to make a saccade either toward, or to the mirror opposite location of a suddenly appearing visual stimulus, depending upon the task instruction, which is often set by the colour of a fixation stimulus (see Fig. 14.1a). Saccades made to the stimulus location are referred to as prosaccades, while those to the opposite location are termed antisaccades. The beauty of this task is that while relatively simple, it does have a number of clear cognitive requirements. First, the participant must process from the instruction stimulus the correct rule to apply on a given trial—look toward or look away. Second, in the antisaccade condition, they must override the almost automatic tendency to look toward the visual stimulus (much more on this later). Third, to make a correct antisaccade, they must process the location of the visual stimulus and use this information to generate a voluntary saccade in the opposite direction. In general, performance in oculomotor tasks is often quantified by saccadic reaction time, and performance accuracy. In this task, there is a very typical profile exhibited by both humans and macaque monkeys. First, reaction times are typically longer on antisaccade than prosaccade trials, presumably because of the time required for the additional cognitive processes required for antisaccades (Fig. 14.1b). Second, performance errors, in which participants look toward rather than away from the visual stimulus in the periphery, are more frequent on antisaccade than prosaccade trials, because of the increased cognitive demands of antisaccade performance.

Antisaccades, of course, are not the only oculomotor paradigm that is performed similarly between human and monkey participants. Indeed, the list is long and encompasses a range of tasks and cognitive processes. Tasks such as visual search, in which the participant is required to saccade to the location of a target stimulus within an array of distracters (Schall & Hanes, 1993), memory-guided saccades, in which the participant is required to generate a saccade to the remembered location of a target stimulus following a delay (Funahashi et al., 1989; Hikosaka & Wurtz, 1983), saccade countermanding, in which participants are required to withhold a saccade in response to a stop-signal (Hanes & Schall, 1995), and change detection, in which the participant is presented with an array of stimuli, and required to saccade to the stim-

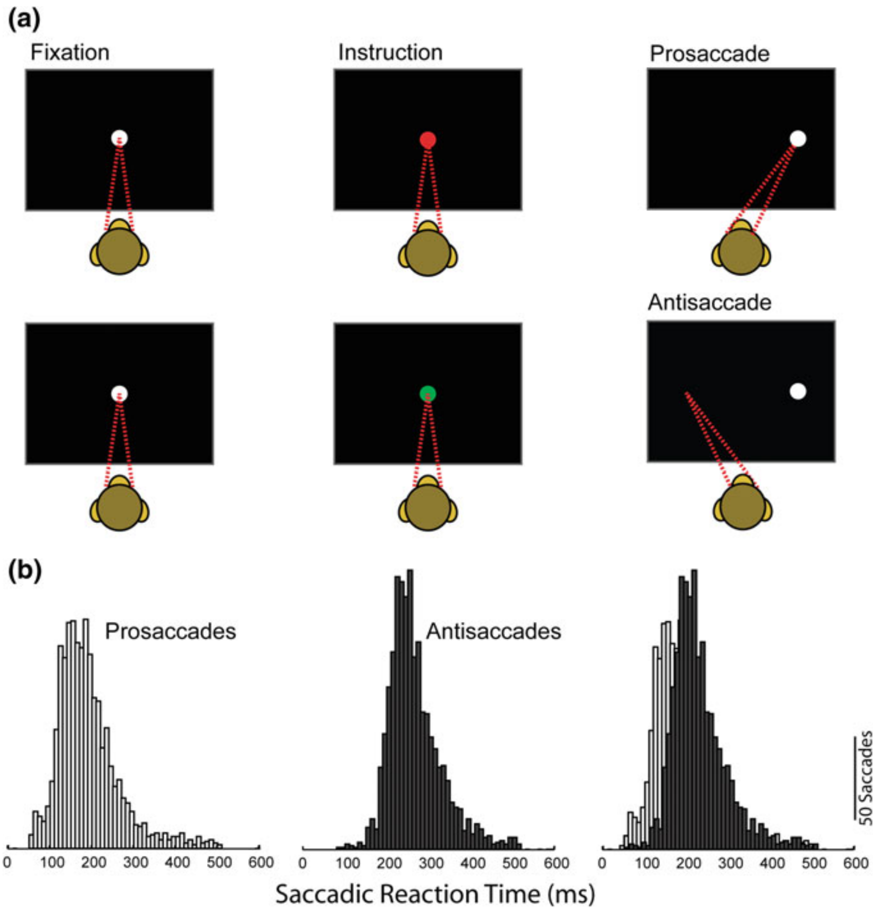


Fig. 14.1 Pro and antisaccades in the rhesus macaque monkey. **a** Animals are required to fixate, and presented with a colour instruction indicating which response will be rewarded. On prosaccade trials, animals are rewarded for a saccade toward the visual stimulus (top panel). On antisaccade trials, a saccade toward the blank space mirror opposite the visual stimulus is rewarded. **b** Histograms of reaction times on pro- (left) and antisaccade (centre) trials. Comparison is presented in rightmost panel. Reaction times on antisaccade trials generally exceeds that on prosaccade trials, because of additional cognitive processes required

ulus which has changed in a subsequent array presented after a delay (Heyselaar, Johnston, & Paré, 2011), variously require attentional, inhibitory, and mnemonic processes. The similarity of performance between the two species suggests a strong functional similarity that, in part, justifies the use of the macaque in systems neuroscience studies of cognition and saccade control.

14.3.2 Humans and Non-human Primates Have Homologous Neural Circuits in Oculomotor Control

A large body of neurobiological evidence suggests strong structural similarities in the circuits responsible for cognition and oculomotor control in humans and non-human primates. We began this chapter by stating that humans are visual animals, and indeed this is the case with non-human primates as well, and is reflected in the structure of their visual system. In macaques and other diurnal primates, the retina contains a well-developed fovea, capable of high-resolution analysis of images and characterized by an absence of rod photoreceptors, dense packing of cone photoreceptors, and an absence of blood vessels. Indeed, the goal of saccadic and smooth pursuit eye movements is to align, and to maintain alignment of this region of the retina with respect to the images of objects of interest. Primates also have an elaborate network of cortical areas responsible for visual analysis. Indeed, roughly 50% of their cortical surface is devoted to various levels of visual processing.

Outside of a well-developed visual system, there are a number of similarities in cortical areas related to cognition and oculomotor control in humans and non-human primates. It is typical within the literature to refer to such areas as *homologous*, meaning that they are shared derived characters between humans and non-human primates, inherited from a common ancestor. In practice, it is difficult to definitively establish homology with respect to brain areas, because we would need to investigate the neural organization of the common ancestor from which areas of interest were derived. This is impossible, since preserved brains are not present in the fossil record. We can however, investigate the brains of existing primates, and use other criteria to establish homology. The most common ones are anatomical, and involve establishing similarities in *cytoarchitecture*—the pattern of organization of neurons in various brain areas, and *connectivity*—the pattern of connections between different brain areas. On these bases, it has been shown that there exists an extensively interconnected network of cortical and subcortical areas that is unique to primates (Fig. 14.2a). This network includes cortical areas such as the subregions of the prefrontal cortex (PFC); the dorsolateral prefrontal cortex (DLPFC), and the premotor cortex (PMC) anterior to the primary motor cortex. Also included are the posterior parietal cortex (PPC), as well as regions of the temporal lobe including the superior temporal sulcus (STS) and inferotemporal cortex (IT). All of these areas share extensive interconnectivity with subcortical structures including the dorsal pulvinar nucleus of the thalamus, as well as the midbrain superior colliculus (SC).

Perhaps unsurprisingly, this set of areas, which has been referred to as the "neurocognitive network", established primarily on the basis of structural considerations, overlaps extensively with the network of areas responsible for control of saccadic eye movements, established using electrophysiological and lesion studies (Fig. 14.2b). The saccade network is composed of a set of interconnected regions encompassing essentially the entire brain, including cerebral cortex, thalamus, basal ganglia, cerebellum, and brainstem. At the cortical level, the saccade network includes prefrontal areas including the DPLFC, and FEF, within the PFC, the supplementary eye fields

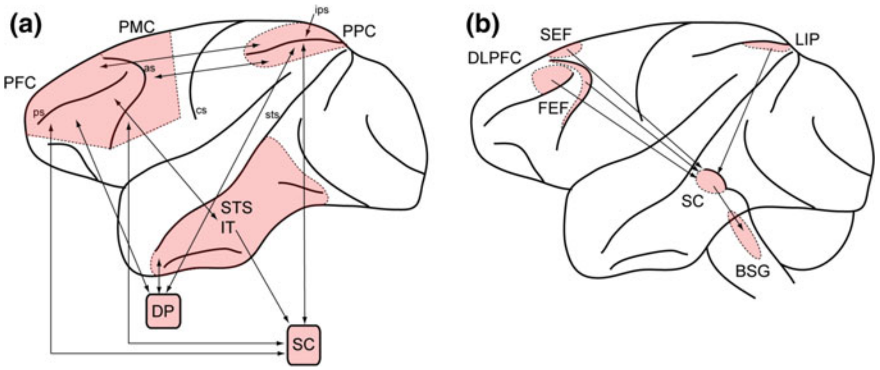


Fig. 14.2 Circuits of the macaque brain involved in cognitive and oculomotor control. **a** Macaque neurocognitive network. Network is composed of a set highly interconnected brain areas present only in primates, including prefrontal cortex (PFC), premotor cortex (PMC), posterior parietal cortex (PPC), cortex within the superior temporal sulcus (STS) and inferotemporal cortex (IT). These cortical areas are reciprocally connected with the dorsal pulvinar nucleus of the thalamus (DP), and the midbrain superior colliculus (SC). **b** Macaque oculomotor circuits. Network controlling eye movements in the macaque includes subregions of the neurocognitive network, including dorsolateral prefrontal cortex (DLPFC), and frontal eye fields (FEF), within the PFC, supplementary eye fields (SEF) within the PMC, and lateral intraparietal area (LIP) within the PPC. All of these areas send direct projections to SC, which in turn projects to preoculomotor neurons within the brainstem saccade generator (BSG) directly responsible for initiation of eye movements. Abbreviations: ps—principal sulcus, as—arcuate sulcus, cs—central sulcus, ips—intraparietal sulcus, sts—superior temporal sulcus

(SEF) within the PMC, and the lateral intraparietal area (LIP) within the PPC. All of these areas send descending projections to the SC, a region critical for saccade production. The SC, in turn, sends projections to structures in the brainstem directly responsible for saccade generation.

Although invasive studies are generally not possible in human participants for obvious reasons, studies of cortical cytoarchitecture carried out in postmortem human brains, as well as task-based and resting-state fMRI studies have revealed networks homologous to those observed in non-human primates. For example, fMRI studies have shown that the region of cortex at the junction of the precentral sulcus and superior frontal sulcus in the human brain is active during saccade tasks, and is functionally connected to posterior parietal cortex, suggesting that it is a likely homologue of the FEF established in non-human primates.

On the evidence above, it seems reasonable to make the argument that there exists a highly-derived, primate-specific network of brain areas responsible for oculomotor control, and that the SC is the centerpiece of this network. Beyond this, there are some characteristics of SC organization that are also primate specific, and are worthy of discussion. Accordingly, we discuss these, and the organization of this critical structure, below.

14.3.3 Organization of Primate Superior Colliculus

The SC resides on the roof of the midbrain and is one of the most evolutionarily conserved neural structures. The SC or its non-mammalian homologue, the optic tectum, are present in all vertebrates. In addition to this phylogenetic ubiquity, there are marked anatomical and physiological similarities in the organization of the structure across species, which are also suggestive of a common evolutionary history. The multilayered organization of this area is conserved across species, as are many of its efferent and afferent connections. Interestingly, however, there are aspects of SC organization that are primate-specific. In primates, each SC contains a map of the contralateral visual field only, while in non-primate mammals, there is at least some representation of the ipsilateral visual field. It is believed that this organization is a result of the more frontally oriented eyes of most primates, and may be related to the stereoscopic vision necessary for viewing and manipulating objects held in front of the face with the forelimb (see Box 1, Primate Origins). A second aspect of SC organization unique to primates pertains to the projections this area receives from the PFC. Only in primates do such projections from this cortical area reach multiple layers of the SC. Although the functional consequences of this organization are currently unknown, it seems reasonable to suggest that these might be involved in more cognitive aspects of eye movements, as the PFC is a relatively evolutionarily new development in the brain.

Box 1—Primate Origins

Primates are characterized by a number of unique features including a large brain, forward-facing eyes, a sophisticated visual system, and grasping hands and feet. It is intriguing to imagine how anthropoid primates like ourselves developed from our tiny tree-dwelling ancestors. How were these characters and the brain shaped by evolution?This is of course an exceptionally complex question, but we do have some clues. In general, it is thought that our earliest ancestors were arboreal. They lived in trees, and may have resembled existing small primates such as the Loris. Many of the adaptations we see today are thought to have resulted from the specialized demands of living in this environment, which has been termed the “fine branch niche”. This ecological niche required the ability to move through it easily, and grasping branches with hands and feet would have been incredibly important, as would stereoscopic vision for judging distances, and a well-developed somatosensory system to provide feedback based on touch. Of course, a large area of cerebral cortex dedicated to these sensory inputs would be important to process and coordinate all this incoming information. Even more intriguing is the theory that the evolution of a grasping foot allowed for free use of the forelimb to grasp and manipulate objects and to be used to capture prey items such as small insects. In the eloquent words of evolutionary biologist Wood Jones, it became a “free organ full of possibilities”. Along with this freedom came the ability to hold items directly in front of the face for visual analysis, and it is possible that this is one reason for the unique retino-collicular organization in currently existing primates. Certainly, the opening up of so many new behavioural possibilities would have been a major factor driving the development of the sophisticated cognitive abilities and cortical organization we see in primates today.

The SC is composed of a total of seven layers. These layers are roughly classified as superficial and deep, with the three outermost layers typically referred to as superficial, and the four remaining layers referred to as deep. The superficial layers primarily represent visual information, and contain a topographic map of the contralateral visual field, organized in retinotopic coordinates, meaning that specific locations on the retina map onto specific SC locations. The deep layers are generally involved in orienting movements of the eyes and head, and share a similar topographic organization.

In primates, the deep layers of the SC contain a saccadic motor map (see Fig. 14.3). Neurons with fixation-related activity are located in the rostrolateral pole. These neurons are tonically active during visual fixation and pause before and during saccades. SC fixation neurons discharge action potentials at a variable rate with a gradual decrease in activity before saccades and an increased post-saccadic discharge. A second broad class, the so-called saccade neurons, have a reciprocal pattern of activity. These neurons are inactive during visual fixation and discharge a high frequency burst of action potentials for saccades to a restricted part of the visual field, which defines their response field. In addition to their saccade-related motor burst, many SC saccade neurons also display low-frequency preparatory activity when there is a high predictability of stimulus appearance, and a high-frequency stimulus-related burst of action potentials. Neurons in the deep layers of the SC exert their control

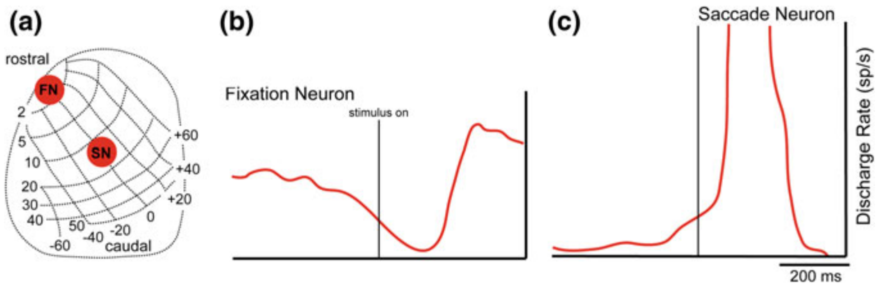


Fig. 14.3 SC motor map and example neurons. **a** SC contains a topographic map, with fixation represented at the rostral pole, and progressively larger amplitude saccades represented caudally. Upward saccades are represented medially, and downward, laterally. **b** Fixation neurons are found in the rostral SC. These neurons discharge tonically during active fixation, and decrease their discharge prior to saccade onset. **c** Saccade neurons exhibit an increase in activity, and discharge a high-frequency burst of action potentials at the time of saccades

of saccadic eye movements by projections to preoculomotor neurons located in the brainstem see Fig. 14.4.

SC output neurons provide direct excitatory inputs to long lead burst neurons (LLBN's) and omnipause neurons (OPN's) within the brainstem reticular formation. LLBN's exhibit activity similar to that of SC saccade neurons, low frequency preparatory activity, and a burst of activity during saccades. OPN's are tonically active during fixation and exhibit a pause in discharge for saccades in all directions. These two types of neurons thus exhibit a complementary, push-pull relationship during saccades, with one driving the remainder of the circuit, and the other gating it. It is believed that these signals are sent to a third class of neurons, the medium-lead burst neurons (MLBN's), which project directly to the motoneurons (MN's) controlling the eye muscles. MLBN's exhibit a discrete burst of activity time-locked to saccade onset, but not low-frequency preparatory activity. To generate a saccade, the agonist eye muscle must contract, and the antagonist must relax. This is driven by a burst of activity in MN's innervating the agonist muscle, and a silencing of activity in those innervating the antagonist. Excitatory and inhibitory MLBN's drive this reciprocal pattern of activity. This is a rather simplified description of this elegant circuit, and there are many excellent reviews that treat it in detail (see Munoz, Dorris, Pare, & Everling, 2000; Sparks, 2002 in the suggested readings).

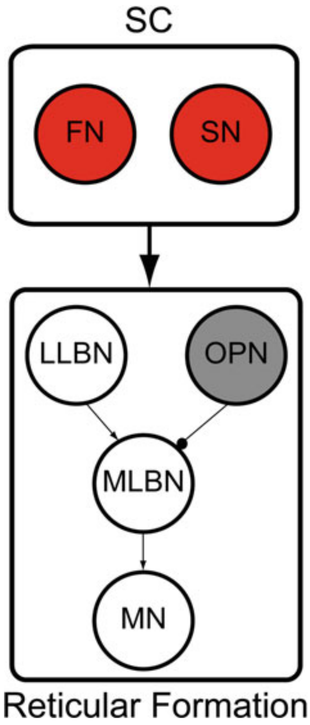


Fig. 14.4 Brainstem saccade circuit. SC fixation and saccade neurons send outputs to LLBN's and OPN's. LLBN's excite and OPN's provide tonic inhibitory control over MLBN's to which they are connected. MLBN's provide inputs to motoneurons (MN) which drive saccades

The activity of SC neurons is shaped by direct and indirect projections from a large number of cortical and subcortical areas. A key question then, is how this shaping takes place. We might ask more specifically what signals are sent from cortical areas of the network to the SC, and what effects these signals have on SC activity. We will detail some experimental approaches to these questions below. First, however, we give a brief overview of the advantages of recording neurons in primates performing oculomotor tasks.

14.4 Neuronal Recording in Behaving Non-human Primates Is a Powerful Tool for Understanding Oculomotor Function

In the sections above, we have discussed the suitability of non-human primates as an appropriate animal model in which to study the neural basis of cognitive and oculomotor control. Functionally, non-human primates have a repertoire of eye movements that differ little from that of humans, and there are striking structural similarities between humans and non-human primates that are not present between primates and other mammals. Having established this, we can turn now to the experimental approach of recording single neurons in behaving animals.

As any scientist knows, the hallmark of a good experiment is control. Single neuron recordings in the behaving primate are arguably the most direct approach to understanding the neural basis of behaviour. Methods for extracellular recordings of single units in conscious animals were first developed by Hubel (1958), and Jasper et al. (1960), and extended to include the trained, behaving rhesus macaque monkeys by Evarts (1966), in his pioneering studies of the neural basis of movement control in the limb motor system. This work was further extended to the visual system by Wurtz (1969), and to the oculomotor system by Wurtz and Goldberg (1972) (for an interesting short review, see Goldberg, 2007). This method combines two powerful tools, behavioral and systems neuroscience. Because non-human primates can be readily trained on oculomotor versions of many tasks, and because reliable, accurate, and non-invasive means of recording eye movements are available, we are able to observe behaviors carried out under strict contingencies. Pairing this with electrophysiological techniques allows us observe the neural events associated with those behaviors as they unfold. Practically, this is achieved offline by temporally aligning neural recordings, task events, and the oculomotor behaviour of the primate (see Box 2—Neural recordings and behaviour). Careful design of the experimental task in order to isolate particular cognitive functions makes it possible to gain substantial insights into the neural basis of specific aspects of cognitive control.

Box 2—Neural Recordings and Behavior

In general, our ability to make sense of the contribution of neural data to behaviour requires us to organize it in a way that makes intuitive sense. The simplest way to do this is to align neural recordings and either the events comprising our experimental task, or the behaviors exhibited by the animal. We can illustrate this by considering a simple example commonly used to characterize the response properties of neurons in many nodes of the oculomotor circuit, the memory-guided saccade, or oculomotor delayed-response task (Fig. 14.5).

In this task, the monkey is required to fixate on a central fixation stimulus for a time period of usually around 500 ms or so. Following this, a visual cue stimulus is briefly flashed at a location in the periphery. The animal is required to maintain fixation during this presentation, and through an ensuing delay period, which is usually 1000 ms or longer. At the end of this delay, the fixation stimulus is extinguished, which serves as a "go" cue to the monkey, who may then saccade to the remembered location of the cue stimulus to obtain a reward. Because this task separates temporally the presentation of the visual cue and the saccade, when we align neural activity (Fig. 14.5) we can evaluate without contamination the visual and saccade-related activity of the neuron under study. In addition, this task has been used extensively to investigate spatial working memory, since the monkey must retain in working memory the location of the cue stimulus. Many neurons in the prefrontal and parietal cortex exhibit visual and saccade-related activity, and show increased activity during the delay period of this task. Because this activity occurs in the absence of a visual stimulus, it has been referred to as "persistent", and since it bridges the temporal gap between cue stimulus presentation and the saccade to the remembered target location, it has been considered a neural signature of working memory.

The value of this approach can be illustrated using an historical example in the study of oculomotor control by the FEF. Many early studies of brain function applied two important tools to determine the function of brain areas. In essence, the approach was push-pull. Either eliminate an area by surgical excision, and characterize the resulting deficits, or experimentally activate the region using electrical stimulation and measure the resulting behavior. In the late nineteenth and early twentieth centuries, David Ferrier (1876), Fritsch and Hitzig (1870) and several others (e.g. Levinsohn, 1909), revealed that stimulation of the frontal lobes of monkeys resulted in contralateral deviations of the eyes. Findings consistent with this were subsequently obtained in humans by neurosurgeons such as Wilder Penfield in the 1930s. Taken together, these results were generally accepted as evidence for a frontal eye field specialized for the initiation of eye movements. Conflicting evidence was obtained in the 1960s, however, when the "classical" view was challenged on the basis of single neuron recordings in the FEF of conscious monkeys carried out at MIT by Bizzi and colleagues (Bizzi, 1968; Bizzi & Schiller, 1970). These early studies demon-

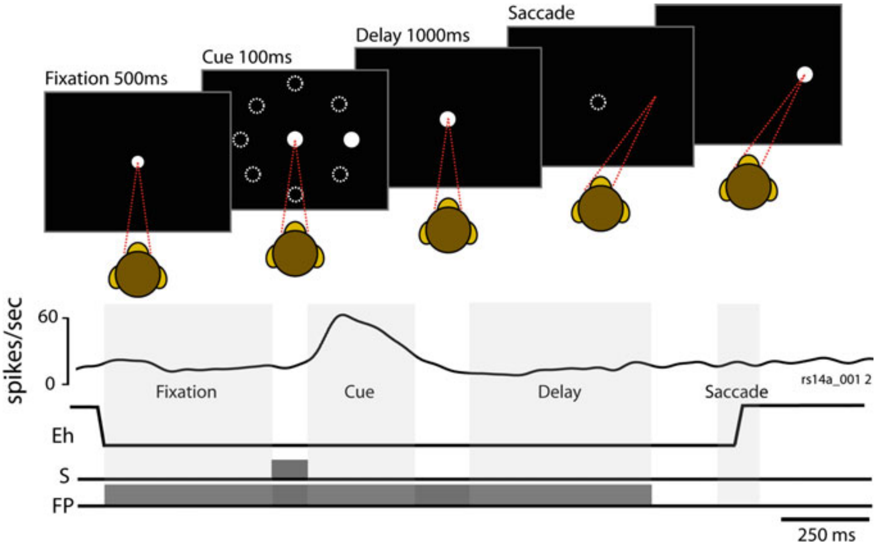


Fig. 14.5 Memory-guided saccade task and associated neural activity. Top panel, monkey is rewarded for making a saccade to the remembered location of a visual cue. Bottom panel, neural activity, eye movements, and task events. Shaded bars show onset and offset of fixation stimulus (FP), and visual cue stimulus (S), Eh represents horizontal eye position. At top, neural activity during each task epoch (shaded boxes). This example neuron exhibited activity related to the visual cue

strated that neurons in this region increased their rate of discharge *after* rather than *prior to* the initiation of voluntary but untrained eye movements, and thus could not be responsible for their initiation. This discrepancy in findings was resolved by a series of investigations revealing a role of the FEF in control of *goal-directed* or *purposive* eye movements by combining recordings of single FEF neurons with strictly controlled oculomotor tasks. These studies showed that many FEF neurons increased their activity and discharged a burst of action potentials prior to the onset of eye movements and thus could be responsible for their advance preparation and initiation (Bruce & Goldberg, 1985).

These studies demonstrated the now well-established visual and saccade-related discharges of FEF neurons in oculomotor tasks. This work provided the foundation for our current understanding of the role of the FEF in processes related to eye movements and attention, as carried out by investigators such as Schall and Moore (Moore & Armstrong, 2003; Schall 2004).

Clearly, we have gained a lot of scientific ground by combining behavioral and electrophysiological tools, and indeed, much of our knowledge of other cortical areas involved in the preparation or regulation of saccades including the DLPFC, SEF, and LIP has been derived using these methods. Similar to the FEF and SC, neurons in all of these areas exhibit responses time-locked to the onset of visual stimuli and saccades, and low-frequency activity preceding saccades within their response

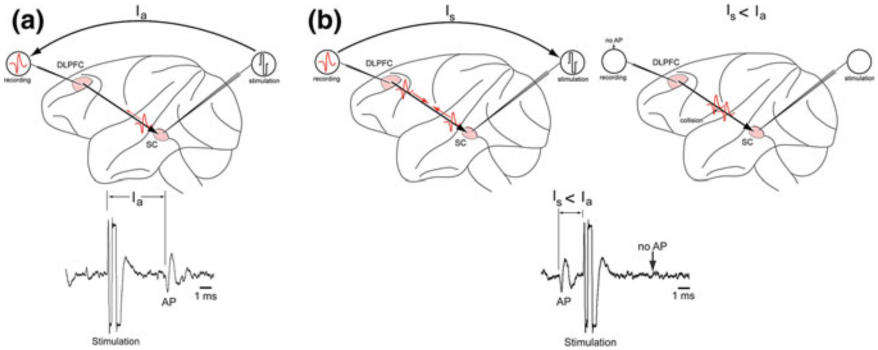


Fig. 14.6 Antidromic identification. **a** Upper panel. Stimulation electrode is placed in SC, and recording electrode in DLPFC. Stimulation pulse is applied to SC, evoked action potential travels antidromically along axon, and is recorded in DLPFC neuron under study. Duration between onset of stimulation pulse and appearance of evoked action potential in DLPFC is the antidromic latency, I_a. Lower panel. Recording record depicting stimulation, and evoked action potential (AP) occurring at I_a. **b** Collision test. Top left. Stimulation pulse is triggered by spontaneously occurring action potential in recorded neuron under study. Latency between action potential occurrence and onset of stimulation pulse is referred to as stimulation latency, I_s. If neuron under study sends direct projection to SC, spontaneous and stimulation-evoked action potentials will be traveling along same axon. Top right, if $I_s < I_a$, and action potentials are traveling along same axon, action potentials will collide, be abolished, and no action potential will be observed in recording. Lower panel, recording record showing spontaneous action potential, and triggered stimulation pulse. Because of collision between spontaneous and stimulation-evoked action potentials, no action potential is observed in recording record subsequent to stimulation pulse

fields. Where these areas seem to differ is in the degree to which their activity can be related to the *initiation* of saccades. This can be tested by combining a saccade countermanding task with neural recordings in each area. The countermanding task requires animals to either make or withhold a saccade on the basis of an infrequently occurring stop signal. For neurons in an area to be involved in saccade initiation, it follows that activity should differ between trials on which a saccade is made versus those on which it is withheld, and that these activity differences should appear before the onset of a saccade—i.e. if activity differences are responsible for launching or withholding a saccade, they must happen *before* the saccade starts. Although DLPFC has not been tested, of the other areas, only the FEF (Hanes, Patterson, & Schall, 1998) and SC meet these criteria. Interestingly, fixation neurons with response characteristics similar to those found in the rostral SC have been found in the FEF (Bruce & Goldberg, 1985; Segraves & Goldberg, 1987), and the drop in activity seen during saccades in these neurons is followed by an abrupt increase when a saccade is withheld (Hanes et al., 1998). This suggests a strong interplay between mechanisms of gaze-shifting, instantiated in the activity of saccade-related neurons, and gaze-holding, instantiated in the activity of fixation neurons, in these two areas.

The other areas of the circuit are thus regarded as having a modulatory role in saccade control, related more to cognitive processes such as working memory, critical for retention of the locations to which saccades may be directed (Funahashi et al., 1989), or visual attention, critical for filtering irrelevant stimuli and selecting saccade targets (Ipata, Gee, Golberg, & Bisley, 2006; Thomas & Paré, 2007). We discuss methods for establishing the relationship between activity thought to be related to cognitive processes in the DLPFC, their effects on SC activity, and the resultant effects on saccade behaviour in following sections of this chapter.

14.4.1 Training Non-human Primates on Oculomotor Tasks

In the section above, we placed substantial emphasis on the importance of behavioral control in the interpretability of results obtained from neural recordings in behaving non-human primates. An obvious question for a researcher from a different subfield might be “how are the animals trained?”. Accordingly, we discuss the process of training below.

Luckily for the experimenter, macaques can be readily trained to perform oculomotor paradigms using the *method of successive approximations*. Oculomotor responses that are progressively more similar to the desired response are rewarded. Typically, small amounts of fluids such as water or fruit juice are used as rewards. Fluids are preferable to solid food rewards, as the chewing movements that would accompany their delivery can interfere substantially with neural recordings by creating noise artifacts. Training in oculomotor paradigms is undoubtedly facilitated by the fact that saccades naturally occur with high frequency, the result being that animals are able to quickly learn contingencies between oculomotor responses and reward. This technique was first applied to training of oculomotor responses in monkeys by Fuchs and Robinson in 1967. In the sections below, we outline successive steps in the progression of behavioral training.

1. *Fixation training.* Since most saccade tasks require the monkey to actively fixate prior to or during the presentation a test array or saccade target, this is a necessary first step in training on all oculomotor paradigms. The simplest method for training animals to fixate is to present a small picture of an interesting object (especially another monkey), or a video clip in a small window, on an otherwise blank display. Because primates are naturally curious, they will usually look at this stimulus and this can be reinforced by providing reward when their eyes are on or at least close to the target. Initially, rewards can be delivered after only short fixations, then, over successive responses the duration may be gradually increased. Following this, the size of the stimulus and area within which eye position must be maintained may be reduced. After this, the picture or video may be replaced with a small fixation spot or cross.
2. *Saccade training.* Once monkeys have learned to fixate a small visual stimulus for a sufficient duration, we can progress to saccade training. Again, this stage is facilitated by the natural curiosity of the animals; they will generally saccade toward a suddenly appearing visual stimulus without specific training. The simplest way to train saccades then is to simply add a peripheral visual stimulus to

the fixation training protocol and modify the reward contingency so that a reward is delivered when the animal actively fixates and subsequently makes a saccade toward the stimulus when it appears.

3. *Specialized Training.* Once the monkey is able to make visually-guided saccades on 85% of trials, training in more specialized tasks may commence. Given our frequent use of the antisaccade task, we are often asked with incredulity "how can you train a monkey to do that?". It seems sensible enough then, to describe as an example the general approach we have used to train monkeys on this task. In one version of the antisaccade task, the central fixation stimulus serves as a behavioral instruction to perform either a pro or an antisaccade. A first step in training then is to establish a contingency between the fixation stimulus and the required saccade. This can be accomplished easily enough by training the animal to follow a simple matching rule. On some trials, a red fixation stimulus may be presented, followed by two simultaneously appearing peripheral stimuli, one of which is red another that is green. Saccades made to the red peripheral stimulus are rewarded, while those to the green one are not. On other trials, a green fixation stimulus would be presented, and saccades to the green peripheral stimulus rewarded. It is usually easiest for monkeys to learn this colour matching rule if trials with one fixation colour are presented in blocks of several trials in a row, followed by several trials with the other fixation colour. Following this, the fixation colours may be presented in a randomly interleaved fashion.

Once monkeys are able to perform this colour matching task, antisaccade training can start. In this phase of training, the luminance and size of one of the coloured peripheral stimuli is gradually reduced until eventually saccades are directed to a blank location. Over a number of sessions, for example, the luminance of the green stimulus may be reduced but saccades to its location continue to be rewarded. Because the animal has previously learned a colour matching rule, it will initially make saccades to the green peripheral stimulus when presented with a green fixation stimulus, and the red peripheral stimulus when presented with a red fixation stimulus. As the luminance and size of the green peripheral stimulus is reduced and it becomes progressively less visible, the animal will gradually learn to make saccades *away* from the location of the red peripheral stimulus when a green fixation stimulus is presented. Once the green peripheral stimulus is no longer visible and the animal is able to reliably make saccades *toward* the red peripheral stimulus when presented with a red fixation stimulus, and *away* from the red peripheral stimulus when presented with a green fixation stimulus, the colour of the red peripheral stimulus may be gradually desaturated. An animal whose performance has been shaped in this manner will make saccades toward the peripheral visual stimulus (prosaccades) when presented with a red fixation stimulus, and away from the peripheral stimulus when presented with a green fixation stimulus (antisaccades). This is obviously a very involved process, and it can take variable amounts of time depending upon the trainability of the participant one is working with. A few weeks to several months is a common time frame.

14.4.2 Neural Circuit Properties Can Be Investigated Using Electrophysiological Techniques

So far, we have seen that the system controlling eye movements in primates is organized as a network composed of a set of highly interconnected cortical and subcortical areas organized roughly around the midbrain SC, and that recording single neuron activity in these areas can be a fruitful approach for understanding their role in the control of eye movements. We wish to turn now to a more specific discussion of approaches we have used in the past ten years to investigate this network.

It follows from the highly interconnected nature of the system under study that attempts to understand its function must take into account not only the neural activity *within* individual nodes of the network, but also characterize the nature of the signals sent *between* them. Antidromic identification is an electrophysiological technique that allows identification of the projection targets of single neurons recorded in a given brain region. Applying microstimulation to a neuron's axon can result in an action potential travelling backwards, up the axon, to the soma. This action potential can be recorded and knowing the site of stimulation, can allow us to conclude that a recorded neuron sends a projection to a specific target region. Combining this with performance in behavioural tasks allows a functional characterization of the signals sent from one region to another, and provides insight into how signals from one area may influence another to produce complex control of eye movements.

In a typical electrophysiological experiment in which the activity of single neurons are extracellularly recorded, microelectrodes—typically constructed of metal such as tungsten—are advanced into the brain using a microdrive system which allows precise control of the electrode depth. Depending upon the nature of the study, the number of electrodes advanced can vary substantially, from one to thirty or more. Signals recorded at the microelectrode tip are filtered, amplified, and recorded along with task and behavioral events using a multichannel recording system. These signals are aligned and displayed on either an oscilloscope or computer monitor depending upon the type of recording system used. Many are commercially available. As an electrode is advanced, the neurophysiologist will listen carefully to the amplified signal through speakers or headphones, until distinct action potentials can be heard amongst the background activity of the brain. A good analogy for this would be trying to listen to a conversation at a busy party. When you are at a distance from the speakers, their voices blend into the general background noise of the room, but as you get closer, the individual words of the speakers can be discriminated. In the case of action potentials, this is manifested as a distinct "popping" sound against a background of static. Isolation is also confirmed by determining visually if waveforms distinct from background activity are present. As the electrode is moved in small increments, typically on the scale of a few micrometers, until the signal can be well-isolated from the activity of other neurons. If the amplified waveforms of electrical activity are consistent in shape and duration, and distinct from background activity, we can be reasonably confident that they represent the extracellular action potentials of a single neuron. The technical and biophysical details of this procedure are beyond the scope of this chapter, but we refer the interested reader to one of the many excellent didactic reviews of this topic (see suggested readings).

The manner in which neurons are sampled can vary depending upon physiological limitations and the aims of the experimenter. There are sampling biases inherent in the extracellular recording technique, such that more active neurons are more likely to be encountered due the fact that they discharge more spontaneous action potentials which the experimenter can detect. Additionally, larger neurons are more easily isolated than smaller ones. This is because action potential discharges of larger neurons involve a greater flow of membrane current. The amplitude of extracellularly recorded action potentials is therefore greater for larger neurons, and may be detected at greater distances from the recording electrode. The exact distance over which action potentials can be detected is difficult to establish, as the magnitude of the extracellular potential varies depending on the size and shape of the soma, and geometry of the dendritic tree, of a particular neuron. We do know that in theory, for a spherical soma, the extracellular potential could be detected at a distance of about 8 times the soma diameter (Lemon, 1984). In practice, this means that larger pyramidal output neurons tend to be sampled preferentially at the expense of smaller local-circuit interneurons. This can be counteracted to some degree by careful experimentation, and an assiduous attempt to non-selectively isolate each potential single neuron that is encountered. Depending upon the goals of the study and the brain region recorded from, one may attempt to characterize the response properties of the neuron and either select or reject it, or simply accept each isolated neuron and record its activity while the animal performs the experimental task. In some of our own studies of PFC neuronal activity, we have sampled non-selectively, since we were interested in characterizing the response properties of any neuron encountered. In others, we used antidromic identification to specifically search for PFC output neurons. This technique involves applying electrical microstimulation to the hypothesized projection target of a candidate neuron, and monitoring that neuron's activity for stimulation-evoked spikes traveling backward along it's axon. In the latter case, each isolated neuron was subjected to a series of tests, and either recorded or rejected. This method, and required tests are described below.

When a neuron's axon is electrically stimulated, an action potential is evoked which travels backward along the axon and invades the soma. Such backward-travelling action potentials are referred to as *antidromic*, in contrast to those traveling from the soma to the axon terminal, which are referred to as *orthodromic*. This stimulation-evoked antidromic action potential can be observed in extracellular recordings. By recording the activity of single neurons in a given brain area, and applying electrical microstimulation to that area's projection target of interest, neurons sending a direct axonal projection to the target area can be identified. Because electrical stimulation can excite neurons either antidromically or transynaptically, and because no single criterion is diagnostic of antidromic activation, a set of criteria must be met by evoked action potentials to distinguish direct, antidromic responses from those evoked indirectly via intervening synapses. These criteria are constant latency, constant threshold, and the collision test.

Constant latency. The latency between onset of stimulation and onset of the evoked action potential should vary only slightly. Latencies of action potentials evoked transynaptically typically vary considerably due to variability in the duration of the synaptic processes intervening between stimulation and the response, while those of antidromic action potentials are affected only by changes in the excitability of the membrane potential of the stimulated neuron.

Constant threshold. The threshold stimulus intensity required for evoking an action potential should remain stable, and increases in the intensity of stimulation should result in minimal changes in response latency.

Collision test. The most convincing criterion for antidromic activation is the collision test, which determines the presence of collision between spontaneously occurring and evoked action potentials. When a spontaneous action potential traveling along an axon in the forward, orthodromic direction encounters an antidromically evoked action potential traveling along the same axon in the opposite direction they will collide and both will be abolished. The collision test exploits this to establish antidromic activation. To apply the test, the occurrence of a spontaneous action potential in a neuron under study is used to trigger application of electrical stimulation of the target region of interest. If the spontaneous, orthodromic action potential, and the stimulation-evoked antidromic action potential are traveling along the same axon they will collide and no evoked action potential will be observed provided that both are traveling along the same axon at the same time.

To illustrate the application of this technique, let us consider an experiment in which we are interested in identifying prefrontal cortical neurons sending a direct projection to the SC. This seems like a reasonable example given the critical role of the SC in gathering cortical signals and controlling saccade initiation. A schematic of this experimental preparation is shown in Fig. 14.6. First, we must place stimulation electrodes in the projection target area of the neurons we are interested in identifying. In this case, the target is the SC. Since we are interested in the influence of cortical outputs on saccade generation, we place our stimulation electrodes in the intermediate layers of the SC.

Next, we advance microelectrodes into the PFC, and isolate a candidate projection neuron. Once we have done so, we then proceed to apply a single microstimulation pulse to the SC, and observe whether an action potential is evoked in our candidate neuron. If this is reliably the case, we then apply the criteria for antidromic identification. First, we determine the threshold current required to evoke the action potential in the candidate PFC neuron by varying the current applied to the SC, and noting the level at which the action potential is evoked 50% of the time. Using this value, we determine whether the constant latency criterion is met by applying stimulation repeatedly and noting the latency between the onset of the stimulation pulse and the evoked action potential in the candidate neuron. If this latency varies by less than about 20%, it is a reasonable assumption that the candidate neuron is antidromically rather than transsynaptically activated. This latency is then noted, and referred to as the *antidromic latency*, or $I_{a.}$

We can then repeat stimulation applications at various current levels to ensure that the criterion of constant threshold is also met.

If the criteria of constant latency and threshold are met, the collision test should be applied. This test is the most definitive method for differentiating antidromic from transynaptic activation. To apply the test, we trigger the stimulation pulse to the SC by a spontaneously action potential from our candidate neuron, and vary the onset latency of the stimulation pulse. We refer to the onset latency between the spontaneous action potential and the microstimulation pulse as I_s. Recall that $I_{a,}$ the antidromic latency, is a measure of the duration required for an action potential to travel the length of the axon of the candidate neuron. If I_s is less than the antidromic latency, the orthodromically traveling spontaneous spike and the stimulation evoked antidromic spike will be traveling along the same axon at the same time, they will collide, and no stimulation-evoked action potential should be observed in the candidate neuron. Conversely, if I_s is greater than the antidromic latency, I_a the orthodromic action potential will pass the stimulated region of the axon before the antidromic spike is evoked, and no collision will result. In this case, an action potential will be observed in the candidate neuron. If a collision is reliably observed, and the criteria of constant latency and threshold are met, the neuron under study may be considered a cortico-tectal projection neuron. We would then proceed to characterize its task-related discharge properties by recording its activity while the monkey participant performs the experimental task of interest.

This technique has been used to characterize the output signals of several nodes of the oculomotor network, particularly between cortical areas and the SC. In a subsequent section, we highlight an example from our own work, in which we investigated the nature of the signals sent from the DLPFC to the SC in the antisaccade task.

14.4.3 Neural Circuit Properties Can Be Investigated by Cryogenic Deactivation

A complementary approach to investigating the network of brain areas and their contribution to cognitive control of eye movements is deactivation of individual elements of the network. There are several methods by which this may be accomplished. Perhaps the oldest is the lesion technique, in which an area of interest is surgically removed, and the impact of this removal on behavior is evaluated. While this unequivocally eliminates any contribution of an area, there are some disadvantages of this method. Often, damage occurs to not only the area removed, but also underlying white matter, so that deficits cannot be unambiguously ascribed to the area of interest. Further, lesions cannot be reversed for obvious reasons, making it implausible to retrain animals on new tasks, and thus requiring that a relatively large number of animals take part in a given research program. For ethical reasons, and because of the large expense of purchasing and housing non-human primates, this is less than optimal. Reversible pharmacological methods are also available. Intracerebral microinjections of the GABA agonist muscimol, or the sodium channel blocker lidocaine have been used in many studies. These methods have the advantage of being reversible, but their effects are often limited to relatively small volumes of an area of interest, and may or may not result in behavioral deficits. This may be counteracted by making multiple injections within a session, but the repeated penetrations required may result in permanent damage.

An alternative method is deactivation of neural tissues using cold. In collaboration with our colleague Stephen Lomber, we have adopted the cryoloop technique as a method of investigating the contribution of subregions of the oculomotor network to behaviour and neural activity. Detailed reviews of this technique have been previously published, and are listed in the suggested readings section. Here, we provide a brief overview.

In this technique, loops fashioned from stainless steel hypodermic tubing are chronically surgically implanted either on the cortical surface or within sulci. Chilled methanol is pumped through the lumens of the loops, and the temperature of the loops is monitored by an attached microthermocouple. A schematic of this arrangement is presented in Fig. 14.7. When the cortical temperature is reduced below approximately 20 °C, synaptic activity is abolished, and neurons in the cooled region cease firing action potentials in response to sensory stimuli.

There are a number of advantages of using this technique to investigate neural networks. First, areas can be cooled and rewarmed, and hence activated and deactivated within a single session. This allows the collection of control and experimental data from the same animal within a single session. Temperatures can be reduced to below the threshold for deactivation within a few minutes, and rewarmed just as quickly. Typically, we have collected control data prior to cooling, and following

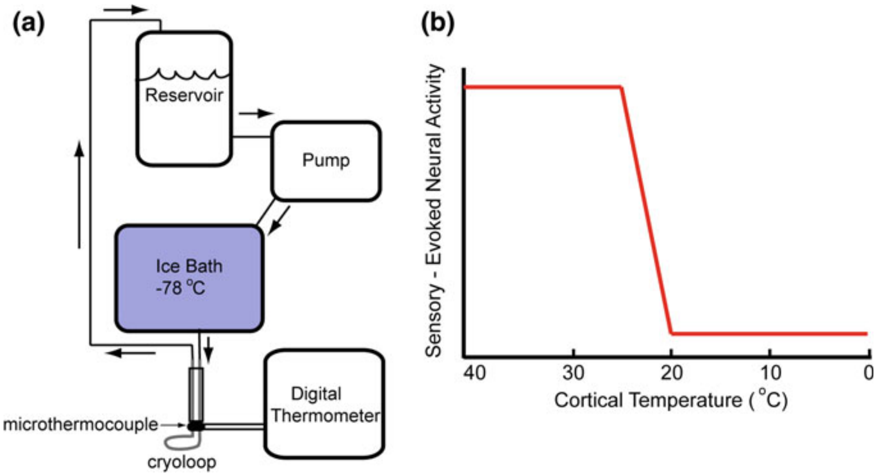


Fig. 14.7 **a** Schematic of experimental setup for cryogenic deactivation studies using the cryoloop technique. Methanol is drawn from a reservoir by a pump, and circulated through a dry ice bath, prior to being circulated through cryoloop and returned to reservoir from which it came. Temperature of the cryoloop is measured by a digital thermometer attached to a microthermocouple. **b** Evoked activity is abolished at temperatures below 20 °C. After Lomber et al. (1999)

rewarming of the cortex, to assess functional recovery. Second, larger regions of cortex can be deactivated, rendering it more likely that behavioural effects will be obtained. In a typical placement in our laboratory, cryoloops measuring roughly 6 × 3 mm are implanted. Taking into account the spread of cooling around the loop, this is sufficient to deactivate a large portion of a cortical area, and sufficient to induce changes in behavior. For example, we have recently investigated the effects of PFC deactivation on a simple saccade selection task. In this task, animals were required to fixate, and subsequently presented with peripheral visual stimuli appearing at various stimulus onset asynchronies (SOAs). A saccade to either target was rewarded, and thus they were free to choose either on any given trial (Fig. 14.8a). In general, saccades were made to left targets when they appeared first, to right targets when they appeared first, and were distributed roughly evenly between left and right targets when they were presented simultaneously. Following deactivation by cortical cooling, we observed shifts in the selection of targets such that those in the hemifield contralateral to the cooled hemisphere were chosen significantly less often (Fig. 14.8b). This change in stimulus selection was accompanied by increases in reaction times to targets in the contralateral hemifield, and decreases for those in the ipsilateral hemifield (Fig. 14.8c).

Perhaps the most significant advantage of this technique is that it permits simultaneous deactivation of one region of a network, and neural recording in another. In this way, it is possible to evaluate the functional contribution of one area of the circuit to activity in another. In a following section, we will discuss our own application of this method to investigate the contribution of the PFC to activity in the SC. More generally however, this combination of techniques is especially power-

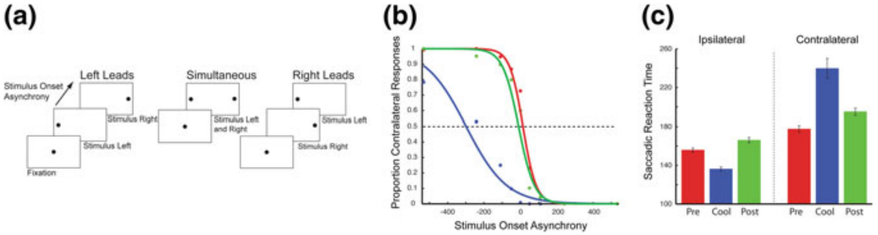


Fig. 14.8 Effects of reversible unilateral cryogenic deactivation on saccade selection, showing bias away from responses to stimuli in the hemifield contralateral to the cooled hemisphere. **a** Free-choice saccade task. Stimuli were presented at various stimulus onset asynchronies, with left stimulus leading, right stimulus leading, or with left and right appearing simultaneously. Monkeys were rewarded for making a saccade to either stimulus. **b** Proportion of saccades to contraversive saccade targets before (pre-cooling, red lines), during (cooling—blue lines), or after (post-cooling, green lines), unilateral cryogenic deactivation of PFC. Negative SOA values represent trials on which the stimulus appeared in the contralateral hemifield first, positive SOA values those on which stimulus appeared in the ipsilateral hemifield first, and zero SOA, both hemifields simultaneously. Leftward shift of blue line indicates fewer contralateral responses during deactivation. **c** Reaction times before (pre—red bars), during (cool—blue bars), and after (post—green bars) unilateral PFC deactivation. Reaction times for stimuli ipsilateral to the cooled hemisphere were reduced during cooling, and those for the contralateral hemifield were increased during cooling

ful, because it allows the simultaneous evaluation of behavior, neural activity, and deactivation-induced changes in both. This can be conceptualized as a combination of correlational and causal techniques.

The comparison of behaviour and the discharge properties of single neurons is a correlational technique. Both behaviour and neural activity are measured simultaneously, and offline, activity is aligned with task and behavioral events. Consider for example, a simple task in which a non-human primate participant is required to fixate, presented with a visual cue stimulus, and required to saccade to the remembered location of this cue following a memory delay. This is typically referred to as an oculomotor delayed response, or memory-guided saccade task. We may align simultaneously recorded PFC activity with the onset of the visual cue, the end of the memory delay, or the onset of the saccadic response, and when we do so, we typically observe cue-related, delay-related, and saccade-related activity (see Box 2). What we cannot be certain of is whether this activity is directly related to the behavior in question. This is because the link between neural activity and task events is correlational in nature—it was not experimentally manipulated, as is required to determine causality. Deactivation techniques allow us to make causal conclusions, because in this case we can directly manipulate neural activity in a given region. By combining these techniques, we can make causal inferences regarding the links between neural activity and behaviour.

14.4.4 Establishing Links Between Neural Activity and Behavior

Obviously, as neuroscientists, we are interested in establishing links between the neural activity we measure, and the sophisticated behaviour we observe in our participants. This line of thought has been especially well-developed in the fields of visual psychophysics and sensory neurophysiology, and more recently with respect to saccade generation by Jeff Schall. When we make statements such as "sustained activity of prefrontal neurons during the delay period of delayed-response tasks is the neural basis of working memory", we are really making a statement about a correspondence between neural activity and a cognitive state. This represents what has been termed a "linking hypothesis" or "linking proposition", and has been explained in detail by Teller (1984). Parker and Newsome (1998) established a set of guidelines for linking neural activity and particular perceptual states, which we can extend to cognitive states. The interested reader is directed to their review, however we transplant two of their criteria to the study of eye movements and cognitive control. First, to be involved in a particular cognitive process, neurons must provide signals relevant to the task at hand—for example, information about the presence or location of a visual stimulus, or activity related to the required response, in our case, an eye movement. Second, deactivation of neurons in a given area must affect the cognitive process in question and thus the performance of the animal in tasks requiring that cognitive process.

In the following section, we aim to synthesize the substantial ground we have covered so far, by providing a selective review of studies investigating cognitive control of oculomotor behavior using the antisaccade task.

14.5 Investigating the Role of Lateral PFC in Inhibitory Saccade Control Using a Systems Neuroscience Approach

No other area of the brain has been more closely associated with cognitive functions than the prefrontal cortex. Of these, perhaps the most prominent one is inhibitory control. As early as the late 1800s David Ferrier, in his early investigations of this area using electrical stimulation, concluded that its function was "inhibitory motor", and in the 1930s the eminent British neurologist Sir Gordon Holmes suggested that the frontal lobes acted to inhibit inappropriate activity in other brain areas.

More recently, this aspect of prefrontal function has been investigated using oculomotor paradigms, perhaps most notably the antisaccade task. As you will recall from previous sections, the antisaccade task requires a participant to look away from the location of a suddenly presented visual stimulus, and is thus thought to require inhibition of what Hess, Burgi, and Bucher in (1946) termed the "visual grasp reflex"—the automatic tendency to look *toward* suddenly presented visual stimuli. In this task, failures of inhibitory control are indexed by erroneous saccades toward the location of the visual stimulus on antisaccade trials, under the assumption that they result from an inability to override this automatic tendency. A second process is also required for correct performance—the ability to generate a voluntary saccade toward the opposite location.

In general, studies investigating performance in this task in human participants have been consistent with an inhibitory function of the PFC. The seminal evidence for this comes from human patients with PFC lesions. Initial evidence for this was obtained in the laboratory of Guitton, Buchtel, and Douglas (1985), who found increased antisaccade error rates in patients who had had large portions of their frontal lobes removed for treatment of intractable epilepsy. In that study, lesions were large and included the FEF. This work was subsequently refined by Pierrot-Deseilligny, Rivaud, Gaymard, and Agid (1991) through studies of patients with more circumscribed lesions of PFC regions. They found that lesions specifically of the middle frontal gyrus (MFG) were associated with antisaccade errors, and that these errors had a spatial bias toward the contralateral hemifield. To account for this, they proposed a model of inhibitory saccade control in which the middle frontal gyrus acts to inhibit the SC. We have termed this the classic "inhibition model" of saccade control by the PFC. fMRI studies in normal human participants are generally consistent with this view. Increased activity in the middle frontal gyrus has been observed on anti- as compared to prosaccade trials in some studies, though this picture is somewhat complex due to variability in task designs. We have reviewed these issues extensively elsewhere (see suggested readings).

Thus far, we have contended that a thorough understanding of the processes of cognitive control relating to eye movements requires investigation of the neural mechanisms instantiating control, and that studies of mechanisms require an invasive approach. We think that our preceding elaboration of the non-human primate model and systems neuroscience techniques has set the stage for a discussion of the application of these techniques to the question of inhibitory saccade control by the PFC. To recap, we have an animal that can be trained to perform pro- and antisaccades, and whose performance is similar to that of humans. We know also that our model animal has a highly similar neural circuit organization to that of humans. This is particularly so in the case of the projections from the frontal lobe to the SC, which differ between primates and non-primate mammals, and are *the* critical pathway proposed for inhibitory saccade control by the inhibition model.

14.5.1 Neural Activity in the SC Differs Between Pro and Antisaccades and Is Consistent with Inhibitory Influence

Given that the SC is a collection and integration point for cortical inputs, and that it directs the generation of saccades by projections to the saccade generator in the brainstem, a discussion of neural activity in this structure during antisaccade performance is a reasonable starting point for our discussion of neurophysiological investigations of the inhibition model.

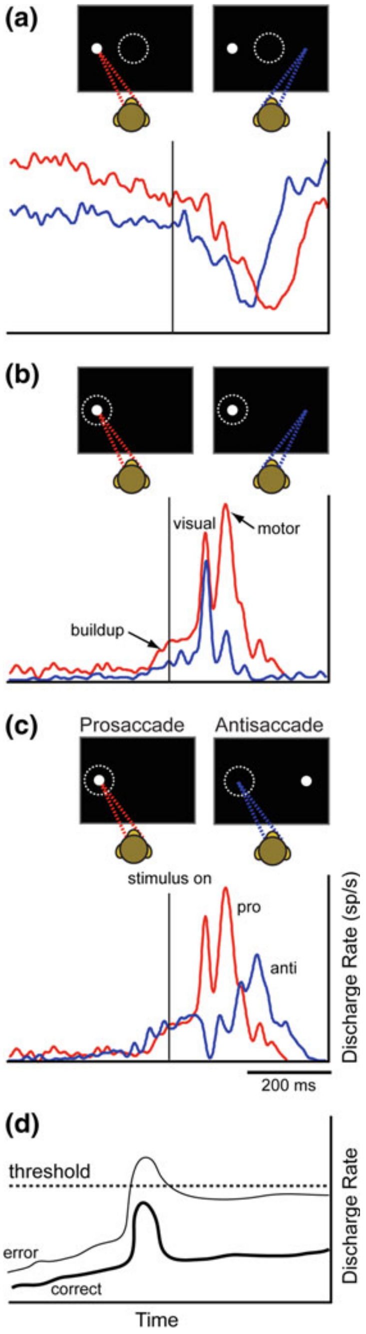
Recall from our earlier section on the SC that neurons in this structure have several properties relevant to our discussion. First, the SC has a topographic organization, with fixation and smaller amplitude saccades encoded in the rostral SC, and progressively larger amplitude saccades encoded in the caudal SC. Second, the response fields of SC neurons encompass a restricted range of the contralateral visual field. Third, SC neurons exhibit a variety of response properties including fixation-related activity, low frequency “buildup” activity preceding the onset of saccades, responses to visual stimuli, and robust saccade-related bursts of activity (see Fig. 14.3). Finally, the link between SC activity and saccade generation may be conceptualized as a complementary relationship between rostral SC neurons with fixation-related activity, and caudal SC neurons with saccade-related activity.

Activity of SC neurons in macaque monkeys performing pro- and antisaccade trials was investigated in the late 1990s, by Munoz and Everling (Everling et al., 1999). In this study, and as we have described above, monkeys were instructed to perform either a pro- or antisaccade by the colour of the central fixation stimulus. In general, we can organize neural activity in this task into epochs aligned with task events. Neural discharges recorded while animals are fixating the instruction stimulus, during which they know which response is required but not where the visual stimulus will be presented, is referred to as “preparatory” activity. A common manipulation in this task is to insert a time gap of a couple of hundred milliseconds between offset of the fixation stimulus and onset of the peripheral stimulus (see also Chap. 2 by Pierce et al.’s chapter on saccades in this volume). Neural activity recorded during this time is referred to as “gap” activity. Discharge time-locked to the onset of a visual stimulus is referred to as “stimulus-related” activity. Many SC neurons exhibit a gradual increase in discharge rates prior to the onset of saccades, and this is referred to as “buildup” activity. Finally, discharge time-locked to saccade onset is referred to as “saccade-related” activity.

This study revealed that SC fixation and saccade-related neurons exhibited complementary activity that differed between pro- and antisaccade trials. The tonic preparatory activity exhibited by fixation neurons was greater on anti- than prosaccade trials (Fig. 14.9a). In contrast, the buildup activity exhibited by a subtype of SC saccade-related neuron exhibited the inverse pattern with lower activity on anti- than prosaccade trials (Fig. 14.9b). Some trials included a “gap” manipulation. Gap activity was characterized by a decline in the tonic activity of fixation neurons and a reduction in the activity difference between pro- and antisaccade trials. Conversely, the steady increase in activity preceding the visual stimulus exhibited by buildup neurons was enhanced, and activity differences between pro- and antisaccade trials were maintained. The level of this pre-stimulus buildup activity was negatively correlated with the reaction times of saccades; greater activity was associated with reduced reaction times. No such relationship was observed for fixation neurons.

An important aspect of SC organization is intercollicular inhibition. Discharges of neurons in one SC are typically accompanied by a suppression of activity in saccade-related neurons in the SC on the opposite side of the brain. This push-pull relationship between activity in the left and right SCs prevents simultaneous programming of incompatible saccades in opposite directions. On prosaccade trials, stimulus-related responses of SC neurons were followed closely by a burst of activity beginning just prior to generation of the saccade into the neuron’s response field. In contrast, on antisaccade trials, which required a saccade away from the visual stimulus present in the response field, stimulus-related activity was attenuated (Fig. 14.9b), and followed by a similarly attenuated burst of saccade-related activity in the opposite SC that accompanied generation of the correct saccade away from the stimulus (Fig. 14.9c). During this time the activity of fixation neurons declined to minimal levels.

One way in which to establish the behavioral relevance of neural signals is to compare activity on trials in which animals perform a task correctly with those on which they made performance errors. If neural activity is behaviorally relevant, activity should differ between these two trial outcomes. Put another way, to be considered behaviorally-relevant, neural activity should predict the outcome of a given trial. To evaluate the link between SC activity and antisaccade performance, activity on erroneous antisaccade trials, in which animals made saccades toward the visual stimulus, was compared with that on correct antisaccade trials. Response errors were associated with a characteristic pattern of task-related neural discharge—an increase in buildup activity, and a burst of action potentials occurring at roughly the same time


(a)
(b)
visual
motor
buildup
(c)
Prosaccade
Antisaccade
stimulus on
pro
anti
Discharge Rate (sp/s)
200 ms
(d)
threshold
error
correct
Discharge Rate
Time

◀**Fig. 14.9** SC activity during pro- and antisaccades. Top panels show conditions associated with neural activity in each panel. Dashed circles denote response fields of example neurons. Red traces, prosaccades, blue, antisaccades. **a** During the preparatory period prior to stimulus onset, activity of fixation neurons is greater on anti-than prosaccade trials. **b** Neuron with buildup, visual stimulus-related activity, and motor burst for saccades. Both buildup and visual activity is reduced on antisaccade trials in which visual stimulus is presented in the response field. This pattern is not apparent on error trials (data not shown). **c** Same example neuron as (**b**), but depicting antisaccade trials on which saccade was made into the neuron's response field. Motor burst is reduced for antisaccade trials. **d** Thresholded accumulator proposed to explain SC results. Activity exceeding threshold triggers a saccade. On correct antisaccade trials, activity must be maintained below this threshold to prevent triggering of an erroneous saccade. Errors result when activity exceeds this level. Modified from Everling et al. (1999)

as the stimulus-related activity on correct trials but of substantially greater magnitude. A subsequent study revealed very similar discharge patterns in FEF neurons, including those antidromically identified as projecting to the SC.

Overall, these findings indicate that two criteria must be met by SC activity to ensure correct performance of antisaccades. First, activity preceding visual stimulus onset in saccade-related neurons must be reduced. Second, stimulus-related activity of SC saccade-related neurons must be attenuated. These conditions may be formalized in a thresholded accumulator model (Fig. 14.9d). In this model, saccade-related activity grows toward a fixed threshold, and a saccade is triggered when activity surpasses this level. For an antisaccade to be correctly executed, pre-stimulus and stimulus-related activity must be suppressed so that they do not exceed threshold and trigger an erroneous prosaccade. This correspondence between attenuated activity and correct antisaccade performance suggests an inhibitory mechanism that suppresses the activity of saccade-related neurons. This mechanism could be a result of activity within two possible types of neural circuits. (1) Activity within a local SC circuit, mediated by fixation neuron activity, (2) a suppressive input from a distant area. For various reasons, the local fixation neuron-based explanation has been deemed unlikely. Altogether, the results of SC recordings in the antisaccade task are suggestive of top-down inhibitory influence on SC activity that enables correct task performance.

14.5.2 *PFC as the Source of Inhibitory Top-Down Control and Potential Mechanisms*

In general, the results described above can be considered consistent with the classic inhibition model—there are differences in SC activity between pro- and antisaccade trials that look like they are shaped by inhibitory control. Combined with the human lesion studies described earlier, it seemed reasonable to suggest that PFC lesions resulted in a removal of suppressive influences on the SC, which in turn resulted in the failures of inhibitory control evident in increased errors on antisaccade trials in human patients. This conceptualization does take a step further in the sense that it proposes and provides some evidence in support of a potential neurophysiological mechanism. We can relate this back to our earlier discussion of linking propositions—if the PFC is deemed the source of inhibitory influences on the SC, we are really devising such a proposition here, namely the following: the cognitive process of inhibitory control corresponds to some physiological state of activity in the PFC that engages a mechanism causing inhibition of saccade-related activity.

This being the case, we need to investigate how inhibitory control might be instantiated by PFC activity. Based on what we know about the circuitry of the oculomotor system, we can come up with some plausible alternatives. First, the PFC could directly inhibit the SC by its direct projections to that area. Second, indirect inhibitory pathways from the PFC through either the FEF or the basal ganglia could result in SC inhibition. From what we know about cortical circuitry, PFC-mediated inhibitory control of the FEF is unlikely. In addition, we as yet have insufficient evidence to definitively implicate the basal ganglia pathway, though it remains a possibility. Accordingly, we focus on the direct projections between the PFC and SC as a potential source of inhibitory control. Here, we run into a problem for the inhibition model. Any model is obviously constrained by the nature of the circuit that instantiates it. We know that the majority of corticotectal fibres—those projecting from the cortex to the SC—target the ipsilateral SC, and originate from pyramidal neurons which are excitatory. This architecture gives us some restrictions, in that PFC-mediated inhibition would have to occur in the ipsilateral SC, and that suppression of SC saccade neurons must be carried out via excitation of either fixation neurons in the rostral SC or inhibitory interneurons in the ipsilateral SC. As mentioned above, activation of fixation neurons is an unlikely mechanism. This suggests that, if the inhibition model is correct, excitatory cortical inputs to SC inhibitory interneurons is the mechanism of suppression of SC activity (see Fig. 14.10).

14.5.3 The Response Properties of PFC Neurons Are not Sufficent to Engage Inhibitory Control

By this point, you may have guessed that we have been leading the reader up to a point in our scientific story where recording the activity of PFC neurons in a behaving macaque monkey is absolutely critical to evaluate the inhibition model, and indeed we have. Obviously, we need to investigate whether PFC activity in the antisaccade task resembles what would be expected of an inhibition signal. What might such a signal be expected to look like? Since the inhibitory model proposes that increased inhibition is present in the PFC on antisaccade trials, we might expect to observe greater neural activity on anti than prosaccade trials. Using a slightly modified version of the task designed to control for the colour of the instruction stimuli, Everling and DeSouza found that this was not the case (Everling & DeSouza, 2005). Approximately half of the PFC neurons sampled exhibited greater activity on antisaccade

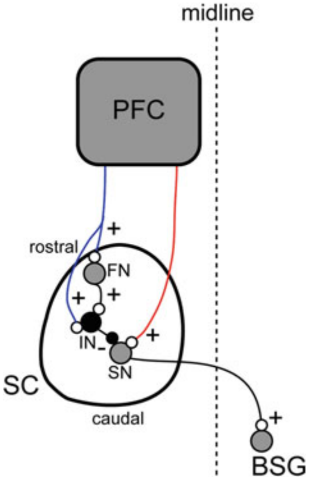


Fig. 14.10 Competing models of PFC-SC interaction. According to the inhibition model (blue lines), PFC sends an excitatory signal to SC neurons, with the net result being a reduction in activity of SC saccade neurons projecting to the brainstem saccade generator (BSG), and reduced motor activity. This is proposed to be mediated either by an excitation of fixation neurons, which excite SC inhibitory interneurons and reduce saccade neuron activity, or direct excitation of SC inhibitory interneurons. Net result of these effects would be a reduction in the activity of SC saccade neurons, and a facilitation of saccade neurons in the contralateral SC. Ultimate result would be a suppression of contraversive saccades, and a facilitation of ipsiversive saccades. Alternative model, red traces. According to this model, PFC neurons send an excitatory projection to saccade neurons, which would facilitate contraversive saccades, and suppress ipsiversive saccades

than prosaccade trials, while the other half had the opposite pattern of activity. It is worth noting however, that these neurons were randomly sampled, and may have included local circuit neurons involved in operations within the PFC in addition to projection neurons reaching targets such as the SC. To address this, we carried out a subsequent study in which we antidromically identified PFC neurons sending a direct projection to the SC, and recorded their activity during pro and antisaccades. Interestingly, we found that PFC neurons projecting to the SC (Johnston & Everling, 2006) did have greater activity on antisaccade trials, consistent with the prediction of the inhibition model. In a subsequent analysis of our earlier data, we sorted the population of recorded PFC neurons into subgroups of pyramidal output neurons and putative local-circuit interneurons using the waveform durations of the extracellularly recorded action potentials. This analysis revealed evidence for opponent-coding of task-selective responses of PFC neurons. Putative pyramidal neurons exhibited greater activity on anti- than prosaccade trials, while putative interneurons showed the inverse pattern of activity. Again, these results were consistent with the inhibition model.

The studies described above revealed that PFC neurons demonstrated the following selective responses: differences in preparatory activity preceding onset of the visual stimulus, and differences in activity following visual stimulus onset. Following stimulus onset, we found that PFC activity was maintained for roughly 130 ms, at which time it increased selectively on antisaccade trials. Such responses certainly resemble what would be expected on the basis of the inhibition model—a sustained increase in activity of PFC output neurons prior to stimulus presentation, followed by a phasic response consistent with a suppression or cancellation signal. This seems especially convincing when compared with the reduced preparatory and stimulus-related activity observed in the SC.

One way to more directly evaluate the possibility of PFC signals suppressing SC activity, as proposed by the inhibition model, is to compare the timing of selective activity. If indeed inhibitory PFC signals influence the SC to enable task performance, changes in the selectivity of PFC neurons should precede those in the SC. To evaluate this, we compared the temporal relationship between the phasic post-stimulus PFC response and SC activity (see Fig. 14.11). A direct comparison between the timing of responses of PFC and SC neurons recorded under the same conditions in our laboratory showed that enhanced PFC signals are generated too late to drive the modulation observed in SC neurons. The reduction of stimulus-related activity in SC neurons begins roughly 54 ms after stimulus onset and the saccade motor burst in the opposite SC follows at about 110 ms. The putative PFC suppression signal we observed commenced about 155 ms after stimulus onset, and based on the average antidromic latency of PFC-SC projection neurons, we know that the transmission time between these areas is about 6 ms, on average. This analysis has our PFC suppression signal reaching the SC about 161 ms after the onset of the peripheral visual stimulus, while in the SC, the saccade-related activity responsible for generating the saccade has started 51 ms *before*. Thus, the putative PFC suppression arrives at the SC too late to account for the attenuated visual responses of SC neurons observed on antisaccade trials, and later even than the onset of the motor burst driving the saccade. This is completely inconsistent with the inhibition model, since the responses of PFC output neurons are too slow to provide a plausible inhibitory signal.

14.5.4 Cryogenic Deactivation of PFC Reveals Changes in SC Activity Inconsistent with the Inhibition Model

So far, we have seen that while the activity of PFC output neurons *appears* consistent with the predictions of the inhibition model, a more detailed comparison of the timing of neural activity in PFC and SC is inconsistent with this proposal. A further, and more direct test of the model would be to deactivate the PFC, and simultaneously record the activity of SC neurons to evaluate the impact of removal of PFC inputs on SC activity. Before we describe these experiments, it is probably a good idea to review the specifics and neural circuitry of the inhibition model, so we can see if the predictions of the model match the findings.

A schematic of the model is presented in Fig. 14.10. As we have seen, the inhibition model proposes that a suppression signal is sent from the PFC to the ipsilateral SC,

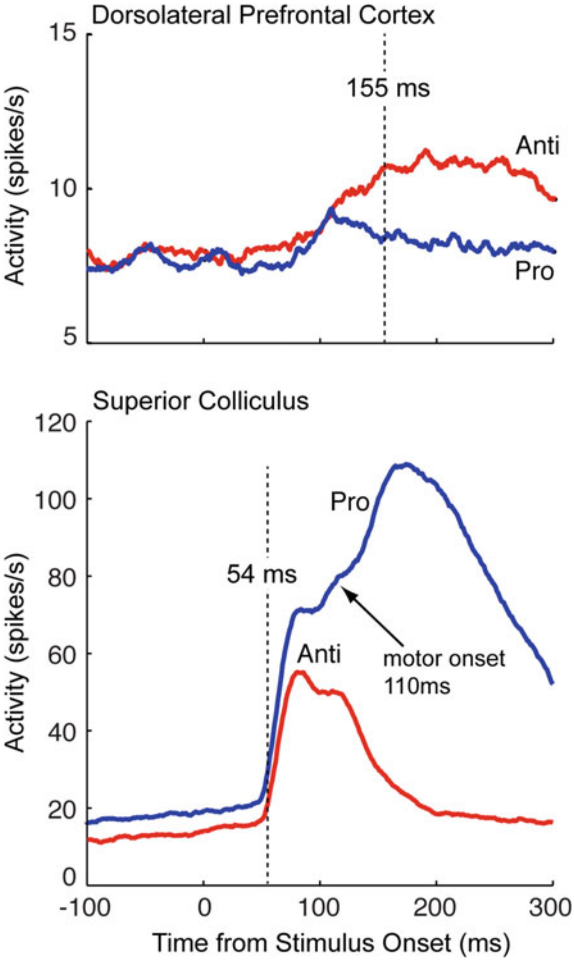


Fig. 14.11 Comparison of timing of neural activity in DLPFC and SC during pro- and antisaccades. Red traces, prosaccades, blue, antisaccades. Top panel. In SC-projecting DLPFC neurons, activity is enhanced on antisaccade trials, differing from prosaccade trials at 155 ms following stimulus onset. Bottom panel, activity in SC differs between pro- and antisaccade trials 54 ms following stimulus onset, and the motor burst on prosaccade trials commences approximately 110 ms following visual stimulus onset. DLPFC activity lags SC activity such that it is logically incapable of providing an effective suppression signal. Modified from Everling and Johnston (2014)

and leads to attenuated activity of SC saccade-related neurons. Because we know from the physiology that this descending signal is excitatory, it must be somehow transformed into an inhibitory signal in the SC itself. This could happen either by excitation of SC fixation neurons or inhibitory interneurons. Since each SC encodes saccades in the contralateral direction, the prediction is of a PFC-mediated inhibition of the ipsilateral SC, which would *suppress contraversive saccades*. Recall also that each SC inhibits the other by intercollicular inhibition, to prevent programming saccades in both directions at the same time. Because activity in the ipsilateral SC would be reduced, a secondary effect would be a reduction of intercollicular inhibition resulting in increased activity of the contralateral SC, and a *facilitation of ipsiversive saccades*. The prediction of the model then is that unilateral deactivations of the PFC should result in increased activity in the ipsilateral SC—*facilitating contraversive saccades*, and decreased activity in the contralateral SC—*suppressing ipsiversive saccades*.

To directly evaluate the model, we placed cryoloops within the region of the macaque brain thought to be homologous to the human MFG—the cortex lining the caudal principal sulcus within the PFC. Note here how the appropriate animal model, and knowledge of homologous structures is critical for establishing the correspondence between studies in humans and non-human primates. We then deactivated the PFC while simultaneously recording the activity of SC saccade-related neurons in monkeys performing pro and antisaccades. Unilateral deactivation of the PFC resulted in changes in SC activity that were inconsistent with the predictions of the inhibition model (Fig. 14.12). Instead of an activity increase in the ipsilateral SC, and decrease in the contralateral SC, we found that the activity of many buildup neurons in the ipsilateral SC decreased, while those in the contralateral SC showed an increase in activity during the prestimulus period on pro- and antisaccade trials—a finding completely opposite to the inhibition model. We also found that the onset of saccade-related activity in the ipsilateral SC was delayed significantly following deactivation. Finally, we found that stimulus-related activity in the contralateral SC increased on antisaccade trials. These changes in neural activity were reflected in changes in behavioral performance as well. Reaction times for contraversive saccades were increased, while those of ipsiversive saccades were decreased. In addition, error rates were increased on trials in which stimuli were presented in the hemifield ipsilateral to the cooled hemisphere and decreased for trials on which the stimulus was presented to the contralateral hemisphere.

The above results are completely at odds with the predictions of the inhibition model. We found that unilateral PFC deactivation lead to activity decreases in the ipsilateral, and increases in the contralateral SC. We found similarly inconsistent results in a further study in which we carried out simultaneous bilateral PFC deactivations and SC recordings. In that study, we observed consistent *decreases* in preparatory activity in both SCs, contrary to the activity *increases* that would be predicted from bilateral removal of an inhibitory input in the inhibition model. Taken together, these findings showed that the macaque homologue of the MFG provides an excitatory drive to SC saccade-related neurons. We proposed that this signal could act to facilitate performance of the voluntary antisaccade, and that rather than being involved

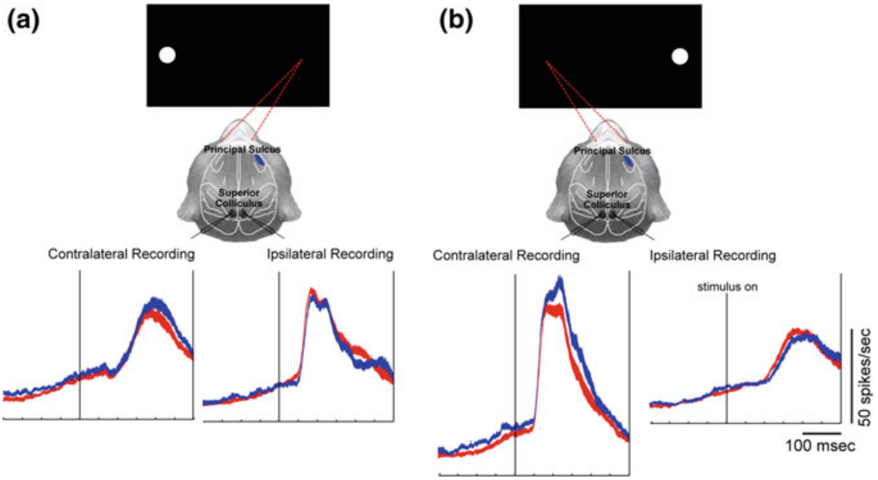


Fig. 14.12 SC recordings during performance of antisaccade trials under unilateral cryogenic deactivation of cortex lining the caudal principal sulcus of DLPFC. Blue traces, cooling, red traces, control. **a** Activity during saccades ipsilateral to the cooled hemisphere. A slight reduction in stimulus-related activity in the ipsilateral SC is observed, accompanied by an increase in saccade-related activity in the contralateral SC. **b** Activity during saccades contralateral to the deactivated hemisphere. A deactivation-induced reduction of saccade-related activity is observed in the ipsilateral SC, accompanied by an increase in stimulus-related activity in the contralateral SC. The net result of these cooling-related changes is a facilitation of saccades ipsiversive to the deactivated hemisphere, and a suppression of saccades contraversive to the deactivated hemisphere

in inhibitory inhibiting inappropriate responses, this effect was more consistent with the notion that PFC activity instead encodes task rules.

14.6 Putting It All Together

In this chapter, we have covered a lot of topics related to the study of eye movements and cognitive control in non-human primates. We have seen that this is an important and valuable model, because non-human primates share functional aspects of eye movements, and structural aspects of brain organization with humans. They can also be trained to perform many of the same tasks used in studies with human participants, and thus systems neuroscience studies in this species can provide direct insights into the neural mechanisms underlying human behaviour in health and disease. This is well-illustrated by our own investigations of the role of the PFC in inhibitory control, which have combined multiple neurophysiological methods and shown that this brain area is not involved in inhibitory control, per se, but rather acts to facilitate behavioural goals. Continued studies in this research tradition promise to provide us with mechanistic details of other cognitive functions that will enhance our understanding of these processes and potentially allow us to develop interventions to improve cognition in cases of neurodegenerative disease and mental illness.

14.7 Suggested Readings

Everling S, Johnston K (2013) Control of the superior colliculus by the lateral prefrontal cortex. Phil Trans R Soc B 368: 20130068.

– *This paper provides a detailed overview of the inhibition model of prefrontal function including lesion and fMRI studies in humans, as well as our experiments using neuronal recordings and cryogenic deactivation.*

Krauzlis RJ (2005) The control of voluntary eye movements: New perspectives. Neuroscientist 11: 124–137.

– *This accessible review discusses and compares the circuitry underlying saccades and smooth pursuit eye movements.*

Lemon R. Methods for neuronal recording in conscious animals. New York: John Wiley & Sons, 1984.

– *Although difficult to find and somewhat outdated in technical sections, this book provides an excellent and readable introduction to the biophysical basis and techniques of extracellular single neuron recording in the behaving animal.*

Lipski J (1981) Antidromic activation of neurons as an analytic tool in the study of the central nervous system. J Neurosci Methods 4: 1–32.

– *Provides an exhaustive discussion of the method of antidromic activation, as well as limitations of the technique, and its various uses in investigating neural circuits.*

Lomber SG, Payne BR, Horel JA (1999) The cryoloop: an adaptable reversible cooling deactivation method for behavioral or electrophysiological assessment of neural function. J Neurosci Methods 86: 179–194.

– *Provides a detailed review of the cryoloop technique.*

Munoz DP, Dorris MC, Paré M, Everling S (2000). On your mark, get set: brainstem circuitry underlying saccadic initiation. Can J Physiol Pharmacol 78: 934–944.

– *A very accessible account of the brainstem mechanisms controlling saccades.*

Munoz DP, Everling S (2004) Look away: The anti-saccade task and the voluntary control of eye movement. Nat Rev Neurosci 5: 218–228.

– *Provides a full review of studies investigating the neural correlates of antisaccade performance in the superior colliculus and frontal eye fields.*

Parker AJ, Newsome WT (1998) Sense and the single neuron: Probing the physiology of perception. Annu Rev Neurosci 21: 227–277.

– *This review describes the principles of making links between neural activity and perceptual phenomena, with many excellent examples. Although it does not focus on eye movements explicitly, it provides an introduction to basic principles.*

Preuss TM (2007) Evolutionary specializations of primate brain systems. In: Rovosa MJ, Dagosto M (eds) Primate Origins: Evolution and Adaptations. Springer, New York

– *This chapter provides an excellent overview of the neural circuits unique to primates and their evolutionary development.*

Schall JD (2004) On building a bridge between brain and behavior. Annu Rev Psychol 55: 23–50.

– *Provides an excellent discussion on linking hypotheses as specifically related to the initiation of eye movements.*

Sparks DL (2002) The brainstem control of saccadic eye movements. Nat Rev Neurosci 3: 952–964.

– *Provides a detailed review of the brainstem mechanisms controlling saccadic eye movements.*

14.8 Questions

(1) Why are invasive studies important to continued progress in answering neuroscientific questions?
(2) What is the difference between correlational and causal methods? How are they complementary?
(3) On what scientific grounds can we justify the use of non-human primates in oculomotor research?
(4) Why is a detailed understanding of neural mechanisms important?
(5) What are the criteria for establishing antidromic activation?

Bibliography

Andersen, R. A., Essick, G. K., & Siegel, R. M. (1987). Neurons of area 7 activated by both visual stimuli and oculomotor behavior. *Experimental Brain Research, 67,* 316–322.

Bizzi, E. (1968). Discharge of frontal eye field neurons during saccadic and following eye movements in unanesthetized monkeys. *Experimental Brain Research, 6,* 69–80.

Bizzi, E., & Schiller, P. H. (1970). Single unit activity in the frontal eye fields of unanesthetized monkeys during eye and head movement. *Experimental Brain Research, 10,* 150–158.

Boch, R. A., & Goldberg, M. E. (1989). Participation of prefrontal neurons in the preparation of visually guided eye movements in the rhesus monkey. *Journal of Neurophysiology, 61,* 1064–1084.

Bruce, C. J., & Goldberg, M. E. (1985). Primate frontal eye fields. I. Single neurons discharging before saccades. *Journal of Neurophysiology, 53,* 603–635.

Diefendorf, A. R., & Dodge, (1908). An experimental study of the ocular reactions of the insane from photographic records. *Brain, 3,* 451–489.

Evarts, E. V. (1966). Pyramidal tract activity associated with a conditioned hand movement in the monkey. *Journal of Neurophysiology, 29,* 1011–1027.

Everling, S., & DeSouza, J. F. (2005). Rule-dependent activity for prosaccades and antisaccades in the primate prefrontal cortex. *Journal of Cognitive Neuroscience, 17,* 1483–1496.

Everling, S., Dorris, M. C., Klein, R. M., & Munoz, D. P. (1999). Role of primate superior colliculus in preparation and execution of anti-saccades and pro-saccades. *Journal of Neuroscience, 19,* 2740–2754.

Ferrier, D. (1876). *The functions of the brain*. London: Smith, Elder & Co.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61,* 331–349.

Fritsch, G., & Hitzig, E. (1870). Electric excitability of the cerebrum. *Arch Anat Physiol Wissen, 37,* 300–332.

Fuchs, A. F. (1967). Saccadic and smooth pursuit eye movements in the monkey. *Journal of Physiology, 191,* 609–631.

Goldberg, M. E. (2007). Studying the visual system in awake monkeys: Two classic papers by Robert H. Wurtz. *Journal of Neurophysiology, 98,* 2495–2496.

Guitton, D., Buchtel, H. A., & Douglas, R. M. (1985). Frontal lobe lesions in man cause difficulties in suppressing reflexive glances and in generating goal-directed saccades. *Experimental Brain Research, 58,* 455–472.

Hallett, P. E. (1978). Primary and secondary saccades to goals defined by instructions. *Vision Research, 18,* 1279–1296.

Hanes, D. P., Patterson, W. F., III, & Schall, J. D. (1998). The role of frontal eye field in countermanding saccades: Visual, movement, and fixation activity. *Journal of Neurophysiology, 79,* 817–834.

Hanes, D. P., & Schall, J. D. (1995). Countermanding saccades in macaque. *Visual Neuroscience, 12,* 929–937.

Hess, W. R., Burgi, S., & Bucher, V. (1946). Motor function of tectal and tegmental area. *Monatsschr Psychiatr Neurol, 112,* 1.

Heyselaar, E., Johnston, K., & Paré, M. (2011). A change detection approach to study visual working memory of the macaque monkey. *Journal of Vision 11,* 11. https://doi.org/10.1167/11.3.11.

Hikosaka, O., & Wurtz, R. H. (1983). Visual and oculomotor functions of monkey substantia nigra pars reticulata. III. Memory-contingent visual and saccade responses. *Journal of Neurophysiology, 49,* 1268–1284.

Hubel, D. H. (1958). Cortical unit responses to visual stimuli in nonanesthetized cats. *American Journal of Ophthalmology, 46*, 110–121 (Discussion 121–112).

Ipata, A. E., Gee, A. L., Golberg, M. E., & Bisley, J. W. (2006). Activity in the lateral intraparietal area predicts the goal and latency of saccades in a free-viewing visual search task. *Journal of Neuroscience, 26,* 3656–3661.

Jasper, H., Ricci, G., & Doane, B. (1960). Microelectrode analysis of cortical cell discharge during avoidance conditioning in the monkey. *International Journal of Electroencephalography and Clinical Neurophysiology Supplement, 131,* 137–156.

Johnston, K., & Everling, S. (2006). Monkey dorsolateral prefrontal cortex sends task-selective signals directly to the superior colliculus. *Journal of Neuroscience, 26,* 12471–12478.

Kojima, S. (1980). Prefrontal unit activity in the monkey: relation to visual stimuli and movements. *Experimental Neurology, 69,* 110–123.

Krauzlis, R. J. (2005). The control of voluntary eye movements: New perspectives. *The Neuroscientist 11,* 124–137.

Lemon R. (1984). *Methods for neuronal recording in conscious animals*. New York: Wiley.

Levinsohn, G. (1909). Uber die Beziehungen der Grosshirnrinde beim Affebn zu den Bewegungen des Auges. *Archiv für Opthalmologie, 71,* 313–378.

Luschei, E. S., & Fuchs, A. F. (1972). Activity of brain stem neurons during eye movements. *Journal of Neurophysiology, 35,* 445–461.

Lynch, J. C., Mountcastle, V. B., Talbot, W. H., & Yin, T. C. T. (1977). Parietal lobe mechanisms for directed visual attention. *Journal of Neurophysiology, 40,* 362–389.

Munoz, D. P., Dorris, M. C., Pare, M., & Everling, S. (2000). On your mark, get set: brainstem circuitry underlying saccadic initiation. *Canadian Journal of Physiology and Pharmacology, 78,* 934–944.

Moore, T., & Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature, 421,* 370–373.

Parker, A. J., Newsome, W. T. (1998). Sense and the single neuron: Probing the physiology of perception. *Annual Review of Neuroscience 21,* 227–277.

Pierrot-Deseilligny, C. H., Rivaud, S., Gaymard, B., & Agid, Y. (1991). Cortical control of reflexive visually-guided saccades. *Brain, 114,* 1473–1485.

Preuss, T. M. (2000). Taking the measure of diversity: Comparative alternatives to the model-animal paradigm in cortical neuroscience. *Brain, Behavior and Evolution, 55,* 287–299.

Robinson, D. A. (1970). Oculomotor unit behavior in the monkey. *Journal of Neurophysiology, 33,* 393–403.

Schall, J. D. (1991). Neuronal activity related to visually guided saccadic eye movements in the supplementary motor area of rhesus monkeys. *Journal of Neurophysiology, 66,* 530–558.

Schall, J. D. (2004). On building a bridge between brain and behavior. *Annual Review of Psychology 55,* 23–50.

Schall, J. D., & Hanes, D. P. (1993). Neural basis of saccade target selection in frontal eye field during visual search. *Nature, 366,* 467–469.

Schlag, J., & Schlag-Rey, M. (1987). Evidence for a supplementary eye field. *Journal of Neurophysiology, 57,* 179–200.

Segraves, M. A., & Goldberg, M. E. (1987). Functional properties of corticotectal neurons in the monkey's frontal eye fields. *Journal of Neurophysiology, 58,* 1387–1419.

Sparks, D. L. (2002). The brainstem control of saccadic eye movements. *Nature Reviews Neuroscience 3,* 952–964.

Teller, D. Y. (1984). Linking propositions. *Vision Research, 24,* 1233–1246.

Thomas, N. W. D., & Paré, M. (2007). Temporal processing of saccade targets in parietal cortex area LIP during visual search. *Journal of Neurophysiology, 97,* 942–947.

Wurtz, R. H. (1969). Visual receptive fields of striate cortex neurons in awake monkeys. *Journal of Neurophysiology, 32,* 727–742.

Wurtz, R. H., & Goldberg, M. E. (1972). Activity of superior colliculus in behaving monkey. III. *Cells Discharging Before Eye Movements, 35,* 575–586.

Part III
Applications in Clinical Settings

Chapter 15
Typical and Atypical Development of Eye Movements

Christoph Klein, Divya Seernani, Chara Ioannou, Yoanna Schulz-Zhecheva, Monica Biscaldi and Michael Kavšek

Contents

Abstract The recording of eye movements is a unique method in the study of development in typical and atypical populations that has gained much valuable knowledge over the decades of its application. In the first part of this chapter, we question why and how we can, or should, study developmental processes, what eye movement recordings have taught us about the typical cognitive development. Moreover, we discuss the strengths and limitations of the paradigms that have been used here. Specifically, we introduce how eye movements reveal what newborns are able to discriminate and what this tells us about their early mental abilities and individual

---

C. Klein (✉) · D. Seernani · C. Ioannou · Y. Schulz-Zhecheva · M. Biscaldi
Department of Child and Adolescent Psychiatry,
University of Freiburg, Freiburg, Germany
e-mail: christoph.klein.kjp@uniklinik-freiburg.de

M. Kavšek
Department of Psychology, University of Bonn, Bonn, Germany

C. Klein
Department of Child and Adolescent Psychiatry, University of Cologne, Cologne, Germany

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_15

differences herein. We also outline how visual attention, closely linked to the orientation of our gaze, develops from infancy onwards. We furthermore explain the development of major eye movement types such as smooth pursuit eye movements, visually-guided saccades and saccades that are voluntarily generated such as those that bring the gaze to a remembered or instructed position in space. We also look at global and local stimulus processing and how that differentiation develops in childhood and adolescence. Finally, we take a closer look at how face processing—as investigated with eye movement recordings—develops and how we use our eye to track occluded objects. In the second part of this chapter, we will provide a definition of "atypical development", explain and discuss the paradigms that have been employed to study atypical development in psychiatric disorders of childhood and adolescence and what has been found in disorders such as autism, attention-deficit hyperactivity disorder, reading disorders and early onset schizophrenia.

15.1 Introduction and Learning Objectives

The study of eye movements is a versatile tool in the study of the human mind and brain and their development. Many important insights into the development of our mental abilities as well as cognitive or emotional/motivational dysfunctions in children and adolescents with psychiatric disorders have been gained through the measurement of eye movements. This chapter provides in its first main part an introduction to eye movement research on typical development with the aims to explain the reasons and ways of studying development; the promises of studying developmental processes using the eye movement technology; and the main findings of developmental research on eye movements and visual attention. Here, the objectives are to gain knowledge about the developmental milestones regarding ocular-motor control in healthy children and adolescents and to critically evaluate the promises and challenges of developmental eye movement research. The second main part of this chapter is dedicated to the study of psychiatric disorders of childhood and adolescence and will explain and discuss the eye movement paradigms that have been developed in this field as well as their main findings. This part should enable the reader to appreciate the different approaches that have been used to study atypical eye movements in neurodevelopmental disorders, summarize and contrast the main findings in the different disorders and provide some guidance to plan her/his own clinical eye movement study.

15.2 Historical Annotations

Among the historical sources of modern eye movement research in general, and eye movement research in child and adolescent development and psychiatry in particular, are doubtlessly Diefendorf and Dodge's (1908) ground breaking eye movements recordings in psychiatric patients, Buswell's (1936) monumental experimental work on visual exploration and Yarbus' (1967) no less monumental basic research on eye movement methodology and human vision. In addition, Holzman and colleagues' (1973) "re-discovery" of eye tracking dysfunctions in patients with schizophrenia paved the way for the science of eye movements becoming one of the most productive fields in psychiatric research.

Regarding typical development, one of the important breakthroughs to the cognitive endowments of neonates was in fact accomplished through the observation of their eye movements. Testing young infants is obviously difficult as they lack "verbal, manual and locomotor response indicators" (Fantz, 1958). Young infants, however, may show preferences for certain stimuli over others or for novel stimuli over familiar ones (Saayman, Ames, & Moffett, 1964). The crucial point here is that both preferential looking and visual orienting to novelty imply *discrimination*, either of the preferred stimulus from the alternative, or of the novel stimulus from the familiar. But if it is possible to determine what a neonate or a young infant can discriminate (e.g., colour, form, number), much can be learned about the neonate's abilities. This new paradigm proved highly productive in the subsequent decades, with one of its most intriguing findings being that individual differences in this early looking behaviour are predictive of individual differences in IQ later in life (Kavšek, 2004). Regarding the typical development of general eye movement control using standard paradigms like the pro- or anti-saccade task or smooth pursuit eye movements, the research field is still fairly "patchy", with only a limited number of researchers pursuing a systematic research programme. Noteworthy, however, is the first cross-sectional life-span developmental study on pro- and anti-saccades, published by Fischer et al. (1997) and providing evidence of different developmental courses of a fixation and a voluntary saccade system.

Regarding atypical development, the body of eye movement research in child and adolescent psychiatry is overall considerably smaller than that of adult psychiatry and gained momentum much later. A few exceptions to the rule are the studies by Pollack and Krieger (1958) who reported normal smooth pursuit in children with schizophrenia, or Dossetor and Papaioannou (1975) and Adler-Grinberg and Stark (1978), who investigated visually-guided saccades in children and adolescents with reading disorders and reported conflicting findings regarding their latency. An early study on eye movements in attention-deficit hyperactivity disorder ("hyperactivity") was published by Bala et al. (1981). They found more saccades during smooth pursuit in patients compared to healthy controls and better performance in medicated compared to non-medicated patients. Rosenhall, Johansson, and Gillberg (1988) were among the first to study eye movements in children and adolescents with autism and reported, among other findings, slower and more hypometric visually-guided saccades and difficulties in following the smooth pursuit task instructions in a majority of their patients. Despite these—and other—early publications, it was not before the nineties of the 20th century that eye movement research in child and adolescent psychiatry gained momentum. Noteworthy here is the year 2002, in which van der Geest, Kemner, Camfferman, Verbaten, and van Engeland (2002) and Klin, Jones, Schultz, Volkmar, and Cohen (2002) reported the findings of studies investigating visual social attention (that is, gaze movements) for static or dynamic social stimuli, respectively, in patients with autism and thus laid the foundations for what is as yet one of the most *cohesive* research venues of eye movement research in child and adolescent psychiatry. The following sections will describe and discuss these and other paradigms and findings in greater detail.

15.3 Typical Development

The study of development is one of the fundamental research paradigms in Psychology, complementing the differential and the experimental methodological approaches (Asendorpf 1999). While its importance is undisputed, we may still ask why this is the case and in what (general) ways we can tackle developmental issues. We also want to clarify how eye movement studies can contribute to developmental research questions and what makes them unique in this regard. So, why should we study typical development at all?

15.3.1 Why Should We Study Typical Development?

There are at least three insights that developmental research can provide that cannot be provided by the other psychological disciplines, which we have called here the historical-epistemological, nomothetic and clinical arguments.

Historical-epistemological argument—We can understand the functioning of a "system" by understanding how it has developed. Outside Psychology, there are many examples for this principle, such as the understanding of the foreign policy of a country by that country's history or the functioning of the lungs by their evolution from the guts. Within Psychology, Behaviourism provided a clear historical-epistemological stance by arguing that the behaviour of an individual can be (fully) explained by knowing its learning history. But also Piaget's stage theory is a good example of this understanding. An example from the eye movement domain is the complex ability to follow another person's gaze to join attention. Developmentally, this begins with a special interest in faces and later the eyes, followed by the ability to discriminate gaze direction, followed by the orientation to objects looked at by another person that becomes more and more differentiated and independent of the distance and location of object (Nation & Penny, 2008).

Nomothetic argument—Variables or processes that develop differently are different. Here, the developmental paradigm adds to similar rationales in the other major psychological disciplines. In Differential Psychology, for instance, the degree of correlation of individual differences quantifies the sameness or differentness of variables; whereas in Experimental Psychology, the experimental dissociation of variables reveals the dissimilarity of the underlying processes. An eye movement-related example is the dissociation of anti-saccadic direction errors with express vs regular latencies through differences in their developmental trajectories (Klein & Fischer, 2005).

Clinical argument—Understanding atypical ("abnormal") development requires knowledge of the typical ("normal") development. This truism is certainly part of most people's understanding of ab-normal processes. Unlike, for instance, in the domain of intelligence research there are however no truly developmental *normative* data but only a couple of large N studies using ad hoc samples (see Sect. 15.3.4.8) that provide some insights into the typical development of selected types of eye movements.

15.3.2 How Can We Study Developmental Processes?

As mentioned before, the term "development" connotes change, and change takes place in time. Proper developmental studies are thus longitudinal. Most published "developmental" studies, however, are cross-sectional. These studies are therefore, strictly speaking, differential research, as individual differences in variables of interest are co-varied with individual differences in chronological age. Also, cross-sectional developmental studies may prove to be a rather imprecise proxy of proper longitudinal studies (Klein, Rauh, & Biscaldi, 2011; Lindenberger et al., 2011). The reason for this somewhat unsatisfactory state of developmental research are the obvious difficulties of conducting longitudinal studies that may entail some of the threats to internal validity which Campbell (1957) has described. Therefore, longitudinal studies—or partial substitutes of it such as cross-sequential studies (Klein et al., 2011)—tracing differential developmental processes are more than rare and usually cover only short time periods (e.g., Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013; Yee, Nuechterlein, & Dawson, 1998). This holds even if only typical participants are involved, let alone including children or adolescents with psychiatric conditions.

Another issue in cross-sectional developmental research is the use of age as the main predictor. Age is obviously a continuous variable as it increases with every moment of time. Nevertheless, in many cross-sectional developmental studies this continuous nature of age is disregarded by the definition of age groups. However, in the absence of valid developmental stage theories (of a given construct and its measurement) such age-grouping is rather arbitrary. Age groups are also a statistical problem as age groups contain within-group variance that is age-dependent and contributes to the error-term in statistical tests such as ANOVA: Age-dependent variance as error-term for the test of age effects is obviously conceptually unconvincing. A continuous coverage of the age range of interest along with regression-based modelling is thus the preferred method (Klein, Foerster, Hartnegg, & Fischer, 2005; Luna, Velanova, & Geier, 2008).

15.3.3 What Can Eye Movement Studies Tell Us About Typical Development?

Eye movement recordings can provide information that is unique and can hardly be provided by other technologies. *First*, eye movement recordings provide access to one of the earliest manifestations of cognitive abilities and individual differences therein, by measuring looking preferences of newborns and infants (see also Sect. 15.3.4.1 below). Likewise, eye movement recordings can be used to measure information processing in individuals with limited mental or communicative abilities (such as understanding instructions, giving verbal answers etc.). *Second*, eye movement recordings enable process analyses of visual attention as they enable us to track how a person processes a visual stimulus or solves a problem. Typical examples are face processing (see also Sect. 15.3.4.7 below), and the solving of certain cognitive problems (e.g., the Tower of London; Hodgson et al., 2000). *Third*, eye movement recordings can be completely non-invasive and non-intrusive, up to the—ethically potentially problematic—point that participants do not even notice that their eye movements are being measured. *Fourth*, given their un-obtrusiveness, eye movement recordings are exquisitely adaptable to all kinds of tasks and many kinds of environments and are thus a highly versatile technology to provide additional information to the main information of interest (such as fMRI brain-metabolic data or event-related potentials). *Fifth*, modern eye trackers filming the eye provide potentially important physiological information "for free" in measuring the frequency of eye-lid closures ("blinks", typically identified as very short stretches of missing data) and the width of the pupil and its changes over time or in response to tasks or stimuli (see also Laeng & Alnaes' chapter on Pupillometry, in this volume). *Sixth*, collecting of eye movement data has become very affordable, with some quality eye trackers or eye-tracking glasses being commercially offered for a few thousand Euros only.

These unique features of eye movement recordings are complemented by some key "virtues" of eye movements themselves, namely importance, soundness, immediacy and richness. While "*importance*" refers to the fact that humans are "visual beings" that gather most of the information required to interact with their environments through the visual system, "*soundness*" points to the ocular-motor system as one of the most thoroughly investigated functional systems of the brain. "*Immediacy*", furthermore, means that through modern video-oculography eye movements can be directly recorded with high spatial and temporal resolution. And "*richness*" simply points to the fact that eye movement recordings provide as their regular output rich sets of meaningfully interpretable parameters, such as the latency, velocity or precision of a saccade (see also the chapter by Pierce et al., in this volume), the different parameters that characterise smooth pursuit (see also chapter by Lencer et al., in this volume) or the intricate patterns of alternating fixations and saccades during visual exploration of scenes (see chapters by Foulsham and Boccignone, in this volume).

Despite these qualities of eye movements and their recordings, looking at the developmental literature covering the ages of 2–4 years reveals a relative scarcity of eye movement studies due to a predominance of observational research at younger and language-based research at older ages. However, while behavioural observations in infants do not differentiate the complexity of ocular-motor behaviour that is present already at that age, language is not yet developed to an extent that would allow the application of methods that can be used in older participants. Eye movement recordings would thus open new avenues of developmental research in this age range in particular, allowing the differentiated observation of (ocular-motor) behaviour using a simple language. Although of course not all ocular-motor paradigms are suited for the testing of infants and young children (e.g., because they involve rather complicated task instructions), many are and would produce data that are directly comparable across broad age ranges.

15.3.4 Typical Development of Eye Movements

The following sections will explain the methods that have been used to study visual attention in infants and what this research has revealed about the early developmental of the human attentional systems. Here, we focus on the development of our ability to integrate features of objects and also on aspects of local and global processing, for instance in face processing. The development of basic eye movements such as smooth pursuit eye movements or visually-guided and voluntarily-generated saccades will be explained as well, including a view on life-span development of ocular-motor control.

As we said before, versatility is one of the features of eye movement recordings, enabling measurement not only in many situations, but also in many kinds of individuals. This is particularly useful in developmental research when very young children, who have limited means for verbal communication (both language comprehension and production), are participants. With these participants, eye movement studies can reveal information about the earliest manifestations of information processing. An important example is the habituation-dishabituation paradigm, which investigates, by means of repetitive or new stimulus exposure, what newborns or very young infants can discriminate. Section 15.3.4.1 will explain how this paradigm works and what its advantages and shortcomings are.

15.3.4.1 Gaze Movements Reveal What Infants Can Discriminate

The traditional view assumes that infants have only rudimentary visual, cognitive, and social-emotional abilities. In the 1960s, however, new experimental methods were introduced that established amazing achievements in the understanding of the psychological abilities of infants by studying the infant's tendency to respond to novel aspects of the world around them. It is a matter of fact, that novelties are interesting, not only for adults but also for infants and even newborns. If a novel stimulus appears in the visual field of a newborn, s/he orients to it. But how does the newborn know that the stimulus is new? This is obviously the case because s/he can discriminate it from what is already familiar and thus not new. Consequently, we can ask what newborns actually discriminate: A blue eye from a brown eye; a smiling face from a surprised face; a "2" from a "Z"; two dots from three dots? And what do the answers to such questions tell us about what is inborn versus acquired in human cognitive development?—These and further questions have been systematically addressed by psychologists employing the habituation-dishabituation method and other discrimination paradigms by simply exploiting the fact that newborns—like adults—show a strong tendency to look at what they identify (or "discriminate") as *new* (see also Box 1). The research interest underlying the investigation of mental faculties in the first year of life is the search for the onset and early development of these capabilities.

Imagine a situation where a mother sits with her baby on a chair in front of a blank screen. A first stimulus that is to be compared later with a second stimulus is presented to her and the baby. Not surprisingly, the baby will look at this stimulus. Depending on the design of the discrimination experiment, the stimulus is presented for either one long period or several shorter periods/trials during the first part of the study. This first part is called the "habituation period" because the baby habituates to the stimulus during its prolonged or repeated exposure trial(s). Typically, with prolonged or repeated exposures, the looking times to the habituation stimulus decrease, which is called habituation. The habituation period ends when the infant's attention to the habituation stimulus has declined to an absolute or relative level (e.g., less than 50% of the initial looking time). At this point, the *dishabituation* period starts, and both the familiar habituation stimulus *and* a novel stimulus are presented either simultaneously or successively. If the infant now looks longer at the novel stimulus than at the familiar stimulus, it is concluded that the infant is able to perceive the difference between the two stimuli, leading to "dishabituation" and a re-orienting or recovery of visual attention. The two visual stimuli to be compared are usually presented for two trials, each of which lasts up to approximately one minute. By dividing the looking time at one of the stimuli by the total length of looking times at both stimuli, a relative preference score can be computed for each participant. For a sample of infants, it can be statistically determined whether the mean relative preference score deviates significantly from chance. This version of the habituation-dishabituation method has been employed in a broad variety of perceptual and cognitive domains. In the *forced-choice novelty preference* variant, which has been used in studies on sensory competencies, the infant is presented with a series of brief trials in both the habituation and the dishabituation period (e.g., Civan, Teller, & Palmer, 2005). Each trial lasts only a few seconds. The experimental stimuli are shown side by side. In each trial of the dishabituation period, the experimenter judges which stimulus is preferred by the infant. The judgments are based on the infant's direction of first fixation, number of looks to each stimulus, and duration of looking time at each stimulus. The judgements are added up to yield the percentage of dishabituation trials in which the infant has preferred the novel stimulus.

The rationale of this ingenious research design is thus simply: If we can identify—via the recovery of attention—what a toddler or newborn is able to discriminate, we can draw conclusions about very early acquired or possibly even innate mental capabilities. Discrimination methods such as the visual habituation-dishabituation paradigm thus provide an essential avenue for investigating a variety of early competencies (although the measurement is for obvious reasons less precise than in studies using eye trackers). And in fact, this paradigm has revealed among other fascinating findings that infants have lightness constancy, that is perceive lightness as constant over variations in illumination (e.g., Chien, Bronson-Castain, Palmer, & Teller, 2006), are able to group visual patterns using form similarity (e.g., Quinn & Bhatt, 2012), conduct mental rotation (e.g., Schwarzer, Freitag, Buckel, & Lofruthe, 2013), have functional colour vision mechanisms (e.g., Kellman & Arterberry, 2006), perceive depth (e.g., Norcia, Anthony, & Holly, 2015), form categorical representations to exemplars of animals such as cats (e.g., Eimas & Quinn, 1994), or perceive—even as newborns—subjective contours (e.g., Valenza & Bulf, 2007). Altogether, these and other findings demonstrate that the investigation of eye movements and fixations has opened a unique venue to the study of early mental capacities in humans.

Having established habituation and dishabituation as mere phenomena in a large number of studies so far, the next question is how to explain habituation and dishabituation as psychological processes. Especially in younger infants, habituation and dishabituation responses might be based on simple adaptation processes—such as selective neural "fatigue"—instead of higher-order cortical information processing (e.g., Bronson, 1974; Dannemiller & Banks, 1983). Neurophysiological research, however, has provided evidence that habituation and dishabituation take place in cortical structures which are responsible for memory and information comparison (e.g., Kavšek, 2013; Reynolds, 2015). These neurophysiological insights, furthermore, support two main models of habituation and dishabituation, the *comparator model* and the *disengagement model*. Both models try to explain the phenomena and, most importantly, individual differences herein. According to the *comparator model*, both visual habituation and dishabituation indicate information processing (e.g., Sokolov, 1963). More specifically, habituation refers to the construction of a mental model of the habituation stimulus. The decrease of looking time during habituation indicates an increasing completion of a memory trace of the habituation stimulus: On each exposure trial, the infant compares the incoming visual stimulus information with the internal stimulus engram and completes it. Over successive exposure trials, the discrepancy between the visual information and the memory trace decreases. As a result, the infant's inspection of the stimulus wanes. When a novel stimulus is presented the dishabituation period begins. Again, the infant compares the stimulus with the memory trace generated during habituation. If the infant detects a difference, s/he will reactivate his/her interest and display a novelty preference. The comparator model implies that individual differences in the speed of habituation reflect individual differences in the speed of constructing the internal stimulus engram. Individual differences in the strength of dishabituation indicate differences in the ability to detect a mismatch between the habituation stimulus and a novel stimulus. Alternatively, the *disengagement model* assumes that duration of looking during habituation depends on an infant's ability to shift the attention away from the stimulus, thus reflecting the ability to *disengage* from the stimulus (e.g., Colombo, 1995).

The processes assumed in the comparator and the disengagement models are complementary events. Richards (2004) described three phases of attentional engagement, which correspond to systematic heart rate changes. When an infant is presented with a stimulus display, heart rate decelerates while the infant orients to the stimulus (up to 5 s). In the subsequent sustained attention phase, heart rate remains decelerated and the stimulus is encoded. In this period, habituation takes place. Finally, attention termination occurs: heart rate increases while the infant continues to look at the stimulus without processing it. In this phase, disengagement takes place (5–7 s). Several studies (e.g., Domsch, Lohaus, & Thomas, 2010) have suggested that, indeed, infant fixation during habituation reflects both stimulus encoding (sustained attention) and disengagement (attention termination).

Box 1: The Natural Preference and the Violation-of-Expectation Paradigms

The habituation-dishabituation method is an extension of the *natural* or *spontaneous preference* paradigm. The goal of the habituation-dishabituation method is to make the infant familiar with a stimulus during habituation. In the subsequent dishabituation period, it is tested whether the infant perceives the differences between the familiar stimulus and a second, novel stimulus. If so, the infant will prefer looking at the novel stimulus. Otherwise, the infant will distribute his/her attention equally across the two stimuli. Often, however, infants prefer looking at one of two stimuli without being habituated to one of them. That is, they display a *natural* or *spontaneous* preference. The most prominent natural preference is a preference for face-like stimuli. For example, even newborn infants prefer looking at a face-like pattern over a scrambled version of the face in which the normal spatial arrangement of the inner elements (eyes, mouth, etc.) is destroyed (e.g., Goren, Sarty, & Wu, 1975).

Both the natural preference and the habituation-dishabituation methods are primarily used to investigate perceptual and visual abilities. In the domain of cognitive, social, and emotional competencies, the *violation-of-expectation* paradigm is employed. This paradigm is an extension of the natural preference and the habituation-dishabituation methods. In the *violation-of-expectation* method, the infant is presented with rather complex events. Both an *expected* and an *unexpected* event are shown to the infant. The outcome of the expected event is consistent with a cognitive, social, or emotional principle. In contrast, the outcome of the unexpected event contradicts that principle. It is predicted that the infant would look longer at the unexpected event, if s/he is able to perceive its inconsistency. In the natural preference case of the violation-of-expectation method, it is tested whether infants use the critical principle. In the habituation-dishabituation case, the expectation is generated by a foregoing (habituation) event. For example, Wynn (1992) presented 5-month-old infants with a habituation event in which a toy disappeared behind a screen. Then, the infants saw how a second toy was placed behind the screen. In the dishabituation events, the screen was lowered. In the expected event, two toys were revealed. In the unexpected event, only one toy was revealed. Wynn found that the participants preferred looking at the unexpected event. She concluded that infants 5 months of age are able to conduct simple arithmetical operations. Other research using the violation-of-expectation method found that infants understand goal-directed intentions (e.g., Csibra, Gergely, Biró, Koós, & Brockbank, 1999), understand that objects continue to exist even when they cannot be seen (e.g., Baillargeon, 1987), and are attracted by individuals who help others, but avoid antisocial individuals (e.g., Hamlin, Wynn, & Bloom, 2007).

To sum up the difference between the two models: Whilst according to the comparator model, infants who display long looking times are slow stimulus encoders, such long-looking infants have difficulties to divert their attention away from the stimulus according to the disengagement model. Notwithstanding these differences, the two models have in common that they investigate basic attentional processes and can be used to study a variety of abilities in the visual, perceptual, emotional, social, and cognitive domains. Furthermore, despite their theoretical differences, the comparator and the disengagement models show that fundamental "dynamic" properties of human visual attention—orienting to novel stimulus information; sustaining attention in order to process novel information; and disengaging from the familiar—are present very early on in human development. As this is only an example of the early manifestations of visual attention we can ask next: How does visual attention develop?

Box 2: Problems of Discrimination Methods

Discrimination methods such as the habituation-dishabituation method assess whether infants are able to perceive the difference between two visual displays or events. When an infant prefers looking at one of two displays, it is concluded that s/he has indeed extracted the difference between them. What, however, if infants look equally at the two experimental displays? If an infant's attention is distributed equally across two displays, it cannot be concluded that s/he does not perceive a difference. Instead, both displays might, for example, be extremely salient for the infant, thereby masking a preference for one of them.

In addition, a significant dishabituation/novelty response in a habituation-dishabituation experiment must be interpreted with caution. The rationale of the habituation-dishabituation paradigm is to assess whether habituation to display A generates a significant dishabituation response to display B. More specifically, it is assumed that a preference for display B results from a comparison between the displays A and B. What, however, if infants look longer at display B than at display A because they notice a difference between the displays that was not considered by the experimenter? Take, for example, Wynn's (1992) violation-of-expectation experiment described in Box 1. According to Wynn, the findings of her experiment show that infants are able to conduct simple numerical operations. However, the assertion of counting processes in infants has been criticized by other studies which established that the outcome of Wynn's (1992) experiment might also be the effect of the infants' ability to notice mere changes in quantity, that is in the space occupied by the (one or two) toys (e.g., Cohen & Marks, 2002).

Another problem worth mentioning is the practical feasibility of discrimination problems. Observation of an infant's gaze behaviour can be difficult because babies often tend to be distracted, restless, or bored. It is then often quite difficult to ascertain whether or not the infant is inspecting the visual display. Usually, two experimenters observe and measure the infants' looking behaviour. Only if the experimenters have obtained similar measurements/looking times, these measurements are regarded as reliable and the data are included in the final data set.

15.3.4.2 Inter- and Intra-individual Differences in Stimulus Processing

Two major challenges of developmental diagnostics are the construction of tests which are suited to assess early cognitive functioning and, moreover, allow the prediction of later cognitive abilities (e.g., during childhood). Tests fulfilling both requirements could be employed to identify and describe infants who are at risk of showing later cognitive developmental delay. According to the comparator model, the habituation-dishabituation method might be a tool to measure early cognitive functioning because both habituation and dishabituation tap basic information processing mechanisms. The application of the habituation-dishabituation paradigm as a diagnostic instrument is based on the assumption made by the comparator model that infant looking contains a dynamic perspective (see Sect. 15.3.1). This can be observed on the level of inter-individual or intra-individual differences (see Table 15.1). On the *inter-individual* level, the comparator model implies that infants with long looking times during habituation are slow stimulus encoders who need a relatively long inspection time to build up an internal engram of the habituation stimulus. In contrast, infants who have short looking times and are thus assumed to reach the habituation criterion very quickly, are fast stimulus encoders. Moreover, infants who display a strong dishabituation response towards a novel stimulus have a high capability of extracting the differences between novel and familiar stimulus information. In contrast, infants with low dishabituation scores have a lower ability to detect stimulus differences. Several experimental observations confirm these basic assumptions. First, short-lookers are infants able to generate a complete stimulus engram within a short time. They show both greater dishabituation than long-lookers (e.g., McCall &

Table 15.1 Inter- and intra-individual differences in infant pattern processing

Stimulus processing abilities	Inter-individual differences	Intra-individual differences
Quantitative differences ↓	Long looking infants and slow habituators and slow stimulus encoders versus short looking infants and fast habituators and fast stimulus encoders ↓	Age-related decrease of looking time and increase of habituation speed and increase of the ability to encode visual patterns ↓
Qualitative differences	• Local versus global stimulus exploration • Smaller global precedence effect in long-lookers	• Age-related trend from local to global stimulus processing

Kagan, 1970) and a higher ability to compare familiar and novel stimulus information. Furthermore, infants at risk for cognitive disadvantages should show diminished habituation and dishabituation performance. Indeed, preterm infants, especially those suffering from risk factors such as intraventricular hemorrhage, display poorer habituation and dishabituation scores than term infants (e.g., Kavšek & Bornstein, 2010). In addition, several studies provide evidence for lower habituation and dishabituation scores in infants with other disorders such as neurological dysfunction and Down syndrome. Finally, habituation and dishabituation are predictive for later cognitive performance in both risk and non-risk samples. More specifically, individual differences in infant looking have been demonstrated to predict individual differences in cognitive variables and academic achievement in childhood and early adolescence (e.g., Bornstein, Hahn, & Wolke, 2013; Rose, Feldman, Jankowski, & van Rossem, 2008). Such inter-individual differences are complemented by *intra-individual* differences such as the decrease of infant looking at visual displays between 10 weeks and 6 months of age, which is considered to reflect an age-related improvement in stimulus processing, according to the comparator model.

These *quantitative* differences in the ability to encode and compare visual displays are complemented by *qualitative* stimulus processing differences that, again, exhibit both intra-individual developmental changes and inter-individual differences. These qualitative differences relate to the question of whether perceiving more the isolated, local details of a visible stimulus or an integrated, coherent global composition. Developmental research has addressed age-related individual differences in this local-global distinction, and eye tracking studies in particular have contributed to this issue by showing how participants of different age groups scan visual patterns.

Several studies on intra-individual differences observed a tendency in very young infants to visually scan only limited portions of visual patterns, especially edges and vertex areas (e.g., Salapatek, 1968). From 2 to 3 months of age onward, infants start to scan more widely over the stimulus area (e.g., Bronson, 1990). The increase in habituation performance observed during the course of the first year of life, hence, *co-varies* with a change in the kind of stimulus exploration. With respect to inter-individual differences, Bronson (1991) found that 3 months-old infants that were found to be slow encoders behaved like 2 weeks-old infants in showing prolonged inspection times and tending to fixate one part of a target. In contrast, fast encoding infants 3 months of age engaged in short inspection times and extensively scanned the complete target (see also Jankowski & Rose, 1997). Overall, by concentrating on the individual features of a stimulus, younger infants tend to be engaged in a local mode of object processing. At 4 months of age at the latest, by extracting the overall stimulus as a whole, infants are able to conduct global object processing. In addition, the *global precedence effect* is more pronounced in short-looking than in long-looking infants 4 months of age (e.g., Colombo, Freeseman, Coldren, & Frick, 1995). The global precedence effect is an adult-like pattern of stimulus processing: the overall configuration is analysed first, followed by an analysis of the local stimulus components.

Habituation and dishabituation performance is thus closely linked not only to quantitative trends in the development of stimulus processing but also to qualitative processing modes. Bronson's (1991) study provides evidence for a link between mode of stimulus perception and habituation performance. A local processing mode prevails in slowly habituating and encoding infants. On the other hand, a configurational mode of object processing can be established in fast encoding infants. These observations provide evidence that some basic visual processing modes are present from early on. Neurophysiological research has scrutinized the brain structures involved in local and global stimulus processing. Colombo (1995) proposes an explanation according to which these findings as well as the global precedence effect result from differential processing modes and speeds of the magnocellular and parvocellular streams which are the major visual pathways of the visual system. The magnocellular stream is tuned to low spatial frequencies and the global characteristics of a stimulus. The parvocellular stream is a high spatial frequencies system which processes local, detailed stimulus characteristics. The magnocellular system transports information faster into higher cortical areas than the parvocellular system, thereby producing the global precedence effect. Moreover, Colombo (1995) points out that the parvocellular stream may mature earlier than the magnocellular stream. As a result, processing of local details develops earlier than processing of global stimulus aspects. Furthermore, long-looking infants/slow habituators have difficulties in conducting global stimulus processing by means of the magnocellular pathway. The local and global processing modes are hence assigned to the parvocellular and magnocellular streams of information processing.

Neurophysiological research, however, has highlighted the contribution of yet another processing mechanism on the local-global processing modes: The ability to group local visual components into unitary, global structures develops into childhood, adolescence and adulthood (e.g., Scherf, Behrmann, Kimchi, & Luna, 2009). Eye movements in (younger) children are strongly influenced by *bottom-up* processes, which reflect the influence of local, salient stimulus features on fixations. The impact of bottom-up variables decreases with age. Moreover, *top-down* processes, which reflect the influence of knowledge and other cognitive processes, become more influential on eye movements (Açik, Sarwary, Schultze-Kraft, Onat, & König, 2010). Compared with older children and adults, younger children display a higher degree of focal fixations, that is, relatively long fixations and saccades of short amplitudes.

Fig. 15.1 Kanizsa Illusory Contours (KIC) as a test of global perception (from Nayar et al., 2015, with permission)

Furthermore, they display a lower degree of ambient fixations, that is of short fixations along with large amplitude saccades. By about 8–10 years of age, adult-like scores are achieved (Helo, Pannasch, Sirri, & Rämä, 2014). According to Helo, Pannasch, Sirri, and Rämä (2014), the dominance of the focal fixation mode in younger children reflects their tendency to process local stimulus features.

Research on perception of subjective contours exemplifies the development of local and global perception. Perception of subjective contours refers to the ability to integrate small pieces of a contour into a larger structure: a shape, for instance, is perceived in the absence of complete physical edges. Looking into the middle of subjective shapes indicates global perception, that is extraction of the illusory shape (see Fig. 15.1). In contrast, fixation of the shape-inducing elements—that is, the parts of the subjective shape which are physically given—indicates perception of the local elements of the stimulus figure without noticing the global shape. For example, Nayar, Franchak, Adolph, and Kiorpes (2015) investigated 3- to 10-year-old children as well as adults. They found local looking patterns in the younger age groups but a global perceptual organization in the older age groups. The younger children looked predominantly at the shape-inducing elements. From 7 to 8 years of age onward, children behaved like adults and tended to predominantly fixate the middle of illusory figures.

Subjective contours have also been investigated using *visual search tasks*. In these tasks, the participant has to identify a critical item mixed in with other items (“distractors”). It is measured how fast the participant finds the critical item. Fast responses are interpreted as indicating an automatic, parallel, pre-attentive processing mode. Slow responses are interpreted as indicating a serial, attention-demanding processing mode. Studies using the visual search paradigm found that both infants and adults are able to detect *real shapes* automatically by employing a parallel processing mode. In contrast, adults, but not infants, displayed an automatic detection mode for *subjective contours*. Instead, infants 6 months of age scanned the visual pattern in a serial mode in order to find the subjective contour (Bulf, Valenza, & Simion, 2009). Even though infants are able to respond to subjective contours, their ability to group the local elements of a target is not adult-like, according to the authors.

15.3.4.3 Development of Attentional Systems

As we know from Experimental Psychology, the term “attention” is a rather multifaceted one (e.g., Posner & Rothbart, 2007; Reynolds, Courage, & Richards, 2013; Ruff & Rothbart, 1996). Visual attention involves a complex interplay between anatomical structures such as the anatomy of the eyes with the visually highly resolving fovea centralis, observable behaviours such as selection of a target stimulus or duration of looking, and latent constructs such as the ability to maintain attention and the disengagement of attention. Moreover, these factors are coordinated by a mixture of internal or “top down” and external or “bottom up” factors. The development of this interplay has been explained to some degree by the development of two basic attention systems, the posterior and the anterior attention systems. The *posterior, reactive system* becomes functional early in the first year of life. It is involved in voluntary disengagement of fixation and in the control of voluntary saccades. It includes the *spatial* “*where*” *orienting network*, which includes areas in the parietal cortex and the “*what*” *object recognition network*, which includes the pathways from the primary visual to the inferior temporal cortex. The spatial “where” orienting network directs attention to locations of objects. The “what” network mediates exploration of object details. The posterior system leads an infant to orient to and to explore novel objects and events; attention is sustained as long as the object or event is novel. These structures control the habituation behaviour of infants: In the course of *habituation*, stimulus exploration diminishes with increasing stimulus processing. The posterior system determines the development of attention/looking duration in the first year of life: between 10 weeks and 6 months of age, duration of looking at a visual display shows a clear decrease (e.g., Courage, Reynolds, & Richards, 2006).

At around 6–12 months of age, the *anterior, executive attention system* becomes functional. It includes frontal brain structures which contribute to voluntary, endogenous control of attention (e.g., Colombo & Cheatham, 2006). One purpose of the executive system is to control the posterior attention system, and gaining this control is crucial to the development of visual attention. Especially, the anterior system is involved in the ability to focus attention in response to potentially relevant information; it helps attending to and exploring interesting stimulus information. Due to the emergence of this system, infants start to inhibit attention to distracters and to maintain sustained attention to interesting displays. The anterior system thereby facilitates object exploration. Functionality of the anterior attention system results in a gradual increase of duration of looking after around 6 months of age. The ability to maintain attention develops through childhood into adolescence (e.g., Aring, Grönlund, Hellström, & Ygge, 2007).

As part of the discussion on the development of attention, there is a debate on the cognitive foundation of fixation behaviour in the first 2–3 months of life. Will memory processes take place in very young infants, that is will infants use the information processing mechanisms postulated by the comparator model of infant habituation-dishabituation? Alternatively, young infants’ visual behaviour might be determined solely by external factors, for example by the saliency of the stimulus. The ability to fixate is present in newborns. From birth to about 10 weeks of age, time spent looking at a stimulus increases, according to several studies (e.g., Slater, Brown, Mattock, & Bornstein, 1996). Young infants sometimes even cannot avert their gaze from the stimulus. This visual behaviour is referred to as *obligatory attention* or *sticky fixation* (e.g., Hood, 1995). Indeed, attention during the first three months of life is assumed to be strongly reflexive, meaning that the infant's attention is stimulus-driven and attracted to salient stimulus features in a reflex-like manner (e.g., Reynolds, 2015). Such results underline the importance of external factors for fixation behaviour in very young infants.

Typically, two procedures have been employed to study the further development of the fixation system: fixation of a target *without distraction* and fixation of a target *with distraction*. While the former is suitable for investigating the basic metrics of fixations, e.g. distance between fixation and a centre of gravity and number of intrusive saccades, the latter introduces an increased cognitive component and is often seen as an inhibition task. Stability of fixation while foveating a stationary object without distraction increases with increasing age through infancy and childhood (e.g., Ygge, Aring, Han, Bolzani, & Hellström, 2005). The number of intrusive saccades and the mean distance of fixations from the center of gravity decreases with increasing age, while the fixation time without intrusive saccades increases during childhood and adolescence (Aring et al., 2007; Paus, Babenko, & Radil, 1990). Adult-like performance is observed between 12 and 15 years of age. In addition, children appear to exhibit more square wave jerks, defined as back-to-back saccades with an intersaccade interval of 80–450 ms, than typically reported in adults (Salman, Sharpe, Lillakas, & Steinbach, 2008). Young infants show considerable fixation scatter, but uninstructed adults can exhibit even greater scatter. Hainline, Harris, and Krinsky (1990) conclude that scatter in infants is limited not only by foveal immaturity; rather, as is the case in adults, attentional factors contribute largely to infants' fixation scatter. The ability to sustain attention and to inhibit responses to distracting information shows a significant increase between 8 and 10 years of age (Paus, 1989). This developmental progress is considered to be related to developing cognitive processes, resulting from the maturation of structures of the frontal cortex (e.g., Paus et al., 1990).

15.3.4.4 Smooth Pursuit Eye Movements

While the previous paragraph dealt with the stability of fixating stationary, nonmoving objects, this paragraph will explain how our ability to fixate smoothly moving objects develops. While the former process entails the inhibition of large-amplitude saccades—but the execution of small-amplitude or "micro-saccades" (see the chapter by Alexander & Martinez-Conde, in this volume)—the fixation of smoothly moving objects requires another eye movement system, the smooth pursuit system (see chapter by Lencer et al., in this volume).

According to several studies, smooth pursuit eye movements can be detected even in the newborn period. This ability, however, is still immature. Two variables play a major role: *size* and *speed* of the moving targets. Smooth pursuit eye movements are restricted to targets moving at low velocities (e.g., Kremenitzer, Vaughan, Kurtzberg, & Dowling, 1979; Lengyel, Weinacht, Charlier, & Gottlob, 1998). Furthermore, studies using small targets show that infants younger than 2 months of age do not display smooth pursuit (e.g., Aslin, 1981). Infants 2 months of age accomplish the "tracking" of fast targets (40°/s) primarily by saccadic eye movements. By 5 months of age, saccadic movements are to a great extent (approximately 70%) replaced by smooth pursuit eye movements (von Hofsten & Rosander, 1997). With small targets, younger infants tend to combine smooth pursuit eye movements with saccades. Nevertheless, the ability to switch from smooth pursuit to saccadic tracking at very high object velocities also improves in the first months of life (Richards & Holley, 1999). This ability facilitates stimulus tracking of fast moving objects. In addition, young infants can sustain smooth pursuit eye movements only for rather short duration and with a very low gain.[1] According to Pieh, Proudlock, and Gottlob (2012), smooth pursuit gain becomes adult-like by the age of 6 months.

Rosander (2007) points out that the rapid changes of smooth pursuit movements in the first months of life depend crucially on development of the prefrontal cortex and cerebellar circuitries. Overall, smooth pursuit eye movements undergo a dramatic development during the first 6 months of life: while the duration of smooth pursuit eye movements increases, the gain increases and the amount of compensatory catch-up saccades decreases (von Hofsten & Rosander, 1997). The *trajectory* of stimulus motion also plays a role: pursuit of vertical targets is inferior to pursuit of horizontal targets not only for adults and children but also for infants (e.g., Grönquist, Gredebäck, & von Hofsten, 2006; Ingster-Moati et al., 2009). Studies on the further development of smooth pursuit have primarily investigated differences in gain. Smooth pursuit eye movements improve through childhood until mid-adolescence, particularly for fast moving targets, and are described to be fully developed in late adolescence. In addition, children are more intra-individually variable than adults (e.g., Katsanis, Iacono, & Harris, 1998).

A smooth pursuit-related phenomenon is the tracking of objects that are *temporarily occluded*. Developmental studies on visual tracking of objects which are temporarily hidden by an occluding surface while moving provide insights into the emergence and development of both gaze behaviour and object representation. In such situations, the observer has to internally represent the object and imagine its trajectory while it is hidden from view. In adults, moving objects elicit smooth pursuit tracking. When the object moves behind an occluding surface, the smooth gradual eye movements are interrupted. Instead, gaze is shifted with saccades across the occluding surface. Once the moving object reappears behind the occluding surface, smooth tracking is continued. Infants younger than 3 months of age, however, cannot predictively track temporarily occluded objects (Rosander & von Hofsten, 2004). From 4 months of age onwards, infants develop the ability to shift the gaze across the occluding surface to the re-appearance side before the object re-emerges (von Hof-

---

[1] Smooth pursuit gain is defined as the ratio of eye velocity to target velocity (see Lencer et al., in this volume). A smooth pursuit gain of 1.0 thus indicates that the eyes move as fast as the target stimulus.

sten, Kochukhova & Rosander, 2007), suggesting that already 4-month-olds maintain a representation of objects and their motion during occlusion. Tajik-Parvinchi, Lillakas, Irving, and Steinbach (2003) presented children aged 4–16 years and adults with a moving object, a part of which was constantly occluded. The participants were instructed to pursue the invisible part of the object. The youngest children (4–6 years of age) were unable to maintain their gaze close to the location of the occluded object part. Instead, they tended to periodically look at the visible part of the moving stimulus. Moreover, these children's pursuit performance was inferior to that of the older participants.

15.3.4.5 Visually-Guided Saccades: Accuracy, Velocity, and Latency

While the previous sections dealt with the maintenance of visual attention and gaze on stationary or moving stimuli as well as in the presence or absence of visual distractors, other studies have investigated how dis-engagement of visual attention and re-engagement to a new visual stimuli develop from childhood into young adulthood and beyond. Visually-guided saccadic eye movements, or "pro-saccades", re-locate the centre of fixation from one stimulus to another. They can be described in terms of the precision or accuracy in reaching the new location, their velocity and, in the case the saccade is a response to the onset of a stimulus or another "signal", by their latency, the reaction time to initiate an eye movement (see chapter by Pierce et al., in this volume).

Saccadic eye movements can be observed right after birth. At that developmental stage, they are, however, typically *hypometric*. Hypometric saccades do not quite reach the location of the target stimulus, and are thus complemented by subsequent saccades or a combination of small saccades and head movements (e.g., Harris, Jacobs, Shawkat, & Taylor, 1993; Roucoux, Culee, & Roucoux, 1983). Bronson (1990) observed substantial improvement in the *accuracy* of stimulus scanning between 2 and 14 weeks of age. Hypometria is also reported for young children in some studies (e.g., Munoz, Broughton, Goldring, & Armstrong, 1998). Alahyane et al. (2016) found a decreased saccade accuracy in toddlers 7–42 month of age compared to adults. More specifically, the toddlers' first saccade amplitude was shorter and more variable. They compensated for this hypometria by making more corrective saccades than the adult participants. In general, the majority of studies report adult-like levels of accuracy in (older) children and adolescents (e.g., Bucci & Seassau, 2014; Salman et al., 2006a). Taken together, the published evidence suggests that saccade accuracy is adult-like by approximately 8 years of age.

Regarding *saccade velocity*, even infants' saccades have been reported to be as fast as those of adults (Garbutt, Harwood, & Harris, 2006). However, infants' saccade velocity depends on the kind of stimulus being processed. According to Hainline, Turkel, and Abramov (1984), infant saccade speed does not deviate from adult saccade speed when viewing textured patterns. In contrast, infants make slower saccades than adults when presented with simple geometric forms. Studies with older participants yielded mixed results. Several of the studies found no age-related effects on peak velocity, the highest velocity reached during a saccade, in childhood, adolescence and adulthood (e.g., Munoz et al., 1998). By contrast, some other studies found a decrease of saccadic peak velocity between childhood and adulthood (e.g., Fioravanti, Inchingolo, Pensiero, & Spanio, 1995).

Regarding the *latency of saccades*, already during the first months of life a rapid decline between 2 and 6 months of age, a general speed-up of efficiency of information processing and an improved ability to disengage from fixation and to shift the gaze to another position has been reported (see Kavšek, 2013). This trend continues further into late adolescence or early adulthood, where mean reaction times reach minimum latencies for horizontal and vertical saccades (Ajrezo, Wiener-Vacher, & Bucci, 2013; Bucci & Seassau, 2014; Fukushima, Hatta, & Fukushima, 2000; Klein & Foerster, 2001; Klein, Rauh, & Biscaldi, 2011; Luna, Garver, Urban, Lazar, & Sweeney, 2004; Regal, Ashmead, & Salapatek, 1983). The same pattern holds for the intra-subject variability of saccadic latencies (that is, the short-term fluctuation of latencies), which is much higher in early infancy than in adulthood (Aslin & Salapatek, 1975; Fukushima et al., 2000). For example, Alahyane et al. (2016) report that response latencies to stimulus appearance were higher and more variable in toddlers than in adults.

15.3.4.6 Voluntarily Generated Saccades

While the dis-engagement and re-engagement of attention from an attended to a new stimulus inducing visually-guided saccades (or “pro-saccades”) is a fairly simple and automatic process that some researchers have even called a “saccadic grasp reflex” (Hess et al., 1946), other tasks have been developed that require a much greater involvement of the anterior attention system than the ones described before. These tasks are first and foremost the anti-saccade task and the memory-guided saccade task.

During the *anti-saccade task*, the same stimuli are presented as during the pro-saccade task described before. The only difference is that participants are instructed to look straight in the opposite, mirror-image position of the peripheral cue (Hallett, 1978). During the *memory-guided saccade task*, participants have to encode the spatial position of a briefly flashed cue whilst keeping their gaze at the central fixation point as long as it remains visible; only after the extinction of the central fixation are they allowed to fixate the remembered position of the previously flashed cue. What both tasks have in common is that they seem to be simple, but in fact comprise of several sub-components: In both tasks, the “reflex-like” saccade to the peripheral cue must be inhibited, and an instruction has to be held in short-term memory to define the landing point of the impending saccade. While in the anti-saccade task, this landing point is “computed” by a transformation of the saccade vector to the mirror-image position of the cue, requiring a stimulus-incompatible response, it is the memorized cue position that defines the saccadic landing point in the memory-guided saccade task. A third commonality of the two paradigms is, hence, that they both challenge functions of the dorsolateral prefrontal cortex by involving the maintenance of an instruction in short-term memory, the inhibition of a response as well as the execution of a response that is stimulus-incompatible (anti-saccade) or memory-based (memory-guided saccade).

It is worth noting that even infants at the age of 4 months can be trained to inhibit automatic “reflexive” saccades to a cue (Johnson, 1995). Of course, infants cannot be verbally instructed to look to a place other than that where the cue stimulus appears. How did Johnson, therefore, succeed in showing inhibition of automatic saccades in infants? Johnson (1995) conducted a series of trials. In each trial, presentation of the cue stimulus was followed by presentation of a salient rotating and colourful target stimulus. It was hypothesized that infants would make a saccade to the cue stimulus in the first trials of the experiment. Over trials, however, due to the saliency of the target stimulus, the infants would learn to inhibit looking at the cue stimulus and to shift their gaze toward the more salient and interesting target. Indeed, these predictions were confirmed by the experimental results and demonstrate in an experimentally elegant manner that already very young infants can indeed inhibit “reflex”-like saccades.

Measuring eye and head latencies, Nakagawa and Sukigara (2007) reported that the likelihood of correctly executed anti-saccades increased with age in infants 3–11 months old. Endogenous control of saccades has made decisive progress around 4 months. This ability might be due to the development of cortical regions, especially the frontal eye fields, which are involved in oculomotor control. Exertion of voluntary inhibition of reflexive eye movements improves with age. Scerif et al. (2005) found correct anti-saccades in toddlers aged 8–38 months, with the number of correctly executed anti-saccades increasing with age. From the age of about 3 years, children are in principle able to perform verbally instructed anti-saccades, even if their performance is poor. During the further course of development, all anti-saccade task parameters improve, including a reduction in direction errors, speed-up of correct anti-saccade latencies, increased likelihood of error corrections and faster correction reaction times (Fischer et al., 1997; Klein, 2001; Klein & Foerster, 2001; Klein et al., 2005; Luna et al., 2004; Munoz et al., 1998). These main findings were largely confirmed by the cross-sequential study by Klein et al. (2011) comparing participants aged 6–18 years at initial testing about one and a half year later.

In the *memory-guided saccade task*, the participant is instructed to remember the location of a visual target. Children need longer time to initiate a correct memory-guided saccade than young and middle-aged adults (e.g., Hikosaka, 1997; Luna et al., 2004; Gonzales, Mon-Williams, & Burge, 2015). Gonzales et al. (2015) compared performance between a condition in which a cue signalled in advance the presentation of a target, and a condition without such a cue. In children, cued saccade latencies were longer than non-cued latencies. In contrast, adults displayed the opposite effect. According to this finding, children’s ability to benefit from advance information is inferior to that in adults. Moreover, the accuracy of the first memory-guided saccade, the distance between the location of the target and the end point of the first saccade, improves throughout childhood and adolescence (Luna et al., 2004). More specifically, while voluntary planning of the eye movement towards the remembered target location is far developed by approximately 15 years of age, the accuracy of the final gaze location matures by about 19 years of age.

The literature reviewed so far suggests that all saccadic functions improve during infancy, childhood, and adolescence. While this summary is correct in principle (and holds for many more functions than discussed here), it neglects the fact that the developmental trajectories for different functions may differ across variables. To give an example, statistical modelling of (cross-sectional) age effects on various parameters of the pro- and anti-saccade tasks shows that the proportion of inter-individual differences in anti-saccadic reaction times or correction reaction times (after direction errors) that can be explained by age is about three times greater than that for pro-saccades, suggesting these two groups of variables undergo substantially different developmental processes between 9 and 27 years of age (Klein et al., 2005). Such differentiations, however, require sampling across a larger age range and a more fine-grained quantification and have been accomplished in only few large-scale eye movement studies (Klein, 2001; Luna et al., 2004).

15.3.4.7 Face Perception

Eye movements are crucial for our orientation in the world, and probably the most important component of our surrounding is other people. We interact with them in multiple ways, and our eye movements are included in most of these interactions. Naturally, facial expression provides many of the relevant cues for social interactions, including emotional cues (facial expression), verbal cues (lip movements), or attentional cues (gaze directions). Eye movements during face processing, therefore, allow us to study the earliest manifestations of social cognition in healthy infants and in infants with neurodevelopmental disorders (see Sect. 15.4 below). In this section we will therefore ask how face perception develops, focusing on two core issues. Research with infants deals with the nature of the early specialization on faces. In addition, the development of face perception is another example for the local-global distinction: perception of the individual components of faces precedes perception of faces as configurations.

The development of face processing starts already in early infancy, but it takes into adulthood to acquire mature expertise in face processing. Faces are most relevant for infants from birth onward, and literally hundreds of studies have dealt with the development of face perception. One major issue here is whether newborn infants are able to perceive and recognize real faces and face-like stimuli (e.g., Fantz 1964; Field, Cohen, Garcia, & Sternberg, 1984). Indeed, from the newborn period onward, infants display a natural visual preference for faces and face-like visual displays. This finding raised the question regarding the mechanisms underlying this early skill. According to one prominent view, the early preference for faces might be the result of *nonspecific perceptual biases*. For example, Turati (2004) points out that newborn infants' preference for faces is influenced by a general (non-specific) constraint to attend to stimuli with more elements in the upper than in the lower part of the overall configuration ("top-heavy bias"). Indeed, experiments using the natural preference method substantiate that newborns prefer depictions of real faces, schematic face-like patterns, and non-face geometric configurations with more elements in the upper than in the lower half (e.g., Cassia, Turati, & Simion, 2004). Shortly after the newborn age, however, this attentional constraint is changed. More specifically, an eye-tracking study conducted by Chien (2011) revealed that the top-heavy bias was missing for geometric and schematic face-like patterns in infants 3 months of age and in adults. However, both adults and infants 3 months of age displayed a preference for upright faces over inverted ones (that is, faces shown upside down). Overall, these findings indicate an early specialization process that is triggered by innate constraints which are tuned to a broad (non-specific) class of visual input. In conjunction with learning experiences, these constraints generate a specialization for natural faces.

The visual input provided by the species-typical environment generates yet another interesting visual phenomenon in the course of the first year of life: research substantiates that young infants are able to discriminate faces from both the own and other races. Again, a specialization takes place in the first year of life in that the ability to discriminate faces is increasingly restricted to own-race faces (e.g., Kelly et al., 2007). Eye-tracking studies support the presence of this perceptual narrowing process. Between 4 and 9 months of age, fixation time toward the internal features of other-race faces but not toward same-race faces decreases (Liu et al., 2011).

The ability to process faces as global, coherent configurations undergoes a substantial development in the first year of life. This development can be shown using the *face inversion effect*. Adults' ability to differentiate upright faces is superior to their ability to differentiate inverted faces. This inversion effect is due to the fact that faces are processed as global, unparsed perceptual configurations, in which the individual, local elements are not represented. The inversion effect can be found in infants 4 months of age: a habituation-dishabituation study conducted by Turati, Sangrigoli, Ruel, and de Schonen (2004) showed that face differentiation is easier for upright than inverted, upside down-oriented faces. Observation of eye movements provided detailed information about the nature of the inversion effect in infancy. In a study conducted by Gallay, Baudouin, Durand, Lemoine, and Lécuyer (2006), infants 4 months of age exhibited no difference in the time needed to habituate to upright and upside down faces. However, the infants explored upright and inverted faces differentially. Upright faces generated more scanning of the internal elements, mainly the mouth and the nose, than did inverted faces. These results suggest that by 4 months of age infants are able to extract configurational facial information. The ability to process upright faces more efficiently increases in the following months (Kato & Konishi, 2013).

Further studies provide evidence that the *eyes* play a crucial role in attracting infants' attention. Looking and eye-tracking studies suggest that a preference for the human eyes is acquired in the course of the first three months of life. For example, infants 3 months of age and older, but not newborn infants, prefer a non-human primate face in which the original eyes are replaced by human eyes over the same non-human primate face with the original primate eyes (Dupierriex et al., 2014). Moreover, eye-tracking has revealed the development of even more refined face perception abilities. Infants 5–8 months of age look more at the eyes than the nose and the mouth of an interaction partner (Smith, Gibilisco, Meisinger, & Hankey, 2013). With increasing age, infants look less at the eyes and more at the mouth of a talker (Tenenbaum, Shah, Sobel, Malle, & Morgan, 2013). This behaviour is most probably produced by an increasing interest in speech. Consistent with these observations, Oakes and Ellis (2013) observed that infants 4.5 and 6.5 months of age focus on the eyes. In contrast, infants 8 and 12.5 months of age scan broader areas of the face. Overall, research provides evidence that faces, especially the eyes and the mouth, are sources for socially relevant information from an early age onward.

Development of face perception continues into childhood. During childhood, a more flexible way of scanning develops, for example observing the mouth when the face makes expressions of speaking/laughing or observing hands when they are engaged in an action (Frank, Vul, & Saxe, 2012). Face processing skills are not adult-like until the age of about 10 years. Again, the development of configurational processing lags behind the development of feature processing. Hu, Wang, Fu, Quinn, and Lee (2014) established a higher number of saccades on whole faces, a higher proportional number of saccades between eyes, mouth, and nose, and a higher proportional fixation percentage on the nose region in Chinese adults than in Chinese children 4–7 years of age. These looking patterns indicate that the integration of individual face elements is more advanced in adults than in children. Especially, scanning of the nose region is believed to indicate a simultaneous processing of several elements of a face; saccades within the features of a face indicate binding of these features into a whole. This study thus suggests that children 4–7 years of age display differential scanning strategies for own-race versus other-race faces. Moreover, the study is consistent with other research by supporting the fact that holistic face processing improves from childhood into adulthood (e.g., Schwarzer, Huber, & Dümmler, 2005).

15.3.4.8 Life-Span Development of Ocular-Motor Control

Developmental ocular-motor research has much focused on the first three decades of life, with a strong emphasis on infancy and childhood. Nevertheless, there is also research on ocular-motor control processes across the life-span, generally supporting the idea of a U-shaped (or inverted U-shaped) developmental trajectory, comprising of a negatively accelerated developmental flank, a peaking (or a trough) in early adulthood, followed by a linearly declining ageing flank. Only a few studies, however, have investigated this *cross-sectional lifespan development* of saccade control. Fischer et al. (1997) tested N = 281 participants aged 8–70 years with pro-saccadic overlap task and the anti-saccadic gap task (*overlap task*: fixation point remains on screen when peripheral cue appears; *gap task*: fixation point is extinguished before the cue appears; see chapter by Pierce et al., in this volume). Among their main findings were a decrease in saccadic reaction times (both tasks) and the proportion of direction errors in the anti-saccade task into the third decade of life, followed by an increase thereafter. The proportions of express saccades (i.e., visually-guided saccades with a latency around 80–130 ms) decreased after the middle of the second decade of life. Factor analyses revealed two factors that loaded the pro-saccade or anti-saccade variables, respectively. The authors interpret their findings in terms of dual control mechanisms, a fixation system tested with the pro-saccade task, and a voluntary saccade system tested with the anti-saccade task and showing larger age effects than the fixation system. This first lifespan developmental study is impressive because of its large sample size. Cross-sectional age effects were, however, primarily described and not statistically modelled.

The second lifespan developmental study was published by Munoz et al. (1998) and tested N = 168 participants aged 5–79 years, administering the gap and overlap conditions of the pro- and anti-saccade tasks. The authors reported a similar developmental pattern as Fischer et al. (1997) although the minimum for saccadic reaction times (both tasks and conditions) were reached somewhat earlier than in Fischer and colleagues' study. The age range was divided into 11 age groups which differed significantly. Also, the anti-saccade effect, that is the difference in saccadic reaction times between the pro- and the anti-saccade task, was greatest in young children and lower in participants older than 15 years. The greatest age effects were found for the intra-subject variability of saccadic reaction times (coefficient of variation). The authors emphasize the relevance of normative developmental data and the value of the anti-saccade task as a test of the maturation (and decline) of frontal lobe functions.

The largest lifespan developmental study published so far tested N = 327 participants aged 9–88 years using the gap condition of the pro-saccade task and the overlap condition of the anti-saccade task (Klein et al., 2005). In contrast to the other two lifespan developmental studies, this study employed regression analysis using age and its inverse to model linear and curvilinear age effects for each of the pro- and anti-saccade task parameters separately to quantify and compare the respective size of the age effects as well as their shape. Analyses were run for the entire sample as well as the sub-samples age 9–27 years and age 28–88 years to differentiate developmental and ageing effects. Across the lifespan, the intra-subject variability of saccadic reaction times showed the greatest age effects ($R^2_{adj} = .40$). Development between 9 and 27 years was characterized by a speed-up of the correction reaction times following direction errors in the anti-saccade task ($R^2_{adj} = .54$), followed by a speed-up of the anti-saccadic latencies ($R^2_{adj} = .47$) and a reduction in the intra-subject variability of pro-saccadic reaction times ($R^2_{adj} = .45$). Ageing between 28 and 88 years was best characterized by a slowing of pro-saccadic reaction times ($R^2_{adj} = .38$) and an increased in their intra-subject variability ($R^2_{adj} = .34$).

Overall, these three large-scale developmental studies suggest that lifespan development of saccade control shows improvements (speed-up and increased stability, both in terms of intra-subject variability of saccadic latencies and errors) between childhood and the third decade of life followed by a slow and linear decay thereafter. Greatest age-related variance is overall provided by the intra-subject variability of saccadic reaction times, suggesting that the developmental gain and the ageing-related loss of stability of saccade control is a developmental marker variable in the ocular-motor system.

15.3.5 Summary Regarding Typical Development

Section 15.3 has introduced the recording of eye movements as a versatile tool to study developmental processes. Starting with reflections about the basic questions that can be addressed with developmental studies and the ways how this can be accomplished, we saw that eye movement research provides methodological virtues that are not provided by other approaches. As an example of this claim we saw that even in newborns and infants, who cannot speak and are limited in comprehending instructions, studying their looking preferences can reveal aspects of their information processing that would remain inaccessible otherwise. We also saw that eye movement recordings provide a unique way of studying the development of visual attention, including visual attention for social stimuli like faces. Furthermore, it became clear that basic ocular-motor processes like fixation, visually-guided saccades, voluntary saccades or smooth pursuit, worth being studied in its own right, can be studied from very early on and at all ages, making eye movement research a distinguished approach in life-span developmental research.

15.4 Atypical Development

15.4.1 What Is "Atypical Development" and Why Is It Important to Study It?

"Atypical development" is development in infancy, childhood and adolescence that deviates in qualitative and/or quantitative ways from what we consider as typical for the vast majority of individuals. Atypical is less judgmental than "abnormal", thus reflecting the insight that individuals who show atypical development do not only have weaknesses or deficits that constitute their disorder, but may show strengths as well or are simply different in certain regards.

In the recently updated version of the Diagnostic and Statistical Manual of Mental Disorders DSM-5 developmental and dimensional aspects of mental disorders are emphasized (APA, DSM-5). The categorical approach of the last decades is no longer considered in line with recent research findings of shared neural substrates and genetic risk factors as well as high comorbidity between disorders. In addition, temperamental risk factors interacting with adverse environmental events have been shown to predispose to mental illness, which manifests itself as a spectrum of psychiatric symptoms with different grades of severity and phenotypes changing during the different stages of life. The DSM-5 summarizes disorders beginning early in life within the group of the Neurodevelopmental Disorders (NDD). These are rather 'genetic' in nature and can be characterized by deficits that are persistent and without the relapses and remissions that are typical for other psychiatric disorders. NDD, however, often show age-related improvements through compensation and adaptive changes. Although not summarized under the DSM-5 rubric of NDD, schizophrenia spectrum disorders and other psychotic disorders are now considered to reflect developmental processes with an early onset, usually at the beginning of the second life decade. Since NDD do not express delays but rather changes in cognitive and behavioural functions and deviations from the typical adaptation to developmental tasks, they are also addressed as an 'atypical' development. The male predominance in almost all NDD indicates the strong genetic component. NDD show a high comorbidity with internalizing and externalizing psychiatric disorders and are considered as 'primary' psychiatric conditions that facilitate mental illness later in life. In line with these considerations is the study of the biological and behavioural underpinnings of NDD, since their complex aetiology and their meaning for psychosocial adaptation processes contributes to the progress in psychiatric management and treatment. Prevalence estimates of NDD are rather high and vary worldwide between about 1% for Autism Spectrum Disorder (Baird et al., 2006; Fombonne, 2009), 5% for ADHD (Polanczyk et al., 2006), and about 7% for reading disorders according to age and IQ discrepancies cut-off criteria (e.g., Shaywitz et al., 1990). While the prevalence of Schizophrenia ranges around 1% world-wide, only about 4 or 1% of all Schizophrenia cases fall sick before the age of 18 or 10 years, respectively (Remschmidt, 2002; Vyas, Kumra, & Puri, 2010). Schizophrenia is considered as a neuro-developmental disorder by many authors although it is not classified as such in the DSM-5.

15.4.2 How Can We Study Eye Movements in Atypically Developing Populations?

The investigation of eye movement processes offers distinct advantages over the study of other forms of behaviour; this holds in particular for the study of "vulnerable" populations such as infants, children or adolescents with NDD. The advantages include the utilisation of the rich body of experimental and neuro-scientific knowledge about the ocular-motor system from experimental, animal, lesion and imaging studies; the precise measurement and detailed parametrisation that eye movement recordings offer; the versatility of the applicability in conjunction with non-invasiveness of the recordings; the naturalness of the tasks themselves; and the potential for process analyses (in addition to "static mapping" of saccades or fixations; see chapter by Boccignone, in this volume). One further important, metric feature of eye movement parameters that is most relevant in all individual differences studies, including clinical ones, is their reliability. The psychometric concept of reliability refers to the stability of inter-individual differences across repeated measurements. Sufficiently high reliability is a crucial requirement in all studies on inter-individual differences, including those along (neuro-)psychiatric dimensions. Fortunately, many eye movements parameters do show high instrumental and moderate to good test-retest reliability (reviewed for saccades in Klein & Berg, 2001; and Klein & Fischer, 2005). Different approaches are being employed in NDD research to study eye movements. These will be explained in the following.

Paradigm 1: Standard ocular motor paradigms—Standard ocular motor paradigms such as the anti-saccade task (see chapter by Pierce et al., in this volume) are frequently being used in clinical research as there is a rich body of experimental, correlative, animal, physiological or clinical research that can be used to interpret potential deficits in such paradigms of NDD patients. One example is Munoz and colleagues' (2003) study on "Altered Control of Visual Fixation and Saccadic Eye Movements in Attention-Deficit Hyperactivity Disorder". These authors administered the anti-saccade task to test ADHD patients' ability to inhibit reflexive and generate voluntary responses, which are both controlled by a functional neuronal system including the frontal cortex. In line with their hypothesis, Munoz et al. (2003) found an increased likelihood of reflex-like pro-saccades to the peripheral cue, pointing to an inhibitory deficit; as well as increased latencies of initiating correct anti-saccades, indicating slowed generation of voluntary eye movements. Overall, the deficits in the anti-saccade task suggest a fronto-striatal dysfunction in ADHD, according to Munoz and colleagues.

The rationales and interpretations in studies like the one of Munoz and colleagues are typically clear and straightforward, demonstrating the strengths of this approach. A potential downside of the approach is its inflexibility: The paradigm is as it is, and only some of its parameters can be changed. So the researcher needs to decide whether it is appropriate to take it "from the shelf". This inflexibility can be overcome to some degree, however, by changing the standard paradigm to "tailor" it to the impairments or biases of specific disorders, as will be explained next.

Paradigm 2: Disorder-specific manipulations of standard ocular motor paradigms—The thoroughly investigated standard paradigms are known to challenge particular psychological functions, for example, response inhibition in the anti-saccade task. Such standard paradigms may be modifiable such that their experimental modification stresses a putative dysfunction of a clinical group more pronounced or more specifically than the original paradigm. An example is the study by Reinholdt-Dunne et al. (2012). Referring to the observation that anxious individuals have an attentional bias for threatening information, including an automatic tendency to direct their attention towards threat as well as an impaired ability to control their attentional response (Eysenck, Derakshan, Santos, & Calvo, 2007), the authors made use of the fact that the anti-saccade task requires directing attention away from a (peripheral) stimulus. Anxious individuals should be impaired at inhibiting glancing to the peripheral stimulus if this stimulus contains threat-related information. And indeed, among the main findings of this study is the attentional bias of high trait-anxious individuals for angry faces (threat-related), but not for fearful or happy faces. This result thus partially confirmed the hypotheses of this study. Furthermore, and given that anxious individuals direct their gaze to *all* kinds of negative faces (including fearful ones) under free viewing conditions, the result of not finding attentional bias for fearful faces suggests that the differential responses to angry versus fearful faces are to some degree task-specific. Disorder-specific manipulations of standard ocular motor paradigms can provide powerful demonstrations of deficits or biases in psychiatric patients if appropriate control conditions (like the happy faces in the previous study) are employed that allow testing the experimental specificity of the reported effects (here: for threat-related stimuli).

Paradigm 3: Visual exploration/free viewing—While paradigms 1 and 2 "constrain" the gaze by requiring eye movement responses to specific stimuli such as the onset of a peripheral cue or a smoothly moving target, it is characteristic of the free viewing or visual exploration paradigm that the gaze is not constrained or only little constrained by a task instruction that specifies what the participant is expected to do during or after the exposure to the stimulus. Showing participants complex visual stimuli such as social scenes, faces, objects etc. is thus an elegant method to investigate *spontaneous* priorities and *spontaneous* temporal-spatial patterns of visual attention. The free viewing paradigm has been applied in many studies with psychiatric patients, including the one by Horley et al. (2003). This study set out with the clinical observation that social phobia is characterised by fear of negative evaluation in social situations and that social phobics tend to avoid direct eye contact. In order to confirm this observation empirically, Horley and colleagues presented faces with different emotional expressions to social phobics and found fewer fixations to neutral and sad faces, generally reduced fixation durations, reduced numbers of fixations to the eyes, in particular for sad faces. Overall, these findings confirmed the clinical observation that individuals with social phobia show reduced eye contact during social interaction.

The free viewing paradigm offers many advantages. *Firstly*, the task itself is rather natural and thus suitable to test individuals that are vulnerable or have difficulties understanding more complex task instructions (including very young children or individuals with low IQ). *Secondly*, this paradigm yields a great number of meaningfully interpretable eye movement parameters such as fixation-related parameters (e.g., individual fixations, fixation clusters, regions of interest, duration, frequency, variability); saccade-related parameters (e.g., lengths of saccades, direction of saccades); and transition-related parameters (e.g., probability of transitions between fixation clusters or regions of interest) as well as their time courses and modifications by experimental procedures. *Thirdly*, not only is the set of potentially showable visual stimuli almost unlimited, these stimuli can also be combined with all kinds of task instructions (to focus on something, to explain something, etc.).

As a special case of the free viewing paradigm, we consider the *preferential looking paradigm* that experimentally manipulates the combination of two simultaneously presented stimuli (e.g., a "social" stimulus on one side, a "non-social" one on the other side of a screen) and measures the gazing preferences of an individual.

Paradigm 4: Co-registration of physiological variables—Eye movement recordings reveal the temporal-spatial patterns of orienting of visual attention. They do not directly show how the brain responds to the seen stimuli (phasic response) or works during the course of the task (tonic response). The co-registration of psychophysiological parameters may provide important additional insights into how the brain processes visual stimuli or tasks. Two psychophysiological measures that are particularly "handy" in eye movement studies are the pupil, indexing workload, emotional valence, arousal, and noradrenergic activity (see chapter by Laeng & Alnaes, in this volume); and the eye blink rate, indexing, among other processes, dopaminergic neurotransmission. An example is the study by Silk et al. (2007). Blunted physiological responses to fearful faces or affective feedback have been reported for children and adolescents with depression and should also show up in the pupillary responses to emotional stimuli. Showing their participants words with different emotional content, the authors found increased pupil dilations in the entire sample following negative words. Depressed children and adolescents compared to controls, however, showed reduced pupillary responses specifically to negative words, this reduction in pupillary response being more marked in patients with more severe depression. The study by Silk et al. (2007) thus showed that the co-registration of physiological responses such as the pupil response can in principle reveal information about the timing (onset latency), intensity (amplitude of the response) and duration of processing that complements the eye movement data. As many physiological parameters are sensitive to a range of processes, the specific interpretation of such findings is dependent upon clear experimentation.

Paradigm 5: Registration of eye movements in naturalistic situations or environments—Almost all eye movement research conducted so far has been run in the lab. This is because the laboratory environment can be fully controlled, and the eye movement findings can be directly related to the specifics of the lab "situation" (including stimuli, tasks, instructions, etc.). This is an obvious advantage. Its potential downside—at least for some kinds of research questions—is that what we observe in the lab (a) may not be (statistically) predictive of behaviour in real life; or (b) may not be observable in real life situations. To give an example: A lack of responsiveness in a gaze cueing experiment in the lab may or may not: (a) be correlated with a particular behaviour in patients with autism (ecological validity as concurrent or predictive validity); or (b) be also found in a real-life interaction between the autistic person and another person (ecological validity as representativeness). So far, only few eye movement studies have addressed the latter topic or measured eye movements in natural setting as such (see chapter by Tatler and colleagues, in this volume). One such study was published by Noris et al. (2012). In a semi-naturalistic interaction between an experimenter and children with autism, the latter were found to show less and shorter fixations of the interacting experimenter than a group of healthy children. Results like this lend strong weight to corresponding laboratory studies assuming explicitly or implicitly that the obtained eye movement findings are predictive or even representative of real-life behaviour of autistic patients.

After the introduction of the five basic paradigms that have been used in psychiatric eye movement research, the next section will explain how these paradigms have been applied to study atypical development.

15.4.3 Atypical Development of Eye Movements

This section focusses on three neurodevelopmental disorders (NDs)—Autism Spectrum Disorders, Attention-Deficit/Hyperactivity Disorder, Reading Disorder—and Early Onset Schizophrenia. While the first three disorders were chosen because they assemble the largest groups of eye movements studies in child and adolescent psychiatry, Early Onset Schizophrenia was added to complement the largest group of eye movements studies in adult psychiatry, that is Schizophrenia (see also chapter by Smyrnis and colleagues, in this volume). The main findings obtained in eye movement studies will be outlined and methodological aspects of studying such patient groups using the eye movement technology will be explained.

15.4.3.1 Autism Spectrum Disorders

The common features of Autism Spectrum Disorders (ASD; or "Autism") are persistent deficits in social interaction and communication (with social functioning being significantly below the given intellectual level) accompanied by restricted, repetitive behaviours, special narrow interests and/or activities. These can vary considerably with regard to stereotyped motor movements, unusual verbal expression or use of toys, 'insistence on sameness' in routines and/or a rigid, inflexible cognitive style. In addition, the DSM-5 has included the atypical sensory experience in the symptomatology of autistic individuals which comprises hyper- or hyposensitivity and/or unusual interest for sensory stimuli. Since the previous categories of autistic disorders (childhood autism, pervasive developmental disorders-not other specified, and Asperger's syndrome) have been challenged by clinical experience and research findings, a spectrum concept has emerged. Different phenotypes within ASD are currently explained by the large range of intellectual levels, language skills, and symptom severity as well as in association with other neurodevelopmental or behavioural disorders including ADHD. While the scope for pharmacological treatment of Autism is rather narrow and focusing on associated behavioural problems such as irritability, tantrums, aggressions, or attention problems, psychotherapeutic intervention aims at reducing maladaptive behaviour and improving psychosocial adjustment by supporting the development of cognitive, language and socialization skills. This latter aspect may include the training of social competence, which, in turn, can involve explicit instructions to look at faces or eyes in social interactions. It is obvious that such treatment components have to be considered when planning or interpreting studies in which eye movements are used to study social attention and the like in Autism.

Different psychological theories have been suggested to account for the symptoms of ASD, among them the ideas that patients with ASD show deficits in mentalising abilities (including "Theory of Mind" and an impaired Mirror Neuron System (MNS), deficits in processing global form ("Central Coherence Theory"; Happé & Frith, 1994), or deficits in executive functions. Eye movement research can contribute to all of these theories as they all involve altered processes of visual attention or eye movement control. Mentalising abilities or global processing abilities can be studied by using static or dynamic visual stimuli displaying, for instance, social interactions. The MNS, based on observational behaviour (e.g., looking at what a conspecific does), is per se amenable to the study of eye movements. And tests like the antisaccade task or the memory-guided saccade task (see chapter by Pierce et al., in this volume) are considered as tests of executive functions (Hutton & Ettinger, 2006).

Regardless of the aforementioned theories, eye movement research on ASD has primarily focused on studying the phenomenology of altered social attention as revealed by gaze control, employing the free-viewing paradigm explained in Sect. 15.4.2. “Social attention” here means to what extent (quantity) and in which ways (quality) individuals with ASD attend to social stimuli such as pictures or videos of one or more people, pictures of faces, stimulus ensembles that combine social with non-social stimuli, gaze cues and the like. The number of stimuli that can be presented is nearly infinite, and the way how these stimuli are presented is also highly diverse: single or repeated presentations; presentations without or with (all kinds of) instructions; presentations for longer or shorter durations; and so on. Likewise, the way free viewing data can be analysed is highly versatile, focusing on the spatial, temporal or spatial-temporal aspects of fixations, saccades or transitions between fixation clusters that are either defined in a top-down or a priori fashion, or derived bottom-up from the eye movement data themselves (see also the chapter by Boccignone, in this volume). Given this richness of the free viewing task, it does not come as a surprise that free viewing research on Autism has produced a great diversity of different approaches. At least three groups of free viewing studies can be differentiated, though, and will be discussed next: (1) studies looking at the relative weights of *attention for social versus non-social stimuli*; (2) studies analysing *gaze following;* and (3) studies investigating the *structure of face processing*. Another group of ASD studies used standard paradigms or modified standard paradigms to study eye movements in ASD. These will be described further below.

(1) *Studies comparing the processing of social versus non-social stimuli*. One of the paradigms that have been used often in the ASD literature, is the preferential looking paradigm, which we have considered as a special case of the free viewing paradigm before. With this paradigm, two stimuli are presented simultaneously on the screen to observe which of these is preferentially looked at. From an experimental point of view, these stimuli should differ only in their content (e.g., social vs non-social), so that differences in looking preferences cannot be attributed to other stimulus features such as perceptual saliency (e.g., one stimulus “pops out” whereas the other one appears rather opaque), size, presentation duration, presentation location etc. As such, “balancing” or “controlling” may be rather difficult to achieve with regard to features that are subtle but potentially influential (such as symmetry, Gestalt closure, complexity and the like), it is reasonable to aggregate across many such stimulus combinations if the to-be-compared stimulus contents do not differ systematically.

In one such study, Pierce et al. (2011) compared looking preferences for social animations versus repetitive geometric patterns and found toddlers with ASD to look more at the patterns than typical toddlers. As “preference” in this paradigm is always *relative* preference, it does make sense to manipulate either of the two stimuli, that is, not only the social, but also the non-social one, as Sasson and Touchstone (2013) did. These authors paired a picture of a face with a picture of an object that was either related or unrelated to the special interest of the participant. In this comparison, reduced fixation of the face stimulus was found *only* when it was combined with the picture of an object that was related to the special interest of the patient. This differentiation was also addressed in a study by Unruh et al. (2016), who tested typically developing adolescents and adolescents with ASD. These authors extended the approach by addressing also the significance of the nature of the non-social object displayed together with the social image. Thus by using a preferential looking task, they displayed arrays containing both social and object images, but in addition differentiated object images between neutral, or "low autism interest" (LAI) images (e.g., clothing, tools, etc.), and images associated with circumscribed interests often found in autistic patients, or "high autism interest" (HAI) images (e.g., trains, clocks etc.). Although these authors did not use stimuli that were chosen to reflect individual circumscribed interests—as done in other studies (e.g., Cascio et al., 2014; Foss-Feig et al., 2016)—they found that in autistic individuals, fixation to faces had significantly shorter durations and began significantly later when these were paired with HAI as compared to LAI objects. The authors therefore concluded, that patients with ASD show an object-related bias of their visual attention that is general, going beyond idiosyncratic preferences for objects but maybe developmentally preceding them.

Other studies also have shown pictures of social interactions and investigated the relative preference of participants to look at face vs non-face parts of the scene. Riby and Hancock (2008), for instance, found adolescents with ASD to look less time at faces than TD controls, a result that was also found by Wilson et al. (2010). A similar result was reported for young adults when looking at faces that were embedded in scenes showing a social interaction, but not when isolated faces were presented (Hanley et al., 2013).

In addition to the "*spatial*" mapping of gaze positions and relative preferences, free viewing tasks enable studying the unfolding of social visual-spatial attention *in time*. Heaton and Freeth (2016), for instance, investigated high functioning adolescents with ASD and matched typically developing controls during free viewing of images for 5 or 15 s which depicted everyday scenes of two types: one group of scenes with one prominent face (centrally positioned) and a second group of scenes with non-prominent faces (e.g., people in the background). The authors presented five innovative analyses, namely: (1) evolution of fixation spread, (2) scan path progression, (3) fixation rate, (4) visual persistence within an area, and (5) reversion to previous fixation locations. They found that the fixation spread was greater in the control group; especially after 2500 ms viewing, the ASD population exhibited slower exploration. Moreover, the evolution of the scan path length suggested that controls explored further areas of the scene whereas the scan paths of ASD participants had a greater tendency to include areas closer to the current fixation. However, there was no difference between groups regarding their fixation rates, and thus the above remarks are more likely to be due to different strategies applied during visual exploration. The authors furthermore reported that visual persistence within an area was increased in the ASD group across both types of stimuli. Although these results indicate a restricted viewing pattern among the ASD population, the authors note that this could also imply a more detail-oriented attention in the respective area and suggest further research of the possible conceptual consequences of this visual behaviour. Differences between autistic individuals and healthy controls were also reported in other studies. Wilson et al. (2010), for instance, found a significant difference in the *latency* of the first fixation of a social stimulus in the display, they did not, however, find differences between autistic individuals and typical controls in the direction of the first fixation. Likewise, Unruh and colleagues (2016) reported that participants with ASD demonstrated longer latency for first fixating at faces when they were displayed together with HAI images. Interestingly enough, this was not the case when they were displayed together with LAI images. Further such results were reported by Freeth, Chapman, Ropar, and Mitchell (2010) and Riby and Hancock (2009b). In sum, the study results summarised in this section suggest that autistic individuals hesitate to attend to social stimuli but do not completely neglect them.

Collectively, these studies make clear that attention for social stimuli may be context-dependent and exhibit altered temporal features in individuals with autism. This conclusion is complemented by the results of a recent meta-analysis, published by Chita-Tegmark (2016) on the basis of 38 analysed studies. This author identified social content, that is the number of people as well as their level of activity and interaction, as the *only* significant predictor of social attention deficits in autism. Other moderator variables, including age, IQ, motion of the stimuli or their ecological validity, as well as the presence of audio input or attention bids or communicative intent all turned out to be non-significant in these studies.

(2) *Studies on gaze following.* Our gaze is not only a "channel" for the intake of visual information, it is also a "signal" to others observing us that conveys what we are attending to, and often also intending to do. If another person then looks at the same object as we ourselves, s/he joins our attention. Individuals with autism are known to show deficits in joint attention, which implies that they do not follow other persons' gazes. A substantial number of studies indeed have shown deficits in gaze following in participants with ASD, pointing to shorter dwell times on looked-at objects (Bedford et al., 2012), less correct gaze following (Falck-Ytter, Fernell, Hedvall, von Hofsten, & Gillberg, 2012), longer latencies of correct gaze following (Freeth et al., 2010) or a more balanced looking at objects that are or are not in the line of sight of a gazing model (Riby, Hancock, Jones, & Hanley, 2013).

In one of these studies, Riby and colleagues (2013) compared *spontaneous* and *cued* gaze-following when confronted with static pictures showing actors gazing at a target item in a complex scene. In one condition ("spontaneous") participants were asked to look at the picture for as long as it remains on screen, in the other condition ("cued") they were asked to detect and name what the actor is looking at. Different areas of interest were defined, including the actor's face and eyes, the correct gazed-at target item as well as a plausible and an implausible target. When cued to identify the correct target and in comparison with the spontaneous gazing condition, typically developing children looked less at the face and the implausible target and more at the correct target. By contrast, children with autism looked more at the face and the eyes of the actor, somewhat less at the implausible target and showed no difference for the plausible and correct targets. The experimental comparison between spontaneous and cued gaze following along with the analysis of the gazing behaviour thus unveils different attentional processes in the two groups that are prompted by the instruction: autistic children understand that they have to look at the face and the eyes in order to identify what someone is looking at, but in contrast to typical children they do not subsequently actually follow the actors gaze but keep looking primarily at the implausible item.

Not all studies, however, have found such differences. A recent study of Billeci et al. (2016) has investigated joint attention in ASD toddlers by differentiating between response to joint attention and initiation of joint attention. As already described in the literature (e.g., Gillespie-Lynch, 2013), *response* to joint attention (resp-JA) refers to the act of gaze or point following, whereas *initiation* of joint attention (ini-JA) refers to the act of directing others' attention with own eye movements or gestures. Billeci and colleagues examined children with ASD and typically developing children by showing them short videos of a female face gazing at the participant, away from the participant or to toys (mainly vehicles) displayed on the desk between the actor and the participant. This setup included one task for response to joint attention and two tasks for initiation of joint attention, one with an expected event and one with an unexpected event. Their main findings were: *Firstly*, that there was no difference between groups in the resp-JA task. *Secondly*, in respect to the ini-JA task with the expected event, ASD participants were found to have longer fixations at the face and also more transitions starting from the face to the target object. In contrast, healthy participants showed more fixations for the non-target object and made more gaze transitions from this object to the face. *Thirdly*, for the ini-JA task with the unexpected event, ASD participants had again longer fixations at the face and made more transitions gazing from the target object to the face. These findings are by and large unexpected and, at the face of it, not supporting the view that gaze movement studies can easily objectivise the known deficits in joint attention in autistic individuals. At the same time, however, they underline that even seemingly simple gaze movement tasks may involve components of visual attention that are not in the focus of the researcher(s) but influence task performance. In the present example, Billeci and colleagues discuss retrospectively whether the longer fixations of the face might represent an impairment in the disengagement of visual attention and the reduced gazing at the non-target object a deficit in divided attention. Only systematic experimentation can elucidate whether such rather basic and general components of visual attention contribute to gazing behaviour of autistic individuals in joint attention tasks. For a detailed discussion of the topic, see the review by Nation and Penny (2008).

(3) *Studies investigating the structure of face processing*. This third group of free viewing studies investigated in what way participants with autism look at what is probably the most relevant visual social "stimulus" of all, the face. In one of the earliest of these studies, Klin et al. (2002b) reported that adolescents with ASD looked less at the eyes and more at the mouth than control subjects. This finding was later also reported by Jones, Carr, and Klin (2008) and in other studies. After

> several more such studies have been published, the overall picture emerges that while the reduced eye fixation seems to be a reliable difference between patients with ASD and controls, the increased mouth fixation has not consistently been reported (Papagiannopoulou, Chitty, Hermens, Hickie, & Lagopoulos, 2014). A further differentiation within the eyes was reported by Guillon, Afzali, Rogé, and Baduel (2015) who found a longer fixation time of the right eye for the ASD group that was significantly different from that of the control group; conversely, controls showed an attentional bias for the left eye that was absent in ASD patients. Yi et al. (2013, 2014), however, reported a bias towards processing the left—rather than the right—eye in patients with Autism. Overall, results are not entirely consistent regarding the relative processing of eyes and mouths in autism (see also review by Falck-Ytter and von Hofsten, 2011; Guillon, Hadjikhani, Baduel, & Rogé, 2014).

Attention to faces has been further investigated through tasks of Facial Emotion Recognition (FER). During such tasks, in their most common form, the participant has to select the emotion displayed from the face that is presented. This task has been widely used in behavioural studies and mostly include the six basic emotions: sadness, happiness, fear, anger, disgust and surprise. Given that this paradigm has been widely used in "behavioural" studies measuring the endpoint of the emotion selection process, and within the context of the present textbook, one can ask what additional insights this paradigm can reveal when eye movement are recorded during facial emotion recognition. More specifically, do autistic individuals look different at emotional faces, and is that the same for all emotions? And to what extend does this looking behaviour relate to their recognition behaviour and, maybe, their symptoms? Two of the latest reviews regarding the above are those of Harms, Martin, and Wallace (2010) and Black et al. (2017). On the whole both reviews conclude that there are indeed atypical gaze patterns in individuals with ASD while recognising emotions in faces and that demographic factors and task demands (duration and complexity of the task) differ between studies. Specifically, for children and adolescents the overall results are not as consistent as expected, and results are more consistent in studies with adult participants. A few studies will be reported here to illustrate the approach.

In Dalton et al. (2005), for instance, fourteen autistic male adolescents and twelve typically developed (TD) control participants were to recognise among 4 emotions (happiness, fear, anger, neutral) displayed from faces for 3 s each. Half of the faces had a direct gaze on the viewer and the other half not. The autistic participants showed significantly shorter fixation times in the eyes of all type of faces and they recognised the emotion of the face with lower accuracy than controls. Moreover, regarding their time to respond, they were slower when the face was emotional (not neutral) and also when its gaze was direct. This study thus shows that group differences in recognition performance can go hand in hand with group differences in looking behaviour.

Specific correlations of eye gaze patterns and performance in FER are mentioned in Bal et al. (2010). In this study, children with ASD and typically developed children were presented emotional faces of the 6 basic emotions. The analysis of the gaze differentiated 3 Areas of Interest (AOIs): Eyes, mouth, and "off" (including the rest of the face). The associations between gaze and performance found for the ASD group were: firstly, that the longer fixation time in the off region was related with the slower response (average recognition latency) when recognising surprise; and secondly, that the longer fixation at the eye region and the shorter at the mouth region had a significant relationship with fewer errors when recognising disgust. Overall, children with ASD made significantly more errors only when recognising anger and as in the previous study they were slower in identifying emotion. Their fixation duration of the eyes had a tendency to be less, but the significant difference reported is that ASD children looked more in the off area when fear was presented.

Another approach to identify the relationship between atypical gaze behaviour and clinical characteristics of participants with ASD is their association with scores in certain diagnostic-behavioural questionnaires. In Tottenham et al. (2014), the analysis focused on angry and neutral faces among a FER task that was conducted with adolescents with ASD and TD participants. In this study, ASD adolescents made significantly less eye movements to the eyes of neutral faces, but were similar to controls when angry faces were presented. Interestingly, the greater perceived threat of the face (as rated on a scale) for neutral faces were associated with fewer eye movements directed toward the eye region. Overall, the last-mentioned group of studies has illustrated that there are different approaches of relating facial emotion recognition performance with gazing behaviour and that the latter has the potential to help interpreting the former.

While the previous sections have explained the different approaches employed within the free viewing paradigm, another, smaller group of studies administered standard eye movement paradigms in patients with ASD (for detailed explanations of such standard paradigms, see the chapters by Pierce et al. and Lencer et al., in this volume). While the rationales of such studies are explained in Sect. 15.4.2 of this chapter, their main findings have been reviewed recently (Johnson, Lum, Rinehart, & Fielding, 2016) and will be summarised next.

Regarding *visually-guided saccades*, saccadic reaction time, peak velocity and saccadic gain have been found to be normal, the only difference being greater variable error of dysmetria in participants with ASD compared to controls (Johnson et al., 2012; Schmitt, Cook, Sweeney, & Mosconi, 2014; Takarae, Minshew, Luna, & Sweeney, 2004). Also, Schmitt et al. (2014) reported increased variability of prosaccadic RT in the autistic participants of their study, but this variable has as yet not been reported in other studies. Regarding *anti-saccades*, increased proportions of direction errors have been found in all eight studies that analysed this parameter, whereas latencies of correct anti-saccades, analysed in three of these studies, were somewhat faster in ASD subjects (as reviewed in Johnson et al., 2016). Given that anti-saccadic latencies and the proportion of direction errors are typically positively correlated (Klein, 2001), this fractionation of the two parameters is unusual and requires further investigation. Only few studies have looked at *memory-guided saccades* in ASD. Goldberg et al. (2002) reported normal latencies along with a greater proportion of suppression errors (or premature saccades). Finally, reduced open-loop and closed-loop gain during *smooth pursuit* (see chapter by Lencer et al., in this volume) has been found in some studies (Takarae et al., 2004; Takarae, Luna, Minshew, & Sweeney, 2008), but not in others (Kemner et al., 2004; Rosenhall et al., 1988).

15.4.3.2 Attention-Deficit/Hyperactivity Disorder

Attention-Deficit/Hyperactivity Disorder (ADHD) is a disorder characterized by persistent symptoms of inattention, hyperactivity and impulsivity with a typical beginning in the first decade of life. The symptoms last longer than 6 months and are not further explained by other neurological or psychiatric conditions. Although the symptoms can differ between patients in their clinical presentation and severity, they ultimately interfere significantly with the social, academic, and—later in life—occupational functioning of the individuals. Various theories have been suggested to explain ADHD (for an overview, see Kuntsi & Klein, 2011), providing starting points for hypothesis generation and ocular motor testing.

The ocular motor literature on ADHD has employed mainly standard ocular motor paradigms like the pro- and anti-saccade task, the memory-guided saccade task as well as fixation and smooth pursuit tasks (excellent reviews have been published by Karatekin (2007), and Rommelse, Van der Stigchel, and Sergeant (2008). In addition to the overarching aim of understanding the processes of ocular-motor control and visual attention in the different disorders better, different types of specific clinical research questions have been addressed in these studies.

(1) Some studies aimed at *improving the clinical diagnosis* or *diagnostic classification* with eye movements variables. Deans et al. (2010) measured the eye movements of children with ADHD or reading disorders during a short reading task. While both clinical groups, compared to healthy controls, were impaired in a range of eye movement parameters, only fixation durations and total reading time revealed slower processing on reading disorder compared to ADHD patients. Classification accuracy showed a low 33 and 59% for ADHD and reading disorder patients, respectively. This is not much, given that much higher classification rates could be achieved in other studies (albeit using different paradigms and differentiating patients with other disorders from controls; e.g., Benson et al., 2012). To the extent that NNDs show aetiological overlap, and display high degrees of co-morbidity or familiarity, high classification accuracy cannot be expected though.

(2) Other studies employed eye movement tasks to *objectivise clinical symptoms*. This approach was chosen for instance by Armstrong and Munoz (2003) who aimed at quantifying impulsive dysfunction in patients with ADHD using an ocular motor "stop signal" or "countermanding" task. During such tasks, the frequent go trials require a response, here: a pro-saccade to a peripheral cue. In rare no-go trials, however, a stop signal presented shortly after the go stimulus requires the countermanding of the prepared and to-be-executed go response. Patients with ADHD, known to be impulsive in their behaviour and showing

deficits in inhibitory control, were less able to countermand the pro-saccade upon presentation of the stop signal than controls, a finding that was interpreted to reflect the increased impulsivity of patients with ADHD. At the face of it, this conclusion is plausible. It would be incorrect though, if impulsive behaviour in real life and the failed inhibition of a prepared saccade in the lab had different causes—such as a high motivational salience of the impulsive act in one situation and a decrease in attention or effort in the other. Showing that the inhibitory deficit in the countermanding task is correlated (a) with comparable deficits in other lab tasks and (b) with direct measures of impulsive behaviour in real life (e.g., ratings, behaviour observations) would help showing that this deficit: (a) is not task-specific but potentially a facet of a greater construct (such as "inhibitory control") and (b) has some ecological validity. Without such indicators of construct validity, direct associations of eye movement abnormalities with clinical symptoms should be read with quite some reservations. Although Armstrong and Munoz (2003) tested adult ADHD patients, the rationale can be directly extended to the study of younger patients.

(3) Still other studies used eye movement tasks *to compare disorders*. One such study was published by Karatekin, Rauh, and Biscaldi (2010) who compared children with ADHD to those with youth-onset psychosis and typically developing controls. While the symptomatology of ADHD and Schizophrenia is obviously rather different, attention impairments, increased intra-subject variability (ISV), and the putative involvement of the dopaminergic transmitter system (albeit in opposite ways) are common to both disorders. It makes sense, therefore, to contrast patients with ADHD or Schizophrenia in potential ocular-motor manifestations of these common abnormalities. In Karatekin and colleagues' study, all participants performed the pro- and anti-saccade task, using an ant or bug as the cue, respectively. The results showed a rather differentiated pattern. With regard to the speed and correctness of saccadic responses, and according to the result description in Table 3 of Karatekin's article, schizophrenic patients compared to controls were slower in all saccadic responses except the latency of erroneous anti-saccades (direction errors) and made more direction errors during the anti-saccade task. ADHD patients, by contrast, only showed less corrections after anti-saccadic direction errors and made these corrections with a longer latency; they also showed longer latencies of correct anti-saccades. The main difference between the clinical groups was thus the greater number of failed error corrections in the ADHD group and the slower pro-saccade latencies in the Schizophrenia group. With regard to the intra-subject variability of reaction times, however, both clinical groups showed a rather consistent increase when compared to healthy controls. Interestingly, the effects sizes for both comparisons of the clinical groups with healthy controls were much higher for anti-saccadic RTs and their ISV than for pro-saccadic RTs and their ISV, with the highest of all effect sizes reported in both groups for the ISV of anti-saccadic RTs. Given that both anti-saccadic latencies and ISV are increased by prefrontal dysfunctions, Karatekin and colleagues' study impressively demon-

strates prefrontal involvement in Schizophrenia and ADHD using two rather simple saccade tasks.

(4) Child and adolescent psychiatry researchers have also used eye movement tasks *to identify potential aetiological factors* in neurodevelopmental disorders. This approach follows the seminal work by Holzman and colleagues' who were the first to report impaired smooth pursuit eye movements in unaffected relatives of schizophrenic patients (Holzman et al., 1974). The rationale behind studying certain functions in the unaffected first-degree relatives of psychiatric patients is that if they show an impaired function without having the disorder, this impaired function may point to a vulnerability factor of the disorder. One such study with ADHD patients was conducted by Rommelse and colleagues (2008). In this study, patients and their unaffected siblings were administered a memory-guided saccade task, one of the standard ocular motor paradigms that yields a rich set of meaningfully interpretable parameters (see also the chapter by Pierce et al., in this volume). While the accuracy of the memory-guided saccade was impaired both in patients and their siblings, its duration was increased only in patients but not their siblings, and its latency was normal both in patients and their siblings. Rommelse et al.'s study is hence interesting in at least three regards. *Firstly*, it demonstrates that parameters derived from one and the same eye movement paradigm—if not from one and the same eye movement, here: the saccade directed at the remembered position of a target—can have fundamentally different meanings, in this case: with regard to the nosology or aetiology of a disorder. Such fine-grained differentiation of eye movement parameters in clinical studies corresponds with the few published large-scale eye movement studies demonstrating the multi-dimensionality of eye movement parameters derived from one or a few eye movement tasks (Fischer et al., 1997; Klein, 2001; Klein & Fischer, 2005). *Secondly*, Rommelse and colleagues' study also shows that eye movement parameters such as the accuracy of the memory-guided saccade may lead to aetiological factors in child and adolescent psychiatric populations, a finding reminiscent of the much larger body of adult psychiatric literature on ocular-motor endophenotypes[2] (for reviews, see Calkins, Iacono, & Ones, 2008; Levy et al., 2008). *Thirdly*, Rommelse and colleagues' study is interesting in comparison with the study by Ross, Heinlein, Zerbe, and Radant (2005), who reported reduced accuracy of the memory-guided saccade in children with Schizophrenia *but not in their relatives*. Accordingly, an ocular motor parameter that may be pointing to an aetiological factor in one disorder is not necessarily aetiologically relevant in another disorder.

(5) Eye movement paradigms have also been used *to assess the effects of pharmacological treatment* (see also chapter by Karpouzian et al., in this volume). In the case of ADHD, methylphenidate (MPh) is the first-choice pharmacological treatment, and some studies have investigated its effects on eye movements in patients with ADHD. One such study, Munoz, Hampton, Moore, and Goldring

---

[2] Related terms are „intermediate phenotype" (on the assumed causal chain between genome and clinical phenotype: intermediate between the two) or "vulnerability marker".

(1999), compared children with ADHD (a) on versus off MPh medication and (b) with healthy controls in pro- and anti-saccade task with gap and overlap conditions. MPh was found to reduce pro- and anti-saccadic RTs and their ISV as well as the proportion of direction errors in the anti-saccade task and improved fixation stability. The results of Munoz et al. (1999) were replicated in the study by Klein et al. (2002) who compared the MPh-on/-off conditions within-subject in boys with ADHD using the pro- and anti-saccade tasks. Even after accounting for some practice effects due to the task repetition under the two MPh conditions, virtually all aspects of saccade control were improved by the drug. Largest effects statistics were found, however, for those parameters that are related to the generation of pro-saccades, both during the pro-saccade task (pro-saccadic RT) and during the anti-saccade task following direction errors, namely the proportion of error corrections and correction reaction times. These studies fit nicely into the broader picture of MPh's beneficial effects on a range of cognitive and ocular-motor functions (as reviewed in Pietrzak, Mollica, Maruff, & Snyder, 2006).

The five approaches outlined before—improve clinical diagnosis, objectivise clinical symptoms, compare disorders, identify potential aetiological factors, and assess treatment effects—are but facets of the overarching aim to understand ocular-motor control and visual attention in child and adolescent psychiatric disorders better. This holds of course not only for the study of ADHD, but for all psychiatric disorders, and it is by no means specific to eye movement research. The recording of eye movements, however, is a distinguished method to do so, and a powerful way to realise its potential is the administration of a *battery of eye movement tasks*, as accomplished for example by Loe, Feldman, Yasui, and Luna (2009). This study recruited typically developing children and adolescents, and those with ADHD, aged 8–13. Both groups were controlled in age and IQ differences. All participants were administered a visually-guided saccade task, an anti-saccade task, a memory-guided saccade task and a fixation task. This allowed the researchers to make comparisons across different tasks. And indeed, one of the interesting results reported by Loe and colleagues was, that inhibitory deficits were found in the anti-saccade task (direction errors), in the memory-guided saccade task (saccades during the delay interval) and in the fixation task (saccade to peripheral stimulus presented at the end of a 200 ms gap in presenting a central fixation cross), suggesting that there is an inhibitory deficit that is not task-specific, but potentially general. Ideally, such conclusions rely on sufficiently high correlations of the different construct indicators (here: of the inhibitory deficit) across tasks. Another advantage of administering a task battery is that a profile of effect sizes of group differences can be made. In the present study, and based on the reported test statistics (F-values), the greatest differences between patients and controls was found for dis-inhibited saccade during the fixation task, while the smallest effect size of group differences were found for the pro-saccadic RTs.

While the previous paragraphs were meant to explain the different approaches and underlying research questions in ocular-motor studies with psychiatric populations, the next sections will provide a summary of what has so far been found and replicated in the study of ocular-motor abnormalities in ADHD. Overall, three results seem to emerge from the ADHD literature. (1) The first consistent finding is the *increased number of premature saccades* in the memory-guided saccade task (Castellanos et al., 2000; Loe et al., 2009; Mostofsky, Lasker, Singer, Denckla, & Zee, 2001; Mostofsky, Lasker, Cutting, Denckla, & Zee, 2001b; Ross et al., 1994a, 1994b; Rommelse et al., 2008) or the pro- or anti-saccade tasks (Karatekin, 2006; Klein, Raschke, & Brandenbusch, 2003). "Premature saccades", in these paradigms, means saccades that occur while participants are supposed to fixate a central fixation, either in anticipation of a peripheral cue (pro- and anti-saccade tasks) or in between the presentation of the peripheral cue and the offset of the central fixation (memory-guided saccade task). In both cases—albeit possibly for different reasons—these premature saccade "intrude" upon what should be stable fixation. Increased numbers of intrusive saccades in patients with ADHD have also been reported for go-nogo or fixation tasks (Castellanos et al., 2000; Gould, Bastain, Israel, Hommer, & Castellanos, 2001; Loe et al., 2009) and are thus apparently not a task-specific phenomenon. Premature saccades seem to occur more often also in adults with ADHD (Adams et al., 2011; Feifel et al., 2004) and are thus possibly pointing to a developmentally stable deficit. Latencies of correct memory-guided saccades have been reported to be normal in ADHD in some (Castellanos et al., 2000; Rommelse et al., 2008; Ross et al., 1994), but not all studies (e.g., Loe et al., 2009). (2) Another consistent finding is the *increased number of direction errors* in the anti-saccade task, which has been replicated several times (e.g., Habeych et al., 2006; Karatekin, 2006; Klein et al., 2003; Loe et al., 2009; Mahone, Mostofsky, Lasker, Zee, & Denckla, 2009; Mostofsky et al., 2001a, b; Munoz et al., 2003; O'Driscoll et al., 2005). The studies that did not report this effect either used an unusual version of the anti-saccade task (Aman et al., 1998); or found just a weaker but non-significant trend in the same direction (Hanish et al., 2006); or used only 10 trials along with younger age of the control participants compared to the ADHD patients (Rothlind et al., 1991). (3) The third consistent finding that shines through is the increased intra-subject variability (ISV) of saccadic reaction times, as defined in most ocular motor studies on ADHD as coefficient of variation (CV; standard deviation divided by mean). This finding has been reported consistently across different standard paradigms and in all of the studies that analysed this parameter (anti: Karatekin et al. 2009; visually-guided: Mahone et al., 2009; Mostofsky et al., 2001a, b; Munoz et al., 2003; memory-guided: Rommelse et al., 2008). This finding is particularly interesting as increased ISV of manual-motor responses has been confirmed as one of the most consistent findings in the entire ADHD literature (Klein, Wendling, Huettner, Ruder, & Peper, 2006; Kofler et al., 2013; Kuntsi & Klein, 2011). Both the increased proportions of premature/intrusive saccades and anti-saccadic direction errors can be considered as inhibitory deficits, while the increase in ISV points to deficient self-regulation.

While there are consistencies in the findings with ADHD patients (see above), there are also inconsistencies in the literature regarding, for instance, pro- and anti-saccadic reaction times or global measure of deviation from smooth pursuit (Bylsma & Pivik, 1989; Castellanos et al., 2000; Jacobsen et al., 1996; adults: Ross, Olincy, Harris, Sullivan, & Radant, 2000). Such—and other—inconsistencies require a discussion of their potential causes. *Firstly*, while there are studies with large samples (e.g., Gould et al., 2001; Klein et al., 2003; Munoz et al., 2003), many studies use small or very small sample sizes. For clinical ocular-motor studies, as for all clinical studies, this endangers the generalisability of their findings given that many psychiatric disorders are aetiologically heterogeneous and participant sampling occurs probably mostly by opportunity. *Secondly*, ocular-motor task performance—and potential impairments herein or developmental effects (see Klein, 2001)—may crucially depend on task settings (e.g., the target velocity in a smooth pursuit task, the maintenance interval in a memory-guided saccade task, or the presence of a gap between fixation and cue in a pro- or anti-saccade task) which are not standardised in the psychiatric literature (Smyrnis et al., 2008). *Thirdly*, ocular-motor task performance and group differences herein may also depend on the length of the test (see Klein et al., 2003) which is also not standardised and in some studies so short that the “test” is hardly more than a “practice block”. Here, accumulating many trials to ascertain test reliability (see Klein & Fischer, 2005) has to be balanced against avoiding time-on-task effects (due to, e.g., practice, fatigue, loss of test motivation) that may differentially affect patients and controls (Klein et al., 2003). Fortunately, recommendations to standardise ocular motor tasks are being developed (Antoniades et al., 2014). Truly normative data for standardised test instruments, as we know from the domains of IQ or neuropsychological testing, are however still missing in ocular-motor research. *Fourthly*, neither typical nor atypical *development* of ocular-motor functions has so far been investigated in proper longitudinal studies. Klein et al. (2011), however, reported findings of a cross-sequential studies in which participants aged 6-18 years at initial testing were retested about 19 months later. Importantly, although the dimensional structure of age-related individual differences (as revealed by an exploratory Principal Components Analysis (PCA) of the cross-sectional data) resembled the dimensional (PCA) structure of the developmental change scores, it was not identical. One of the difference between the cross-sectional and the longitudinal data related to the *factorial splitting* of pro- and anti-saccadic RTs alongside a *factorial unifying* of the ISV of pro- and anti-saccadic RT in the longitudinal data that was not found for the cross-sectional data. In other words, while developmental changes in saccadic RT turned out to be to some extent task-specific with regard to the pro- and anti-saccade tasks, developmental changes in the ISV of saccadic RT were task-independent and thus potentially pointing to a more fundamental developmental process. This pattern was not apparent in the cross-sectional data. Results such as these underline the importance of longitudinal studies to investigate how our ocular-motor system is “build up” during infancy, childhood and adolescence; what functions or processes can be differentiated through studying this developmental process; and how patients with neurodevelopmental disorders differ herein.

Box 3: First Time Working with Oculomotor Paradigms in Child and Adolescent Psychiatry (CAP)?—What to Keep in Mind!

Getting comfortably seated—The testing environment itself is an important starting point to think about the question ‘is this CAP friendly?’. For collecting eye data, it is important to maintain an optimal screen-to-participant distance. For speeded motor responses, the participant should also be able to comfortably place their forearms on the table. So the apparatus should be arranged in a way that the participant’s face is in the middle of the screen at a distance approximately 70–80 cm (depending on the eye tracker), with forearms still comfortably reaching the response buttons. Given that children and adolescents achieve varying heights at different speeds in their development, it is ideal to keep these items in the lab—screen, chair, table and (potentially) chin rest—adjustable. Another helpful provision is to have a chair that does not move or swirl when someone is seated. For smaller children, a foot step may be useful too. These arrangements are particularly useful for children with hyperactivity or a tendency to move and fidget. In any event, it is important to ascertain a comfortable sitting position for the participant before starting with instructions, calibration and testing.

Explaining the Equipment—Once the participant is comfortably seated, and before delving into instructions, it is conducive to explain the environment to the participant. This is especially good with younger children who like to touch and grasp things to learn about them. By running them through what everything in the room is, what it does, and why they shouldn’t touch it, the experimenters can ensure not only their attention but also direct their behaviour into favourable positions. A neat way to do this is with a child friendly explanation about what is expected of them in the lab. For example, one could tell a young participant that „the red light in the tracker is what locates your eyes. Touching the tracker would mean it has fingerprints on it, just like a mobile screen, and then it can’t locate the eyes any more. We don’t want that! What we do want, is to reply as fast and accurately as possible to the game on screen. So let’s put our hands close to the response buttons instead“.

Explaining the Tasks—Such child-friendly explanations can continue into the instructions phase. An orienting cross, for example, can be explained as a start point. One has to stay on the cross with their eyes or the trial after is void, just like in a race. If one is not at the starting point, they cannot participate in the race. In general, the content of task instructions needs to be broken down for children for better understanding, in particular if the task is complex and time bound.

When designing experiments for children, one has many opportunities to make tasks child-friendly. Age appropriate tasks fall neatly into stories as well. For example, a pro-saccade task can involve looking at stars (or something beautiful) and an anti-saccade task looking away from insects (or something ugly). It is more concrete to understand ‘look at stars and look away from insects', as compared to 'look at the blue circle, look away from the yellow circle'.

Once clear on instructions, the participants may be first allowed to practice without a time restriction. Once the experimenter is certain of task understanding, a time restriction can be added and finally the practice block mimicking what the main block will contain. This round about procedure adds a mere 2–3 min to the existing paradigm, but ensures participant understanding.

While the content of instructions remains unchanged across participants, the speed of administering instructions may have to be altered for each developmental group. Some groups may need more processing time, slower instructions and more pauses. Others may need shorter pauses and faster speaking to sustain attention. Such variations, however, should be kept to a minimum to ascertain standardisation of testing conditions.

When in doubt if the participant understood the task or not, a good strategy is to ask them to repeat the instructions. While this would work with most children, some clinical groups (with Autism for instance) will not want to talk or look at the experimenter. In these cases, showing them illustrations of the task and asking questions where they can point, such as 'what will you press/where will you look when this comes on screen', is a good second or complementary choice.

Presence of Experimenter—While leaving the participant alone in the room ensures that there are no external influences on performance, the experimenter's presence around the participant in CAP is useful for two main reasons: Firstly, it maintains an environment of discipline and safety; secondly, the experimenter can reinforce instructions ('don't forget to look at the starting point', 'please keep still') quickly and effectively during the experiment when needed. However, as the experimenter is hardly ever "blind" with regard to diagnostic status of the participant tested, such interventions should be delivered parsimoniously and according to a standardised protocol. In many cases, they will help patients more than typically developing children to stay on task and thus work—if anything—towards the Null hypothesis.

Length of tasks—While a sufficient number of trials is required to ascertain psychometric test reliability, this requirement needs to be balanced against potential fatigue and time-on-tasks effects that may affect patients or younger participants more than healthy or older participants. The same applies to the overall duration of the testing session, for which around 20–40 min in infants and 60–90 min in children and adolescents may be acceptable. Inserting breaks, providing some diversity of the different task and establishing a good personal contact with the participants can help them to stay motivated and concentrated for longer periods of testing time.

One's own scan paths: Insightful surprise!—Finally, a lovely thing to do at the end of a testing session is to show the participants their scan paths. This comes as a pleasant surprise to most children, opens them up to talking about their experience during the testing and asking the experimenters questions. A small act of showing participants their scan paths gathers qualitative feedback, while widening scientific curiosity. Adding a short (≈30 s) free viewing of a child-appropriate scenery to the test battery will be sufficient to provide this nice surprise at the end of a testing session.

15.4.3.3 Reading Disorders

Impairments of reading are defined in DSM-5 as an inaccurate or slow and effortful word reading not better explained by intellectual disabilities, neurological disorders, impairments in visual and auditory acuity or lack of adequate school achievement. The deficits include for example letter by letter reading, exchange of similarly shaped letters, frequent regressions during reading of sentences. They are usually accompanied by difficulties in the segmentation of words in their sound components and impairments in the association of phonemes with letter symbols. Reading disorders belong to the group of specific neurodevelopmental disorders (Developmental Dyslexia) and are frequently accompanied by disorders of the written expression (involving low spelling and grammar accuracy and disorganization of written sentences). There is some discussion about the cut-off necessary to identify a reading disorder, which is usually defined according to a significant discrepancy between reading performance and IQ and/or a reading performance below age expectations. The rates of comorbidity with internalizing and externalizing disorders, especially ADHD, are very high and emphasize the impact of specific learning impairments on psychosocial functioning.

Early in the 1980s there has been an increasing interest in studying eye movements in reading. This increased interest was certainly also inspired by the ‘emerging methodological and technical advancements in the eye movement research’ in these years (see chapter by Hyönä & Kaakinen, in this volume). The fact that abilities in extracting word information and in processing the linguistic properties of a written text can be optimally explored by means of eye movement analysis during reading has led to the hypothesis that an abnormal pattern of alternating saccades and fixations could be responsible for emerging reading problems in children at the first stages of the literacy development. Pavlidis was the first one to claim that abnormal eye movement patterns during reading are probably the ‘key’ to dyslexia (Pavlidis, 1981) and his theory has been vehemently discussed for nearly two decades (Black, Collins, De Roach, & Zubrick, 1984; Brown et al., 1983; Eden, Stein, Wood, & Wood, 1994; Fischer & Weber, 1990; Olson, Kliegl, & Davidson, 1983). Finally, late in the 1990s, it was concluded that children with a specific reading disorder just mimic reading patterns of a lower reading developmental stage (longer fixations, shorter saccades and more regression saccades during reading of a sentence; Rayner, 1998) but show normal RT, velocity and amplitude of saccades in visually-guided tasks with single visual stimuli and normal fixation patterns for sequentially-arranged visual targets (Black et al., 1984; Brown et al., 1983; Olson et al., 1983). These findings along with the fact that children with dyslexia have definitively non-ocular-motor problems to segment a spoken word into its phoneme components and to convert phonemes into graphemes seemed to put an end to the discussion and seemed to finish further oculomotor research on dyslexia (Rommelse et al., 2008). Henceforth, the obvious deficits in the ability to manipulate phonological material have supported the hypothesis of a fundamental phonological impairment, which has been for a long time the most influential theory of developmental dyslexia (Vellutino, 1987).

However, Biscaldi et al. (1998, 2000) analysed saccade parameters acquired in standard oculomotor paradigms from populations with and without reading impairments and clearly found an increased intra-subject variability (ISV) of the SRT (broader SRT distributions with an increased density of anticipatory and late saccades) and an increased number of corrected or uncorrected direction errors in the anti-saccade task. They highlighted the importance of eye movement measurements in non-linguistic tasks as a strong tool to study reading disabilities and interpreted the results in the frame of an inhibition deficit in dyslexia probably due to a delay in the maturation of the prefrontal cortex. In this line and in addition to the phonological deficit, an important etiological aspect of dyslexia (or at least subgroups of it) could also be found in abnormalities of visual attention and fixation control.

Given that eye movement research has come to a seeming end a decade ago, we may, nevertheless, ask why eye movements are still relevant in research on reading disorders (RD). As outlined above, there are two main fields of previous eye movement research in RD. *Firstly*, concerning the pattern of eye movements during reading of words and texts (linguistic tasks); and, *secondly*, concerning saccadic behaviour in standard oculomotor paradigms, i.e., visually guided saccade tasks (non-linguistic tasks). These will be discussed next.

Ad 1.: Even if it is clear that, during reading, the increased number of fixations and regressive saccades in dyslexic children resembles that of younger unexperienced readers and thus reflects the developmental delay in reading performance, there are some unsettled issues in this field. First of all, the old studies compared only parameters such as the number of fixations and regressions during reading of sentences in small samples of children within a limited age-range. Yet, while the number of innovative studies about foveal word and parafoveal text processing in developing readers with normal abilities is growing, only little is known about the development of these functions in children with a specific RD. In other words, it is not clear whether eye movements of children with RD (with regards to the optimal landing position in words and the perceptual spam) develop in a similar manner like those of typically developing children evolving from an unexperienced to a skilled reading performance. Moreover, the new eye tracking methodology, which allows oculomotor analysis during reading of longer and complex passages of text, is not only of potential interest for readers with ‘poor text comprehension’ in general, but also for investigating specific linguistic processing deficits in individuals with dyslexia. Finally, the potential role of Intra-Subject Variability (ISV) on reading performance is completely unknown. This parameter could provide a measure of intra-individual stability (or instability) of reading performance in those with RD. Since increased ISV is a robust finding in ADHD and previous literature gives hints that ISV might be increased in (a subgroup of) RD (Biscaldi et al., 1998), the parameter could detect attentional fluctuations during reading in a comorbid ADHD plus RD group. This finding would be relevant for implementation of individualized learning therapy tools.

Ad 2.: Past research has vehemently argued against a basic visual or auditory perception problem (Ramus, 2003) and a theory of oculomotor dysfunctions in dyslexia (Rayner, 1998). However, two main findings from studies on saccadic eye movements of dyslexic children in standard oculomotor tasks—not involving linguistic processing—might add evidence to the hypothesis of a visual/attentional deficit in the aetiology of dyslexia, besides a phonological impairment. On the one hand is the finding of an impaired performance in the anti-saccade task (Biscaldi et al., 2000), which could be due to a disturbed visual-spatial neuronal network involving top-down attentional processes (Vidyasagar, 2013) and/or to an inhibition and working memory deficit similar to that found in ADHD (the latter hypothesis being particularly important with regard to the high comorbidity between RD and ADHD). On the other hand, as Rommelse et al. (2008) stated in their review, there is evidence for an increased number of saccades in visually-guided and smooth pursuit tasks in dyslexic children, in spite of replication problems due to methodological differences between studies. The findings vary from elevated numbers of anticipatory and microsaccades during fixation as well as of forward and regressive saccades while scanning a horizontally arranged array of visual stimuli up to an increased amount of intrusive saccades during smooth pursuit tasks. The last finding was finally put forwards by Eden et al. (1994) but had not been replicated by Ygge, Lennerstrand, Rydberg, Wijecoon, and Pettersson (1993). After these contradictory results no further study was published on this issue. Rommelse et al. (2008) argued that an increased saccade production had been found across different and hardly comparable tasks. Therefore, although the finding is of potential interest, its meaning in the context of specific RD still remains unclear.

Given the previous two arguments, we may ask whether the linguistic theories of dyslexia are really irreconcilable with theories of visual and oculomotor deficits?—At last, the hypothesis of a phonological impairment as the only explanation for understanding dyslexia has been recently challenged by Peterson and Pennington (2015). In their review, Peterson and Pennington (2015) put forwards the ‘dysconnection’ theory of dyslexia, which postulates an ‘intact but less accessible phonetic representation’ (Boets, 2014), a finding supported by several brain imaging studies (Boets, 2014; Ramus, 2014; Ramus & Szenkovits, 2008). At the same time, the authors go further and provide a framework for a multiple deficit account of dyslexia by arguing that the ability to extract information from phonemes and graphemes and to match them is both auditory and visual in nature. This complex cross-modal or multi-sensorial task might be impaired in dyslexia as well, especially for sentence reading. Thus, the main point seems to find the appropriate behavioural studies to address the issue of a multifunctional deficit in dyslexia. A good example are findings that cognitive-linguistic processes involving short-term memory and speeded performance (i.e., rapid serial naming, rapid temporal auditory processing, and fast visual detection) can also predict reading difficulties (McGrath et al., 2011; Fischer, Hartnegg, Mokler, 2000; Tallal, 2000). If it holds true that phoneme and grapheme processing is both auditory and visual, also studying eye movements with appropriate tasks is systematically involved in testing linguistic theories of dyslexia.

Which findings and task paradigms are worthy to be pursued in the future?—The above mentioned theories on perceptual deficits, impairments of rapid information processing, and, most peculiarly for the aim of this chapter, crowding of visual information (Facoetti et al., 2010) provide a strong link to potential oculomotor dysfunctions in dyslexia. Especially the 'crowding' effect, defined as decreased foveal letter identification due to interference of peripheral letters, seems to affect subgroups of dyslexics and may be linked to a visual attentional problem (Facoetti et al., 2010). Unfortunately, there are as yet no studies that analyse crowding during reading in relation to eye movements.

The fact that oculomotor research in reading is well established should actually promote further research bringing forwards the previous uncompleted findings. Here are some suggestions:

1. In line with the recommendations put forwards by Rommelse et al. (2008) there is a need for investigating saccades in standard oculomotor tasks of pro- and anti-saccades in gap and overlap fixation paradigms as well as in tasks of smooth-pursuit. The aim will be to replicate past results of unstable fixation using controlled task parameters and testing large samples taking into account ADHD comorbidity. Biscaldi et al. (1998) observed increased standard deviations of SRT but not of mean SRT in dyslexics compared to normally reading subjects. Increased intra-subject variability (ISV) measured by different parameters (Feige et al., 2013) is one of the most stable finding and probably an endophenotype of ADHD (Karalunas, Geurts, Konrad, Bender, and Nigg, 2014; Klein et al., 2006). These parameters could be useful as measure of fixation stability in reading and non-reading tasks, especially in subgroups with high ADHD comorbidity. In contrast, it has been sufficiently proved that parameters such as velocity, amplitude, and mean RT of saccades do not differ between dyslexic and typically developing children in visually-guided tasks (see summary in Rommelse et al., 2008). These findings do not need replication.
2. Developmental aspects are a neglected issue in oculomotor research of neurodevelopmental disorders in general, and of RD in particular (Luna et al., 2008). Longitudinal studies can provide additional information, especially about the age-dependent 'true' difference in performance due to developmental factors (Klein et al., 2011). In Biscaldi et al. (2000) the anti-saccade task was applied for the first time in a cross-sectional sample aged 7–17 years with reading deficits. The results clearly showed that the inhibition problem was even worse in adolescents than in children, hence probably not due to a mere developmental delay. Yet, this assumption should be validated in longitudinal studies.
3. Klein et al. (2010) additionally showed in a healthy population that parameters of anti-saccades specifically correlate with components of cognition such as fluid

intelligence, working-memory or inhibitive performance. It would be interesting to identify the impaired cognitive components in individuals with dyslexia and to correlate them with parameters from saccade tasks.

4. Literature on eye movements during reading gives suggestions for further studies with psychiatric populations by using specific parameters of 'best landing position', 'regressive saccades', 'perceptual span', 'foveal and parafoveal processing' and well established models of eye guidance during reading. In this context, advanced and refined parameters of binocular convergence and divergence in reading and non-linguistic tasks should be studied in relation to crowding and attentional problems (Bucci & Seassau, 2012; Jainta & Kapoula, 2011). This approach is not only interesting with respect to reading disorders. Even though some studies in ADHD use reading tasks and even use a comparison group of specific reading disorders, binocular convergence and divergence are usually not measured. Interesting for these studies could be also the investigation of visual search and its comparison with binocular control during sentence reading tasks (Bucci and Seassau 2012; Seassau, Gérard, Bui-Quoc, & Bucci, 2014) as well as of eye tracking during free exploration (Kapoula et al., 2009).
5. Finally, a further suggestion is to use gaze movement as a therapeutic tool, i.e., using gaze movement as feedback to train participants to read better. Neither has been explored and either would be a good asset to the literature, when studied.

In conclusion, the pattern of eye movements in word and text reading has been studied for a long time and it has been postulated that the abnormal pattern in RD simply reflects problems in the development of linguistic processes rather than an immature oculomotor system (see chapter by Hyönä & Kaakinen, in this volume). Yet, there are several old und new robust findings in the field (abnormal anti-saccade performance, instable fixation, irregularity in smooth-pursuit, instable binocular control, and crowding just to mention a few of them) that should be further investigated using comparable eye movement paradigms, well characterized clinical groups, and new available technical and analysis tools of eye tracking. The fine and complicated relationship between linguistic and visual/attentional processes as well as visual-motor coordination during reading has to be further explored in experienced and inexperienced readers as well as in readers with specific neurodevelopmental disorders. The aetiology of neurodevelopmental disorders can be understood when we try to highlight which characteristics, deficits, and peculiarities in information processing and cognition are common to and which differ between the different phenotypes. Eye movement studies are a powerful potential tool to do this.

15.4.3.4 Early Onset Schizophrenia

The schizophrenia spectrum is characterized by thought disorders, perceptual abnormalities, and diminished emotional expression. In the DSM-5 the strong classification in clinical subtypes and Schneider's concept of first-rank symptoms have been replaced by a dimensional approach considering symptom domains with variable psychopathological severity. The key symptom groups are defined as delusions, hallucinations, and disorganized thinking (derailed up to incoherent speech) as well as abnormal motor behavior (ranging from disorganized motor activity to agitation or catatonic behavior) and negative symptoms (blunted affect, emotional and social withdrawal up to disturbance of volition) (DSM-5, APA). Placing the schizophrenia spectrum straight after the Neurodevelopmental Disorders (NDD) section, the DSM-5 underlies the developmental character of schizophrenic diseases, that is, peak age at onset in the twenties, usually with slow and gradual development of symptoms and cognitive abnormalities preceding psychotic symptoms. Especially Early Onset Schizophrenia (EOS; Petruzzelli, Margari, Bosco, Craig, Palumbi, & Margari, 2018), although rare among children with an incidence rate rising in the adolescence (estimated 0.23%), definitively shows similarities to some characteristics of the NDD, such as male prevalence and a frequent association with developmental delays of intellectual, language, and motor functions, which precede the gradual emerging of psychotic symptoms. Evidence of emotional and behavioral abnormalities can be found during early childhood also. Diagnosis of EOS can be challenging and the outcome has been estimated to be poorer than in later-onset psychosis. In the early onset psychosis, perceptual abnormalities such as hallucinations are not predominant and less specific. Clinical subtypes with emotional and social withdrawal (negative symptoms) or conceptual disorganisation are more frequent than during adulthood, and symptom development is linked to the co-occurrence of neurodevelopmental deficits.

The investigation of eye movement dysfunctions in EOS is interesting for at least two reasons. *Firstly*, it is generally assumed that increased genetic load leads to the early onset of the disorder. As different eye movement dysfunctions—including those of smooth pursuit eye movements, anti-saccades and memory-guided saccades—have been suggested to reflect endophenotypes of schizophrenia (for reviews see, Calkins et al., 2008; Gooding and Basso, 2008; and the chapter by Smyrnis et al., in this volume) and are thus putatively related to the genetic aetiology of the disorder, they are particularly well suited to investigate potentially unique patterns of inheritance in EOS. *Secondly*, given that schizophrenia is considered a neurodevelopmental disorder by many researchers, the potential relationship of endophenotypic eye movement dysfunctions with the rather heterogeneous course of the disorder in patients at a developmentally dynamic age is worth being investigated. However, given the low prevalence of EOS, the mere amount of studies that have tested these patients is rather limited.

To summarize the main findings for standard ocular-motor paradigms: Regarding smooth pursuit, different studies have reported reduced gain (that is, a lower velocity of the eye compared to the target) or increased root mean square error (RMSE; a global measure of eye tracking deviance; see chapter by Lencer et al., in this volume; e.g., Jacobsen et al., 1996; Kumra et al., 2001). Other studies have reported increased numbers of anticipatory or leading saccades (e.g., Jacobsen et al., 1996; Kumra et al., 2001) or saccades intruding into smooth pursuit (Mather, 1985) and suggest overall impaired smooth pursuit in EOS. Dis-inhibited—and thus "intruding"—saccades were also reported for the memory-guided saccade (MGS) task (Ross et al., 2005). Concerning visually-guided saccades, normal latencies and precision have been reported in one study (Mather, 1985), with these results being "echoed" in another study showing normal latencies and accuracies of MGS (Ross et al., 2005). Increased inaccuracies of MGS were, however, reported by White, Mous, and Karatekin (2014). Ross et al. (2005), anecdotally, also reported that younger EOS patients in particular had problems to accomplish the anti-saccade task and that this task was therefore dropped from the testing protocol. Although the nature of this problem is not described, it certainly includes the lack of control over "reflex-like" glances at the peripheral cue. If these reflex-like glances were followed by corrective saccades into the opposite hemi-field (confirming that the participant knows what s/he is supposed to do), they could be interpreted as dis-inhibited pro-saccades and add to the (limited) evidence of inhibitory failures in ocular-motor control in EOS reported for smooth pursuit.

While the previous paragraph has summarised some of the main findings, the following will take a more detailed look at a few studies that have addressed questions that go beyond assessing ocular-motor performance in EOS patients.

In line with the first argument above for studying eye movement dysfunctions in EOS patients, Ross et al. (1999) investigated the inheritance of anticipatory saccades during smooth pursuit by comparing the likelihoods of such distortions in the parents of (a) childhood-onset and (b) adult-onset schizophrenic patients. In line with the assumption that EOS is associated with increased inheritance of the gene(s) coding for endophenotypes such as anticipatory saccades during smooth pursuit, the authors reported greater likelihood that both parents show this kind of smooth pursuit impairment in childhood-onset compared to adulthood-onset schizophrenia. This so-called "bilenial inheritance" (i.e., both parents rather than just one parent showing the abnormality) that was found for anticipatory saccades during smooth pursuit was not found for schizophrenia itself, suggesting that childhood-onset schizophrenia may involve a greater genetic risk due to the inheritance of schizophrenia-risk genes from both parents. Further confirmation of the potential endophenotypic status of impaired smooth pursuit was the observation that the non-psychotic children of schizophrenic patients showed more such anticipatory saccades than normal children and less than EOS patients. Finding an expression of a variable in unaffected first-degree relatives of index cases that is in-between the cases and healthy controls is what can be expected from an endophenotype of a neuropsychiatric disorder. The same reasoning, however, can be employed to argue *against* considering MGS impairments as an endophenotype of schizophrenia, as childhood-onset schizophrenia cases, but not the children of adult schizophrenic patients, showed an increased number of premature saccades during this task in Ross and colleagues' (2005) study.

Concerning free viewing tasks in children and adolescents with schizophrenia, only very little research has been done. Even when considering the low prevalence of EOS this is surprising given the intriguing findings that have been reported for adults with schizophrenia. One such study, showing the potential diagnostic value of eye movement parameters, was published by Benson et al. (2012). In this study, a free viewing, a fixation and a smooth pursuit task discriminated adult schizophrenic patients and healthy controls with nearly perfect accuracy. In fact, the fixation dispersal during free viewing was restricted in schizophrenic participants and turned out to be the best single discriminator of the two groups. Moreover, Benson and colleagues reported that this scanning style generalised across all their stimuli; it was thus not only found for emotional or social stimuli but rather reflected a general processing style of the patients. Similar results were later reported by Morita et al. (2017) who used the scanpath length from their free viewing task for the discriminant analysis. One of the few free viewing studies in children and adolescents with schizophrenia is that of Karatekin and Arsanow (1999). The researchers presented thematic pictures (drawings from children books) for 8 s each which had an ambiguous or complex theme and enabled them to ask 3 questions reflecting different aspects of attention: (1) a global question that was intended to represent the top-down control of selective visual attention under self-guided behaviour (e.g., "What is happening in this picture?"); (2) a focal question where specific, discrete areas had to be attended in order for the participant to answer (e.g., "How old do you think each person is?"); and (3) a counting question where the visual attention had as a purpose items in areas other than those used for the focal question to be counted (e.g., "How many animals are there?"). All the questions were posed before the stimuli were presented. For each question, they measured the proportion of time spent viewing relevant and irrelevant regions (as each picture was divided into regions and each region was assigned a rating of relevance for each question), the average fixation duration and the sum of distances between consecutive fixations. The results revealed that EOS participants did not differ from controls in the basic control of eye movements (mainly demonstrated through the last two questions) or in the initial stages of search. Furthermore, there was no evidence that patients with EOS performed an excessive scanning. However, they looked at fewer relevant, but not more irrelevant regions than normal controls and, most importantly, they had significantly longer fixations than controls in the global question where there was the least constriction of the visual behaviour. On the basis of these findings, the authors concluded that children with EOS stared more and showed deficient top-down control of visual attention when the condition was not constraining their gaze behaviour and was more under self-guided action.

15.4.4 Summary Regarding Atypical Development

Starting with an introduction and discussion of five basic paradigms in studying eye movements in atypically developing populations, Sect. 15.4 focused on four neurodevelopmental disorders—autism, attention-deficit/hyperactivity disorder, reading disorder and early-onset schizophrenia—and explained what eye movement research has contributed to the understanding of these disorders. We saw that standard paradigms like pro-, anti-, or memory-guided saccades can provide valuable information about deficits that may or may not be specific to individual disorders, whilst the free viewing paradigm can reveal peculiarities in visual attention for various kinds of social or non-social stimuli. Altogether, the studies presented in Sect. 15.4 made clear that the recording of eye movements is a highly versatile approach in the study of neurodevelopmental disorders, capable of yielding unique results that cannot be yielded by other technologies. Among the study designs that should be used more often are the combination of different tasks in the same study and the comparison of different clinical groups using the same task batteries. Finally, employing designs with a longitudinal component is highly demanded, albeit difficult to implement.

15.5 Suggested Readings

Guillion, Q., Hadjikhani, N., Baduel, S., Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. Neuroscience and Biobehavioural Reviews, 42, 279–297.

Johnson, B. P., Lum, J. A. G., Rinehart, N. J., Fielding, J. (2016). Ocular motor disturbances in autism spectrum disorder: Systematic review and comprehensive meta-analysis. *Neuroscience and Biobehavioral Reviews*, 69, 260–279.

Karatekin C. (2007). Eye tracking studies of normative and atypical development. Developmental Reviews, 27, 283–348.

Klein, C. & Ettinger, U. (2008). A hundred years of eye movement research in psychiatry. *Brain and Cognition*, Volume 68 (special issue).

Luna, B., Velanova, K., Geier, C. F. (2008). Development of eye movement control. *Brain and Cognition*, 68, 293–308.

Papagioannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. Social Neuroscience, 9, 610–632.

Rommelse, N. N. J., Van der Stigchel, S., & Sergeant, J. A. (2008). A Review on Eye Movement Studies in Childhood and Adolescent Psychiatry. *Brain and Cognition, 68,* 391–414.

15.6 Questions Students Should Be Able to Answer

- What is the rationale behind the habituation-dishabituation paradigm? Which advantages and disadvantages does it have?
- What kind of paradigms can be used to study development?
- What do we know about the typical development of the saccade system?
- Compare the development of the anterior and posterior attention systems. How can their differential developmental trajectories be observed in the gaze movements of infants?
- What can eye/gaze movements tell us about development that other methods cannot?
- If you were on board a longitudinal project with typically developing children and adolescents, what paradigms would you choose, and why?

- How can we use gaze movements to study comorbidity? What can it tell us?
- If you were designing paradigms for children with ASD, what would be important for you to know?
- Which factors may reduce the validity of eye movement tasks in atypically developed populations?
- How should you form research questions in child and adolescent psychiatry?
- What research question do you think is still unaddressed in eye movements in childhood onset schizophrenia.
- Can you think of paradigms not covered in this chapter that would benefit from using eye-tracking in child and adolescent psychiatry?

Bibliography

Açik, A., Sarwary, A., Schultze-Kraft, R., Onat, S., & König, P. (2010). Developmental changes in natural viewing behavior: Bottom-up and top-down differences between children, young adults and older adults. *Frontiers in Psychology, 1,* 207. https://doi.org/10.3389/fpsyg.2010.00207.

Ajrezo, L., Wiener-Vacher, S., & Bucci, M. P. (2013). Saccades improve postural control: A developmental study in normal children. *PLoS ONE, 8*(11), e81066. https://doi.org/10.1371/journal.pone.0081066.

Alahyane, N., Lemoine-Lardennois, C., Tailhefer, C., Collins, T., Fagard, J., & Dore-Mazars, K. (2016). Development and learning of saccadic eye movements in 7- to 42-month-old children. *Journal of Vision, 16*(1), 6. https://doi.org/10.1167/16.1.6.

Antoniades, C., Ettinger, U., Gaymard, B., Gilchrist, I., Kristjansson, A., Kennard, C., et al. (2014). An internationally standardised antisaccade protocol. *Vision Research, 84,* 1–5.

Aring, E., Grönlund, M. A., Hellström, A., & Ygge, J. (2007). Visual fixation development in children. *Graefe's Archive for Clinical and Experimental Ophthalmology, 245*(11), 1659–1665. https://doi.org/10.1007/s00417-007-0585-6.

Asendorpf, J. (1999). *Psychologie der Persönlichkeit. Grundlagen*. Berlin Heidelberg: Springer.

Aslin, R. N. (1981). Development of smooth pursuit in human infants. In D. Fisher, R. Monty, & J. Senders (Eds.), *Eye movements: Cognition and visual perception* (pp. 31–51). Hillsdale/N.J: Lawrence Erlbaum Associates Inc.

Aslin, R. N., & Salapatek, P. (1975). Saccadic localization of visual targets by the very young human infant. *Perception and Psychophysics, 17,* 293–302. https://doi.org/10.3758/BF03203214.

Baillargeon, R. (1987). Object permanence in 31/2- and 41/2-month-old infants. *Developmental Psychology, 23,* 655–664.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: The Special Needs and Autism Project (SNAP). *The lancet, 368*(9531), 210–215.

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders, 40*(3), 358–370.

Bedford, R., Elsabbagh, M., Gliga, T., Pickles, A., Senju, A., Charman, T., et al. (2012). Precursors to social and communication difficulties in infants at-risk for autism: Gaze following and attentional engagement. *Journal of Autism and Developmental Disorders, 42*(10), 2208–2218. https://doi.org/10.1007/s10803-012-1450-y.

Benson, P. J., Beedie, S. A., Shephard, E., Giegling, I., Rujescu, D., & St, Clair D. (2012). Simple viewing tests can detect eye movement abnormalities that distinguish schizophrenia cases from controls with exceptional accuracy. *Biological Psychiatry, 72,* 716–724. https://doi.org/10.1016/j.biopsych.2012.04.019.

Berlyne, D. E. (1958). The influence of albedo and complexity of stimuli on visual fixation in the human infant. *British Journal of Psychology, 49,* 315–318.

Billeci, L., Narzisi, A., Campatelli, G., Crifaci, G., Calderoni, S., Gagliano, A., … Comminiello, V. (2016). Disentangling the initiation from the response in joint attention: An eye-tracking study in toddlers with autism spectrum disorders. *Translational Psychiatry, 6*(5), e808. https://doi.org/10.1038/tp.2016.75.

Biscaldi, M., Fischer, B., & Hartnegg, K. (2000). Voluntary saccadic control in dyslexia. *Perception, 29*, 509–521.

Biscaldi, M., Gezeck, S., & Stuhr, V. (1998). Poor saccadic control correlates with dyslexia. *Neuropsychologia, 36*(11), 1189–1202.

Biscaldi, M., Rauh, R., Müller, C., Irion, L., Saville, C. W., Schulz, E., et al. (2015). Identification of neuromotor deficits common to autism spectrum disorder and attention deficit/hyperactivity disorder, and imitation deficits specific to autism spectrum disorder. *European Child and Adolescent Psychiatry, 24*(12), 1497–1507.

Black, M. H., Chen, N. T., Iyer, K. K., Lipp, O. V., Bölte, S., et al. (2017). Mechanisms of facial emotion recognition in autism spectrum disorders: Insights from eye tracking and electroencephalography. *Neuroscience and Biobehavioral Reviews, 80,* 488–515. https://doi.org/10.1016/j.neubiorev.2017.06.016.

Boets, B. (2014). Dyslexia: Reconciling controversies within an integrative developmental perspective. *Trends of Cognitive Science, 18*(10), 501–503. https://doi.org/10.1016/j.tics.2014.06.003. Epub 2014 Jul 14, Review.

Bornstein, M. H., Hahn, C.-S., & Wolke, D. (2013). Systems and cascades in cognitive development and academic achievement. *Child Developmen, 84*(1), 154–162. https://doi.org/10.1111/j.1467-8624.2012.01849.x.

Bronson, G. W. (1990). Changes in infants' visual scanning across the 2- to 14-week age period. *Journal of Experimental Child Psychology, 49,* 101–125.

Bronson, G. W. (1991). Infant differences in rate of visual encoding. *Child Development, 62,* 44–54. https://doi.org/10.1111/j.1467-8624.1991.tb01513.x.

Brown, B., Haegerstrom-Portnoy, G., Adams, A. J., Yingling, C. D., Galin, D., Herron, J., & Marcus, M. (1983). Predictive eye movements do not discriminate between dyslexic and control children. *Neuropsychologia, 21*(2), 121–128.

Bucci, M. P., & Seassau, M. (2012). Saccadic eye movements in children: A developmental study. *Experimental Brain Research, 222*(1–2), 21–30. https://doi.org/10.1007/s00221-012-3192-7.

Bucci, M. P., & Seassau, M. (2014). Vertical saccades in children: A developmental study. *Experimental Brain Research, 232*(3), 927–934. https://doi.org/10.1007/s00221-013-3805-9.

Bulf, H., Valenza, E., & Simion, F. (2009). The visual search of an illusory figure: A comparison between 6-month-old infants and adults. *Perception, 38,* 1313–1327.

Bylsma, F. W., & Pivik, R. T. (1989). The effects of background illumination and stimulant medication on smooth pursuit eye movements of hyperactive children. *Journal of Abnormal Child Psychology, 17*(1), 73–90.

Calkins, M. E., Iacono, W. G., & Ones, D. S. (2008). Eye movement dysfunction in first-degree relatives of patients with schizophrenia: A meta-analytic evaluation of candidate endophenotypes. *Brain and Cognition, 68,* 436–461.

Cascio, C. J., Foss-Feig, J. H., Heacock, J., Schauder, K. B., Loring, W. A., Rogers, B. P., et al. (2014). Affective neural response to restricted interests in autism spectrum disorders. *Journal Child Psychology Psychiatry, 55*, 162–171. https://doi.org/10.1111/jcpp.12147 Cermak,

Castellanos, F. X., Marvasti, F. F., Ducharme, J. L., Walter, J. M., Israel, M. E., Krain, A., ... & Hommer, D. W. (2000). Executive function oculomotor tasks in girls with ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 39*(5), 644–650.

Chien, S. H.-L. (2011). No more top-heavy bias: Infants and adults prefer upright faces but not top-heavy geometric or face-like patterns. *Journal of Vision, 11*(6), 1–14. https://doi.org/10.1167/11.6.13.

Chien, S. H.-L., Bronson-Castain, K., Palmer, J., & Teller, D. Y. (2006). Lightness constancy in 4-month-old infants. *Vision Research, 46*(13), 2139–2148.

Chita-Tegmark, M. (2016). Attention allocation in ASD: A review and meta-analysis of eye-tracking studies. *Review Journal of Autism and Developmental Disorders, 3*(3), 209–223. https://doi.org/10.1007/s40489-016-0077-x.

Civan, A., Teller, D. Y., & Palmer, J. (2005). Relations among spontaneous preferences, familiarized preferences, and novelty effects: Measurement with forced-choice techniques. *Infancy, 7*(2), 111–142.

Cohen, L. B., & Marks, K. S. (2002). How infants process addition and subtraction events. *Developmental Science, 5*(2), 186–201.

Black, J. L., Collins, D. W. K., De Roach, J. N., & Zubrick, S. (1984). A detailed study of sequential saccadic eye movements for normal-and poor-reading children. *Perceptual and Motor Skills, 59*(2), 423–434.

Colombo, J. (1995). On the neural mechanisms underlying developmental and individual differences in visual fixation in infancy: Two hypotheses. *Developmental Review, 15*(2), 97–135.

Colombo, J., & Cheatham, C. L. (2006). The emergence and basis of endogenous attention in infancy and early childhood. *Advances in Child Development and Behavior, 34,* 283–322.

Colombo, J., Freeseman, L. J., Coldren, J. T., & Frick, J. E. (1995). Individual differences in infant fixation duration: Dominance of global versus local stimulus properties. *Cognitive Development, 10*(2), 271–285.

Connolly, A. J., Rinehart, N. J., Johnson, B., Papadopoulos, N., & Fielding, J. (2016). Voluntary saccades in attention-deficit/hyperactivity disorder: Looking into the relationship between motor impairment and autism spectrum disorder symptoms. *Neuroscience, 334,* 47–54.

Courage, M. L., Reynolds, G. D., & Richards, J. E. (2006). Infants' attention to patterned stimuli: Developmental change from 3 to 12 months of age. *Child Development, 77*(3), 680–695.

Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition, 72*(3), 237–267.

Dalton, K. M., Nacewicz, B. M., Johnstone, T., Schaefer, H. S., Gernsbacher, M. A., Goldsmith, H. H., et al. (2005). Gaze fixation and the neural circuitry of face processing in autism. *Nature Neuroscience, 8,* 519–526.

Deans, P., O'Laughlin, L., Brubaker, B., Gay, N., & Krug, D. (2010). Use of eye movement tracking in the differential diagnosis of attention deficit hyperactivity disorder (ADHD) and reading disability. *Psychology, 2010*(1), 238–246. https://doi.org/10.4236/psych.2010.14032.

Domsch, H., Lohaus, A., & Thomas, H. (2010). Influences of information processing and disengagement in infant's looking behavior. *Infant & Child Development, 19*(2), 161–174.

Dowiasch, S., Marx, S., Einhauser, W., & Bremmer, F. (2015). Effects of aging on eye movements in the real world. *Frontiers in Human Neuroscience, 9,* 46. https://doi.org/10.3389/fnhum.2015.00046.

Dupierrix, E., de Boisferon, A. H., Méary, D., Lee, K., Quinn, P. C., Di Giorgio, E., … Pascalis, O. (2014). Preference for human eyes in human infants. *Journal of Experimental Child Psychology, 123,* 138–146. https://doi.org/10.1016/j.jecp.2013.12.010.

Eden, G. F., Stein, J. F., Wood, H. M., & Wood, F. B. (1994). Differences in eye movements and reading problems in dyslexic and normal children. *Vision Research, 34*(10), 1345–1358.

Eenshuistra, R. M., Ridderinkhof, K. R., & van der Molen, M. W. (2004). Age-related changes in antisaccade task performance: Inhibitory control or working-memory engagement? *Brain and Cognition, 56*(2), 177–188. https://doi.org/10.1016/j.bandc.2004.02.077.

Eenshuistra, R. M., Ridderinkhof, K. R., Weidema, M. A., & van der Molen, M. W. (2007). Developmental changes in oculomotor control and working-memory efficiency. *Acta Psychologica, 124*(1), 139–158. https://doi.org/10.1016/j.actpsy.2006.09.012.

Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development, 65,* 903–917.

Elison, J. T., Sasson, N. J., Turner-Brown, L. M., Dichter, G., & Bodfish, J. W. (2012). Age trends in visual exploration of social and nonsocial information in children with autism. *Research in Autism Spectrum Disorder, 6,* 842–851. https://doi.org/10.1016/j.rasd.2011.11.005.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336.

Facoetti, A., Trussardi, A. N., Ruffino, M., Lorusso, M. L., Cattaneo, C., Galli, R., Molteni, M., & Zorzi, M. (2010). Multisensory spatial attention deficits are predictive of phonological decoding skills in developmental dyslexia. *Journal of Cognitive Neuroscience, 22*(5), 1011–1025. https://doi.org/10.1162/jocn.2009.21232.

Falck-Ytter, T., Fernell, E., Hedvall, A. L., von Hofsten, C., & Gillberg, C. (2012). Gaze performance in children with autism spectrum disorder when observing communicative actions. *Journal of Autism and Developmental Disorders, 42*(10), 2236–2245. https://doi.org/10.1007/s10803-012-1471-6.

Falck-Ytter, T., & von Hofsten, C. (2011). How special is social looking in ASD: A review. *Progress in Brain Research, 189,* 209–222. https://doi.org/10.1016/B978-0-444-53884-0.00026-9.Falkmer.

Fantz, R. L. (1958). Pattern vision in young infants. *Psychological Record, 8,* 43–47.

Fantz, R. L. (1964). Visual experience in infants: Decreased attention to familiar patterns relative to novel ones. *Science, 146*(3644), 668–670.

Fischer, B., & Weber, H. (1990). Saccadic reaction times of dyslexic and age-matched normal subjects. *Perception, 19*(6), 805–818.

Fischer, B., Hartnegg, K., & Mokler, A. (2000). Dynamic visual perception of dyslexic children. *Perception, 29*(5), 523–530.

Fioravanti, F., Inchingolo, P., Pensiero, S., & Spanio, M. (1995). Saccadic eye movement conjugation in children. *Vision Research, 35*(23–24), 3217–3228. https://doi.org/10.1016/0042-6989(95)00152-5.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research, 65*(6), 591.

Foss-Feig, J. H., McGugin, R. W., Gauthier, I., Mash, L. E., Ventola, P., & Cascio, C. J. (2016). A functional neuroimaging study of fusiform response to restricted interests in children and adolescents with autism spectrum disorder. *Journal of Neurodevlopmental Disorder, 8*, 1. https://doi.org/10.1186/s11689-016-9149-6 Gillespie-Smith.

Frank, M. C., Vul, E., & Saxe, R. (2012). Measuring the development of social attention using free-viewing. *Infancy, 17*(4), 355–375. https://doi.org/10.1111/j.1532-7078.2011.00086.x.

Freeth, M., Chapman, P., Ropar, D., Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? *Evidence from Eye-Tracking*, 534–547. https://doi.org/10.1007/s10803-009-0893-2.

Fukushima, J., Hatta, T., & Fukushima, K. (2000). Development of voluntary control of saccadic eye movements. I. Age-related changes in normal children. *Brain and Development, 22*(3), 173–180. https://doi.org/10.1016/s0387-7604(00)00101-7.

Gallay, M., Baudouin, J.-Y., Durand, K., Lemoine, C., & Lécuyer, R. (2006). Qualitative differences in the exploration of upright and upside-down faces in four-month-old infants: An eye-movement study. *Child Development, 77*(4), 984–996.

Garbutt, S., Harwood, M. R., & Harris, C. M. (2006). Infant saccades are not slow. *Developmental Medicine and Child Neurology, 48*(8), 662–667. https://doi.org/10.1017/S0012162206001393.

Geurts, H. M., & Embrechts, M. (2008). Language profiles in ASD, SLI, and ADHD. *Journal of Autism and Developmental Disorders, 38*(10), 1931.

Gillespie-Lynch, K. (2013). Response to and initiation of joint attention: Overlapping but distinct roots of development in autism? *OA Autism, 1*(2), 13.

Gonzalez, C. C., Mon-Williams, M., & Burke, M. R. (2015). Children and older adults exhibit distinct sub-optimal cost-benefit functions when preparing to move their eyes and hands. *PLoS ONE, 10*(2), e0117783. https://doi.org/10.1371/journal.pone.0117783.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*(1), 20–25.

Goren, C., Sarty, M., & Wu, P. (1975). Visual following and pattern discrimination of facelike stimuli by newborn infants. *Pediatrics, 56,* 544–549.

Gould, T. D., Bastain, T. M., Israel, M. E., Hommer, W., & Castellanos, F. X. (2001). Altered performance on an ocular fixation task in attention-deficit/hyperactivity disorder. *Biological Psychiatry, 2001*(50), 633–635.

Grönqvist, H., Gredebäck, G., & von Hofsten, C. (2006). Developmental asymmetries between horizontal and vertical tracking. *Vision Research, 46*(11), 1754–1761. https://doi.org/10.1016/j.visres.2005.11.007.

Guillon, Q., Afzali, M. H., Rogé, B., & Baduel, S. (2015). The importance of networking in autism gaze analysis, 1–14. https://doi.org/10.1371/journal.pone.0141191.

Guillon, Q., Hadjikhani, N., Baduel, S., & Rogé, B. (2014). Visual social attention in autism spectrum disorder: Insights from eye tracking studies. *Neuroscience and Biobehavioral Reviews*. https://doi.org/10.1016/j.neubiorev.2014.03.013.

Hainline, L., Harris, C. M., & Krinsky, S. J. (1990). Variability of refixations in infants. *Infant Behavior and Development, 13,* 321–342.

Hainline, L., Turkel, J., Abramov, I., Lemerise, E., & Harris, C. M. (1984). Characteristics of saccades in human infants. *Vision Research, 24,* 1771–1780.

Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature Letters, 450,* 557–560. https://doi.org/10.1038/nature06288.

Hanley, M., McPhillips, M., Mulhern, G., & Riby, D. M. (2013). Spontaneous attention to faces in Asperger syndrome using ecologically valid static stimuli. *Autism International Journal Research Practise, 17*(6), 754–761. https://doi.org/10.1177/1362361312456746.

Harms, M. B., Martin, A., & Wallace, G. L. (2010). *Neuropsychology Review, 20,* 290. https://doi.org/10.1007/s11065-010-9138-6.

Harris, C. M., Jacobs, M., Shawkat, F., & Taylor, D. (1993). The development of saccadic accuracy in the first seven months. *Clinical Vision Sciences, 8*(1), 85–96.

Heaton, M. F., & Freeth, (2016). Reduced visual exploration when viewing photographic scenes in individuals with Autism Spectrum Disorder. *Journal of Abnormal Psychology, 125*(3), 399–411.

Helo, A., Pannasch, S., Sirri, L., & Rämä, P. (2014). The maturation of eye movement behavior: Scene viewing characteristics in children and adults. *Vision Research, 103,* 83–91. https://doi.org/10.1016/j.visres.2014.08.006.

Hikosaka, O. (1997). Changes and disorders in voluntary saccades during development and aging. *No To Hattatsu, 29*(3). https://doi.org/10.11251/ojjscn1969.29.213.

Holzman, P. S., Proctor, L. R., Levy, D. L., et al. (1974). Eye-tracking dysfunctions in schizophrenic patients and their relatives. *Archives of General Psychiatry, 1974*(31), 143–151.

Hood, B. M. (1995). Shifts of visual attention in the human infant: A neuroscience approach. In C. Rovee-Collier & L. Lipsitt (Eds.), *Advances in infancy research* (Vol. 9, pp. 163–216). Norwood, NJ: Ablex.

Hu, C., Wang, Q., Fu, G., Quinn, P. C., & Lee, K. (2014). Both children and adults scan faces of own and other races differently. *Vision Research, 102,* 1–10. https://doi.org/10.1016/j.visres.2014.05.010.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: A critical review. *Psychophysiology, 43,* 302–313.

Ingster-Moati, I., Vaivre-Douret, L., Bui Quoc, E., Albuisson, E., Dufier, J.-L., & Golse, B. (2009). Vertical and horizontal smooth pursuit eye movements in children: a neuro-developmental study. *European Journal of Paediatric Neurology: EJPN: Official Journal of the European Paediatric Neurology Society, 13*(4), 362–366. https://doi.org/10.1016/j.ejpn.2008.07.003.

Irving, E. L., Steinbach, M. J., Lillakas, L., Babu, R. J., & Hutchings, N. (2006). Horizontal saccade dynamics across the human life span. *Investigative Ophthalmology & Visual Science, 47*(6), 2478–2484. https://doi.org/10.1167/iovs.05-1311.

Jacobsen, L. K., Hong, W. L., Hommer, D. W., Hamburger, S. D., Castellanos, F. X., Frazier, J. A., ... & Rapoport, J. L. (1996). Smooth pursuit eye movements in childhood-onset schizophrenia: Comparison with attention-deficit hyperactivity disorder and normal controls. *Biological Psychiatry, 40*(11), 1144–1154.

Jainta, S., & Kapoula, Z. (2011). Dyslexic children are confronted with unstable binocular fixation while reading. *PLoS One, 6*(4), e18694. https://doi.org/10.1371/journal.pone.0018694.

Jankowski, J. J., & Rose, S. A. (1997). The distribution of visual attention in infants. *Journal of Experimental Child Psychology, 65,* 127–140.

Johnson, B. P., Lum, J. A., Rinehart, N. J., & Fielding, J. (2016). Ocular motor disturbances in autism spectrum disorders: Systematic review and comprehensive meta-analysis. *Neuroscience and Biobehavioral Reviews, 69,* 260–279.

Johnson, B. P., Rinehart, N. J., Papadopoulos, N., Tonge, B., Millist, L., White, O., et al. (2012). A closer look at visually guided saccades in autism and Asperger's disorder. *Frontiers in Integrative Neuroscience, 6,* 99. https://doi.org/10.3389/fnint.2012.00099.

Johnson, M. H. (1995). The inhibition of automatic saccades in early infancy. *Developmental Psychobiology, 28*(5), 281–291.

Jones, W., Carr, K., & Klin, A. (2008). Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiarty, 65*(8), 946–954. https://doi.org/10.1001/archpsyc.65.8.946.

Kapoula, Z., Ganem, R., Poncet, S., Gintautas, D., Eggert, T., Brémond-Gignac, D., & Bucci, M. P. (2009). Free exploration of painting uncovers particularly loose yoking of saccades in dyslexics. *Dyslexia, 15*(3), 243–259. https://doi.org/10.1002/dys.377.

Karalunas, S. L., Geurts, H. M., Konrad, K., Bender, S., & Nigg, J. T. (2014). Annualresearch review: Reaction time variability in ADHD and autism spectrumdisorders: Measurement and mechanisms of a proposed trans-diagnosticphenotype. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 55*(6), 685–710.

Karatekin, C. (2006). Improving antisaccade performance in adolescents with attentiondeficit/hyperactivity disorder (ADHD). *Experimental Brain Research, 174*(2), 324.

Karatekin, C. (2007). Eye tracking studies of normative and atypical development. *Developmental Review, 27*(3), 283–348.

Karatekin, C., & Asarnow, R. F. (1999). Exploratory eye movements to pictures in childhood-onset schizophrenia and attention-deficit/hyperactivity disorder (ADHD). *Journal of Abnormal Child Psychology, 27,* 35–49.

Karatekin, C., Bingham, C., & White, T. (2009). Oculomotor and pupillometric indices of pro-and antisaccade performance in youth-onset psychosis and attention deficit/hyperactivity disorder. *Schizophrenia Bulletin, 36*(6), 1167–1186.

Kato, M., & Konishi, Y. (2013). Where and how infants look: The development of scan paths and fixations in face perception. *Infant Behavior & Development, 36*(1), 32–41. https://doi.org/10.1016/j.infbeh.2012.10.005.

Katsanis, J., Iacono, W. G., & Harris, M. (1998). Development of oculomotor functioning in preadolescence, adolescence, and adulthood. *Psychophysiology, 35*(1), 64–72. https://doi.org/10.1111/1469-8986.3510064.

Kavšek, M. (2013). The comparator model of infant visual habituation and dishabituation: Recent insights. *Developmental Psychobiology, 55*(8), 793–808. https://doi.org/10.1002/dev.21081.

Kavšek, M., & Bornstein, M. H. (2010). Visual habituation and dishabituation in preterm infants: A review and meta-analysis. *Research in Developmental Disabilities, 31*(5), 951–975. https://doi.org/10.1016/j.ridd.2010.04.016.

Kellman, P. J., & Arterberry, M. E. (2006). Infant visual perception. In D. Kuhn, R. Siegler, W. Damon, & R. Lerner (Eds.), *Handbook of child psychology. Vol 2: Cognition, perception, and language* (6th ed., pp. 109–160). Hoboken, NJ: Wiley.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy: Evidence of perceptual narrowing. *Psychological Science, 18*(12), 1084–1089. https://doi.org/10.1111/j.1467-9280.2007.02029.x.

Kelly, D. J., Walker, R., & Norbury, C. F. (2013). Deficits in volitional oculomotor control align with language status in autism spectrum disorders. *Developmental Science, 16*(1), 56–66.

Klein, C., & Berg, P. (2001). Four-week test-retest stability of individual differences in the saccadic CNV, two saccadic task parameters, and selected neuropsychological test. *Psychophysiology, 38,* 704–711.

Klein, C., & Feige, B. (2005). An independent components analysis (ICA) approach to the study of developmental differences in the saccadic contingent negative variation. *Biological Psychology, 70*(2), 105–114. https://doi.org/10.1016/j.biopsycho.2004.11.016.

Klein, C., & Fischer, B. (2005a). Developmental fractionation and differential discrimination of the anti-saccadic direction error. *Experimental Brain Research, 165,* 132–138.

Klein, C., & Fischer, B. (2005b). Instrumental and test-retest reliability of saccadic measures. *Biological Psychology, 68,* 201–213.

Klein, C. H., Raschke, A., & Brandenbusch, A. (2003). Development of pro–and antisaccades in children with attention–deficit hyperactivity disorder (ADHD) and healthy controls. *Psychophysiology, 40*(1), 17–28.

Klein, C., Rauh, R., & Biscaldi, M. (2010). Cognitive correlates of anti-saccade task performance. *Experimental Brain Research, 203*, 759–764.

Klein, C., Rauh, R., & Biscaldi, M. (2011). Patterns of change in ocular motor development. *Experimental Brain Research, 210*(1), 33–44. https://doi.org/10.1007/s00221-011-2601-7.

Klein, C., Fischer, B., Hartnegg, K., Heiss, W. H., & Roth, M. (2000). Optomotor and neuropsychological performance in old age. *Experimental Brain Research, 135*(2), 141–154. https://doi.org/10.1007/s002210000506.

Klein, C., Foerster, F., Hartnegg, K., & Fischer, B. (2005). Lifespan development of pro- and anti-saccades: Multiple regression models for point estimates. *Developmental Brain Research, 116,* 113–123.

Klin, A., Jones, W., Schultz, R., Volkmar, F., & Cohen, D. (2002). Visual fixation patterns during viewing of naturalistic social situations as predictors of social competence in individuals with autism. *Archives of General Psychiatry, 59,* 809–816. https://doi.org/10.1001/archpsyc.59.9.809.

Kremenitzer, J. P., Vaughan, H. G., Kurtzberg, D., & Dowling, K. (1979). Smooth-pursuit eye movements in the newborn infant. *Child Development, 50,* 442–448.

Lengyel, D., Weinacht, S., Charlier, J., & Gottlob, I. (1998). The development of visual pursuit during the first months of life. *Graefe's Archive for Clinical and Experimental Ophthalmology, 236*(6), 440–444. https://doi.org/10.1007/s004170050103.

Klein, C., Wendling, K., Huettner, P., Ruder, H., & Peper, M. (2006). Intra-Subjectvariability in attention-deficit hyperacivity disorder. *Biological Psychiatry, 8*(5), 906–913.

Levy, D. L., Bowman, E. A., Abel, L., Krastoshecsky, O., Krause, V., & Mendell, N. R. (2008). Does performance on the standard antisaccade task meet the co-familiarity criterion for an endophenotype? *Brain and Cognition, 68,* 462–476.

Liu, S., Quinn, P. C., Wheeler, A., Xiao, N., Ge, L., & Lee, K. (2011). Similarity and difference in the processing of same- and other-race faces as revealed by eye tracking in 4- to 9-month-olds. *Journal of Experimental Child Psychology, 108*(1), 180–189. https://doi.org/10.1016/j.jecp.2010.06.008.

Loe, I. M., Feldman, H. M., Yasui, E., & Luna, B. (2009). Oculomotor performance identifies underlying cognitive deficits in attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*(4), 431–440.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development, 75*(5), 1357–1372. https://doi.org/10.1111/j.1467-8624.2004.00745.x.

Luna, B., Thulborn, K. R., Munoz, D. P., Merriam, E. P., Garver, K. E., Minshew, N. J., et al. (2001). Maturation of widely distributed brain function subserves cognitive development. *Neuroimage, 13*(5), 786–793. https://doi.org/10.1006/nimg.2000.0743.

Luna, B., Velanova, K., & Geier, C. F. (2008). Development of eye-movement control. *Brain and Cognition, 68*(3), 293–308.

Macchi Cassia, V., Turati, C., & Simion, F. (2004). Can a nonspecific bias toward top-heavy patterns explain newborns' face preference. *Psychological Science, 15*(6), 379–383.

McGrath, L. M., Pennington, B. F., Shanahan, M. A., Santerre-Lemmon, L. E., Barnard, H. D., Willcutt, E. G., Defries, J. C., & Olson, R. K. (2011). A multiple deficit model of reading disability and attention-deficit/hyperactivity disorder: Searching for shared cognitive deficits. *Journal of Child Psychology and Psychiatry, 52*(5), 547–557. https://doi.org/10.1111/j.1469-7610.2010.02346.x.

Mahone, E. M., Mostofsky, S. H., Lasker, A. G., Zee, D., & Denckla, M. B. (2009). Oculomotor anomalies in attention-deficit/hyperactivity disorder: Evidence for deficits in response preparation and inhibition. *Journal of the American Academy of Child and Adolescent Psychiatry, 48,* 749–756. https://doi.org/10.1097/CHI.0b013e3181a565f1.

Matsuda, T., Matsuura, M., Ohkubo, T., Ohkubo, H., Matsushima, E., Inoue, K., et al. (2004). Functional MRI mapping of brain activation during visually guided saccades and antisaccades: Cortical and subcortical networks. *Psychiatry Research, 131*(2), 147–155. https://doi.org/10.1016/j.pscychresns.2003.12.007.

McCall, R. B., & Kagan, J. (1970). Individual differences in the infant's distribution of attention to stimulus discrepancy. *Developmental Psychology, 2*(1), 90–98.

Morita, K., Miura, K., Fujimoto, M., Yamamori, H., Yasuda, Y., Iwase, M., et al. (2017). Eye movement as a biomarker of schizophrenia: Using an integrated eye movement score. *Psychiatry and Clinical Neurosciences, 71,* 104–114.

Moschner, C., & Baloh, R. W. (1994). Age-related changes in visual tracking. *Journal of Gerontology, 49*(5), M235–M238.

Mostofsky, S. H., Lasker, A. G., Singer, H. S., Denckla, M. B., & Zee, D. S. (2001a). Oculomotor abnormalities in boys with Tourette syndrome with and without ADHD. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(12), 1464–1472.

Mostofsky, S. H., Lasker, A. G., Cutting, L. E., Denckla, M. B., & Zee, D. S. (2001b). Oculomotor abnormalities in attention deficit hyperactivity disorder: A preliminary study. *Neurology, 57*(3), 423–430.

Munoz, D. P., Armstrong, I. T., Hampton, K. A., & Moore, K. D. (2003). Altered control of visual fixation and saccadic eye movements in attention-deficit hyperactivity disorder. *Journal of Neurophysiology, 90*(1), 503–514.

Munoz, D. P., Broughton, J. R., Goldring, J. E., & Armstrong, I. T. (1998). Age-related performance of human subjects on saccadic eye movement tasks. *Experimental Brain Research, 121*(4), 391–400. https://doi.org/10.1007/s002210050473.

Munoz, D. P., Hampton, K. A., Moore, K. D., & Goldring, J. E. (1999). Control of purposive saccadic eye movements and visual fixation in children with attention-deficit hyperactivity disorder. In W. Becker, H. Deubel, & T. Mergner (Eds.), *Current oculomotor research. Physiological and psychologicial aspects*. New York: Kluwer Academic/Plenum Publishers.

Muszkat, M., de Mello, C. B., Muñoz, P. D. O. L., Lucci, T. K., David, V. F., de Oliveira Siqueira, J., & Otta, E. (2015). Face scanning in autism spectrum disorder and attention deficit/hyperactivity disorder: Human versus dog face scanning. *Frontiers in Psychiatry, 6*.

Nakagawa, A., & Sukigara, M. (2007). Infant eye and head movements toward the side opposite the cue in the anti-saccade paradigm. *Behavioral and Brain Functions, 3*.

Nayar, K., Franchak, J., Adolph, K., & Kiorpes, L. (2015). From local to global processing: The development of illusory contour perception. *Journal of Experimental Child Psychology, 131,* 38–55. https://doi.org/10.1016/j.jecp.2014.11.001.

Noiret, N., Vigneron, B., Diogo, M., Vandel, P., & Laurent, E. (2016). Saccadic eye movements: What do they tell us about aging cognition? *Neuropsychology, Development, and Cognition.*

Section B, Aging, Neuropsychology and Cognition, 1–25. https://doi.org/10.1080/13825585.2016.1237613.

Norcia, A. M., & Gerhard, H. E. (2015). Development of three-dimensional perception in human infants. *Annual Review of Vision Science, 1,* 569–594. https://doi.org/10.1146/annurev-vision-082114-035835.

Noris, B., Nadel, J., Barker, M., Hadjikhani, N., & Billard, A. (2012). Investigating gaze of children with ASD in naturalistic settings. *PLOS ONE, 7*.

Oakes, L. M., & Ellis, A. E. (2013). An eye-tracking investigation of developmental changes in infants' exploration of upright and inverted human faces. *Infancy, 18*(1), 134–148. https://doi.org/10.1111/j.1532-7078.2011.00107.x.

Olincy, A., Ross, R., Youngd, D., & Freedman, R. (1997). Age diminishes performance on an antisaccade eye movement task. *Neurobiology of Aging, 18*(5), 483–489. https://doi.org/10.1016/S0197-4580(97)00109-7.

Olson, R. K., Kliegl, R., Davidson, B. J. (1983). Dyslexic and normal readers' eye movements. *Journal of Experimental Psychology: Human Perception and Performance, 9*(5), 816–825.

Papagiannopoulou, E. A., Chitty, K. M., Hermens, D. F., Hickie, I. B., & Lagopoulos, J. (2014). A systematic review and meta-analysis of eye-tracking studies in children with autism spectrum disorders. *Social Neuroscience*, 1–23. https://doi.org/10.1080/17470919.2014.934966 PMID: 24988218 23.

Paus, T. (1989). The development of sustained attention in children might be related to the maturation of frontal cortical functions. *Acta Neuorbiologiae Experimentalis, 49*(1), 51–55.

Paus, T., Babenko, V., & Radil, T. (1990). Development of an ability to maintain verbally instructed central gaze fixation studied in 8- to lo-year-old children. *International Journal of Physiology, 10,* 53–61. https://doi.org/10.1016/0167-8760(90)90045-F.

Pavlidis, G. T. (1981). Do eye movements hold the key to dyslexia? *Neuropsychologia, 19*(1), 57–64.

Peltsch, A., Hemraj, A., Garcia, A., & Munoz, D. P. (2011). Age-related trends in saccade characteristics among the elderly. *Neurobiology of Aging, 32*(4), 669–679. https://doi.org/10.1016/j.neurobiolaging.2009.04.001.

Peterson, R. L., & Pennington, B. F. (2015). Developmental dyslexia. *Annual Review of Clinical Psychology, 11*, 283–307.

Petruzzelli, M. G., Margari, L., Bosco, A., Craig, F., Palumbi, R., & Margari, F. (2018). Early onset first episode psychosis: dimensional structure of symptoms, clinical subtypes and related neurodevelopmental markers. *European Child and Adolescent Psychiatry, 27,* 171–179.

Pieh, C., Proudlock, F., & Gottlob, I. (2012). Smooth pursuit in infants: Maturation and the influence of stimulation. *The British journal of ophthalmology, 96*(1), 73–77. https://doi.org/10.1136/bjo.2010.191726.

Pierce, K., Conant, D., Hazin, R., & Stoner, R. (2016). *For Autism, 68*(1), 101–109. https://doi.org/10.1001/archgenpsychiatry.2010.113.Preference.

Pierrot-Deseilligny, C., Müri, R. M., Ploner, C. J., Gaymard, B., Demeret, S., & Rivaud-Pechoux, S. (2003). Decisional role of the dorsolateral prefrontal cortex in ocular motor behaviour. *Brain, 126*(6), 1460–1473. https://doi.org/10.1093/brain/awg148.

Pietrzak, R. H., Mollica, C. M., Maruff, P., & Snyder, P. J. (2006). Cognitive effects of immediate-release methylphenidate in children with attention-deficit/hyperactivity disorder. *Neuroscience and Biobehavioral Reviews, 30,* 1225–1245.

Posner, M. I., & Rothbart, M. K. (2007). *Educating the human brain*. Washington, DC: American Psychological Association.

Quinn, P. C., & Bhatt, R. S. (2012). Grouping by form in young infants: Only relevant variability promotes perceptual learning. *Perception, 41*(12), 1468–1476. https://doi.org/10.1068/p7306.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction? *Current Opinion in Neurobiology, 13*(2), 212–218, Review.

Ramus, F. (2014). Neuroimaging sheds new light on the phonological deficit in dyslexia. *Trends of Cognitive Science, 18*(6), 274–275. https://doi.org/10.1016/j.tics.2014.01.009.

Ramus, F., & Szenkovits, G. (2008). What phonological deficit? *Quarterly Journal of Experimental Psychology, 61*(1), 129–141, Review.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin, 124*(3), 372–422, Review.

Regal, D. M., Ashmead, D. H., & Salapatek, P. (1983). The coordination of eye and head movements during early infancy: A selective review. *Behavioural Brain Research, 10,* 125–132.

Reinholdt-Dunne, M. L., Mogg, K., Benson, V., Bradley, B. P., Hardin, M. G., Liversedge, S. P., ... & Ernst, M. (2012). Anxiety and selective attention to angry faces: An antisaccade study. *Journal of Cognitive Psychology, 24*(1), 54–65.

Remschmidt, H. (2002). Early-onset schizophrenia as a progressive-deteriorating developmental disorder: Evidence from child psychiatry. *Journal of Neural Transmission, 109,* 101–117.

Reynolds, G. D. (2015). Infant visual attention and object recognition. *Behavioural Brain Research, 285,* 34–43. https://doi.org/10.1016/j.bbr.2015.01.015.

Reynolds, G. D., Courage, M. L., & Richards, J. E. (2013). The development of attention. In D. Reisberg (Ed.), *Oxford library of psychology. The oxford handbook of cognitive psychology* (pp. 1000–1013). New York, NY: Oxford University Press.

Riby, D. M., Hancock, P. J. B., Jones, N., & Hanley, M. (2013). Spontaneous and cued gaze-following in autism and Williams syndrome Spontaneous and cued gaze-following in autism and Williams syndrome.

Riby, D. M., & Hancock, P. J. B. (2008). Viewing it differently: Social scene perception in Williams syndrome and autism. *Neuropsychologia, 46*(11), 2855–2860. https://doi.org/10.1016/j.neuropsychologia.2008.05.003.

Richards, J. E. (2004). The development of sustained attention in infants. In M. Posner (Ed.), *Cognitive neuroscience of attention* (pp. 342–356). New York: Guilford Press.

Richards, J. E., & Holley, F. B. (1999). Infant attention and the development of smooth pursuit tracking. *Developmental Psychology, 35*(3), 856–867.

Rommelse, N. N., Van der Stigchel, S., & Sergeant, J. A. (2008). A review on eye movement studies in childhood and adolescent psychiatry. *Brain and Cognition, 68*(3), 391–414.

Rommelse, N. N., Geurts, H. M., Franke, B., Buitelaar, J. K., & Hartman, C. A. (2011). A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience and Biobehavioral Reviews, 35*(6), 1363–1396.

Rosander, K. (2007). Visual tracking and its relationship to cortical development. *Progress in Brain Research, 164,* 105–122. https://doi.org/10.1016/S0079-6123(07)64006-0.

Rosander, K., & von Hofsten, C. (2002). Development of gaze tracking of small and large objects. *Experimental Brain Research, 146*(2), 257–264. https://doi.org/10.1007/s00221-002-1161-2.

Rosander, K., & von Hofsten, C. (2004). Infants' emerging ability to represent occluded object motion. *Cognition, 91*(1), 1–22. https://doi.org/10.1016/S0010-0277(03)00166-5.

Rose, S. A., Feldman, J. F., Jankowski, J. J., & van Rossem, R. (2008). A cognitive cascade in infancy: Pathways from prematurity to later mental development. *Intelligence, 36*(4), 367–378.

Rosenhall, U., Johansson, E., & Gillberg, C. (1988). Oculomotor findings in autistic children. *The Journal of Laryngology & Otology, 102*(5), 435–439.

Ross, R. G., Heinlein, S., Zerbe, G. O., & Radant, A. (2005). Saccadic eye movement task identifies cognitive deficits in children with schizophrenia, but not in unaffected child relatives. *Journal of Child Psychology and Psychiatry, 46*(12), 1354–1362.

Ross, R. G., Olincy, A., Harris, J. G., Sullivan, B., & Radant, A. (2000). Smooth pursuit eye movements in schizophrenia and attentional dysfunction: Adults with schizophrenia, ADHD, and a normal comparison group. *Biological Psychiatry, 48*(3), 197–203.

Ross, R. G., Olincy, A., Harris, J. G., Radant, A., Adler, L. E., Compagnon, N., et al. (1999). The effects of age on a smooth pursuit tracking task in adults with schizophrenia and normal subjects. *Biological Psychiatry, 46*(3), 383–391. https://doi.org/10.1016/S0006-3223(98)00369-2.

Roucoux, A., Culee, C., & Roucoux, M. (1983). Development of fixation and pursuit eye movements in human infants. *Behavioural Brain Research, 10,* 133–139.

Ruff, H. A., & Rothbart, M. K. (1996). *Attention in early development: Themes and variations*. New York: Oxford University Press.

Saayman, G., Ames, E. W., & Moffett, A. (1964). Response to novelty as an indicator of visual discrimination in the human infant. *Journal of Experimental Child Psychology, 1*(2), 189–198.

Salapatek, P. (1968). Visual scanning of geometric figures by the human newborn. *Journal of Comparative and Physiological Psychology, 66*(2), 247–258.

Salman, M. S., Sharpe, J. A., Eizenman, M., Lillakas, L., Westall, C., To, T., et al. (2006). Saccades in children. *Vision Research, 46*(8–9), 1432–1439. https://doi.org/10.1016/j.visres.2005.06.011.

Salman, M. S., Sharpe, J. A., Lillakas, L., & Steinbach, M. J. (2008). Square wave jerks in children and adolescents. *Pediatric Neurology, 38*(1), 16–19. https://doi.org/10.1016/j.pediatrneurol.2007.08.011.

Sasson, N. J., Dichter, G. S., & Bodfish, J. W. (2012). Affective responses by adults with autism are reduced to social images but elevated to images related to circumscribed interests. *PLoS ONE, 7,* e42457.

Sasson, N. J., Turner-Brown, L. M., Holtzclaw, T. N., Lam, K. S., & Bodfish, J. W. (2008). Children with autism demonstrate circumscribed attention during passive viewing of complex social and nonsocial picture arrays. *Autism Res., 1,* 31–42. https://doi.org/10.1002/aur.4.

Sasson, N. J., & Touchstone, E. W. (2013). Visual attention to competing social and object images by preschool children with autism spectrum disorder. Journal Autism Developmental Disorder (Ci). http://dx.doi.org/10.1007/s10803-013-1910-z. Sasson.

Scerif, G., Karmiloff-Smith, A., Campos, R., Elsabbagh, M., Driver, J., & Cornish, K. (2005). To look or not to look? Typical and atypical development of oculomotor control. *Journal of Cognitive Neuroscience, 17*(4), 591–604.

Scharre, J. E., & Creedon, M. P. (1992). Assessment of visual function in autistic children. *Optometry and Vision Science, 69*(6), 433–439.

Scherf, K. S., Behrmann, M., Kimchi, R., & Luna, B. (2009). Emergence of global shape processing continues through adolescence. *Child Development, 80*(1), 162–177. https://doi.org/10.1111/j.1467-8624.2008.01252.x.

Schmitt, L. M., Cook, E. H., Sweeney, J. A., & Mosconi, W. W. (2014). Saccadic abnormalities in autism spectrum disorder indicate dysfunctions in cerebellum and brainstem. *Molecular Autism, 5,* 47.

Schwarzer, G., Freitag, C., Buckel, R., & Lofruthe, A. (2013). Crawling is associated with mental rotation ability by 9-month-old infants. *Infancy, 18*(3), 432–441. https://doi.org/10.1111/j.1532-7078.2012.00132.x.

Schwarzer, G., Huber, S., & Dümmler, T. (2005). Gaze behavior in analytical and holistic face processing. *Memory & Cognition, 33*(2), 344–354. https://doi.org/10.3758/BF03195322.

Seassau, M., Gérard, C. L., Bui-Quoc, E., & Bucci, M. P. (2014). Binocular saccade coordination in reading and visual search: A developmental study in typical reader and dyslexic children. *Frontiers in Integrative Neuroscience, 8*, 85. https://doi.org/10.3389/fnint.2014.00085.

Smith, N. A., Gibilisco, C. R., Meisinger, R. E., & Hankey, M. (2013). Asymmetry in infants' selective attention to facial features during visual processing of infant-directed speech. *Frontiers in Psychology, 4,* Article 601. https://doi.org/10.3389/fpsyg.2013.00601.

Smyrnis, N. (2008). Metric issues in the study of eye movements in psychiatry. *Brain and Cognition, 68,* 341–358.

Sokolov, E. N. (1963). *Perception and the conditioned reflex*. Oxford: Pergamon Press.

South, M., Ozonoff, S., & McMahon, W. M. (2005). Repetitive behavior profiles in asperger syndrome and high-functioning autism. *Journal Autism Development Disorder, 35*, 145–158. https://doi.org/10.1007/s10803-004-1992-8 Sterling.

Stanley-Cary, C., Rinehart, N., Tonge, B., White, O., & Fielding, J. (2011). Greater disruption to control of voluntary saccades in autistic disorder than Asperger's disorder: Evidence for greater cerebellar involvement in autism? *The Cerebellum, 10*(1), 70–80.

Sweeney, J. A., Rosano, C., Berman, R. A., & Luna, B. (2001). Inhibitory control of attention declines more than working memory during normal aging. *Neurobiology of Aging, 22*(1), 39–47. https://doi.org/10.1016/S0197-4580(00)00175-5.

Tajik-Parvinchi, D. J., Lillakas, L., Irving, E., & Steinbach, M. J. (2003). Children's pursuit eye movements: A developmental study. *Vision Research, 43,* 77–84.

Takarae, Y., Luna, B., Minshew, N. J., & Sweeney, J. A. (2008). Patterns of visual sensory and sensorimotor abnormalities in autism vary in relation to history of early language delay. *Journal of the International Neuropsychological Society, 14*(6), 980–989.

Takarae, Y., Minshew, N. J., Luna, B., & Sweeney, J. A. (2004a). Oculomotor abnormalities parallel cerebellar histopathology in autism. *Journal of Neurology, Neurosurgery and Psychiatry, 75*(9), 1359–1361.

Takarae, Y., Minshew, N. J., Luna, B., Krisky, C. M., & Sweeney, J. A. (2004b). Pursuit eye movement deficits in autism. *Brain, 127*(12), 2584–2594.

Tallal, P. (2000). The science of literacy: From the laboratory to the classroom. *Proceedings of the National Academy of Sciences, 97*(6), 2402–2404.

Tedeschi, G., Di Constanzo, A., Allocca, S., Quattrone, A., Casucci, G., Russo, L., et al. (1989). Age-dependent changes in visually guided saccadic eye movements. *Functional Neurology, 4*(4), 363–367.

Tenenbaum, E. J., Shah, R. J., Sobel, D. M., Malle, B. F., & Morgan, J. L. (2013). Increased focus on the mouth among infants in the first year of life: A longitudinal eye-tracking study. *Infancy, 18*(4), 534–553. https://doi.org/10.1111/j.1532-7078.2012.00135.x.

Tottenham, N., Hertzig, M. E., Gillespie-Lynch, K., Gilhooly, T., Millner, A. J., & Casey, B. J. (2014). Elevated amygdala response to faces and gaze aversion in autism spectrum disorder. *Social Cognitive and Affective Neuroscience, 9,* 106–117. https://doi.org/10.1093/scan/nst050.

Turati, C. (2004). Why faces are not special to newborns: An alternative account of the face preference. *Current Directions in Psychological Science, 13*(1), 5–8.

Turati, C., Sangrigoli, S., Ruel, J., de Schonen, S., & de Schonen, S. (2004). Evidence of the face inversion effect in 4-month-old infants. *Infancy, 6*(2), 275–297.

Unruh, K. E., Sasson, N. J., Shafer, R. L., Whitten, A., Miller, S. J., Turner-Brown, L., et al. (2016). Social orienting and attention is influenced by the presence of competing nonsocial information in adolescents with autism. *Frontiers in Neuroscience, 10,* 1–12. https://doi.org/10.3389/fnins.2016.00586.

van der Geest, J. N., Kemner, C., Camfferman, G., Verbaten, M. N., & van Engeland, H. (2002). Looking at images with human figures: Comparison between autistic and normal children. *Journal Autism Development Disorder, 32,* 69–75.

Valenza, E., & Bulf, H. (2007). The role of kinetic information in newborns' perception of illusory contours. *Developmental Science, 10*(4), 492–501.

Vellutino, F. R. (1987). Dyslexia. *Scientific American, 256*(3), 34–41.

Vidyasagar, T. R. (2013). Reading into neuronal oscillations in the visual system: Implications for developmental dyslexia. *Frontiers in Human Neuroscience, 27*(7), 811. https://doi.org/10.3389/fnhum.2013.00811.

von Hofsten, C., Kochukhova, O., & Rosander, K. (2007). Predictive tracking over occlusions by 4-month-old infants. *Developmental Science, 10*(5), 625–640. https://doi.org/10.1111/j.1467-7687.2007.00604.x.

von Hofsten, C., & Rosander, K. (1997). Development of smooth pursuit tracking in young infants. *Vision Research, 37*(13), 1799–1810.

Vyas, N. S., Kumra, S., & Puri, B. K. (2010). What insights can we gain from studying early-onset schizophrenia? The neurodevelopmental pathway and beyond. *Expert Review of Neurotherapeutics, 10*(8), 1243–1247. https://doi.org/10.1586/ern.10.109.

White, T., Mous, S., & Karatekin, C. (2014). Memory-guided saccades in youth-onset psychosis and attention deficit hyperactivity disorder (ADHD). *Early Intervention in Psychiatry, 8*(3), 229–239.

Wilson, C. E., Brock, J., & Palermo, R., (2010). Attention to social stimuli and facial identity recognition skills in autism spectrum disorder. *Journal Intellectual Disability Research JIDR, 54*(12), 1104–1115, http://dx.doi.org/10.1111/j.1365-2788.2010.01340.x. Wolf.

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749–750.

Yee, C. M., Nuechterlein, K. H., & Dawson, M. E. (1998). A longitudinal analysis of eye tracking dysfunction and attention in recent-onset schizophrenia. *Psychophysiology, 35*(4), 443–451. https://doi.org/10.1111/1469-8986.3540443.

Ygge, J., Aring, E., Han, Y., Bolzani, R., & Hellström, A. (2005). Fixation stability in normal children. *Annals of the New York Academy of Sciences, 1039,* 480–483. https://doi.org/10.1196/annals.1325.049.

Ygge, J., Lennerstrand, G., Rydberg, A., Wijecoon, S., & Pettersson, B. M. (1993). Oculomotor functions in a Swedish population of dyslexic and normally reading children. *Acta Ophthalmologica, 71*(1), 10–21.

Yi, L., Fan, Y., Quinn, P. C., Feng, C., Huang, D., Li, J., et al. (2013). Abnormality in face scanning by children with autism spectrum disorder is limited to the eye region: Evidence from multimethod analyses of eye tracking data. *Journal of Vision, 13,* 1–13. https://doi.org/10.1167/13.10.5. PMID: 23929830.

Yi, L., Feng, C., Quinn, P. C., Ding, H., Li, J., Liu, Y., et al. (2014). Do individuals with and without autism spectrum disorder scan faces differently? A new multi-method look at an existing controversy. *Autism Research, 7,* 72–83. https://doi.org/10.1002/aur.1340. PMID: 24124133.

Chapter 16
Eye Movements in Psychiatry

Nikolaos Smyrnis, Isabelle Amado, Marie-Odile Krebs and John A. Sweeney

Contents

N. Smyrnis (✉)
Psychiatry Department, National and Kapodistrian University of Athens,
Medical School, Eginition Hospital, Athens, Greece
e-mail: smyrnis@med.uoa.gr

Laboratory of Cognitive Neuroscience and Sensorimotor Control, University Mental Health,
Neurosciences and Precision Medicine Research Institute “COSTAS STEFANIS”,
Athens, Greece

I. Amado · M.-O. Krebs
Department of Cognitive Remediation and Rehabilitation, Psychiatry,
Centre Hospitalier Sainte Anne, Inserm U894, Paris, France

J. A. Sweeney
Department of Psychiatry and Behavioral Neuroscience,
University of Cincinnati, Cincinnati, USA

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_16

Abstract In the last half century, a large body of research literature has concentrated on oculomotor function in psychiatric syndromes and especially schizophrenia spectrum disorders. The emerging picture from these studies is that specific deficits measured in specific oculomotor function tasks such as smooth eye pursuit, antisaccades and memory saccades are evident in patients with psychosis and probably to a lesser degree in their first degree relatives. Such oculomotor function deficits have helped develop a better understanding of the genetic and neurobiological substrate of psychiatric syndromes, bearing in mind that these functions have a complex genetic and neurobiological substrate that remains to be fully understood. Some of these functions such as visually guided and predictive saccades are sensitive to medication status and the natural course of psychiatric syndromes. This fact could potentially lead to the future development of oculomotor biological markers for treatment response and outcome predictions in these syndromes.

16.1 Introduction and Learning Objectives

Eye movement paradigms such as smooth pursuit eye movements (SPEM) and saccades have been extensively used to study psychiatric disorders in the last 50 years. The aims of this large body of scientific endeavor have been to address the utility of these paradigms to be used (a) as biomarkers for dissociating complex psychiatric syndromes, their symptomatology, their course and their response to treatment, (b) as endophenotypes in the study of the genetic factors underlying the aetiopathology of psychiatric disorders and finally (c) as probes for the study of the neuronal basis of psychiatric disorders.

These aims remain a major challenge in the field of psychiatry. In contrast to all other medical specialties modern psychiatry (at the time this chapter was written) lacks reliable biological markers (quantitative and objectively measured parameters) that can be used in the diagnosis and differential diagnosis of psychiatric syndromes, in the prognosis of the course of these disorders as well as in the selection and effectiveness of different treatments. Although we know that genetic factors play a major role in the aetiopathology of psychiatric syndromes and especially psychotic spectrum disorders we are still struggling with diagnostic uncertainty in our effort to bridge the gap between genes and clinical manifestations of these syndromes. The use of specific endophenotypes such as oculomotor function tests that can bridge this gap is a promising tool in genetic studies of these disorders. Finally, albeit the impressive methodological advances in all fields of neuroscience we still don't understand the neurobiological substrate underlying psychiatric disorders. In this effort the use of oculomotor function deficits in psychiatric disorders combined with our increasing knowledge on the neural substrate of these functions could shed light into the complex neurobiology of these disorders.

In this chapter, we will introduce the general approach of studying eye movements in psychiatry and present selected, major findings from this large literature using representative examples. The focus is on schizophrenia, given the substantial body of work on this psychiatric disorder, but other diagnostic categories are also considered briefly. For reasons of conciseness we also focus on the most widely studied and most informative paradigms in this context, i.e. saccade and smooth pursuit paradigms. For studies of fixations and scanpaths in psychiatric disorders, we refer the reader to recent overviews of that work published elsewhere (Armstrong, Olatunji, 2012; Beedie, 2011; Toh et al., 2011).

The chapter is organized in sections discussing different eye movement paradigms (SPEM, saccades). Each section is further divided into subsections that address the following basic sets of questions: (a) What are the differences between patients and controls in the particular eye movement paradigm and what is the sensitivity of these differences in dissociating patients from controls? (b) Are these differences specific to different psychiatric syndromes? (c) Are these differences related to specific psychopathology states and medication status? (d) Are the differences observed between patients and controls in the particular eye movement paradigm hereditary and what is the nature of the genetic factors influencing these differences? (e) What is the underlying neurobiological substrate of the differences between patients and controls in the particular eye movement paradigm?

The learning objectives of this chapter are as follows: first, having read the chapter you will have familiarized yourself with the scientific rationale underlying the study of eye movements in patients with psychiatric disorders. Moreover, you will have gained an overview of the key historical developments in this work and an insight into the current findings. More generally, having read this chapter you will be aware of the key methodological issues in the study of eye movements in psychiatric patient populations and you will be able to critically evaluate current developments in this field of research.

16.2 Historical Annotations

The history of the study of eye movements in psychiatric disorders parallels the development of fundamental theoretical conceptions that have shaped the history of modern psychiatry (see "A History of Psychiatry: from the era of asylum to the age of Prozac" by Shorter, 1997). At the beginning of the 20th century, Kraepelin and others laid the foundations of clinical psychiatry by defining the psychopathology of the major psychiatric syndromes (Kraepelin et al. 1919). The prevailing theoretical conception was that these syndromes were the result of brain lesions yet to be identified. This original biological approach to psychiatry was influenced by the paradigm of neurology, and led to the study of biological markers in an effort to characterize the underlying neurological deficits.

In 1908, Diefendorf and Dodge recorded, for the first time, saccadic and smooth pursuit eye movements in a sample of institutionalized mentally ill patients using "*a photochronograph*". They observed deficits in patients with "*dementia praecox*", the clinical term for what would now be referred to as schizophrenia, relative to healthy controls in smooth pursuit but not in saccades.

In the years following this discovery, the prevailing theoretical dogma in psychiatry shifted from a biological to a psychological perspective, and the study of eye movements was temporarily all but abandoned [with the exception of studies by Couch and Fox (1934), Silverman and Gaarder (1967)]. It was the accidental discovery of antipsychotic medications in the 1950s that led to a new shift back towards biological models of major mental illnesses in psychiatry. This shift was greatly accelerated by developments in brain imaging, which conclusively demonstrated significant alterations in both brain anatomy and physiology in schizophrenia patients (Wang et al., 2015). Brain imaging research also differentiated brain regions involved in the control of pursuit and saccadic eye movements (Rosano et al., 2002), which then was used to guide hypothesis-driven studies clarifying the neural basis of eye movement abnormalities in psychiatric disorders (Lencer et al., 2011).

In a 1973 landmark study, Holzman and his coworkers recorded smooth pursuit eye movements in schizophrenia patients and healthy controls using electrooculography (EOG) and reported pronounced deficits in their performance. In a subsequent seminal study, Holzman et al. (1974) reported similar deficits in first-degree relatives of patients with schizophrenia. This finding introduced the eye movement paradigm as a tool of modern biological psychiatric genetics research.

Another milestone was the first report of pronounced antisaccade impairments in schizophrenia patients by Fukushima et al. (1988). That same study also related these deficits to abnormalities in frontal lobe structure, firmly rooting the study of oculomotor deficits in psychiatry in the neurological tradition. A related step forward was work by Park showing disturbances in memory guided saccades that also were attributed to prefrontal cortex alterations and seen in unaffected family members of patients (Park, Holzman, & Goldman-Rakic, 1995).

Since these early studies, eye movement paradigms have been widely used to study cognitive function in patients with schizophrenia as well as all other major psychiatric syndromes. Comprehensive recent reviews were published in 2008 in a special issue of Brain and Cognition on “A hundred years of eye movement research in psychiatry” (Klein & Ettinger, 2008) that celebrated the centenary of the Diefendorf and Dodge (1908) discovery (see, e.g., Calkins, Iacono, & Ones, 2008; Gooding & Basso, 2008; Levy et al., 2008; O’Driscoll & Callahan, 2008; Rommelse et al., 2008). These studies indicate a high prevalence of eye movement alterations in schizophrenia and other psychotic disorders, that these alterations run in families of affected patients, and that despite well known effects of Central Nervous System (CNS) active drugs on eye movements (Reilly, Lencer, Bishop, Keedy, & Sweeney, 2008) psychiatric drug therapies minimally affect the magnitude of the oculomotor abnormalities (Hill, Reilly, Harris, Khine, & Sweeney, 2008).

16.3 Smooth Pursuit Eye Movements

Smooth pursuit eye movements (SPEM) are slow eye movements that serve to maintain the image of moving objects on the fovea (see chapter by Lencer et al., in this volume). Different paradigms have been studied that focus on a variety of subprocesses in the control of SPEM. There is a large literature on the neural substrate of SPEM in both humans and animals (see chapter by Lencer et al., in this volume). Thus the study of SPEM in psychiatry is of particular interest allowing the link between deficits in behavior and brain function. Indeed, SPEM is the most widely studied eye movement paradigm in psychiatry.

16.3.1 What Are the Differences from Healthy Controls? What Is the Diagnostic Sensitivity?

Since the pioneering work of Holzman, Proctor, and Hughes (1973), Holzman et al. (1974), a large number of studies have confirmed deficits in SPEM in patients suffering from schizophrenia. Deficits have also been observed in other major psychiatric disorders such as bipolar disorder, major depression, anxiety and personality disorders (O'Driscoll & Callahan, 2008). SPEM dysfunction has come to be regarded as an intermediate phenotype for schizophrenia (Calkins et al., 2008; Lencer et al., 2015; Levy et al., 2010).

A variety of different SPEM paradigms have been used in psychiatry research. Most studies have used a number of cycles of continuous target motion following a sinusoid, triangular or trapezoid speed pattern at one or more levels of frequency or mean speed to measure the closed-loop sustained pursuit performance (Smyrnis 2008; see chapter by Lencer et al., in this volume). Some studies have used the step ramp task to measure the initial open loop pursuit (initial acceleration of the eyes to catch the moving target after the step) as well as the subsequent closed-loop pursuit (Smyrnis, 2008; Sweeney et al., 1998; see chapter by Lencer et al., in this volume). Finally, a few studies have measured the predictive component of pursuit performance by manipulating the stimulus so as to remove visual feedback during pursuit (Smyrnis, 2008; see chapter by Lencer et al., in this volume). Yet others have used dual task paradigms to disengage attention from pursuit in order to study performance in a less controlled setting in patients and in unaffected relatives (Klein and Andresen, 1991; Rosenberg et al., 1997; Smyrnis, 2008; Tomer, Mintz, Levy, & Myslobodsky, 1981).

In the early years of SPEM studies in psychiatric disorders, overall SPEM performance was measured using qualitative ratings (estimations of pursuit quality from poor to excellent using a qualitative scale and a template of pursuit records) by experienced researchers (Levy, Holzman, Matthysse, & Mendell, 1994). Then, quantitative global measures of overall pursuit performance were introduced, such as the Root Mean Square (RMS) error, log signal/noise ratio and total saccade rate that estimate the accuracy with which the eye trace (eye position as a function of time) reproduces the target trace (Smyrnis, 2008; see chapter by Lencer et al., in this volume). Global quantitative measures are still used to characterize overall SPEM performance in psychiatric disorders, although they cannot dissociate between different patterns of

SPEM dysfunction such as for example slowing of pursuit resulting in catch up saccades or presence of intrusive saccades in SPEM with normal pursuit speed (Abel and Ziegler, 1988; Clementz and Sweeney, 1990; Smyrnis, 2008). More specific quantitative measures that better characterize the smooth pursuit response, developed in neurology and animal physiology laboratories, have been used in more recent studies in psychiatry. These approaches aim to measure specific and more biologically informative aspects of eye movement performance, including parameters such as open-loop acceleration (acceleration of the eye during the initial 100 ms of target movement), closed-loop pursuit gain (ratio of eye speed to target speed during sustained pursuit), and subtypes of saccades during pursuit that include compensatory saccades (catch up saccades and back up saccades that reduce tracking error) and intrusive saccades (anticipatory saccades and square wave jerks that increase tracking error) (Smyrnis, 2008; Sweeney et al. 1997; see chapter by Lencer et al., in this volume).

All early studies using qualitative ratings of global SPEM performance in schizophrenia showed deficits compared to healthy controls (Fig. 16.1) (Levy et al., 1994) and a meta-analysis showed that this measure had a large effect size, although with large confidence intervals (O'Driscoll & Callahan, 2008). Studies using quantitative global measures of SPEM performance also revealed deficits in patients with schizophrenia (Gooding et al., 1994) with similarly large effect sizes and confidence intervals in a meta-analysis (O'Driscoll & Callahan, 2008). Among specific quantitative measures of SPEM performance, decreased closed loop gain is the one that is most replicated in the literature (Abel, Friedman, Jesberger, Malki, & Meltzer, 1991; Arolt, Teichert, Steege, Lencer, & Heide, 1998; Clementz, Sweeney, Hirt, & Haas,

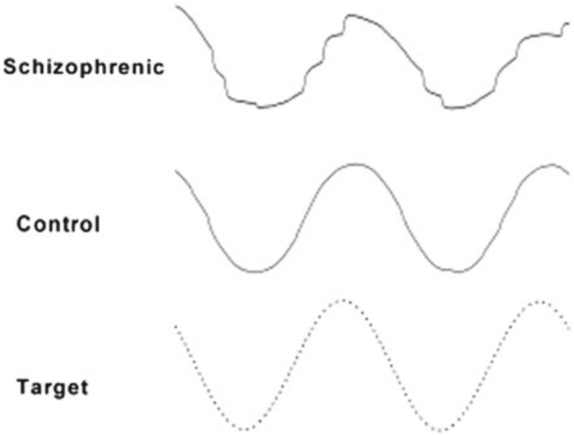


Fig. 16.1 Illustrative tracings of smooth pursuit eye movements of a schizophrenia patient (top panel) and of a normal control (middle panel), following a 0.4 Hz sine wave target (bottom panel, dotted line). The record of the schizophrenic patient shows many irregularities that suggest low gain pursuit with frequent saccadic eye movements. The normal control group record is smooth (reprinted with permission from Holzman, 2000)

1990; Levy et al., 2000; Sweeney et al., 1993; Thaker, Cassady, Adami, Moran, & Ross, 1996) and had the largest effect size and narrowest confidence interval in a meta-analysis (O'Driscoll & Callahan, 2008).

Some studies have reported that patients with schizophrenia have a specific deficit in the initial phase of smooth eye pursuit (open-loop pursuit) (Clementz et al. 1996b; Clementz & McDowell, 1994; Lencer et al., 2010; Sweeney et al., 1994, 1998, 1999). These studies have used a specific paradigm, the step ramp task, in which the visual target jumps from a central fixation points towards a peripheral location and then moves in the opposite direction in constant speed (ramp). Other studies have concentrated on predictive pursuit where pursuit does not rely on visual feedback from the moving stimulus (see chapter by Lencer et al., in this volume) proposing that a specific reduction in residual gain during predictive pursuit in schizophrenia implicates a specific deficit in predictive mechanisms (Hong et al., 2008; Lencer et al., 2004; Thaker et al., 1996; Thaker, Ross, Buchanan, Adami, & Medoff, 1999).

Several research groups assessed motion processing in schizophrenia by evaluating the accuracy of saccades to moving targets (Clementz et al. 1996b; Lencer et al., 2004; Sweeney et al., 1998, 1999; Thaker et al., 1999) and showed no deficit in schizophrenia (Sweeney et al., 1998, 1999). In a recent study, Spering, Dias, Sanchez, Schütz, and Javitt (2013) compared medication-treated patients with schizophrenia and controls in a paradigm where participants had to make a judgment whether a small dot moving linearly across a computer screen would hit a stationary line. In halve of the trials subjects were allowed to pursue the moving dot while in the other halve they had to fixate the stationary line. Both normal controls and patients performed equally well when they had to judge the dot movement with eyes fixated. When allowed to pursue the moving dot, normal controls improved their performance while no such improvement was observed in patients. These results clearly confirmed that motion processing was intact in patients while the use of extra-retinal cues (such as the efference copy signal from the smooth eye pursuit) to improve performance was impaired. Finally, Nkam et al. (2010) used predictable (pure sinusoidal) and unpredictable (pseudorandom) target motion paradigms showing that patients with schizophrenia had SPEM deficit only when target motion was predictable. In contrast to this, Lencer et al. (2010) found no deficit for patients in the predictive phase of smooth pursuit in the step ramp task. In the meta-analysis of O'Driscoll and Callahan (2008) open-loop pursuit, predictive pursuit as well as other specific measures such as catch up and anticipatory saccades yielded smaller effect sizes representing degree of difference between patients and controls compared to closed-loop pursuit gain and global measures of pursuit performance. In the most recent study using the largest patient sample size, Lencer et al. (2015) showed that both open loop and closed-loop pursuit gain dissociated patients from controls.

Most early studies using qualitative ratings of SPEM performance followed the dichotomy introduced by Holzman et al. (1973) between normal and abnormal SPEM records to divide patients and their first-degree relatives into those with pathological SPEM and those with normal SPEM (Levy, Holzman, Matthysse, & Mendell, 1993). With the introduction of quantitative measures, the effort to categorize patients according to SPEM performance relied on mixture model analysis that divided the continuous quantitative measure distribution into a normal and a deviant one (Clementz, Grove, Iacono, & Sweeney, 1992; Iacono, Moreau, Beiser, Fleming, & Lin, 1992; Ross, Thaker, Buchanan, & Lahti, 1996; Sweeney et al., 1993). A clear dissociation of the distribution of RMS error measured in SPEM, distinguished patients and their first-degree relatives from healthy controls. Although the categorization of patients based on SPEM performance could help in the effort to reduce the heterogeneity of the clinical syndrome and serve as a clinically useful tool, it has not been verified for specific quantitative SPEM measures (Thaker et al., 1993), though because of their limited neurobiological specificity global measures have not been applied in more recent studies with much larger sample sizes (Lencer et al., 2010, 2015).

In conclusion, SPEM deficit is present in schizophrenia patients. This deficit seems to be global, affecting more than one components of the pursuit system. This may be the reason why global indices of SPEM are more powerful in dissociating patients from controls (Table 16.1). While global measures of pursuit performance are advantageous for aggregating deficits in different pursuit systems, measures of specific components of pursuit may have advantages for teasing apart different neurophysiological alterations and genetic factors contributing to SPEM disturbances (Lencer et al., 2014). The sensitivity of smooth eye pursuit deficit for dissociating patients from healthy controls is not high enough to be clinically relevant for diagnostic purposes with individual patients, though may prove to be useful for patient subtyping and pathophysiology research. The challenge to general application in clinical diagnosis is that the distributions of SPEM measures in healthy and patient populations are heavily overlapping, and healthy individuals with highly deviant SPEM performance are commonly found especially in the case of large populations (Smyrnis et al., 2007).

16.3.2 What Is the Diagnostic Specificity?

In the pioneering study reporting SPEM deficits in schizophrenia, Holzman et al. (1973) also measured SPEM in patients with mania and other non-psychotic patients in an effort to explore the diagnostic specificity of this deficit. Some early studies using qualitative measures of SPEM performance showed that this deficit was not specific to schizophrenia but was also observed in patients with major affective disorders such as bipolar disorder and major depression, although it was much more common in schizophrenia (Levy et al., 1993; Sweeney et al., 1994). Later on, also studies using quantitative measures of pursuit performance such as closed-loop gain clearly showed that SPEM deficit is also present in patients with affective disorders (Abel et al., 1991; Kathmann, Hochrein, Uwer, & Bondy, 2003). Studies using the step-ramp pursuit task showed that more similar deficits both for closed-loop and initial open loop pursuit were observed for schizophrenia and affective disorders (Lencer et al., 2004, 2010; Sweeney et al., 1998, 1999), though these studies involved patients with relatively severe illness and often psychotic symptoms which

Table 16.1 Evidence for oculomotor function deviance in psychiatric disorders

	SCZ	BD	MD	OCD
Smooth eye pursuit				
Closed loop gain	↓*	↓*	↓/-	↓/-
Open loop gain	↓*	↓*	↓/-	–
Predictive gain	↓	?	?	?
Catch up saccades	↑/-	↑/-	↑/-	–
Intrusive saccades	↑/-	↑/-	↑/-	↑/-
Visually guided saccades				
Accuracy	- /↓	–	- /↓	–
Mean latency	- /↓	–	- /↑	–
Latency ISV	↑	?	?	–
Antisaccades				
Error rate	↑*	↑*	↑/-	↑/-
Mean latency	↑/-*	↑/-*	↑/-	↑/-
Latency ISV	↑	?	?	–
Accuracy	↓	?	?	?
Memory guided saccades				
Accuracy	↓	–	↓/-	?
Anticipatory saccades	↑/-	?	?	?
Mean latency	↑/-	↑/-	↑/-	?
Latency ISV	?	?	?	?
Predictive saccades				
Accuracy	↓	↓	↓	?

Note *SCZ* Schizophrenia, *BD* Bipolar Disorder, *MD* Major Depression, *OCD* Obsessive Compulsive Disorder. *Symbols* ↓: decrease compared to controls, ↑: increase compared to controls, -: no difference from controls, ?: no evidence, *: evidence from large samples

may account for this difference regarding diagnostic specificity of pursuit alterations. In the most recent study using the largest sample sizes in the literature, it was clearly shown that disorders of the psychotic spectrum (schizophrenia, schizoaffective disorder and psychotic bipolar disorder) shared a similar deficit in SPEM performance (Lencer et al., 2015).

These findings suggest that the SPEM deficit is not specific to schizophrenia but is present in all psychotic disorders extending to major depression (Table 16.1). This finding is in line with the current view of psychotic spectrum disorders using a dimensional approach (Clementz et al., 2015; van Os & Kapur, 2009). According to this approach psychosis is manifested in different symptom dimensions such as positive symptoms, negative symptoms, mania, depression and cognitive impairment. Differential contribution of such dimensions characterizes the taxonomically defined syndromes such as schizophrenia (prominent positive and negative symptoms), bipolar disorder (prominent manic symptoms) and major depression (prominent depression symptoms) but there can be any mixture of the different dimensions for each patient. In this approach then cognitive impairment can be present independently of other symptom dimensions, so if SPEM deficit is an expression of the cognitive impairment dimension one should expect to find it in all of the taxonomically defined psychotic spectrum disorders.

Fewer studies have investigated SPEM in anxiety disorders such as obsessive-compulsive disorder (OCD). In a recent literature review, Jaafari et al. (2011) concluded that the pursuit deficits in OCD patients are not observed in a systematic fashion as is the case in schizophrenia. A few studies have shown abnormalities in smooth pursuit function (Gambini, Abbruzzese, & Scarone, 1993; Sweeney, Palumbo, Halper, & Shear, 1992). Significantly lower gain and increased frequency of anticipatory saccades were found in OCD patients compared to control subjects (Clementz et al., 1996a). However, according to other studies, OCD patients' performance on smooth eye pursuit was not significantly impaired (Farber, Clementz, & Swerdlow, 1997; Lencer et al., 2004; Nickoloff, Radant, Reichler, & Hommer, 1991). In two studies comparing SPEM performance in patients with schizophrenia and OCD it was observed that OCD patients in contrast to schizophrenia patients had no deficit in the pursuit mechanism as measured by gain (Damilou, Apostolakis, Thrapsanioti, Theleritis, & Smyrnis, 2016; Lencer et al., 2004). Interestingly, however, in the recent study of Damilou et al. (2016), OCD patients but not schizophrenia patients produced more intrusive saccades during pursuit than healthy controls, paralleling a previous report (Sweeny, Palumbo, Halper, & Shear, 1992). This finding suggests a double dissociation between psychosis and OCD, with the former affecting the pursuit system and the latter affecting the suppression of the saccadic system during pursuit.

In conclusion, a SPEM deficit characterizes schizophrenia and affective psychoses (major depression and bipolar disorder) and extends to the psychotic spectrum traits in the population of healthy individuals, as will be shown in the next session, whilst in anxiety disorders such as OCD the SPEM function is largely spared.

16.3.3 What Is the Relation to Clinical Characteristics and Medication Status?

SPEM abnormalities are minimally influenced by the duration of symptoms or medication status; instead, they are considered a trait characteristic in schizophrenia and affective disorders (Flechtner, Steinacher, Sauer, & Mackert 2002; Hutton et al., 1998; Kallimani et al., 2009; Lencer et al., 2010; Sweeney et al., 1994). While all studies show that the deficit in SPEM persists over time, some significant but subtle change with treatment and recovery from untreated acute illness have been noted (Lencer et al., 2008).

Importantly, SPEM abnormalities have also been found in subjects within the schizophrenia spectrum that have not developed the full blown schizophrenia syndrome. van Tricht et al. (2010) examined ultra-high risk subjects (UHR) with a proneness to develop psychosis later in the life span, compared to healthy subjects. They found that saccadic rates during pursuit were higher in the UHR group, with no difference in gain. Other studies confirmed SPEM deficits for individuals presenting with clinically diagnosed schizotypal personality disorder (Brenner, McDowell, Cadenhead, & Clementz, 2001; Clementz et al., 1993), and individuals having high scores in psychometric schizotypy questionnaires without a clinical diagnosis of schizotypal personality disorder (Holahan & O'Driscoll, 2005; Kattoulas et al., 2011a, b; O'Driscoll, Lenzenweger, & Holzman, 1998; Smyrnis et al., 2007; Van Kampen & Deijen, 2009).

In conclusion, SPEM deficits in psychotic disorders are present in drug naive patients and patients experiencing a first psychotic episode and remain relatively stable during the course of the disorder. They are also not related to symptom severity. Thus SPEM deficit seems like a stable trait unrelated to the clinical characteristics and minimally influenced by medication status of patients with psychotic disorders.

16.3.4 What Role Do Heritability and Genetic Factors Play?

Since the pioneering work of Holzman et al. (1973), SPEM abnormalities have been found in many studies assessing healthy first-degree relatives of individuals with schizophrenia (Clemetnz, Sweeny, Hirt & Haas, 1991; Lencer, Trillenberg-Krecker, Schwinger, & Arolt, 2003; Rybakowski & Borkowska 2002; Sporn et al., 2005) but not all (Boudet et al., 2005). SPEM co-segregates with the disorder in multiply affected families (Hong et al., 2006; Iacono et al., 1992). Lencer et al. (2000) compared families with sporadic cases of schizophrenia, families with multiple cases, and families without schizophrenia cases and found similarities in SPEM abnormalities in families with both sporadic and multiple schizophrenia cases.

In the meta-analysis reported by Calkins et al. (2008), relatives of schizophrenia patients were found to be deviant in their smooth pursuit functioning mainly in closed-loop gain compared to healthy controls. Concerning open loop gain, results are more controversial, with impairment reported in some studies (see for example Clementz, Reid, McDowell, & Cadenhead, 1995) but no deficit found in others (see for example Thaker et al., 2003). Recently, Ivleva et al. (2014) assessed a group of patients with schizophrenia and their relatives as well as a group of patients with bipolar disorder and their relatives. They found that both groups of patients had lower maintenance and predictive pursuit gain compared with controls with no difference between patient groups. Concerning relatives, no differences on any of the SPEM measures were found for both groups compared to controls. In the study with the largest sample size to date, Lencer et al. (2015) showed that relatives of patients with psychotic disorders with and without psychosis spectrum personality traits were impaired on initial eye acceleration, the most direct sensorimotor pursuit measure, but not on pursuit gain measures. There were no differences in pursuit measures between relatives of different psychotic disorder proband groups. It thus seems that the SPEM deficit is a trait that is familial in major psychotic disorders (i.e. schizophrenia and affective disorders). The degree to which this alteration represents a disturbance in predictive pursuit versus sensorimotor transformation remains an active area of research.

Molecular genetic studies have attempted to link the SPEM deficit to specific polymorphisms associated with psychosis. Haraldsson et al. (2009) as well as Park et al. (2009) found no association between SPEM dysfunction and a functional polymorphism related to COMT. Another gene that has been shown to dissociate patients with schizophrenia from controls is *neuregulin 1* (*NRG1*). However, Haraldsson et al. (2009) found no association in an Icelandic group of patients between the *NRG1* risk haplotype and SPEM. In contrast, Smyrnis et al. (2011) showed that SPEM performance was modulated by the schizophrenia related *NRG1* polymorphism in a large sample of healthy controls while Kim et al. (2012) found no association between four variants of the *NRG1* gene and SPEM function. Recently Bae et al. (2014) reported a tendency toward associations between *ERBB4* polymorphisms (the human receptor tyrosine-protein kinase gene, which mediates neuregulin 1 signaling) and SPEM abnormality in a Korean population of patients with schizophrenia. In a genome wide association study investigating the influence of common genetic variation loci across the genome to oculomotor performance function indexes in a large sample of young healthy men, SPEM function was not related to common genetic variation (Hatzimanolis et al., 2015). There was though a significant association between schizophrenia related loci as these have been identified from the psychosis genetics consortium study (Lentz et al., 2014) and SPEM performance (Hatzimanolis et al., 2015). In a recent study of first episode schizophrenia patients not yet receiving antipsychotic medication Lencer et al. (2014) showed an association of the dopamine related DRD2 gene (expressed in basal ganglia) to deficits in pursuit initiation which reflects a primarily motor function. On the other hand, deficits in predictive and sustained pursuit related to cognitive SPEM control showed an association to the glutamatergic GRM3 genotype (expressed in association cortex). This latter observation suggests that genetic association studies may require a more detailed parsing of oculomotor behavior in order to link different pursuit deficits to specific neurochemical processes and their related genes.

The conclusion from these studies is that SPEM deficits in psychiatric disorders, especially those of the psychotic spectrum, are shared with their first-degree relatives (Table 16.2). One the other hand, molecular genetic studies have failed so far to consistently reveal specific genetic loci that are responsible for the heritability of the SPEM deficit in psychiatric disorders. This is probably due to the fact that the SPEM deficit is a complex phenotype itself (maybe as complex as the psychiatric syndromes themselves (Iacono et al., 2017) thus there is need for much larger sample sizes (in the magnitude of thousands of individuals) in order to reveal the molecular genetic basis of the SPEM deficit more reliably.

Table 16.2 Heritability of oculomotor function parameters in schizophrenia

	Studies with 1st degree relatives	Homozygote twin studies	Molecular genetics studies
Smooth eye pursuit			
Closed loop gain	↓/- *		No
Open loop gain	↓/- *		No
Predictive gain	↓/-		?
Catch up saccades	↑/-		?
Intrusive saccades	↑/-		?
Visually guided saccades			
Accuracy	- *	–	No
Mean latency	- *	–	No
Latency ISV	?	?	?
Antisaccades			
Error rate	↑/- *	–	No
Mean latency	- *	–	No
Latency ISV	?	?	?
Accuracy	↓/-	↓	?
Memory Saccades			
Accuracy	↓	?	?
Anticipatory saccades	↓	?	?
Mean latency	–	?	?
Latency ISV	?	?	?
Predictive Saccades			
Accuracy	↓/-	?	?

Note *Symbols* ↓: decrease compared to controls, ↑: increase compared to controls, -: no difference from controls, ?: no evidence, *: evidence from large samples, No: no definite association with genetic polymorphisms

16.3.5 What Is the Neurobiological Substrate?

The neural pathways underlying SPEM are well described and include regions responsible for both motion processing and prediction, such as frontal eye fields (FEFs), parietal eye fields (PEFs), supplementary eye fields (SEFs), cerebellar vermis and the motion-sensitive area V5 (see chapter by Lencer et al. in this volume). How are SPEM abnormalities in psychiatric patients reflected at the neural level? A number of studies have attempted to answer this question.

In a sample of first-degree relatives of schizophrenia patients, activation (measured using positron emission tomography, PET) of the right FEF correlated with SPEM performance and relatives with qualitatively abnormal pursuit failed to activate this area while normally activating posterior parietal areas (such as area V5) (O'Driscoll et al., 1999). Reduced activation of the right FEF in patients with schizophrenia compared to healthy controls was also observed in a subsequent fMRI study (Tregellas et al., 2004). In that same study, increased activation of the bilateral hippocampi and the right fusiform area was observed for patients compared to controls. A specific decease in activation of motion processing area V5 of the right hemisphere in patients with schizophrenia compared to controls has also been reported (Lencer, Nagel, Sprenger, Heide, & Binkofski, 2005, 2011, see Box 1).

Box 1: Imaging the motion processing and pursuit system in psychosis

In a recent fMRI study, Lencer et al. (2011) studied both passive motion processing and smooth eye pursuit in untreated patients with schizophrenia and bipolar disorder and a group of healthy controls. During the passive visual motion processing task, both patients with schizophrenia (A in Fig. 16.2) and patients with bipolar disorder (B in Fig. 16.2) showed reduced activation in the posterior parietal projection fields of motion-sensitive extrastriate area V5 such as the superior parietal lobule (area 1 in Fig. 16.2), the anterior intraparietal sulcus (area 2 in Fig. 16.2) and the precuneus (area 3 in Fig. 16.2), but not in V5 itself. This observation might suggest reduced bottom-up transfer of visual motion information from extrastriate cortex to visuomotor systems in parietal association cortex. During active pursuit, both patients with schizophrenia (Fig. 16.3 A1) and patients with bipolar disorder (Fig. 16.3 A2) showed enhanced activation in anterior intraparietal sulcus (area 3 in Fig. 16.3), posterior superior temporal sulcus (area 4 in Fig. 16.3), dorsolateral prefrontal cortex (area 2 in Fig. 16.3) and dorsomedial thalamus (area 1 in Fig. 16.3), possibly resulting from increased demands on predictive pursuit systems due to a reduced availability of visual motion information about target speed and tracking error.

Using SPEM with and without target blanking, Nagel et al. (2007) observed that a reduced pursuit velocity gain during continuous target presentation in schizophrenic patients correlated with decreased neuronal activation of the right ventral premotor cortex as well as with increased activation of the left dorsolateral prefrontal cortex, the right thalamus and crus II of the left cerebellar hemisphere. During smooth pursuit intervals with target blanking, cerebellar activations were decreased and activation of the anterior cingulum, the superior temporal gyrus and the FEF bilaterally were increased for patients compared to controls (Keedy, Ebens, Keshavan, & Sweeney, 2006). Recently, Krishna et al. (2014) examined functional connectivity between posterior regions and frontal regions in patients with schizophrenia and healthy controls, using EEG with eye tracking tasks. They found a normal activation in patients in the frontal region while maintaining the representation of the target motion on line. However, during the pursuit response to repeated target motion, patients showed a lower frontal/posterior functional connectivity.

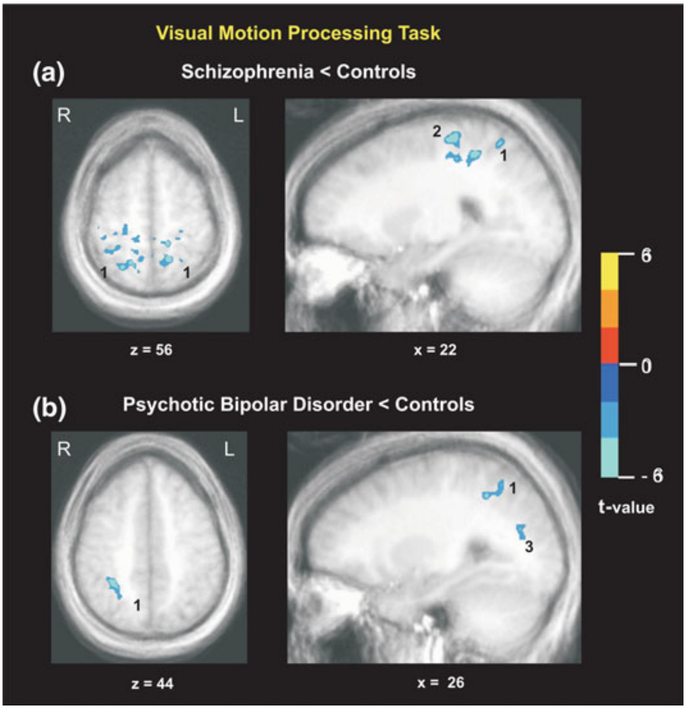


Fig. 16.2 Comparisons of activation during passive visual motion processing between controls (N = 20) and untreated patients with **a** schizophrenia (N = 24) and **b** psychotic bipolar disorder (N = 13). Activations in blue show regions with less activation in patients compared to controls: 1 = superior parietal lobule, 2 = anterior intraparietal sulcus, 3 = precuneus. The color bar indicates t-values (reprinted with permission from Lencer et al., 2011)

The emergent picture from these studies on the neurobiological substrate of the SPEM deficit in schizophrenia is not entirely consistent, with activation differences between patients and controls arising in different brain areas and with different activation patterns dissociating patients from controls for the same pursuit related brain areas (Table 16.3). One possible explanation for these somewhat contradictory results might be the fact that all these studies were performed on small samples. Given the large heterogeneity of patient characteristics the need for larger sample sizes, the use of drug naive patient samples and an effort for replication of findings in different samples is evident before one could reach any definite conclusions on the neurobiological substrate of SPEM deficits in psychiatric disorders.

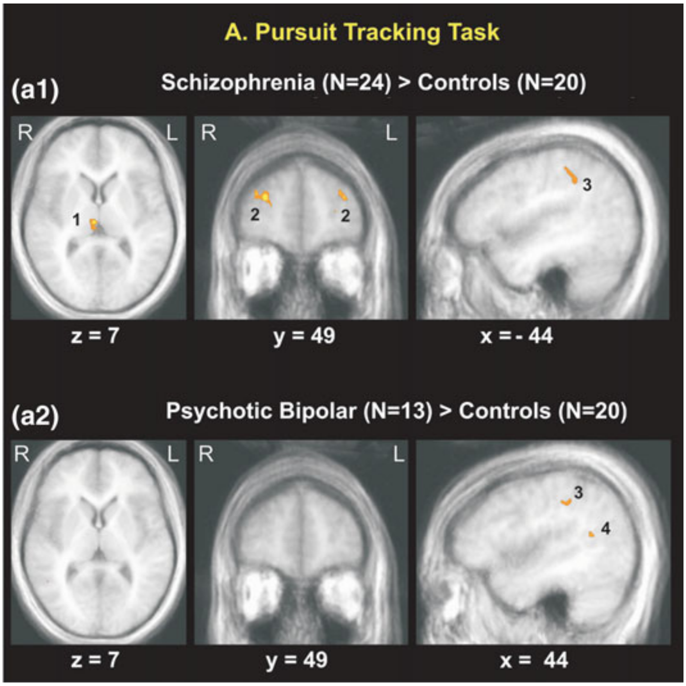


Fig. 16.3 **a** Comparisons of activation during pursuit tracking between controls (N = 20) and untreated patients with (**a1**) schizophrenia (N = 24), and (**a2**) psychotic bipolar disorder (N = 13). Activations in red show regions with higher activations in patients compared to controls (reprinted with permission from Lencer et al., 2011)

Table 16.3 Evidence for differences in brain activation for patients with schizophrenia versus healthy controls in oculomotor function tests

Brain area	SPEM	VGS	AS	MS
Molecular genetics studies	↓/-/↑	↓/-	↓/-	↓
Supplementary eye field		–	↓/-	↓
Dorsolateral prefrontal cortex	↑	–	↓/-	↑
Area V5	↓/-			
Parietal eye field		–	↓/-	↓
Anterior cingulum	↑		↓	
Hippocampus	↑			
Putamen		–	↓/-	
Caudate		–	↓/-	
Cerebellum	↓			

Note *SPEM* smooth pursuit eye movements, *VGS* visually guided saccades, *AS* antisaccades, *MS* memory saccades, *PS* predictive saccades. *Symbols* ↓: decrease compared to controls, ↑: increase compared to controls, -: no difference from controls

16.4 Saccades

Saccades are rapid eye movements that serve to shift the gaze in order to foveate objects such as during visual exploration, reading or in response to changes in the periphery (see chapter by Pierce et al., in this volume). As with SPEM, a large variety of different paradigms using saccadic eye movements have been developed that probe different cognitive processes. There is wealth of knowledge from studies using both animals and humans on the neural substrate of the saccadic eye movement system and the different cognitive processes engaged in different saccadic paradigms (see chapter by Pierce et al. in this volume). Thus, in the same vain as with the study of SPEM, the study of saccades in psychiatry offers the possibility to probe specific cognitive processes and their potential deficits and discover the neural substrate of these deficits.

16.4.1 What Are the Differences from Healthy Controls? What Is the Diagnostic Sensitivity?

In early studies of saccadic eye movements in schizophrenia it was found that the mean latency for visually guided saccades of medicated schizophrenia patients was normal, a finding which casts doubt on the hypothesis of psychomotor slowing (slower reaction times in all volitional movement tasks) in schizophrenia (Gale & Holzman 2000; Iacono, Tuason, & Johnson, 1981; Levin, Jones, Stark, Merrin, & Holzman, 1982). Other characteristics of visually guided saccades in these patients also appeared normal (Levin et al., 1982). In a large review of the literature, Gooding and Basso (2008) also reported normal visually guided saccades in schizophrenia.

However, Crawford, Haeger, Kennard, Reveley, and Henderson (1995) observed a deficit in the spatial accuracy of saccades in unmedicated patients. Curtis, Calkins, and Iacono (2001) observed the same deficit in acutely ill and remitted patients. Reilly and colleagues (Reilly, Harris, Keshavan, & Sweeney, 2005; Reilly, Harris, Khine, Keshavan, & Sweeney, 2008) found reduced latency of prosaccades in drug naïve patients with schizophrenia that was no longer present after six weeks of risperidone treatment. In a recent study, Krebs et al. (2010) confirmed the reduction of latency for visually guided saccades in drug naïve patients and further showed that patients made more anticipatory saccades, fewer regular saccades and had shorter latencies for express saccades than controls. Although patients with schizophrenia do not manifest slower reaction times for visually guided saccades, their saccadic latency distributions differ from those of normal controls and the intra-subject variability of saccadic latencies is larger in these patients (Karantinos et al., 2014; Smyrnis et al., 2009; Theleritis, Evdokimidis, & Smyrnis, 2014).

In 1988, Fukushima and coworkers for the first time applied the antisaccade task (see chapter by Pierce et al. in this volume) in schizophrenia patients and healthy controls. They showed a robust deficit, namely that patients made twice as many erroneous prosaccades towards the target when instructed to look away from those targets. These results were similar with findings that had been reported in patients with frontal lobe lesions (Fukushima et al., 1988; Fukushima, Fukushima, Morita, & Yamashita, 1991). Although patients made more errors in the antisaccade task, these errors were as effectively corrected as in healthy controls, thus the percentage of uncorrected errors in patients was similar to that of controls. This is an important observation, suggesting that the deficient behavior observed in patients is not due to nonspecific lack of motivation or misunderstanding of the task, but rather appeared linked to a disturbance in voluntary inhibitory control.

These initial studies triggered a large body of literature using the antisaccade task in psychiatric patient populations (Gooding & Basso, 2008). The percentage of trials with erroneous prosaccades is the most widely used and most reliable measure of antisaccade performance (Ettinger et al., 2003). Healthy adults typically fail to suppress erroneous prosaccades toward the target on about 20–25% of trials, although there is a very large inter-subject variability in young healthy adults ranging from 0% to above 90% error rate as has been shown in a sample of 2000 young healthy men (Evdokimidis et al. 2002). Also, the mean error rates in groups of patients with schizophrenia vary from approximately 20 to 75% (Gooding & Basso 2008; Smyrnis 2008). The difference though in error rate between schizophrenia patients and controls remains a robust finding that has been extensively replicated (Allen, Lambert, Johnson, Schmidt, & Nero, 1996; Clementz, McDowell, & Zisook, 1994; Crawford et al., 1998; Curtis et al., 2001; Gooding et al., 1997; Reuter, Rakusan, & Kathmanna, 2005; Sereno & Holzman 1995; Thaker et al., 1988; see Gooding & Basso, 2008 for review; Harris et al., 2009a). Patients with schizophrenia produce fewer correct responses on the antisaccade tasks, regardless of the version of the paradigm (Gooding & Basso, 2008). Further, the magnitude of the difference between patients and controls increases with an overlap condition, in which the fixation point remains illuminated when the peripheral dot appears (McDowell & Clementz, 1997; McDowell et al., 1999). The magnitude of these deficits is stable over time (Harris, Reilly, Keshavan, & Sweeney, 2006).

Two recent large multicenter studies included the antisaccade task as an intermediate endophenotype for psychosis. In the study of Reilly et al. (2014) a group of 267 patients with schizophrenia had an increase in antisaccade error rate of 23.7% compared to 244 healthy controls and the effect size of the group difference was 1.05. Even after correction for a generalized cognitive deficit the effect size was 0.73, indicating a robust difference between patients and controls. In another multicenter study including 345 patients and 517 community controls the difference in error rate between patients and controls was 20% and the effect size was 1.07 (Millard et al., 2016; Radant et al., 2007). It should be made clear though that this difference is still on a group level and cannot be used on an individual basis to dissociate patients from controls, precluding the use of the antisaccade task as a diagnostic tool in clinical practice. In an effort to standardize the antisaccade paradigm in order to compare results from different studies and pool data across different studies in meta-analyses, a harmonization paradigm for the saccade-antisaccade task to be used in clinical studies has been proposed recently by a group of researchers in the field (Antoniades et al., 2013).

In addition to error rate, the antisaccade task yields several other measures that provide important insights into the integrity of the cognitive and neural mechanisms involved in decision making (see chapter by Pierce et al.; Hutton & Ettinger, 2006; Smyrnis 2008). These include the reaction times (RT) of antisaccades and error prosaccades, the time taken to correct errors (the time between an erroneous prosaccade and subsequent corrective antisaccade), the percentage of errors that are corrected, and various spatial accuracy measures including the amplitude of antisaccades and error prosaccades and the final eye position of correct responses (see chapter by Pierce et al., in this volume; Hutton & Ettinger, 2006; Smyrnis 2008).

Most studies that have measured the latency of correct antisaccades have shown that the mean latency of correct antisaccades is increased in patients with schizophrenia (Barton et al., 2002; Curtis et al., 2001; Damilou et al., 2016; Fukushima et al., 1988, 1990), although some studies did not find any difference (Caldani et al., 2017; Clementz et al., 1994; Ettinger et al., 2004; Hutton, Joyce, Barnes, & Kennard, 2002). An increase in latencies in untreated first episode patients has been reported (Harris et al., 2006). In studies where it was measured, the mean latency of erroneous prosaccades in the antisaccade task did not differ between patients and controls (Curtis et al., 2001; Damilou et al., 2016; McDowell & Clementz, 1997). The mean latency of corrective antisaccades after an error prosaccade was found to be increased in patients (Damilou et al., 2016). Although most studies have shown differences in antisaccade mean latencies, other characteristics of the latency distributions have not been studied. In a recent study, Damilou et al. (2016) showed that the intra-subject variability of correct antisaccade as well as error prosaccade latencies is increased in patients with schizophrenia.

The introduction of a short time gap (usually 200 ms) between the extinction of the central fixation target and the onset of the peripheral target in the saccade and antisaccade task results a decrease of saccadic reaction time and an increase in the error rate in the antisaccade task (gap effect; see chapter by Pierce et al., in this volume). Two studies reported that the introduction of a gap between fixation offset and target presentation resulted in a similar gap effect in error rate and antisaccade reaction time in patients compared to controls (McDowell & Clementz 1997; Reilly et al., 2008). In another study half of the patients did not show a gap effect in antisaccade latency while the gap effect in error rate was again no different for patients and controls (Smyrnis et al., 2004). The percentage of corrected errors in the antisaccade task is not different between patients and controls (Gooding & Tallent, 2001; Polli et al., 2008). Several studies have measured the spatial accuracy of correct antisaccades and have reported that patients with schizophrenia have reduced spatial accuracy when mirror movements are required in the antisaccade task (Ettinger et al., 2004; Karoumi, Ventre-Dominey, Vighetto, Dalery, & d'Amato, 1998; McDowell et al., 1999).

In the memory-guided saccade task (MGS), subjects have to withhold a saccade to a visually presented target briefly for a few seconds while retaining in spatial working memory the location of the target. Then a saccade to the memorized target location has to be performed. The task is a well-established, translational measure of visuospatial working memory (Funahashi, Bruce, & Goldman-Rakic, 1989). Using this paradigm, Park and Holzman (1992) showed that memory-guided saccades were less accurate and were performed with prolonged reaction times by patients with schizophrenia when compared to healthy controls suggesting a deficit in spatial working memory in the saccadic system. In later studies the spatial accuracy deficit in memory guided saccades was confirmed (Landgraf, Amado, Bourdel, Leonardi, & Krebs, 2008; McDowell & Clementz 1996; Ross et al., 1998). Some studies also confirmed the increase in reaction time (McDowell & Clementz 1996; McDowell et al., 2001) although not all (Ross et al., 1998). Another measure of performance in this task is the percentage of anticipatory saccades to the target location during the memory delay period that was found to be increased for patients compared to controls in some studies (McDowell & Clementz 1996; McDowell et al., 2001) while not all (Ross et al., 1998). Recently, the percentage of anticipatory saccades has been found increased in individuals at ultrahigh risk for developing psychosis. The increase was specific to the subgroup of such ultrahigh risk individuals that also presented with neurobiological soft signs (such motor coordination deficit), indicating that this subgroup probably expressed a neurodevelopmental vulnerability to psychosis (Caldani et al., 2017). The reduced accuracy of memory guided saccades has been confirmed in never treated first episode patients (Reilly, Harris, Keshavan, & Sweeney, 2006; Reilly, Harris, Khine, Keshavan, & Sweeney, 2007). In those same studies the administration of antipsychotic medications (risperidone) was shown to further deteriorate performance of these patients compared to healthy controls.

In the predictive saccade task, the visual target alternates with a constant frequency between two positions in the horizontal plane, eliciting anticipated saccades and allowing the analysis of externally elicited non-anticipated and internally elicited anticipated saccades during the same task (Amado, Landgraf, Bourdel, Leonardi, & Krebs, 2008). McDowell and Clementz (1996) observed a reduction in saccadic amplitude during this task in patients with schizophrenia regardless of reaction time, suggesting that the alterations found were not due to a lack of preparation.

In conclusion, several saccadic eye movement paradigms have been used in investigations of patients with schizophrenia (Table 16.1). Differences in performance were observed in patients compared to controls in a variety of different parameters measured in each one of these tasks. In terms of robustness of these effects, it seems that the increase in antisaccade error rate is the best replicated finding in this large literature with the largest effect size in dissociating patients from controls at the group level.

16.4.2 What Is the Diagnostic Specificity?

Visually guided saccades were studied in patients with mood disorders in comparison to schizophrenia and in most studies the accuracy and mean latency of these saccades was found to be normal in these patients as it was also observed in schizophrenia (Crevits, Van den Abbeele, Audenaert, Goethals, & Dierick, 2005; Sereno & Holzman, 1993; see Gooding & Basso, 2008). However, in some studies, patients with major depression showed decreased accuracy of visually guided saccades compared to controls (Fig. 16.2) (Curtis et al., 2001; Harris et al., 2009; Sweeney et al., 1998), consistent with multiple lines of evidence indicating cerebellar alterations in depressed patients. In one study (Mahlberg, Steinacher, Mackert, & Flechtner, 2001), unmedicated patients with major depression had increased mean latencies of visually guided saccades compared to controls. When patients with bipolar disorder were investigated in visually guided saccadic tasks no difference from controls was detected for both accuracy and mean saccadic latency (Crawford et al., 1995; Gooding, Mohapatra, & Shea, 2004; Reilly et al., 2014; Tien, Ross, Pearlson, & Strauss, 1996). Visually guided saccades have also been studied in patients with OCD and most studies reported no difference from controls in mean reaction times and accuracy (Nickolof et al., 1991; Tien, Pearlson, Machlin, Bylsma, & Hoehn-Saric, 1992; Spengler et al., 2006). In a recent study, Theleritis et al. (2014) also confirmed that these patients present normal reaction time distributions with no increase in the variability of saccadic reaction times in the visually guided saccade task dissociating them from patients with schizophrenia that showed a significant increase in the variability of saccadic reaction times.

The initial studies focusing on the antisaccade task performance in schizophrenia used samples of patients with affective disorders such as depression and bipolar disorder as control psychiatric populations to document the specificity of the antisaccade performance deficits in schizophrenia. In one such study, Fukushima et al. (1991) reported that only patients with schizophrenia showed increased antisaccade error rates and increased latencies of antisaccades while patients with affective disorders were no different from controls. Later studies using larger sample sizes though showed that patients with major depression (especially severe depression with psychotic features) as well as patients with bipolar disorder (again, especially those with psychotic features) also have increased error rates in the antisaccade task compared to controls (Harris et al., 2009a; Martin et al., 2007; Tien et al., 1996; Sweeney et al., 1998). In the study of Harris et al. (2009a), patients with non-psychotic major depression showed no difference in antisaccade performance compared to controls. However, patients with major depression with psychotic features showed similar deficits as patients with schizophrenia and patients with bipolar disorder. Concerning the mean latency for antisaccades, the study of Crevits et al. (2005) showed increased mean latency for acutely ill, depressed patients while in the studies of Sweeney et al. (1998) and Harris et al. (2009a) there was no difference for patients with depression or patients with bipolar disorder compared to controls. In the largest study to date, the antisaccade error rate was compared between groups of 267 schizophrenia patients, 150 schizoaffective patients and 202 psychotic bipolar patients and 244 controls (Reilly et al., 2014). All patient groups had significantly elevated error rates compared to controls, and the schizophrenia patients had significantly elevated error rates compared to both schizoaffective and psychotic bipolar patients. These results confirm that a deficit in antisaccade performance spans across the psychotic disorders spectrum although with decreasing magnitude as one moves from the psychotic (schizophrenia) to the affective (depression) dimensions of psychosis. Unfortunately, large scale studies such as the one previously mentioned report only on error rate and not on other antisaccade task variables (latencies, accuracy) and their differences among the different diagnostic groups of major psychotic disorders.

The antisaccade performance has also been investigated in patients with OCD with mixed results. Two studies reported an increase in antisaccade error rate for these patients while the antisaccade latency was found to be similar to that of healthy controls (Rosenberg et al., 1997; Tien et al., 1992). However, in two subsequent studies, the exact opposite results were obtained, namely an increase in mean latency of antisaccades and no difference in error rate compared to controls (Hardeman et al., 2006; Maruff, Purcell, Tyler, Pantelis, & Currie, 1999). In a more recent study, Kloft et al. (2011) reported that patients with OCD display normal performance in the antisaccade task while another study by Lennertz et al. (2012) reported that these patients showed both an increase in mean antisaccade latency and error rate compared to healthy controls. Studies by Rosenberg et al. showed elevated antisaccade error rates but not antisaccade latency or accuracy in pediatric OCD, and no alterations at all in prosaccade metrics (Rosenberg, Averbach, et al., 1997; Rosenberg, Sweeney, et al., 1997). This profile suggests a selective association between antisaccade errors indicating inhibitory failures with behavioral inflexibility that characterizes OCD as has been demonstrated in individuals with autism with prominent repetitive behavior problems (Mosconi et al., 2009). Two studies compared the antisaccade task performance of OCD patients to that of schizophrenia patients. In the first study, Spengler et al. (2006) showed a dissociation of schizophrenia and OCD in the sense that schizophrenia patients had increases in both mean antisaccade latency and error rate while OCD patients displayed an increase in latency with normal error rates. In a more recent study, Damilou et al. (2016) showed that both patient groups had similar increases in error rate and antisaccade mean latencies while the dissociation between the patient groups was in the variability of error prosaccade latencies that was increased only for the schizophrenia patients group.

In the first study of memory-guided saccades in schizophrenia patients, Park and Holzman (1992) used a group of bipolar patients to address the issue of diagnostic specificity. The bipolar group performed the memory-guided saccades with the same accuracy as controls but the latencies of memory saccades were prolonged as in schizophrenia. Normal accuracy of memory-guided saccades was also observed in a later study which, in addition, showed no increase in the latency of memory saccades for these patients (Crawford et al., 1995). In another study, patients with major depression had reduced accuracy of memory saccades with no increase in mean latency (Sweeney et al., 1998). In a later study it was shown that depressed patients with melancholic features showed increased mean latency of memory guided saccades while the accuracy was not different from that of control subjects (Winograd-Gurvich, Georgiou-Karistianis, Fitzgerald, Millist, & White, 2006).

In the study of Winograd-Gurvich et al. (2006), anticipatory saccades in a predictive saccade task were hypometric (did not reach the target) for depressed patients compared to controls. Predictive saccade performance was studied in a group of drug naive patients with bipolar disorder and it was observed that the accuracy of anticipatory saccades was significantly reduced compared to controls and drug naive patients with schizophrenia. This effect remained after treatment initiation for these patients (Keedy, Reilly, Bishop, Weiden, & Sweeney, 2014).

In conclusion, there is evidence that saccadic eye movement abnormalities are not restricted to patients with schizophrenia but are also observed in patients with other psychotic disorders and other psychiatric conditions such as OCD (Table 16.1). The pattern of these abnormalities seems to be qualitatively similar for patients with psychotic disorders where a deficit in volitional saccadic control can be observed which quantitatively decreases in the psychotic spectrum from schizophrenia to affective disorders. The relation of saccadic abnormalities to other psychiatric conditions such as OCD is much less consistent across paradigms and studies.

16.4.3 What Is the Relation to Clinical Characteristics and Medication Status?

Visually guided saccade performance was tested in patients with schizophrenia experiencing their first episode of psychosis while these patients were receiving antipsychotic medication and it was shown that both the mean latency and accuracy of the saccades were not different from controls (Hutton et al., 2002). Similarly, the mean latency and saccadic gain of visually guided saccades did not differ between patients experiencing acute psychosis and patients in the remission phase (Curtis et al., 2001). However, Reilly et al. (2005) performed a longitudinal study of antipsychotic naïve patients with schizophrenia measuring mean saccadic latency and accuracy of visually guided saccades before and after antipsychotic treatment. They observed that these patients had faster mean latencies compared to healthy controls before treatment, and were as accurate as healthy controls. Interestingly, the speeding of prosaccade latencies was associated with an increased rate of antisaccade errors, probably because it was harder for stop signals to prevent saccades since they were started earlier (Reilly et al., 2008). After treatment with antipsychotics, the mean latency increased to the level of healthy controls while accuracy became worse than that of healthy controls, suggesting that antipsychotic medication (especially second generation antipsychotics such as risperidone) have a distinct effect on saccade performance for these patients (Sweeney et al., 1997). In another study it was observed that the latency of visually guided saccades did not differ between two phases, one in which the patients were medication free and one in which they were receiving antipsychotic medication (Kallimani et al., 2009), suggesting that the reduction of saccadic latency might be observed only in medication naïve or perhaps more acutely ill patients with schizophrenia. This specific decrease of prosaccade latency for medication naïve patients with schizophrenia was not affected by manipulations of the central fixation (gap overlap conditions) (Reilly et al., 2008). It would be interesting to investigate further the reaction time distribution characteristics in patients before and after initial treatment with antipsychotics. There has not been observed a significant correlation of saccadic eye movement parameters with symptom severity or particular symptom dimensions of schizophrenia (Curtis et al., 2001; Kallimani et al., 2009; Reilly et al., 2005). Finally, one study reported a correlation of increasing levels of schizotypy in a sample of healthy adults to decreasing accuracy of visually guided saccades (Ettinger et al., 2005).

An increase in antisaccade error rate has been reported in patients with recent onset and/or first-episode schizophrenia (Broerse, Crawford, & Den Boer, 2002; Ettinger et al. 2004; Grootens et al., 2008, Hutton et al., 2002, 2004). Antisaccade mean latency was found to be increased in these patients in some studies (Grootens et al., 2008) while not in others (Ettinger et al., 2004, Hutton et al., 2002). Similarly, the amplitude gain of antisaccades was found to be decreased in some studies (Ettinger et al., 2004) while not in others (Hutton et al., 2002). Harris et al. (2006) conducted a longitudinal study of antisaccade performance on medication naïve patients with schizophrenia that were followed for a year after the initiation of antipsychotic treatment. Both the antisaccade error rate and the antisaccade mean latency were significantly elevated in these patients compared to controls before receiving medication and then gradually decreased over the course of treatment but remained significantly elevated compared to controls even at the end of the follow-up period of one year. The same results were obtained in a later study in which medication naïve patients were tested at baseline and then after six weeks of risperidone treatment (Hill et al., 2008). Antisaccade error rate and mean latency were also found to remain stable over time in patients with schizophrenia independent of changes in clinical state (Curtis et al., 2001; Gooding et al., 2004) as well as medication status (Kallimani et al., 2009). However, another study reported an increase in error rate during periods of acute illness because of psychosis-related impairments in attention (Reilly et al., 2008).

In ultra high risk subjects, antisaccade error rate was found to be higher than in controls, with a trend toward an increase in error rate in those who later made a transition to psychosis (Nieman et al., 2007). However, Caldani et al. (2017) found no difference in antisaccade errors in subjects having an ultra high risk for psychosis compared to controls. Antisaccade error rate but not antisaccade mean latency was also found to be elevated in individuals that had very high scores in self reporting questionnaires measuring schizotypy (O'Driscoll et al., 1998; Smyrnis et al., 2003). Although one study showed a correlation of increasing levels of schizotypy and antisaccade error rate (Ettinger et al., 2005) in healthy individuals, another study showed that the correlation was not significant and the increase in error rate was only evident for a small subgroup of individuals with very high scores of schizotypy (Smyrnis et al., 2003). In another study that compared antisaccade performance in individuals with schizotypal personality disorder identified by clinical interviews to controls there was no significant difference in error rate, antisaccade latency and antisaccade accuracy (Brenner et al., 2001).

Finally, in most studies, there was no significant correlation of particular symptom clusters, presence of high neurological soft signs scores or symptom severity, and clinical course on antisaccade performance measures in patients with schizophrenia (Caldani et al., 2017; Gooding et al., 2004; Harris et al., 2006; Kallimani et al., 2009). However, one study reported that patients identified having a deficit syndrome (an excess of negative symptoms that endure in time) had increased antisaccade latencies compared to non-deficit syndrome patients (Nkam et al., 2001).

The effects of medication on memory-guided saccades were tested in medication naïve patients with schizophrenia at baseline and then after six weeks of treatment with risperidone. It was shown that the accuracy of memory-guided saccades increased from baseline to retest in the healthy control group while it decreased in the patient group (Reilly et al., 2006). In one study where patients with schizophrenia were divided into high and low negative symptom scores there was no difference in memory saccade accuracy between the two groups. However, there was a significant increase in accuracy variability for the high negative scores group and a correlation of increased total symptom severity to decreased accuracy in the memory guided saccade task (Winograd-Gurvich et al., 2008).

In recent-onset, untreated, or drug-naive schizophrenia patients, Krebs et al. (2001) found decreased accuracy on anticipated saccades in the predictive saccade task compared to controls. In a later study Krebs et al. (2011) observed shorter latencies of predictive saccades in drug naïve patients compared to controls. Harris used this task to study untreated first episode schizophrenia patients before and 6 weeks after antipsychotic treatment, and observed a significant reduction in the gain of predictive saccades after treatment but no alteration before or after treatment in the ability to reduce saccade latencies as target appearance could be anticipated (Harris et al., 2009b). In a recent study Keedy et al. (2014) reported that drug naive patients showed no difference from controls in predictive saccade performance at baseline while there was a significant reduction of predictive saccade accuracy with no effect on latency after 4–6 weeks of antipsychotic treatment. This evidence suggests that acute D2 blockade with antipsychotic medications may induce a parkinsonian-like reduction in voluntary saccade gain (hypometric saccades) made without sensory guidance. In two studies (Harris et al., 2009b; Winograd-Gurvich et al., 2008) predictive saccade performance did not vary between groups of patients with predominant negative symptoms and other patients and did not correlate with symptom severity.

In conclusion, performance of visually guided saccades, memory-guided saccades and predictive saccades may be affected by the medication status and/or clinical symptoms of the patients while the antisaccade task performance deviance (especially the error rate) is independent of clinical state, medication status and psychopathological manifestations of the disorder. A deviance in antisaccade task performance is also observed in healthy individuals that have schizotypal personality traits as well as individuals at ultra high risk for developing psychosis, providing further evidence for the trait characteristics of the antisaccade performance deviance in schizophrenia.

16.4.4 What Role Do Heritability and Genetic Factors Play?

Family studies have measured the performance of first-degree relatives of patients with schizophrenia in visually guided saccades. Largely, they reported no difference from controls for mean latency and accuracy of these saccades (Crawford et al., 1998; Curtis et al., 2001; Ettinger et al., 2004). In a meta-analysis on the familiar effects of oculomotor function parameters in schizophrenia, Calkins et al. (2008) showed that the effect size of the difference between relative and healthy controls for both saccade latency and accuracy was about zero. Still, it remains an open question whether other parameters of the latency distribution such as intra-subject variability that dissociate patients and controls can also dissociate relatives from controls.

The most widely used saccadic eye movement paradigm in heritability and family studies in schizophrenia is the antisaccade task. Initial studies showed that first-degree relatives of patients with schizophrenia made significantly more antisaccade errors than healthy controls (Clementz et al., 1994; Katsanis et al., 1997). The study of Crawford et al. (1998) used a larger sample of 50 relatives and showed that when the group of relatives was considered as a whole there was no difference in error rate from healthy controls. When the group of relatives was divided based on the antisaccade performance of the schizophrenic probands, then the relatives of probands with a high error rate also had a significantly elevated error rate compared to healthy controls. In that same study the relatives were not dissociated from controls for mean antisaccade latency and antisaccade accuracy. In a later study with an even larger sample size of 116 relatives, Curtis et al. (2001) (Fig. 16.4) showed that the error rate of relatives was significantly elevated compared to controls but also significantly less elevated compared to patients, while mean antisaccade latency for relatives was not different from controls. In a later study, a difference between relatives and controls was observed only for the accuracy of antisaccades but not for error rate and mean antisaccade latency (Ettinger et al., 2004). The same effect of decrease in accuracy but not error rate was observed in a study where healthy monozygotic twins of patients with schizophrenia were compared to control monozygotic twins (Ettinger et al., 2006). In the meta-analysis of Calkins et al. (2008), the effect size for the difference in error rate between relatives of patients with schizophrenia and healthy controls was calculated to be 0.42 while for mean antisaccade latency it was 0.39.

The results of two large, recent consortium studies on endophenotypes for schizophrenia that were published recently have not resolved the discrepancy. The Bipolar Schizophrenia Network on Intermediate phenotypes (B-SNIP) consortium compared the antisaccade error rate in large groups of patients with schizophrenia, bipolar disorder and schizoaffective disorder as well as their first-degree relatives. The three relatives groups had similar error rates that were not different among the groups (25–28%), but were significantly elevated compared to the healthy control group (18.5%) and significantly less elevated compared to the probands groups, leading to the conclusion that antisaccade error rate deficit is heritable across the psychosis spectrum disorders (Reilly et al., 2014). Another study (Radant et al., 2010) reported the results of the Consortium on the Genetics of Schizophrenia (COGS), again on a very large sample of patients and relatives. Although again the relatives' performance in terms of error rate was intermediate between patients and controls, when adjusting for confounding factors such as age and gender, the difference between relatives and controls was no longer significant as it was in the B-SNIP study. In that COGS sample, a different analysis was performed assessing whether a significant proportion of inter-individual variance in error rate can be attributed to heritability (belonging

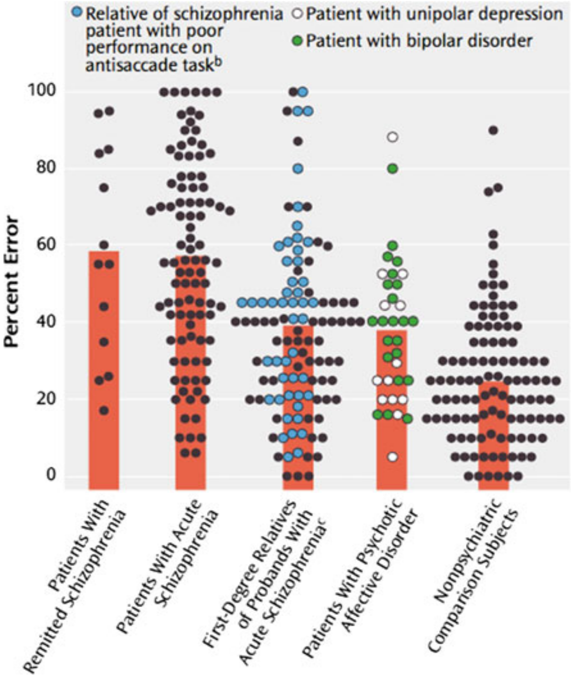


Fig. 16.4 Antisaccade Performance of Patients With Remitted or Acute Schizophrenia, First-Degree Relatives of Probands With Acute Schizophrenia, Patients With Psychotic Affective Disorder, and Nonpsychiatric Comparison Subjects[a] (reprinted with permission from Curtis et al., 2001). [a]The bars represent the group means. As determined by Student- Newman-Keuls follow-up tests that used an alpha level of 0.05, both schizophrenia groups made significantly more errors than the patients with psychotic affective disorder and the first-degree relatives, both of which groups, in turn, made significantly more errors than the nonpsychiatric comparison subjects. No other group differences were significant. [b]Poor performance on the antisaccade task was defined as performance two standard deviations worse than that of the mean for the nonpsychiatric comparison subjects. [c]Relatives with no lifetime psychotic diagnosis

to the same family). It was shown that antisaccade performance was significantly heritable ($h^2 = 0.42$) (Greenwood et al., 2007). In a more recent report including the full COGS data set, the authors found that a combined measure of antisaccade performance (error rate) and prepulse inhibition (PPI) resulted in a heritability estimate ranging from 22 to 29% (Seidman et al., 2015). These results suggest that antisaccade performance heritability might be independent of schizophrenia in the sense that the level of performance in one family member is a predictor of performance of the other family members in a similar fashion in healthy individuals and patients with schizophrenia. Thus patients as well as control subjects with a poor performance in the antisaccade task have relatives with poor performance explaining the specificity of the results of the Crawford et al. (1998) study.

Following the same rationale as for SPEM, a number of molecular genetic studies have attempted to link the antisaccade performance deficit with specific single nucleotide polymorphisms associated with psychosis and more specifically with prefrontal cortical function. The most widely investigated such polymorphism is a G-to-A transition at codon 158 of the COMT gene, resulting in a valine (*val*)-to-methionine (*met*) substitution. This substitution yields three different *val158 met* genotypes (met/met, val/met and val/val) that result in the COMT enzyme having increasing potency in binding to dopamine leading to decreasing levels of dopamine in the prefrontal cortex. In initial studies, a relation was found between COMT genotype and performance in the Wisconsin Card Sorting Test, probing prefrontal function in healthy humans (Egan et al., 2001; Melhotra et al., 2002). Antisaccade error rate was found not to relate to the COMT *val158 met* polymorphism in healthy young males (Stefanis et al., 2004). A series of single nucleotide polymorphisms on the gene regulator for protein G subtype 4 (RGS4) was also found to be related to schizophrenia and prefrontal cortical activity and relation of these polymorphisms to oculomotor task performance was investigated in a sample of healthy young adults (Kattoulas et al., 2012). In that study a specific relation of two of these polymorphisms was observed only for antisaccade error rate and the intra-subject variability of latency for correct antisaccades but not for smooth eye pursuit and sustained fixation task variables.

In a follow up of their initial study on memory-guided saccades, Park et al. (1995) examined the performance in this task in first-degree relatives of patients with schizophrenia. The relatives were significantly less accurate than healthy controls in this task while they were also significantly more accurate than patients. The response latency of memory saccades did not differ between relatives and controls while they were both significantly faster compared to patients. Subsequent studies (McDowell et al., 1996, 2001) confirmed that first-degree relatives of patients had decreased accuracy of memory-guided saccades compared to controls and they also showed an increased percentage of anticipatory errors during the delay period in this task that was observed in relatives as well as patients compared to controls. This failure of inhibition of unwanted saccades during the delay period for first-degree relatives of patients with schizophrenia was also confirmed by a later study (Landgraf et al., 2007). In their meta-analysis, Calkins et al. (2008) showed that the effect size for the difference between relatives and controls was 0.56 for the anticipatory error rate, 0.66 for the accuracy of memory saccades and 0.06 for the latency of memory saccades. These effect sizes were the largest ones compared to all other variables measured for smooth eye pursuit, visually guided saccade antisaccade and visual fixation tasks but the authors correctly warn that the number of studies included in the memory saccade paradigm is very small. Assessing siblings of patients with schizophrenia in predictive saccades, Amado et al. (2008) showed that siblings also demonstrated a deficit in accuracy compared to controls.

A genome wide association study investigated the influence of common genetic variation loci across the genome on oculomotor performance function indexes in a large sample of young healthy men (Hatzimanolis et al., 2015). Antisaccade error rate and the intra-subject variability of correct antisaccade latency resulted in a significantly increased proportion of associations with single nucleotide polymorphism loci compared to that expected by change. However, the small sample size (1079 subjects) did not allow the dissociation of true positive from false positive associations. There was no association of the antisaccade performance loci and schizophrenia related loci found in the psychosis genetics consortium study (Lencz et al., 2014) suggesting that the genetic substrate of antisaccade function and schizophrenia either do not overlap or have a small overlap that would require much larger samples in order to be detected.

In conclusion, the heritability of saccadic task performance deficits in schizophrenia is a complex phenomenon yet to be clarified (Table 16.2). The most widely studied paradigm is the antisaccade task and especially the error rate in this task has produced mixed results for heritability. It seems that antisaccade performance is significantly heritable in healthy individuals as well as in patients, but whether and how this heritability is directly related to the disorder is still an open question. Molecular genetic studies trying to link this heritability with specific genes have yielded some positive results but much remains to be done. Finally, there is some evidence for the heritability of memory-guided saccade deficits in schizophrenia but larger samples are needed to verify this effect. It seems that the antisaccade and probably the memory saccade are complex phenotypes with complex heritability involving probably many genes and gene interactions. The genetic link of these phenotypes to schizophrenia and other psychiatric disorders remains to be investigated in depth.

16.4.5 What Is the Neurobiological Substrate?

The neural substrate for saccadic eye movements has been extensively studied in animals using electrophysiological methods such as single neuron recording (see chapter by Johnson & Everling, in this volume) and in humans studying the effects of lesions (see chapter by Müri et al., in this volume), as well as in healthy humans using electrophysiological (EEG/MEG) and imaging (structural MRI, fMRI) methods (see chapter by Pierce et al., in this volume). Thus we have detailed information on the role of specific brain areas in the production of saccadic eye movements and the specific role of particular brain areas in the performance of particular oculomotor paradigms. Relevant knowledge concerns, for example, the role of the frontal and PEF as well as the superior colliculus for the production of visually guided saccades, the additional specific role of the dorsolateral prefrontal cortex and caudate nucleus in the production of antisaccades, the role of the dorsolateral prefrontal cortex for memory-guided saccades, the role of the FEF for the production of predictive saccades and the role of the SEF for the production of sequential saccades. Thus the saccadic eye movement dysfunction in schizophrenia can serve as an excellent tool for the study of the neurobiological substrate of major psychiatric disorders using noninvasive structural (MRI) and functional (fMRI, PET) imaging as well as electrophysiological recordings (EEG/MEG)

An initial study using PET imaging (Nakashima et al., 1994) reported decreased activation of the frontal eye field in patients with schizophrenia compared to controls during performance of a visually guided saccade task. In a subsequent study using fMRI, there was no difference in activation between patients with schizophrenia and control subjects in all activated areas including, the PEF, the FEF and the SEF (McDowell et al., 2002). However, such differences were prominent in a study of unmedicated first episode schizophrenia patients who have latency differences suggesting reduced corticofugal modulation of subcortical oculomotor systems (Keedy et al., 2006). Finally, in a study comparing healthy controls with siblings of patients with schizophrenia, there was again no group difference in all areas activated in the visually guided saccade task (Raemaekers, Vink, van den Heuvel, Kahn, & Ramsey, 2006). One study used diffusion tensor imaging to measure white matter integrity of the anterior cingulum in patients with schizophrenia and healthy controls and then compared differences in measures of white matter integrity to performance in saccadic eye movement tasks (Manoach et al., 2007). The authors report reduced white matter integrity of the anterior cingulum buddle in patients, but more importantly, they showed that the specific reduction of white matter integrity of the buddle underlying the anterior cingulate cortex, the frontal eye field and the posterior parietal cortex correlated with increased latency for visually guided saccades and antisaccades.

A series of studies investigated neural activation during the antisaccade task, comparing patients with schizophrenia and healthy controls. Again, the results are somewhat conflicting. The study of McDowell et al. (2002) reported that when contrasting activation in the antisaccade task versus the visually guided saccade task, patients had decreased activation in the right superior frontal gyrus (area 9 of the dorsolateral prefrontal cortex) compared to controls. In another study, Raemakers et al. (2002) used the contrast between the combination of an antisaccade task and a "no go" saccade task (where subjects were instructed to keep fixating at the center when the peripheral target appeared and the central fixation target was turned off) and the visually guided saccade task trials. Using this contrast, they demonstrated that patients showed reduced activation relative to controls in the putamen and the caudate body bilaterally. In a subsequent study contrasting performance in antisaccade trials to simple visual fixation, Tu, Yang, Kuo, Hsieh, and Su (2006) showed that patients compared to controls failed to activate the lentiform nucleus and thalamus bilaterally as well as the left inferior frontal gyrus and the left inferior parietal lobule.

Adding to these findings are the results of a subsequent study (Camchong, Dyckman, Austin, Clementz, & McDowell, 2008) assessing activation during antisaccade and memory saccade trials contrasted to visual fixation in patients with schizophrenia, their first-degree relatives and healthy controls. There were no differences in activation between the two saccade tasks but there were group differences. In this study, patients with schizophrenia did not differ from relatives and healthy controls in the activation of area 9 of the frontal cortex and the striatum, contrasting a study of the same group regarding the prefrontal cortex (McDowell et al., 2002) and two studies reporting deficient activation only in the striatum for patients with schizophrenia (Raemaekers et al., 2002) as well as their first-degree relatives (Raemaekers et al., 2006). In this later study, deficient activation only in the patient group was observed bilaterally in the lateral part of the FEF (but not the medial part) and the SEF while deficient activation of both the patient and the relative group was observed bilaterally for Brodmann area 10 of the prefrontal cortex, anterior cingulate gyrus, cuneus, insula and middle occipital gyrus.

Finally, the study of Dyckman et al. (2011) concentrated on the characteristics of the hemodynamic response function during antisaccade trials. It showed that in patients with schizophrenia, compared to healthy controls the hemodynamic response in the FEF was slower to peak and return to baseline but reached the same peak activation level. The authors suggested that this delayed and persistent response might explain perseverative errors of patients during the antisaccade task. The study of Ettinger et al. (2004) examined volumetric differences between patients with schizophrenia and healthy controls and showed that increases of caudate nucleus volume related to hypometric gain of antisaccades only for patients (see Box 2).

Box 2: Imaging the error processing network in schizophrenia

Error performance in cognitive tasks activates a network of error related areas in the brain. These areas are thought to form two distinct networks one comprising of the dorsal anterior cingulate cortex, the striatum and the brainstem (possibly substantia nigra) that has been suggested to perform an enforcement learning based on error commission and the second comprising of the rostral anterior cingulate cortex, the insula and the amygdala that has been suggested to appraise the motivational and emotional aspects of error commission. Polli et al. (2008) used the antisaccade task to study error related activation in these two networks in patients with schizophrenia and healthy controls. They showed that antisaccade errors compared to correct antisaccade responses led to increased activation in the dorsal anterior cingulate cortex (see red lines in plots of Left Hemisphere LH dACC and Right Hemisphere RH dACC in Fig. 16.5), striatum and brainstem areas of healthy controls while there was no increased activation for antisaccade errors compared to correct responses for patients with schizophrenia (see blue lines in plots of Left Hemisphere LH dACC and Right Hemisphere RH dACC in Fig. 16.5). This suggested that the reinforcement learning network after an error commission was dysfunctional in patients. Similarly the controls activated the emotional error processing areas after an error commission such as the rostral anterior cingulate cortex and the insula (see red lines in plots of Right Hemisphere, RH rACC and RH Insula on Fig. 16.5) more compared to correct antisaccades while again the patients failed to increase the activation of these areas during error commission, suggesting a dysfunction of the emotional error processing network.

Klein, Heinks, Andresen, Berg, and Moritz (2000) measured event related cortical potentials during the performance of visually guided saccades and antisaccades in patients with schizophrenia and healthy controls. They observed a vertex predominant contingent negative variation (CNV) potential that was larger preceding antisaccades compared to visually guided saccades in healthy subjects. This augmentation of the CNV for antisaccades was absent in patients with schizophrenia. In

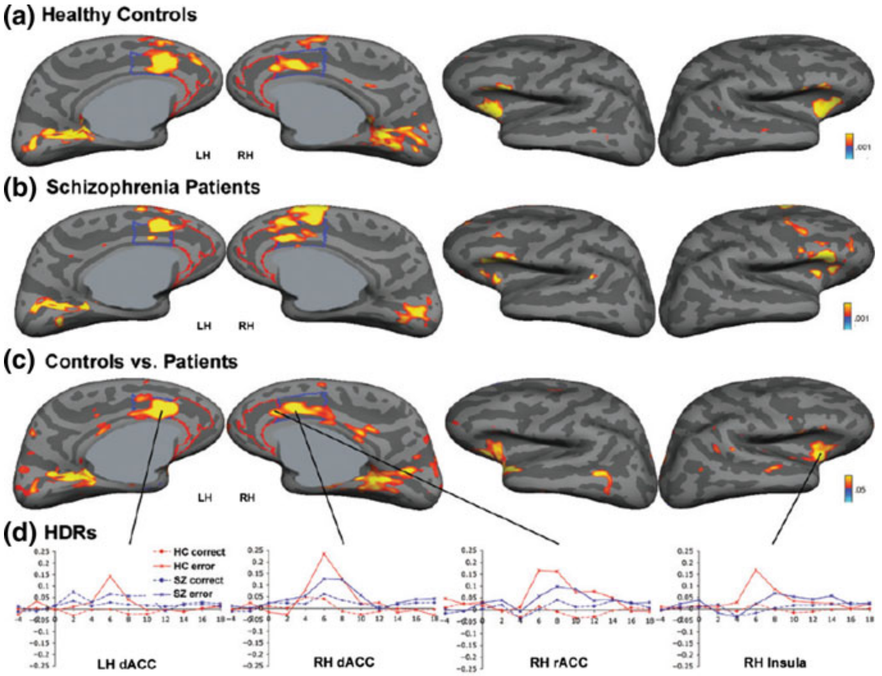


Fig. 16.5 Cortical activation for the contrast of error versus correct antisaccade at 6 s in **a** healthy controls, **b** schizophrenia patients and **c** the group comparison displayed on the lateral and medial views of the inflated cortical surface. rACC is outlined in red and dACC in blue. The grey masks cover non-surface regions in which activity is displaced. **d** The HDR time course graphs are displayed for vertices with peak activation in selected regions for each group and each condition. Time in seconds is on the x-axis and percent signal change relative to the fixation baseline is on the y-axis (reprinted with permission from Polli et al., 2008)

a subsequent study, Kirenskaya et al. (2013) did not confirm this difference in CNV between patients and controls but showed another difference in an early component of the CNV response.

The study of Camchong, Dyckman, Chapman, Yanasak, and McDowell (2006) examined activation differences between patients with schizophrenia and healthy controls in a memory guided saccade task and demonstrated that patients showed reduced activation relative to controls in the left medial FEF, the SEF bilaterally and the left inferior parietal lobule. Also healthy controls activated regions that were not activated in the patient group such as the right Brodmann 9 area and bilateral Brodmann 10 areas of the prefrontal cortex, the bilateral insula, the right putamen the left thalamus. In a fMRI study of prosaccades and predictive saccades, patients showed reduced activation in the dorsal neocortical visual attention network prior to treatment. Activation deficits were significantly reduced post-treatment. For the predictive saccade task, patients' dorsolateral prefrontal cortex (DLPFC) was unimpaired prior to treatment but showed significantly reduced activation after treatment, suggesting that antipsychotic medications may reduce hypofunction in attention systems related to eye movement control, but increase alterations in DLPFC—consistent with effects seen in the memory guided saccade task (Keedy et al., 2015).

In conclusion, neuroimaging and electrophysiological studies have shown differences in brain activity in areas related to saccadic eye movement control between patients with schizophrenia and healthy controls and maybe also between the first degree relatives of patients and controls (Table 16.3). These differences though were not verified in the same regions between different studies and many conflicting results have emerged. These results point again to the heterogeneity of the schizophrenia syndrome and maybe to the complexity of the neural networks involved in the production of saccades especially when considering volitional saccade tasks such as the antisaccade and the memory-guided saccade task.

16.5 Conclusions and Future Directions

As stated in the introduction we tried to address in this chapter a set of basic questions regarding the use of oculomotor function paradigms in psychiatric research. Focusing mainly on schizophrenia, the best studied psychiatric disorder in this field we will try to provide concluding answers to these questions given currently available knowledge:

(a) *What are the differences between patients and controls in the particular eye movement paradigm and what is the sensitivity of these differences in dissociating patients from controls?*

A large body of literature confirms that a SPEM deficit is present in schizophrenia patients. This deficit seems to be global, affecting more than one component of the pursuit system. The most robust measure of pursuit deficit in schizophrenia is the decrease in gain that has been replicated in large sample sizes and has the largest effect size compared to other specific measures of pursuit performance. Differences in performance were observed in patients with schizophrenia compared to controls in a variety of saccadic eye movement tasks using different parameters measured in each one of these tasks. In terms of robustness of these effects, the increase in antisaccade error rate is the best replicated finding in this large literature with the largest effect size in dissociating patients from controls at the group level. Future studies using large sample sizes and standardized paradigms are needed to provide definite evidence regarding the deficits in schizophrenia in this large variety of saccadic eye movement tasks. The sensitivity of all oculomotor function deficits to dissociate patients from healthy controls is not high enough to be clinically relevant for diagnostic purposes, and the differences are observed only at the group level. The fact that schizophrenia and other psychosis spectrum disorders are probably phenomenologically related syndromes with multiple underlying pathophysiological mechanisms could lead to the hypothesis that specific sub-groups of these individuals would manifest specific oculomotor function deficits. The definition of such specific vulnerable groups could be a future target of research.

(b) *Are these differences specific to different psychiatric syndromes?*

The same qualitative deficit in SPEM that is observed in schizophrenia has also been found in affective psychoses (major depression and bipolar disorder) and extends to the psychotic spectrum traits in the population of healthy individuals, with decreasing magnitude as one moves away from the schizophrenia syndrome to affective disorders. In contrast to psychotic spectrum disorders, in anxiety disorders such as OCD, the SPEM function seems to be spared, though some data suggest atypical intrusive saccades during pursuit in this disorder. Similarly, saccadic eye movement abnormalities are not restricted to patients with schizophrenia but are also observed for other patients with affective psychotic disorders and other psychiatric conditions such as anxiety disorders (for example OCD). The pattern of these abnormalities seems to be qualitatively similar for patients with psychotic disorders, where a deficit in volitional saccadic control can be observed, which quantitatively decreases in the psychotic spectrum from psychosis to affective disorders. The relation of saccadic abnormalities to other psychiatric conditions such as OCD is much less consistent across paradigms and studies. Future studies should address the issue of differences in oculomotor function among psychiatric syndromes and their significance in delineating the neurophysiological basis of these syndromes.

(c) *Are these differences related to specific psychopathology states and medication status?*

The SPEM deficits in psychotic disorders are present for drug naive patients and patients experiencing a first psychotic episode and remain stable during the course of the disorder. They are, for the most part, not related to symptom severity. The performance of visually guided saccades, memory-guided saccades and predictive saccades may be affected by the medication status and/or clinical symptoms of the patients, while the antisaccade task performance deviance (especially the error rate) is independent of clinical state, medication status and psychopathological manifestations of the disorder. It can be concluded that specific oculomotor function deficits have trait characteristics and are independent from the natural course of psychotic disorders and medication status while other oculomotor function deficits may to a degree vary over time, being more sensitive to medication status, course of the disorder and clinical manifestations. Future studies will determine if oculomotor functions could be used as biological markers for predicting the functional outcome of these disorders, or the clinical response to various treatments.

(d) *Are the differences observed between patients and controls in the particular eye movement paradigm hereditary and what is the nature of the genetic factors influencing these differences?*

There is converging evidence that SPEM deficits in psychiatric disorders, especially those of the psychotic spectrum, are heritable. The heritability of saccadic task performance deficits in schizophrenia and other disorders of the psychotic spectrum has also received support, although still not conclusive. The most widely studied measure has been the antisaccade error rate which has been shown to be heritable in healthy individuals as well as psychotic patients, but whether and how this heritability is directly related to psychotic disorders and especially schizophrenia is unclear. Molecular genetic studies have failed so far to reveal specific genetic loci that underlie the heritability of the SPEM and saccadic eye movement deficits in psychiatric disorders. SPEM function is a complex phenotype and future studies should rely on much larger sample sizes (in the magnitude of thousands of individuals) in the effort to map the genetic loci that contribute to SPEM and saccadic function and dysfunction in psychiatric syndromes.

(e) *What is the underlying neurobiological substrate of the differences between patients and controls in the particular eye movement paradigm?*

Although the neurobiological substrate for the performance of SPEM and saccadic eye movement tasks is well specified, the neurobiological substrate of the SPEM and saccadic eye movement control deficit in schizophrenia and other psychiatric disorders is not entirely clear. In part this is due to divergences in findings across studies in the same regions. One possible explanation for these somewhat contradictory results might be the fact that all these studies were performed on small samples, medication status was rarely controlled, acute illness level of patients was not consistent and studies varied in analytic procedures. Given the large heterogeneity of patient characteristics the need for larger sample sizes with standardized oculomotor paradigms and procedures, the use of drug naive patient samples and an effort for replication of findings in different samples will be important in future studies. On the other hand, these results point to the complexity of the neural networks involved in the production of oculomotor functions such as SPEM and volitional saccade tasks such as the antisaccade and the memory-guided saccade task and the need for refined tools for assessing the activity of these networks using neuroimaging and electrophysiological recording techniques.

16.6 Suggested Readings

Holzman P.S. (2000). Eye movements and the search for the essence of schizophrenia. *Brain Research Reviews 31,* 350–356.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: a critical review. *Psychophysiology*, *43*(3), 302–313.

Gooding, D. C., & Basso, M. A. (2008). The tell-tale tasks: a review of saccadic research in psychiatric patient populations. *Brain and Cognition*, *68*(3), 371–390.

O'Driscoll, G. A., & Callahan, B. L. (2008). Smooth pursuit in schizophrenia: a meta-analytic review of research since 1993. *Brain and Cognition*, *68*(3), 359–370.

Smyrnis, N. (2008). Metric issues in the study of eye movements in psychiatry. *Brain and Cognition*, *68*(3), 341–358.

16.7 Questions that Students Should Be Able to Answer

1. What is the major difference in pursuit performance between schizophrenia patients and healthy controls?
2. What parameters dissociate smooth pursuit performance in schizophrenia, affective disorders (bipolar, major depression) and obsessive compulsive disorder?
3. What parameters of smooth eye pursuit performance can be considered as endophenotypes of schizophrenia and why?
4. What brain areas are associated with the smooth eye pursuit deficit in schizophrenia?
5. What is the major difference in visually guided saccade performance between schizophrenia patients and healthy controls?
6. What is the major difference in antisaccade performance between schizophrenia patients and healthy controls?
7. What is the major difference in memory saccade performance between schizophrenia patients and healthy controls?
8. What is the major difference in predictive saccade performance between schizophrenia patients and healthy controls?
9. What parameters dissociate antisaccade pursuit performance in schizophrenia, affective disorders and obsessive compulsive disorder?
10. What parameters of antisaccade performance can be considered as endophenotypes and why?
11. What parameters of memory saccade performance can be considered as endophenotypes and why?
12. What brain areas are associated with the antisaccade pursuit deficit in schizophrenia?
13. What brain areas are associated with the memory saccade deficit in schizophrenia?

References

Abel, L. A., Friedman, L., Jesberger, J., Malki, A., & Meltzer, H. Y. (1991). Quantitative assessment of smooth pursuit gain and catch-up saccades in schizophrenia and affective disorders. *Biological Psychiatry, 29*(11), 1063–1072.

Abel, L. A., & Ziegler, A. S. (1988). Smooth pursuit eye movements in schizophrenics—What constitutes quantitative assessment? *Biological Psychiatry, 24*(7), 747–761.

Allen, J. S., Lambert, A. J., Johnson, F. Y., Schmidt, K., & Nero, K. L. (1996). Antisaccadic eye movements and attentional asymmetry in schizophrenia in three Pacific populations. *Acta Psychiatrica Scandinavica, 94*(4), 258–265.

Amado, I., Landgraf, S., Bourdel, M. C., Leonardi, S., & Krebs, M. O. (2008). Predictive saccades are impaired in biological nonpsychotic siblings of schizophrenia patients. *Journal of Psychiatry & Neuroscience: JPN, 33*(1), 17–22.

Antoniades, C., Ettinger, U., Gaymard, B., Gilchrist, I., Kristjánsson, A., Kennard, C. … Tarnowski, A. (2013). An internationally standardised antisaccade protocol. *Vision research*, *84*, 1–5.

Arolt, V., Teichert, H. M., Steege, D., Lencer, R., & Heide, W. (1998). Distinguishing schizophrenic patients from healthy controls by quantitative measurement of eye movement parameters. *Biological Psychiatry, 44*(6), 448–458.

Armstrong, T., & Olatunji, B. O. (2012). Eye tracking of attention in the affective disorders: A meta-analytic review and synthesis. *Clinical Psychology Review, 32*(8), 704–723.

Barton, J. J., Cherkasova, M. V., Lindgren, K., Goff, D. C., Intriligator, J. M., & Manoach, D. S. (2002). Antisaccades and task switching. *Annals of the New York Academy of Sciences, 956*(1), 250–263.

Beedie, S. (2011). Atypical scanpaths in schizophrenia: Evidence of a trait- or state-dependent phenomenon? *Journal of Psychiatry & Neuroscience, 36*(3), 150–164.

Brenner, C. A., McDowell, J. E., Cadenhead, K. S., & Clementz, B. A. (2001). Saccadic inhibition among schizotypal personality disorder subjects. *Psychophysiology, 38*(3), 399–403.

Broerse, A., Crawford, T. J., & Den Boer, J. A. (2002). Differential effects of olanzapine and risperidone on cognition in schizophrenia? A saccadic eye movement study. *The Journal of Neuropsychiatry and Clinical Neurosciences, 14*(4), 454–460.

Boudet, C., Bocca, M. L., Chabot, B., Delamillieure, P., Brazo, P., Denise, P., & Dollfus, S. (2005). Are eye movement abnormalities indicators of genetic vulnerability to schizophrenia? *European Psychiatry, 20*(4), 339–345.

Caldani, S., Bucci, M. P., Lamy, J. C., Seassau, M., Bendjemaa, N., Gadel, R., et al. (2017). Saccadic eye movements as markers of schizophrenia spectrum: exploration in at risk mental states. *Schizophrenia Research, 181,* 30–37.

Calkins, M. E., Iacono, W. G., & Ones, D. S. (2008). Eye movement dysfunction in first-degree relatives of patients with schizophrenia: A meta-analytic evaluation of candidate endophenotypes. *Brain and Cognition, 68*(3), 436–461.

Camchong, J., Dyckman, K. A., Austin, B. P., Clementz, B. A., & McDowell, J. E. (2008). Common neural circuitry supporting volitional saccades and its disruption in schizophrenia patients and relatives. *Biological Psychiatry, 64*(12), 1042–1050.

Camchong, J., Dyckman, K. A., Chapman, C. E., Yanasak, N. E., & McDowell, J. E. (2006). Basal ganglia-thalamocortical circuitry disruptions in schizophrenia during delayed response tasks. *Biological Psychiatry, 60*(3), 235–241.

Clementz, B. A., Farber, R. H., Lam, M. N., & Swerdlow, N. R. (1996a). Ocular motor responses to unpredictable and predictable smooth pursuit stimuli among patients with obsessive-compulsive disorder. *Journal of Psychiatry and Neuroscience, 21*(1), 21.

Clementz, B. A., Grove, W. M., Iacono, W. G., & Sweeney, J. A. (1992). Smooth-pursuit eye movement dysfunction and liability for schizophrenia: Implications for genetic modeling. *Journal of Abnormal Psychology, 101*(1), 117.

Clementz, B. A., Iacono, W. G., & Grove, W. M. (1996b). The construct validity of root-mean-square error for quantifying smooth-pursuit eye tracking abnormalities in schizophrenia. *Biological Psychiatry, 39*(6), 448–450.

Clementz, B. A., & McDowell, J. E. (1994). Smooth pursuit in schizophrenia: Abnormalities of open-and closed-loop responses. *Psychophysiology, 31*(1), 79–86.

Clementz, B. A., McDowell, J. E., & Zisook, S. (1994). Saccadic system functioning among schizophrenia patients and their first-degree biological relatives. *Journal of Abnormal Psychology, 103*(2), 277.

Clementz, B. A., Reid, S. A., McDowell, J. E., & Cadenhead, K. S. (1995). Abnormality of smooth pursuit eye movement initiation: specificity to the schizophrenia spectrum? *Psychophysiology, 32*(2), 130–134.

Clementz, B. A., & Sweeney, J. A. (1990). Is eye movement dysfunction a biological marker for schizophrenia? A methodological review. *Psychological bulletin, 108*(1), 77.

Clementz, B. A., Sweeney, J. A., Hamm, J. P., Ivleva, E. I., Ethridge, L. E., Pearlson, G. D. … Tamminga, C. A. (2015). Identification of distinct psychosis biotypes using brain-based biomarkers. *American Journal of Psychiatry*, *173*(4), 373–384.

Clementz, B. A., Sweeney, J. A., Hirt, M., & Haas, G. (1990). Pursuit gain and saccadic intrusions in first-degree relatives of probands with schizophrenia. *Journal of Abnormal Psychology, 99*(4), 327.

Clementz, B. A., Sweeney, J. A., Hirt, M., & Haas, G. (1991). Phenotypic correlations between oculomotor functioning and schizophrenia-related characteristics in relatives of schizophrenic probands. *Psychophysiology, 28*(5), 570–578.

Crawford, T. J., Haeger, B., Kennard, C., Reveley, M. A., & Henderson, L. (1995). Saccadic abnormalities in psychotic patients. I. Neuroleptic-free psychotic patients. *Psychological Medicine, 25*(03), 461–471.

Crawford, T. J., Sharma, T., Puri, B. K., Murray, R. M., Berridge, D. M., & Lewis, S. W. (1998). Saccadic eye movements in families multiply affected with schizophrenia: The Maudsley family study. *American Journal of Psychiatry, 155*(12), 1703–1710.

Crevits, L., Van den Abbeele, D., Audenaert, K., Goethals, M., & Dierick, M. (2005). Effect of repetitive transcranial magnetic stimulation on saccades in depression: A pilot study. *Psychiatry Research, 135*(2), 113–119.

Curtis, C. E., Calkins, M. E., Grove, W.M., Feil, K. J., & Iacono, W. G. (2001). Saccadic disinhibition in acute and remitted schizophrenia and their first-degree biological relatives. *American Journal of Psychiatry,* 158, 100–106.

Damilou, A., Apostolakis, S., Thrapsanioti, E., Theleritis, C., & Smyrnis, N. (2016). Shared and distinct oculomotor function deficits in schizophrenia and obsessive compulsive disorder. *Psychophysiology, 53,* 796–805.

Diefendorf, A. R., & Dodge, R. (1908). An experimental study of the ocular reactions of the insane from photographic records. *Brain, 31*(3), 451–489.

Dyckman, K. A., Lee, A. K., Agam, Y., Vangel, M., Goff, D. C., Barton, J. J., & Manoach, D. S. (2011). Abnormally persistent fMRI activation during antisaccades in schizophrenia: a neural correlate of perseveration? *Schizophrenia Research*, *132*(1), 62–68.

Egan, M. F., Goldberg, T. E., Kolachana, B. S., Callicott, J. H., Mazzanti, C. M., Straub, R. E. … Weinberger, D. R. (2001). Effect of COMT Val108/158 Met genotype on frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of Sciences, 98*(12), 6917–6922.

Ettinger, U., Kumari, V., Chitnis, X. A., Corr, P. J., Crawford, T. J., Fannon, D. G. … Sharma, T. (2004). Volumetric neural correlates of antisaccade eye movements in first-episode psychosis. *American Journal of Psychiatry, 161*(10), 1918–1921.

Ettinger, U., Kumari, V., Crawford, T. J., Davis, R. E., Sharma, T., & Corr, P. J. (2003). Reliability of smooth pursuit, fixation, and saccadic eye movements. *Psychophysiology, 40*(4), 620–628.

Ettinger, U., Kumari, V., Crawford, T. J., Flak, V., Sharma, T., Davis, R. E., & Corr, P. J. (2005). Saccadic eye movements, schizotypy, and the role of neuroticism. *Biological Psychology, 68*(1), 61–78.

Evdokimidis, I., Smyrnis, N., Constantinidis, T., Stefanis, N., Avramopoulos, D., Paximadis, C. … Stefanis, C. (2002). The antisaccade task in a sample of 2,006 young men. *Experimental Brain Research, 147*(1), 45–52.

Farber, R. H., Clementz, B. A., & Swerdlow, N. R. (1997). Characteristics of open-and closed-loop smooth pursuit responses among obsessive-compulsive disorder, schizophrenia, and nonpsychiatric individuals. *Psychophysiology, 34*(2), 157–162.

Flechtner, K. M., Steinacher, B., Sauer, R., & Mackert, A. (2002). Smooth pursuit eye movements of patients with schizophrenia and affective disorder during clinical treatment. *European Archives of Psychiatry and Clinical Neuroscience, 252*(2), 49–53.

Fukushima, J., Fukushima, K., Chiba, T., Tanaka, S., Yamashita, I., & Kato, M. (1988). Disturbances of voluntary control of saccadic eye movements in schizophrenic patients. *Biological Psychiatry, 23*(7), 670–677.

Fukushima, J., Fukushima, K., Morita, N., & Yamashita, I. (1991). Disturbances in the control of saccadic eye movement and eye-head coordination in schizophrenics. *Journal of Vestibular Research: Equilibrium & Orientation*.

Funahashi, S., Bruce, C. J., & Goldman-Rakic, P. S. (1989). Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *Journal of Neurophysiology, 61*(2), 331–349.

Gale, H. J., & Holzman, P. S. (2000). A new look at reaction time in schizophrenia. *Schizophrenia Research, 46*(2), 149–165.

Gambini, O., Abbruzzese, M., & Scarone, S. (1993). Smooth pursuit and saccadic eye movements and Wisconsin Card Sorting Test performance in obsessive-compulsive disorder. *Psychiatry Research, 48*(3), 191–200.

Gooding, D. C., & Basso, M. A. (2008). The tell-tale tasks: A review of saccadic research in psychiatric patient populations. *Brain and Cognition, 68*(3), 371–390.

Gooding, D. C., Mohapatra, L., & Shea, H. B. (2004). Temporal stability of saccadic task performance in schizophrenia and bipolar patients. *Psychological Medicine, 34*(05), 921–932.

Gooding, D. C., & Tallent, K. A. (2001). The association between antisaccade task and working memory task performance in schizophrenia and bipolar disorder. *The Journal of Nervous and Mental Disease, 189*(1), 8–16.

Greenwood, T. A., Braff, D. L., Light, G. A., Cadenhead, K. S., Calkins, M. E., Dobie, D. J. … Mintz, J. (2007). Initial heritability analyses of endophenotypic measures for schizophrenia: the consortium on the genetics of schizophrenia. *Archives of General Psychiatry, 64*(11), 1242–1250.

Grootens, K. P., van Luijtelaar, G., Buitelaar, J. K., van der Laan, A., Hummelen, J. W., & Verkes, R. J. (2008). Inhibition errors in borderline personality disorder with psychotic-like symptoms. *Progress in Neuro-psychopharmacology and Biological Psychiatry, 32*(1), 267–273.

Haraldsson, H. M., Ettinger, U., Magnusdottir, B. B., Sigmundsson, T., Sigurdsson, E., Ingason, A., & Petursson, H. (2009). COMT val 158 met genotype and smooth pursuit eye movements in schizophrenia. *Psychiatry Research, 169*(2), 173–175.

Hardeman, H. H., Ramsey, N. F., Raefmakers, M., Van Megen, H. J., Denys, D. A., Westenberg, H. G., & Kahn, R. S. (2006). Saccadic abnormalities in psychotropic-naive obsessive–compulsive disorder without co-morbidity. *Psychological Medicine, 36*(09), 1321–1326.

Harris, M. S., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2006). Longitudinal studies of antisaccades in antipsychotic-naive first-episode schizophrenia. *Psychological Medicine, 36*(4), 485.

Harris, M. S., Reilly, J. L., Thase, M. E., Keshavan, M. S., & Sweeney, J. A. (2009a). Response suppression deficits in treatment-naive first-episode patients with schizophrenia, psychotic bipolar disorder and psychotic major depression. *Psychiatry Research, 170*(2), 150–156.

Harris, M. S., Wiseman, C. L., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2009b). Effects of risperidone on procedural learning in antipsychotic-naive first-episode schizophrenia. *Neuropsychopharmacology, 34*(2), 468–476.

Hatzimanolis, A., Bhatnagar, P., Moes, A., Wang, R., Roussos, P., Bitsios, P. … Stefanis, N. C. (2015). Common genetic variation and schizophrenia polygenic risk influence neurocognitive performance in young adulthood. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 168*(5), 392–401.

Hill, S. K., Reilly, J. L., Harris, M. S., Khine, T., & Sweeney, J. A. (2008). Oculomotor and neuropsychological effects of antipsychotic treatment for schizophrenia. *Schizophrenia Bulletin, 34*(3), 494–506.

Holzman, P. (2000). Eye movements and the search for the essence of schizophrenia. *Brain Research Reviews, 31*(2–3), 350–356.

Holahan, A. L. V., & O'Driscoll, G. A. (2005). Antisaccade and smooth pursuit performance in positive-and negative-symptom schizotypy. *Schizophrenia Research, 76*(1), 43–54.

Holzman, P. S., Proctor, L. R., & Hughes, D. W. (1973). Eye-tracking patterns in schizophrenia. *Science, 181*(4095), 179–181.

Holzman, P. S., Proctor, L. R., Levy, D. L., Yasillo, N. J., Meltzer, H. Y., & Hurt, S. W. (1974). Eye-tracking dysfunctions in schizophrenic patients and their relatives. *Archives of General Psychiatry, 31*(2), 143–151.

Hong, L. E., Mitchell, B. D., Avila, M. T., Adami, H., McMahon, R. P., & Thaker, G. K. (2006). Familial aggregation of eye-tracking endophenotypes in families of schizophrenic patients. *Archives of General Psychiatry, 63*(3), 259–264.

Hong, L. E., Turano, K. A., O'Neill, H., Hao, L., Wonodi, I., McMahon, R. P. … Thaker, G. K. (2008). Refining the predictive pursuit endophenotype in schizophrenia. *Biological Psychiatry, 63*(5), 458–464.

Hutton, S. B., Crawford, T. J., Puri, B. K., Duncan, L. J., Chapman, M., Kennard, C. … Joyce, E. M. (1998). Smooth pursuit and saccadic abnormalities in first-episode schizophrenia. *Psychological Medicine, 28*(03), 685–692.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: a critical review. *Psychophysiology, 43*(3), 302–313.

Hutton, S. B., Joyce, E. M., Barnes, T. R. E., & Kennard, C. (2002). Saccadic distractibility in first-episode schizophrenia. *Neuropsychologia, 40*(10), 1729–1736.

Iacono, W. G., Moreau, M., Beiser, M., Fleming, J. A., & Lin, T. Y. (1992). Smooth-pursuit eye tracking in first-episode psychotic patients and their relatives. *Journal of Abnormal Psychology, 101*(1), 104.

Iacono, W. G., Tuason, V. B., & Johnson, R. A. (1981). Dissociation of smooth-pursuit and saccadic eye tracking in remitted schizophrenics: An ocular reaction time task that schizophrenics perform well. *Archives of General Psychiatry, 38*(9), 991–996.

Ivleva, E. I., Moates, A. F., Hamm, J. P., Bernstein, I. H., O'Neill, H. B., Cole, D. … Tamminga, C. A. (2014). Smooth pursuit eye movement, prepulse inhibition, and auditory paired stimuli processing endophenotypes across the schizophrenia-bipolar disorder psychosis dimension. *Schizophrenia Bulletin, 40*(3), 642–652.

Jaafari, N., Baup, N., Bourdel, M. C., Olié, J. P., Rotge, J. Y., Wassouf, I. … Krebs, M. O. (2011). Neurological soft signs in OCD patients with early age at onset, versus patients with schizophrenia and healthy subjects. *The Journal of Neuropsychiatry and Clinical Neurosciences, 23*(4), 409–416.

Kallimani, D., Theleritis, C., Evdokimidis, I., Stefanis, N. C., Chatzimanolis, I., & Smyrnis, N. (2009). The effect of change in clinical state on eye movement dysfunction in schizophrenia. *European Psychiatry, 24*(1), 17–26.

Karantinos, T., Tsoukas, E., Mantas, A., Kattoulas, E., Stefanis, N. C., Evdokimidis, I., & Smyrnis, N. (2014). Increased intra-subject reaction time variability in the volitional control of movement in schizophrenia. *Psychiatry Research, 215*(1), 26–32.

Karoumi, B., Ventre-Dominey, J., Vighetto, A., Dalery, J., & d'Amato, T. (1998). Saccadic eye movements in schizophrenic patients. *Psychiatry Research, 77*(1), 9–19.

Kathmann, N., Hochrein, A., Uwer, R., & Bondy, B. (2003). Deficits in gain of smooth pursuit eye movements in schizophrenia and affective disorder patients and their unaffected relatives. *American Journal of Psychiatry, 160*(4), 696–702.

Katsanis, J., Kortenkamp, S., Iacono, W. G., & Grove, W. M. (1997). Antisaccade performance in patients with schizophrenia and affective disorder. *Journal of Abnormal Psychology, 106*(3), 468.

Kattoulas, E., Evdokimidis, I., Stefanis, N. C., Avramopoulos, D., Stefanis, C. N., & Smyrnis, N. (2011a). Predictive smooth eye pursuit in a population of young men: II. Effects of schizotypy, anxiety and depression. *Experimental Brain Research, 215*(3–4), 219–226.

Kattoulas, E., Smyrnis, N., Stefanis, N. C., Avramopoulos, D., Stefanis, C. N., & Evdokimidis, I. (2011b). Predictive smooth eye pursuit in a population of young men: I. Effects of age, IQ, oculomotor and cognitive tasks. *Experimental brain research, 215*(3–4), 207–218.

Kattoulas, E., Stefanis, N. C., Avramopoulos, D., Stefanis, C. N., Evdokimidis, I., & Smyrnis, N. (2012). Schizophrenia-related RGS4 gene variations specifically disrupt prefrontal control of saccadic eye movements. *Psychological Medicine, 42*(04), 757–767.

Keedy, S. K., Ebens, C. L., Keshavan, M. S., & Sweeney, J. A. (2006). Functional magnetic resonance imaging studies of eye movements in first episode schizophrenia: smooth pursuit, visually guided saccades and the oculomotor delayed response task. *Psychiatry Research: Neuroimaging, 146*(3), 199–211.

Keedy, S. K., Reilly, J. L., Bishop, J. R., Weiden, P. J., & Sweeney, J. A. (2014). Impact of antipsychotic treatment on attention and motor learning systems in first-episode schizophrenia. *Schizophrenia Bulletin*, sbu071.

Kim, J. H., Park, B. L., Pasaje, C. F. A., Bae, J. S., Park, C. S., Cha, B. … Choi, I. G. (2012). Lack of associations of neuregulin 1 variations with schizophrenia and smooth pursuit eye movement abnormality in a Korean population. *Journal of Molecular Neuroscience, 46*(3), 476–482.

Kirenskaya, A. V., Kamenskov, M. Y., Myamlin, V. V., Novototsky-Vlasov, V. Y., & Tkachenko, A. A. (2013). The antisaccade task performance deficit and specific CNV abnormalities in patients with stereotyped paraphilia and schizophrenia. *Journal of Forensic Sciences, 58*(5), 1219–1226.

Klein, C., & Andresen, B. (1991). On the influence of smoking upon smooth pursuit eye movements of schizophrenics and normal controls. *Journal of Psychophysiology, 5*(4), 361–369.

Klein, C., Heinks, T., Andresen, B., Berg, P., & Moritz, S. (2000). Impaired modulation of the saccadic contingent negative variation preceding antisaccades in schizophrenia. *Biological Psychiatry, 47*(11), 978–990.

Kloft, L., Kischkel, E., Kathmann, N., & Reuter, B. (2011). Evidence for a deficit in volitional action generation in patients with obsessive-compulsive disorder. *Psychophysiology, 48*(6), 755–761.

Kraepelin, E., Barclay, R. M., & Robertson, G. M. (1919). Dementia praecox.

Krebs, M. O., Bourdel, M. C., Cherif, Z. R., Bouhours, P., Lôo, H., Poirier, M. F., & Amado, I. (2010). Deficit of inhibition motor control in untreated patients with schizophrenia: Further support from visually guided saccade paradigms. *Psychiatry Research, 179*(3), 279–284.

Krebs, M. O., Gut-Fayand, A., Amado, I., Daban, C., Bourdel, M. C., Poirier, M. F., & Berthoz, A. (2001). Impairment of predictive saccades in schizophrenia. *Neuroreport, 12*(3), 465–469.

Landgraf, S., Amado, I., Bourdel, M. C., Leonardi, S., & Krebs, M. O. (2008). Memory-guided saccade abnormalities in schizophrenic patients and their healthy, full biological siblings. *Psychological Medicine, 38*(06), 861–870.

Lencer, R., Bishop, J. R., Harris, M. S., Reilly, J. L., Patel, S., Kittles, R. … Sweeney, J. A. (2014). Association of variants in DRD2 and GRM3 with motor and cognitive function in first-episode psychosis. *European Archives of Psychiatry and Clinical Neuroscience, 264*(4), 345–355.

Lencer, R., Keedy, S. K., Reilly, J. L., McDonough, B. E., Harris, M. S., Sprenger, A., & Sweeney, J. A. (2011). Altered transfer of visual motion information to parietal association cortex in untreated first-episode psychosis: Implications for pursuit eye tracking. *Psychiatry Research: Neuroimaging, 194*(1), 30–38.

Lencer, R., Nagel, M., Sprenger, A., Heide, W., & Binkofski, F. (2005). Reduced neuronal activity in the V5 complex underlies smooth-pursuit deficit in schizophrenia: evidence from an fMRI study. *Neuroimage, 24*(4), 1256–1259.

Lencer, R., Nagel, M., Sprenger, A., Zapf, S., Erdmann, C., Heide, W., & Binkofski, F. (2004). Cortical mechanisms of smooth pursuit eye movements with target blanking. An fMRI study. *European Journal of Neuroscience, 19*(5), 1430–1436.

Lencer, R., Reilly, J. L., Harris, M. S., Sprenger, A., Keshavan, M. S., & Sweeney, J. A. (2010). Sensorimotor transformation deficits for smooth pursuit in first-episode affective psychoses and schizophrenia. *Biological Psychiatry, 67*(3), 217–223.

Lencer, R., Sprenger, A., Harris, M. S., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2008). Effects of second-generation antipsychotic medication on smooth pursuit performance in antipsychotic-naive schizophrenia. *Archives of General Psychiatry, 65*(10), 1146–1154.

Lencer, R., Sprenger, A., Reilly, J. L., McDowell, J. E., Rubin, L. H., Badner, J. A. … Clementz, B. A. (2015). Pursuit eye movements as an intermediate phenotype across psychotic disorders: Evidence from the B-SNIP study. *Schizophrenia Research*, *169*(1), 326–333.

Lencer, R., Trillenberg-Krecker, K., Schwinger, E., & Arolt, V. (2003). Schizophrenia spectrum disorders and eye tracking dysfunction in singleton and multiplex schizophrenia families. *Schizophrenia Research, 60*(1), 33–45.

Lencz, T., Knowles, E., Davies, G., Guha, S., Liewald, D. C., Starr, J. M. … Reinvang, I. (2014). Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: A report from the Cognitive Genomics consorTium (COGENT). *Molecular Psychiatry, 19*(2), 168–174.

Lennertz, L., Rampacher, F., Vogeley, A., Schulze-Rauschenbach, S., Pukrop, R., Ruhrmann, S. … Wagner, M. (2012). Antisaccade performance in patients with obsessive–compulsive disorder and unaffected relatives: further evidence for impaired response inhibition as a candidate endophenotype. *European Archives of Psychiatry and Clinical Neuroscience, 262*(7), 625–634.

Levin, S., Jones, A., Stark, L., Merrin, E. L., & Holzman, P. S. (1982). Saccadic eye movements of schizophrenic patients measured by reflected light technique. *Biological Psychiatry*.

Levy, D. L., Holzman, P. S., Matthysse, S., & Mendell, N. R. (1993). Eye tracking dysfunction and schizophrenia: A critical perspective. *Schizophrenia Bulletin, 19*(3), 461–536.

Levy, D. L., Holzman, P. S., Matthysse, S., & Mendell, N. R. (1994). Eye tracking and schizophrenia: A selective review. *Schizophrenia Bulletin, 20*(1), 47–62.

Levy, D. L., Lajonchere, C. M., Dorogusker, B., Min, D., Lee, S., Tartaglini, A. … Mendell, N. R. (2000). Quantitative characterization of eye tracking dysfunction in schizophrenia. *Schizophrenia Research, 42*(3), 171–185.

Mahlberg, R., Steinacher, B., Mackert, A., & Flechtner, K. M. (2001). Basic parameters of saccadic eye movements–differences between unmedicated schizophrenia and affective disorder patients. *European Archives of Psychiatry and Clinical Neuroscience, 251*(5), 205–210.

Manoach, D. S., Ketwaroo, G. A., Polli, F. E., Thakkar, K. N., Barton, J. J., Goff, D. C. … Tuch, D. S. (2007). Reduced microstructural integrity of the white matter underlying anterior cingulate cortex is associated with increased saccadic latency in schizophrenia. *Neuroimage, 37*(2), 599–610.

Martin, L. F., Hall, M. H., Ross, R. G., Zerbe, G., Freedman, R., & Olincy, A. (2007). Physiology of schizophrenia, bipolar disorder, and schizoaffective disorder. *American Journal of Psychiatry, 164*(12), 1900–1906.

Maruff, P., Purcell, R., Tyler, P., Pantelis, C., & Currie, J. (1999). Abnormalities of internally generated saccades in obsessive–compulsive disorder. *Psychological Medicine, 29*(06), 1377–1385.

McDowell, J. E., Brown, G. G., Paulus, M., Martinez, A., Stewart, S. E., Dubowitz, D. J., & Braff, D. L. (2002). Neural correlates of refixation saccades and antisaccades in normal and schizophrenia subjects. *Biological Psychiatry, 51*(3), 216–223.

McDowell, J. E., & Clementz, B. A. (1996). Ocular-motor delayed-response task performance among schizophrenia patients. *Neuropsychobiology, 34*(2), 67–71.

McDowell, J. E., & Clementz, B. A. (1997). The effect of fixation condition manipulations on antisaccade performance in schizophrenia: Studies of diagnostic specificity. *Experimental Brain Research, 115*(2), 333–344.

McDowell, J. E., & Clementz, B. A. (2001). Behavioral and brain imaging studies of saccadic performance in schizophrenia. *Biological Psychology, 57*(1), 5–22.

McDowell, J. E., Myles-Worsley, M., Coon, H., Byerley, W., & Clementz, B. A. (1999). Measuring liability for schizophrenia using optimized antisaccade stimulus parameters. *Psychophysiology, 36*(1), 138–141.

Millard, S. P., Shofer, J., Braff, D., Calkins, M., Cadenhead, K., Freedman, R. … Lazzeroni, L. C. (2016). Prioritizing schizophrenia endophenotypes for future genetic studies: An example using data from the COGS-1 family study. *Schizophrenia Research*, *174*(1), 1–9.

Mosconi, M. W., Kay, M., D'cruz, A. M., Seidenfeld, A., Guter, S., Stanford, L. D., & Sweeney, J. A. (2009). Impaired inhibitory control is associated with higher-order repetitive behaviors in autism spectrum disorders. *Psychological Medicine, 39*(09), 1559–1566.

Nagel, M., Sprenger, A., Nitschke, M., Zapf, S., Heide, W., Binkofski, F., & Lencer, R. (2007). Different extraretinal neuronal mechanisms of smooth pursuit eye movements in schizophrenia: An fMRI study. *Neuroimage, 34*(1), 300–309.

Nickoloff, S. E., Radant, A. D., Reichler, R., & Hommer, D. W. (1991). Smooth pursuit and saccadic eye movements and neurological soft signs in obsessive-compulsive disorder. *Psychiatry Research, 38*(2), 173–185.

Nieman, D., Becker, H., van de Fliert, R., Plat, N., Bour, L., Koelman, H. … Linszen, D. (2007). Antisaccade task performance in patients at ultra high risk for developing psychosis. *Schizophrenia Research, 95*(1), 54–60.

Nkam, I., Bocca, M. L., Denise, P., Paoletti, X., Dollfus, S., Levillain, D., & Thibaut, F. (2010). Impaired smooth pursuit in schizophrenia results from prediction impairment only. *Biological Psychiatry, 67*(10), 992–997.

Nkam, I., Thibaut, F., Denise, P., Van Der Elst, A., Segard, L., Brazo, P. … Vasse, T. (2001). Saccadic and smooth-pursuit eye movements in deficit and non-deficit schizophrenia. *Schizophrenia Research, 48*(1), 145–153.

O'Driscoll, G. A., Benkelfat, C., Florencio, P. S., Wolff, A. L. V., Joober, R., Lal, S., & Evans, A. C. (1999). Neural correlates of eye tracking deficits in first-degree relatives of schizophrenic patients: A positron emission tomography study. *Archives of General Psychiatry, 56*(12), 1127–1134.

O'Driscoll, G. A., & Callahan, B. L. (2008). Smooth pursuit in schizophrenia: A meta-analytic review of research since 1993. *Brain and Cognition, 68*(3), 359–370.

O'Driscoll, G. A., Lenzenweger, M. F., & Holzman, P. S. (1998). Antisaccades and smooth pursuit eye tracking and schizotypy. *Archives of General Psychiatry, 55*(9), 837–843.

Park, S., & Holzman, P. S. (1992). Schizophrenics show spatial working memory deficits. *Archives of General Psychiatry, 49*(12), 975–982.

Park, S., Holzman, P. S., & Goldman-Rakic, P. S. (1995). Spatial working memory deficits in the relatives of schizophrenic patients. *Archives of General Psychiatry, 52*(10), 821–828.

Park, B. L., Shin, H. D., Cheong, H. S., Park, C. S., Sohn, J. W., Kim, B. J. … Choi, I. G. (2009). Association analysis of COMT polymorphisms with schizophrenia and smooth pursuit eye movement abnormality. *Journal of Human Genetics, 54*(12), 709–712.

Polli, F. E., Barton, J. J., Thakkar, K. N., Greve, D. N., Goff, D. C., Rauch, S. L., et al. (2008). Reduced error-related activation in two anterior cingulate circuits is related to impaired performance in schizophrenia. *Brain, 131*(4), 971–986.

Radant, A. D., Dobie, D. J., Calkins, M. E., Olincy, A., Braff, D. L., Cadenhead, K. S. … Light, G. A. (2007). Successful multi-site measurement of antisaccade performance deficits in schizophrenia. *Schizophrenia Research, 89*(1), 320–329.

Radant, A. D., Dobie, D. J., Calkins, M. E., Olincy, A., Braff, D. L., Cadenhead, K. S. … Gur, R. C. (2010). Antisaccade performance in schizophrenia patients, their first-degree biological relatives, and community comparison subjects: Data from the COGS study. *Psychophysiology, 47*(5), 846–856.

Raemaekers, M., Jansma, J. M., Cahn, W., Van der Geest, J. N., van der Linden, J. A., Kahn, R. S., & Ramsey, N. F. (2002). Neuronal substrate of the saccadic inhibition deficit in schizophrenia investigated with 3-dimensional event-related functional magnetic resonance imaging. *Archives of General Psychiatry, 59*(4), 313–320.

Raemaekers, M., Vink, M., van den Heuvel, M. P., Kahn, R. S., & Ramsey, N. F. (2006). Effects of aging on BOLD fMRI during prosaccades and antisaccades. *Journal of Cognitive Neuroscience, 18*(4), 594–603.

Reilly, J. L., Frankovich, K., Hill, S., Gershon, E. S., Keefe, R. S., Keshavan, M. S. … Sweeney, J. A. (2014). Elevated antisaccade error rate as an intermediate phenotype for psychosis across diagnostic categories. *Schizophrenia Bulletin, 40*(5), 1011–1021.

Reilly, J. L., Harris, M. S., Keshavan, M. S., & Sweeney, J. A. (2005). Abnormalities in visually guided saccades suggest corticofugal dysregulation in never-treated schizophrenia. *Biological Psychiatry, 57*(2), 145–154.

Reilly, J. L., Harris, M. S., Keshavan, M. S., & Sweeney, J. A. (2006). Adverse effects of risperidone on spatial working memory in first-episode schizophrenia. *Archives of General Psychiatry, 63*(11), 1189–1197.

Reilly, J. L., Harris, M. S., Khine, T. T., Keshavan, M. S., & Sweeney, J. A. (2007). Antipsychotic drugs exacerbate impairment on a working memory task in first-episode schizophrenia. *Biological Psychiatry, 62*(7), 818–821.

Reilly, J. L., Harris, M. S., Khine, T. T., Keshavan, M. S., & Sweeney, J. A. (2008a). Reduced attentional engagement contributes to deficits in prefrontal inhibitory control in schizophrenia. *Biological Psychiatry, 63*(8), 776–783.

Reilly, J. L., Lencer, R., Bishop, J. R., Keedy, S., & Sweeney, J. A. (2008b). Pharmacological treatment effects on eye movement control. *Brain and Cognition, 68*(3), 415–435.

Reuter, B., Rakusan, L., & Kathmanna, N. (2005). Poor antisaccade performance in schizophrenia: an inhibition deficit? *Psychiatry Research, 135*(1), 1–10.

Rosano, C., Krisky, C. M., Welling, J. S., Eddy, W. F., Luna, B., Thulborn, K. R., & Sweeney, J. A. (2002). Pursuit and saccadic eye movement subregions in human frontal eye field: A high-resolution fMRI investigation. *Cerebral Cortex, 12*(2), 107–115.

Rosenberg, D. R., Averbach, D. H., O'hearn, K. M., Seymour, A. B., Birmaher, B., & Sweeney, J. A. (1997). Oculomotor response inhibition abnormalities in pediatric obsessive-compulsive disorder. *Archives of General Psychiatry, 54*(9), 831–838.

Rosenberg, D. R., Sweeney, J. A., Squires-Wheeler, E., Keshavan, M. S., Cornblatt, B. A., & Erlenmeyer-Kimling, L. (1997b). Eye-tracking dysfunction in offspring from the New York High-Risk Project: Diagnostic specificity and the role of attention. *Psychiatry Research, 66*(2), 121–130.

Ross, R. G., Olincy, A., Harris, J. G., Radant, A., Adler, L. E., & Freedman, R. (1998). Anticipatory saccades during smooth pursuit eye movements and familial transmission of schizophrenia. *Biological Psychiatry, 44*(8), 690–697.

Ross, D. E., Thaker, G. K., Buchanan, R. W., & Lahti, A. C. (1996). Association of abnormal smooth pursuit eye movements with the deficit syndrome in schizophrenic patients. *The American Journal of Psychiatry, 153*(9), 1158.

Rybakowski, J. K., & Borkowska, A. (2002). Eye movement and neuropsychological studies in first-degree relatives of schizophrenic patients. *Schizophrenia Research, 54*(1), 105–110.

Seidman, L. J., Hellemann, G., Nuechterlein, K. H., Greenwood, T. A., Braff, D. L., Cadenhead, K. S. … Green, M. F. (2015). Factor structure and heritability of endophenotypes in schizophrenia: Findings from the Consortium on the Genetics of Schizophrenia (COGS-1). *Schizophrenia Research,163*(1–3), 73–79.

Sereno, A. B., & Holzman, P. S. (1993). Express saccades and smooth pursuit eye movement function in schizophrenic, affective disorder, and normal subjects. *Journal of Cognitive Neuroscience, 5*(3), 303–316.

Sereno, A. B., & Holzman, P. S. (1995). Antisaccades and smooth pursuit eye movements in schizophrenia. *Biological Psychiatry, 37*(6), 394–401.

Smyrnis, N. (2008). Metric issues in the study of eye movements in psychiatry. *Brain and Cognition, 68*(3), 341–358.

Smyrnis, N., Evdokimidis, I., Mantas, A., Kattoulas, E., Stefanis, N. C., Constantinidis, T. S. … Stefanis, C. N. (2007). Smooth pursuit eye movements in 1,087 men: Effects of schizotypy, anxiety, and depression. *Experimental Brain Research, 179*(3), 397–408.

Smyrnis, N., Evdokimidis, I., Stefanis, N. C., Avramopoulos, D., Constantinidis, T. S., Stavropoulos, A., & Stefanis, C. N. (2003). Antisaccade performance of 1,273 men: Effects of schizotypy, anxiety, and depression. *Journal of Abnormal Psychology, 112*(3), 403.

Smyrnis, N., Karantinos, T., Malogiannis, I., Theleritis, C., Mantas, A., Stefanis, N. C. … Evdokimidis, I. (2009). Larger variability of saccadic reaction times in schizophrenia patients. *Psychiatry Research*, *168*(2), 129–136.

Smyrnis, N., Kattoulas, E., Stefanis, N. C., Avramopoulos, D., Stefanis, C. N., & Evdokimidis, I. (2011). Schizophrenia-related neuregulin-1 single-nucleotide polymorphisms lead to deficient smooth eye pursuit in a large sample of young men. *Schizophrenia Bulletin, 37*(4), 822–831.

Smyrnis, N., Malogiannis, I. A., Evdokimidis, I., Stefanis, N. C., Theleritis, C., Vaidakis, A. … & Stefanis, N. C. (2004). Attentional facilitation of response is impaired for antisaccades but not for saccades in patients with schizophrenia: Implications for cortical dysfunction. *Experimental Brain Research, 159,* 47–54.

Spengler, D., Trillenberg, P., Sprenger, A., Nagel, M., Kordon, A., Junghanns, K. … Lencer, R. (2006). Evidence from increased anticipation of predictive saccades for a dysfunction of fronto-striatal circuits in obsessive–compulsive disorder. *Psychiatry Research, 143*(1), 77–88.

Spering, M., Dias, E. C., Sanchez, J. L., Schütz, A. C., & Javitt, D. C. (2013). Efference copy failure during smooth pursuit eye movements in schizophrenia. *The Journal of Neuroscience, 33*(29), 11779–11787.

Sporn, A., Greenstein, D., Gogtay, N., Sailer, F., Hommer, D. W., Rawlings, R. … Weinberger, D. R. (2005). Childhood-onset schizophrenia: smooth pursuit eye-tracking dysfunction in family members. *Schizophrenia Research, 73*(2), 243–252.

Stefanis, N. C., Van Os, J., Avramopoulos, D., Smyrnis, N., Evdokimidis, I., Hantoumi, I., et al. (2004). Variation in catechol-o-methyltransferase val158 met genotype associated with schizotypy but not cognition: A population study in 543 young men. *Biological Psychiatry, 56*(7), 510–515.

Sweeney, J. A., Bauer, K. S., Keshavan, M. S., Haas, G. L., Schooler, N. R., & Kroboth, P. D. (1997). Adverse effects of risperidone on eye movement activity: A comparison of risperidone and haloperidol in antipsychotic-naive schizophrenic patients. *Neuropsychopharmacology, 16*(3), 217–228.

Sweeney, J. A., Clementz, B. A., Escobar, M. D., Li, S., Pauler, D. K., & Haas, G. L. (1993). Mixture analysis of pursuit eye-tracking dysfunction in schizophrenia. *Biological Psychiatry, 34*(5), 331–340.

Sweeney, J. A., Clementz, B. A., Haas, G. L., Escobar, M. D., Drake, K., & Frances, A. J. (1994). Eye tracking dysfunction in schizophrenia: characterization of component eye movement abnormalities, diagnostic specificity, and the role of attention. *Journal of Abnormal Psychology, 103*(2), 222.

Sweeney, J. A., Luna, B., Haas, G. L., Keshavan, M. S., Mann, J. J., & Thase, M. E. (1999). Pursuit tracking impairments in schizophrenia and mood disorders: Step-ramp studies with unmedicated patients. *Biological Psychiatry, 46*(5), 671–680.

Sweeney, J. A., Luna, B., Srinivasagam, N. M., Keshavan, M. S., Schooler, N. R., Haas, G. L., & Carl, J. R. (1998). Eye tracking abnormalities in schizophrenia: evidence for dysfunction in the frontal eye fields. *Biological Psychiatry, 44*(8), 698–708.

Sweeney, J. A., Palumbo, D. R., Halper, J. P., & Shear, M. K. (1992). Pursuit eye movement dysfunction in obsessive-compulsive disorder. *Psychiatry Research, 42*(1), 1–11.

Thaker, G., Kirkpatrick, B., Buchanan, R. W., Ellsberry, R., Lahti, A., & Tamminga, C. (1988). Oculomotor abnormalities and their clinical correlates in schizophrenia. *Psychopharmacology Bulletin, 25*(3), 491–497.

Thaker, G. K., Avila, M. T., Hong, E. L., Medoff, D. R., Ross, D. E., & Adami, H. M. (2003). A model of smooth pursuit eye movement deficit associated with the schizophrenia phenotype. *Psychophysiology, 40*(2), 277–284.

Thaker, G. K., Cassady, S., Adami, H., Moran, M., & Ross, D. E. (1996). Eye movements in spectrum personality disorders: Comparison of community subjects and relatives of schizophrenic patients. *American Journal of Psychiatry, 153*(3), 362–368.

Thaker, G. K., Ross, D. E., Buchanan, R. W., Adami, H. M., & Medoff, D. R. (1999). Smooth pursuit eye movements to extra-retinal motion signals: deficits in patients with schizophrenia. *Psychiatry Research, 88*(3), 209–219.

Theleritis, C., Evdokimidis, I., & Smyrnis, N. (2014). Variability in the decision process leading to saccades: A specific marker for schizophrenia? *Psychophysiology, 51*(4), 327–336.

Tien, A. Y., Pearlson, G. D., Machlin, S. R., Bylsma, F. W., & Hoehn-Saric, R. (1992). Oculomotor performance in obsessive-compulsive disorder. *American Journal of Psychiatry, 149*(5), 641–646.

Tien, A. Y., Ross, D. E., Pearlson, G., & Strauss, M. E. (1996). Eye movements and psychopathology in schizophrenia and bipolar disorder. *The Journal of Nervous and Mental Disease, 184*(6), 331–338.

Tregellas, J. R., Tanabe, J. L., Miller, D. E., Ross, R. G., Olincy, A., & Freedman, R. (2004). Neurobiology of smooth pursuit eye movement deficits in schizophrenia: An fMRI study. *American Journal of Psychiatry, 161*(2), 315–321.

Toh, W. L., Rossell, S. L., & Castle, D. L. (2011) Current visual scanpath research: A review of investigations into the psychotic, anxiety, and mood disorders. *Comprehensive Psychiatry, 52*(6), 567–579.

Tomer, R., Mintz, M., Levy, A., & Myslobodsky, M. (1981). Smooth pursuit pattern in schizophrenic patients during cognitive task. *Biological Psychiatry, 16*(2), 131–144.

Tu, P. C., Yang, T. H., Kuo, W. J., Hsieh, J. C., & Su, T. P. (2006). Neural correlates of antisaccade deficits in schizophrenia, an fMRI study. *Journal of Psychiatric Research, 40*(7), 606–612.

van Kampen, D., & Deijen, J. B. (2009). SPEM dysfunction and general schizotypy as measured by the SSQ: A controlled study. *BMC Neurology, 9*(1), 27.

van Tricht, M. J., Nieman, D. H., Bour, L. J., Boerée, T., Koelman, J. H. T. M., de Haan, L., & Linszen, D. H. (2010). Increased saccadic rate during smooth pursuit eye movements in patients at ultra high risk for developing a psychosis. *Brain and Cognition, 73*(3), 215–221.

Wang, Z., Meda, S. A., Keshavan, M. S., Tamminga, C. A., Sweeney, J. A., Clementz, B. A. … Pearlson, G. D. (2015). Large-scale fusion of gray matter and resting-state functional MRI reveals common and distinct biological markers across the psychosis spectrum in the B-SNIP cohort. *Frontiers in Psychiatry, 6,* 174.

Winograd-Gurvich, C., Fitzgerald, P. B., Georgiou-Karistianis, N., Millist, L., & White, O. (2008). Inhibitory control and spatial working memory: A saccadic eye movement study of negative symptoms in schizophrenia. *Psychiatry Research, 157*(1), 9–19.

Winograd-Gurvich, C., Georgiou-Karistianis, N., Fitzgerald, P. B., Millist, L., & White, O. B. (2006). Self-paced and reprogrammed saccades: differences between melancholic and non-melancholic depression. *Neuroscience Research, 56*(3), 253–260.

Chapter 17
Eye Movements in Neurology

René Müri, Dario Cazzoli and Thomas Nyffeler

Contents

Abstract Eye movement disorders play an important role in clinical neurology, since they are part of many neurological diseases. Furthermore, from a scientific point of view, the assessment of the patterns of eye movement disturbances after

---

R. Müri (✉)
Department of Neurology, University of Bern, Bern, Switzerland
e-mail: rene.mueri@insel.ch

D. Cazzoli
Gerontechnology and Rehabilitation Group, ARTORG Center for Biomedical Engineering Research, University of Bern, Bern, Switzerland

T. Nyffeler
Clinic for Neurology and Neurorehabilitation, Luzerner Kantonsspital, Luzern, Switzerland

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_17

focal brain lesions is an excellent way to study the cortical and subcortical networks involved in eye movement control. The present chapter deals with the clinical aspects of eye movements disturbances in patients with neurological disorders. The clinical examination of eye movements is explained and put in relation to the underlying neuroanatomical structures. Moreover, the eye movement disturbances typically observed in common neurological diseases (such as ischemic stroke, multiple sclerosis, and neurodegenerative diseases) are discussed.

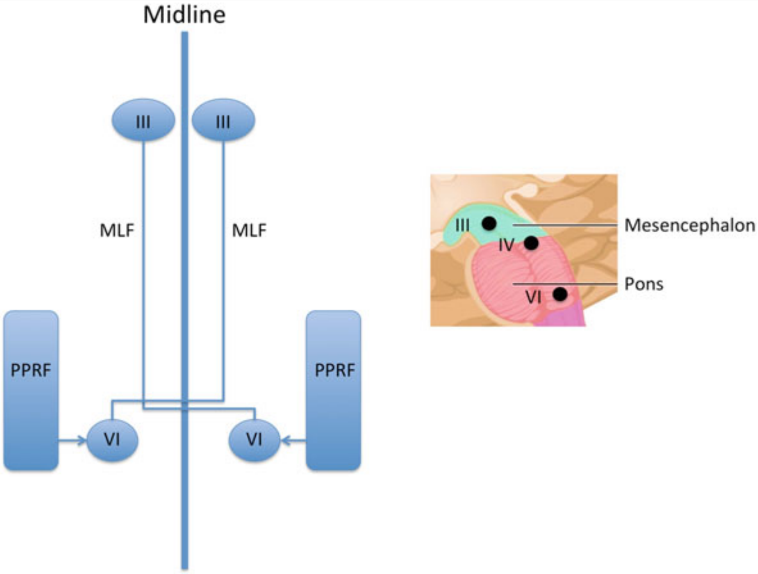
17.1 Introduction and Learning Objectives

Eye movements are very relevant in clinical neurology. Since many structures in the brainstem, cerebellum, basal ganglia, and cortex are involved in eye movement control, diseases affecting these structures are often also associated with disturbances of eye movements. Therefore, the careful examination of eye movements is important. The clinical examination of eye movements is simple, but the interpretation of the findings requires some experience.

Saccades are quick conjugate eye movements that shift the gaze from one object to another. During clinical examination, the examiner usually holds two objects, and the patient is asked to look at one and then at the other target object. Horizontal and vertical saccades are tested separately. Three saccadic parameters are important: latency, accuracy, and velocity. The latency of a saccade can be defined as the time elapsed between the command to move the eyes to the beginning of the eye movement. Minor prolongations of saccadic latencies are usually difficult to detect in clinical examination; however, gross difficulties in initiating saccades are easily detectable. Some patients are not able to generate saccades on command (i.e., voluntary saccades). In this case, it should be tested whether they are able to produce saccades towards suddenly appearing objects or unexpected sounds (i.e., reflexive saccades). The second important parameter is saccadic accuracy, i.e., the precision of the amplitude of a saccade with respect to a given target. In general, large saccades (i.e., >15°) undershoot the target position by about 10%, and are followed by single, corrective saccades that bring the gaze on the target. Greater undershoots, sometimes accompanied by multiple corrective saccades, indicate pathologic hypometria. By contrast, hypermetria is observed when saccades show an excessively large amplitude, and are followed by corrective saccades in the reverse direction. The third important parameter is saccade velocity. Saccade velocity is dependent on saccade amplitude (the so-called main sequence). Slowing of saccadic velocity occurs within several neurological disorders, but may also be observed in fatigued healthy individuals (Fig. 17.1).

Box 1: How the Brainstem Controls Horizontal Eye Movements

For conjugate horizontal eye movements, the brain has to innervate the medial rectus muscle (innervated by the cranial nerve III) of one eye, which moves the eyeball towards the nose (adduction), and, simultaneously, the lateral rectus muscle (innervated from the cranial nerve VI) of the other eye, which moves the eyeball away from the nose (abduction). This is accomplished by the medial longitudinal fasciculus (MLF), which connects both cranial nerve nuclei.



Abbreviations: *PPRF* Paramedian pontine reticular formation; *VI* Abducens nucleus; *MLF* Medial longitudinal fasciculus; *III* Oculomotor nucleus; *IV* Trochlear nucleus.

The abducens nucleus contains two types of neurons, the motor neurons and the interneurons, connecting via the MLF the subnucleus for the rectus medialis muscle of the oculomotor nucleus. A premotor command (e.g., from the right PPRF to the right VI nucleus), activates both the right lateral rectus muscle and (via interneurons and MLF the subnucleus) the left medial rectus muscle, resulting in a conjugate eye movement towards the right side. A lesion of the MLF results in a restricted adduction of one eye, called internuclear ophthalmoplegia (INO).

Furthermore, the abducens nucleus is the final common pathway, i.e., here the premotor information for saccades coming from the PPRF, and the premotor information for smooth pursuit coming from the dorsolateral pontine nuclei converges (see also Box 2 and Box 3). A lesion of the PPRF results in an ipsilesional saccade palsy (with intact smooth pursuit), whereas an isolated lesion of the abducens nucleus results in an abolition of both saccades and smooth pursuit movements.

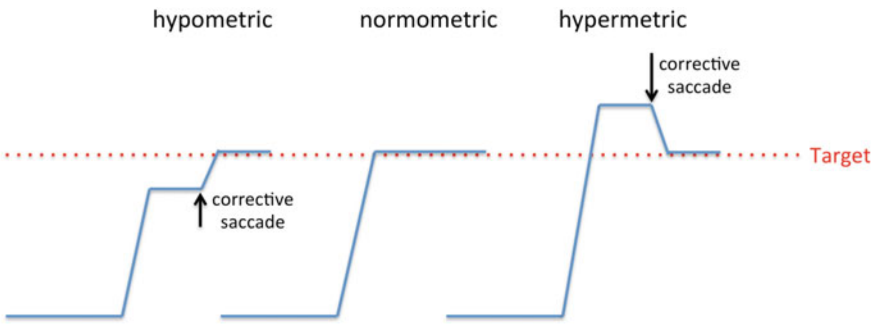
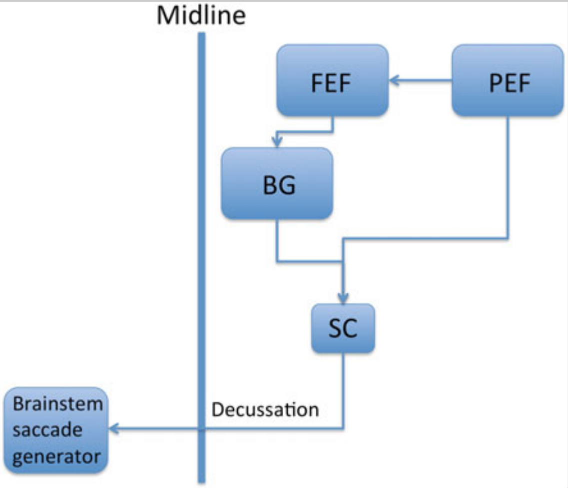


Fig. 17.1 Illustration of saccadic amplitudes. Clinical examination may show the following three possibilities in terms of the amplitude accuracy of the executed saccade: hypometric, normometric, or hypermetric. Large saccades (i.e., with an amplitude of more than 15° between the fixation point and the target) are usually hypometric, and a corrective saccade of about 10% of the original amplitude is considered to be normal. Normometric saccades (i.e., the landing point of the saccade and the position of the target are identical), are usually observed with small amplitudes (i.e., <10°). Hypermetric saccades are generally pathologic, and often indicate a cerebellar problem

Smooth pursuit eye movements, which allow to closely follow with the gaze an object in motion, are usually tested with an object moving sinusoidally from left to right, and from up to down (with a frequency of about 0.5–1 Hz). Tired persons will either lag behind the target and make small saccades, or jump ahead of the target, anticipating its localization. In both cases, the patients must be urged to concentrate on keeping their gaze locked to the target. If pursuit is perfect, then the eye speed will match the target speed. The 'pursuit gain' is calculated as the ratio of the eye speed over the target speed, which will equal one in case of perfect pursuit. Disturbed pursuit is characterized by the interruption of the smooth, following eye movement by small 'catch-up' saccades, which occur when the position of the eyes is too far behind the target. These saccades are corrective in nature, and are a sign of a decreased smooth pursuit gain.

Optokinetic nystagmus (OKN) is elicited by the motion of large portions of the visual field (in contrast to the motion of small targets, which generates pursuit responses). OKN is usually tested with large, striped or textured, moving objects, such as OKN drums. The patients are asked to "look at each stripe as it goes by". This generates a jerk nystagmus, with the slow phase in the same direction as the stripe motion. If the patient is unable to generate the optokinetic slow phase, the eyes will remain stationary. In humans, deficits in optokinetic slow phases are usually associated with abnormalities in smooth pursuit. However, asymmetries in smooth eye movements are sometimes more easily appreciable during OKN than during sinusoidal smooth pursuit.

Box 2: Organization of Saccade Triggering Saccades are controlled by cortical, subcortical, and brainstem structures, as depicted in the scheme below. The main cortical structures involved are the parietal eye field (PEF), within the intraparietal sulcus, and the frontal eye field (FEF), located within the superior pre-central sulcus, near the caudal end of the superior frontal sulcus (for further details, see chapters by Jamadar & Johnson and by Pierce et al.).



Abbreviations: *FEF* frontal eye field; *PEF* parietal eye field; *BG* basal ganglia; *SC* superior colliculus.

It is generally accepted that the PEF is more involved in the control of reflexive saccades, whereas the FEF is more important for voluntary saccades. Saccades can be triggered either by the FEF or by the PEF. Both direct and indirect pathways connect these regions to the superior colliculus, and reach—after decussating—the other side of the brainstem saccade generator, in particular the paramedian pontine reticular formation (PPRF; see also Box 1 for further details). Of clinical importance is the fact that the pathways cross the midline at brainstem level. Thus, cortical or basal ganglia lesions will result in mainly contralesional saccade deficits. In contrast, lesions in the region of the brainstem saccade generator will result in ipsilesional deficits.

The vestibulo-ocular reflex (VOR) is more difficult to observe clinically. Normally, the VOR keeps the eyes stable in space despite head motion. However, if the head is moving in concert with a moving target, it may be desirable to keep the eyes stabilized within the orbits in order to follow the target only by head movements. This is called VOR-cancellation, and can be tested with the patient sitting on a swivel chair. The patient has to place his/her outstretched arms together, interlacing his/her fingers, and is instructed to fixate his/her thumbs, while he/she is rotated “en bloc” with the chair (i.e., the head of the patient moves “en bloc” with the rest of his/her body, and thus also in concert with the moving target, represented by his/her thumbs). If the VOR-cancellation is intact, the eyes will remain stationary in the orbit. An impairment of the VOR-cancellation will cause a lag of the eyes behind the target, which provokes catch-up saccades, creating a nystagmus.

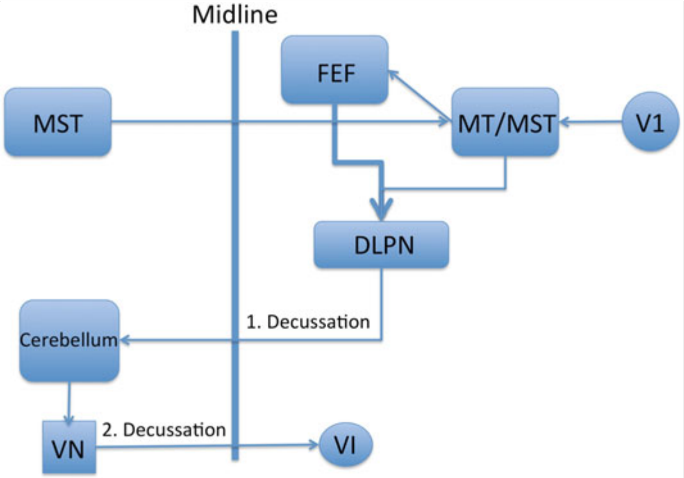
Several neurological disorders are associated with eye movement disturbances and, in many cases, the clinical examination of the different eye movements (such as saccades, smooth pursuit, or vestibular-induced eye movements) can lead to an appropriate diagnosis.

Objective eye movement recordings using video-oculography, infrared oculography, or electro-oculography are an important adjunct in diagnosing, documenting, and managing a variety of neurological diseases. These include peripheral and central vestibular imbalances, as well as brainstem, cerebellar, and cortical lesions, which may result in several eye movement abnormalities, such as slow or dysmetric (i.e., with an incorrect amplitude) saccades, inaccurate pursuit, and various forms of acquired nystagmus. Furthermore, objective eye movement recordings are able to detect changes that are too small to be resolved by direct observation during clinical examination. For instance, the documentation of an acquired nystagmus (i.e., involuntary eye movements, comprising a smooth drift followed by a backward saccade) with very small amplitudes, which may not be readily visible even in a careful clinical examination, can provide an explanation for some worrisome symptoms reported by the patients. These include blurred vision and oscillopsia (a visual disturbance in which visual objects appear to oscillate; images can no longer be held steady on the retina due to ocular instability).

Box 3: How the Brain Controls Smooth Pursuit Eye Movements (A Simplified View)

The control of smooth pursuit eye movements is complex. Visual information, coming from the striate cortex (V1), reaches the middle temporal area (MT) and the medial superior temporal area (MST). The frontal eye field is also involved in smooth pursuit control. The dorsolateral pontine nuclei are the premotor structures in the brainstem. Note that there is a twofold decussation of the pathways in the brainstem (for further details see the chapter by Lencer et al., in this volume).

Clinically important is that an isolated lesion of the FEF provokes an ipsilesional reduction of the gain of smooth pursuit, and a contralesional hypometria of saccades.



Abbreviations: *V1* Striate cortex; *MT* Middle temporal area; *MST* Medial superior temporal area; *FEF* Frontal eye field; *DLPN* Dorsolateral pontine nuclei; *VN* Vestibular nucleus

The analysis of eye movements in stroke is useful for the clinical localization of the lesion, especially in brainstem and cerebellar strokes. Given the complex neuroanatomy of the brainstem, with the cranial nerve nuclei controlling the eye movement muscles and the supranuclear structures controlling conjugate eye movements, the patterns of deficits may appear confusing at first. Therefore, a careful examination of the eye movements is mandatory. Even with modern, high-resolution magnetic resonance imaging (MRI) techniques, small lesions within the brainstem, provoking distinct eye movement disorders, may not be visible. The clinical presentation of an oculomotor syndrome (such as those discussed in the following paragraphs) may thus be helpful for determining lesion localization and for the neurological diagnosis.

In addition to aiding clinical diagnosis and management, the study of oculomotor disturbances in patients with focal cerebral lesions is also a way to understand the organization of eye movement control in the brain. The study of the consequences of focal cerebral lesions remains the best way to determine whether a given area is crucial or not in the oculomotor control during a given task. In fact, even small lesions within critical cerebral areas can lead to a significant deficit. For instance, the schema of cortical control of saccade planning and execution (depicted in Fig. 17.5) represents a synthesis of the results of many different studies in patients with focal cerebral lesions, involving distinct cortical regions that are critical for eye movement control.

The analysis of eye movements may also be used as a (bio)marker in different neurological (such as in neurodegeneration) or psychiatric diseases. For such diseases, distinctive patterns of eye movement disturbances may be found. As for the future, there is also hope that eye movements may act as a (bio)marker of given diseases already in the preclinical stages.

Learning Objectives

The learning objectives of this chapter are: (1) to learn important aspects of the clinical-neurological examination of eye movements in humans; (2) to understand the anatomical organization of eye movement control in humans; (3) to understand the differences in eye movement control between the brainstem and the cortex; (4) to know the clinical presentation of eye movement disorders in common neurological diseases.

17.2 Historical Annotations

In 1879, Emile Javal, a French ophthalmologist, observed that people do not read smoothly across a page, but rather pause on some words, while moving quickly through other words. He also used the term ‘saccade’ for the first time. The neurologists of the 19th century were interested in the relationship between eye movement disturbances and pathological findings. Furthermore, since localizing a brain lesion in living patients (i.e., as opposed to post-mortem studies) was only possible by means of clinical observations, at that time a very accurate examination was essential to reach a correct diagnosis.

In the early 1900s, Edmund Huey built a device capable of tracking eye movements during reading. This device was quite intrusive und uncomfortable, since it required readers to wear lenses that had only a small opening, and a pointer attached to them. More comfortable, non-intrusive eye tracking technologies arrived with Charles H. Judd’s eye movement camera, which recorded the movements of the eyes on film. Using these techniques, Guy Thomas Buswell studied and analysed eye movements in reading. Alfred L. Yarbus developed small suction devices to be attached to the eye. Some of them were attached only to the sclera, leaving the visual field unobstructed, and reflecting light onto a photosensitive surface. Thanks to the new infrared and video-based oculography systems, the measurement of eye movements thrived during the 1970s and 1980s. In the 1970s, the eye trackers became less intrusive, provided better accuracy, and were able to separate eye from head movements. Advances in computer technology in the 1980s allowed to perform eye tracking in real time. Nowadays, remote eye tracking systems even allow contact-free eye movement measurements.

17.3 Eye Movements in Stroke

A central nervous system infarction is defined as the cell death in the brain, spinal cord, or in the retina that is attributable to ischemia. An ischemic stroke specifically refers to a central nervous system infarction that is accompanied by overt neurological symptoms. Stroke is the third most common cause of death in developed countries, exceeded only by coronary heart disease and cancer. Worldwide, about 15 million people suffer from a stroke each year. One-third of these patients die, and one-third are left permanently disabled. In Europe, the incidence of stroke varies from 101.1 to 239.3 per 100,000 in men, and from 63.0 to 158.7 per 100,000 in women (EROS Investigators, 2009). According to the WHO, there is an up to ten-fold difference in age-adjusted and sex-adjusted mortality rates and burden between low- and high-income countries, with considerably higher rates for low-income countries (Mackay & Mensah, 2004).

Strokes in the posterior circulation territory provoke by definition more often eye movement disturbances. This is because the posterior circulation (with the vertebral arteries, and the basilar artery with its branches) irrigates the brainstem and the cerebellum, which are key regions for oculomotor control. The nuclei of the oculomotor nerves and the premotor structures are located at different levels of the brainstem (see Fig. 17.2). Strokes in these regions may provoke typical syndromes, as discussed below.

On the other hand, the anterior circulation irrigates the hemispheres. Strokes in these territories result in more subtle deficits of eye movements, since—at the hemispheric level—the representation of the supranuclear oculomotor control shows some redundancy.

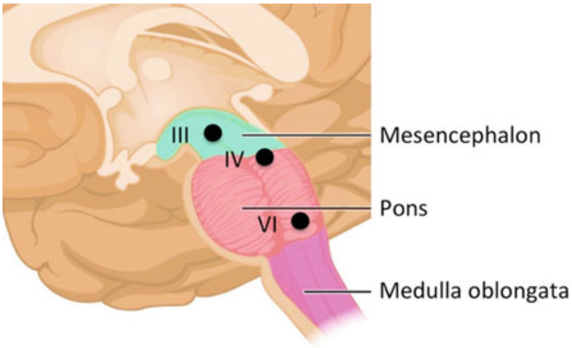


Fig. 17.2 Schematic representation of the different parts of the brainstem. The oculomotor nucleus (III) is located in the mesencephalon, the trochlear nucleus (IV) in the upper pons, and the abducens nucleus (VI) in the lower pons. The supranuclear structure that controls conjugate horizontal eye movements is the parapontine reticular formation (PPRF), which is located at the level of the abducens nucleus in the pons. The supranuclear structure that controls conjugate vertical eye movements is the rostral interstitial nucleus of the medial longitudinal fasciculus (riMLF), which is located at the level of the oculomotor nucleus in the mesencephalon

17.3.1 Mesencephalic Syndromes

Numerous eponymous brain stem vascular syndromes have been described, which are characterized on the basis of their particular combination of cranial nerve palsies and additional neurological signs (such as cerebellar signs and long tract signs). These oculomotor syndromes may be supranuclear, nuclear, and infranuclear (or fascicular) (Bogousslavsky & Meienberg, 1987). Supranuclear syndromes may be provoked by lesions at the level of the rostral interstitial nucleus of the MLF (riMLF), which is located at the level of the upper pole of the red nucleus. The riMLF is the premotor structure that controls vertical saccades, i.e., the final common pathway for these saccades, and the generator of the vertical saccade pulses. Small vascular lesions of this structure are rare. The most important neuroanatomical structures within the brain stem are depicted in Fig. 17.2.

Vascular lesions of the oculomotor nerve complex may result in an *isolated nuclear syndrome* (Meienberg & Müri, 1992). Clinically, this consists of an ipsilesional (i.e., on the same side of the lesion) palsy of the muscles innervated by the oculomotor nerve, and of a contralesional (i.e., on the opposite side with respect to the lesion) paresis of the elevation of the eye. Furthermore, bilateral dilatation of the pupil (mydriasis) and bilateral ptosis of the eyelid may occur.

Lesions affecting the oculomotor nerve fascicles result in ipsilesional oculomotor paralysis. Such paralysis may be isolated or, more often, combined with other neurological deficits. Very small lesions may affect only a part of the nerve fascicles in which fibres going to the pupil and to the inferior rectus muscle are located, above those of the single muscles innervated by the oculomotor nerve (i.e., superior rectus muscle, levator palpebrae muscle, medial rectus muscle, and inferior oblique muscle).

Weber's syndrome (named after Hermann David Weber, who first described the condition in 1863) consists of an ipsilesional palsy of the oculomotor nerve, and of a contralesional hemiplegia or hemiparesis, due to a concomitant lesion of the descending motor pyramidal tract. Figure 17.3 shows the anatomical relations at the level of the nucleus and of the fascicle of the third nerve, with important structures such as the red nucleus, the pyramidal tract, and the cerebellar peduncle (i.e., one of the fibre tracts that connects with the cerebellum). In case of an involvement of the red nucleus, or of the superior cerebellar peduncle, the resulting syndrome is called *Claude-Nothnagel's syndrome* (named after Henri Charles Jules Claude, who first described the condition in 1912; fascicular or nuclear), consisting of an ipsilesional palsy of the oculomotor nerve and of a contralesional ataxia.

Another syndrome, called *Benedikt's syndrome*, consists of an ipsilesional oculomotor palsy and of contralesional involuntary movements, such as tremor and chorea, due to the lesion of the red nucleus and of the substantia nigra.

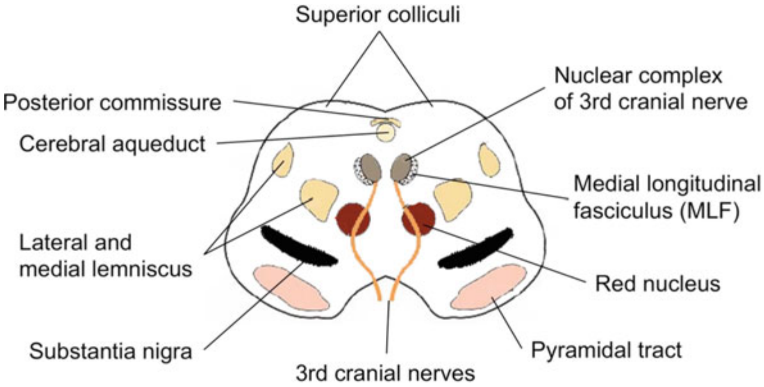


Fig. 17.3 Schematic representation of a cross section through the rostral mesencephalon, at the level of the oculomotor nuclei

17.3.2 Pontine Syndromes

Many important prenuclear and internuclear structures are localized in the pons. Furthermore, the fourth and the sixth nuclei lie also in the pons. In contrast to the third nerve, both the fourth and the sixth nerve innervate only one eye muscle (the superior oblique muscle, and the lateral rectus muscle, respectively). Isolated fourth nerve palsies due to a lesion within the brainstem are very rare, since the nucleus lies in a very dorsal position within the brainstem, and the fascicle is short (see Fig. 17.4). A further particularity of the fourth nerve is that the fascicle crosses the midline within the brainstem, i.e., the left nucleus innervates the right superior oblique muscle. Due to this anatomical localization, there are no eponymous syndromes known.

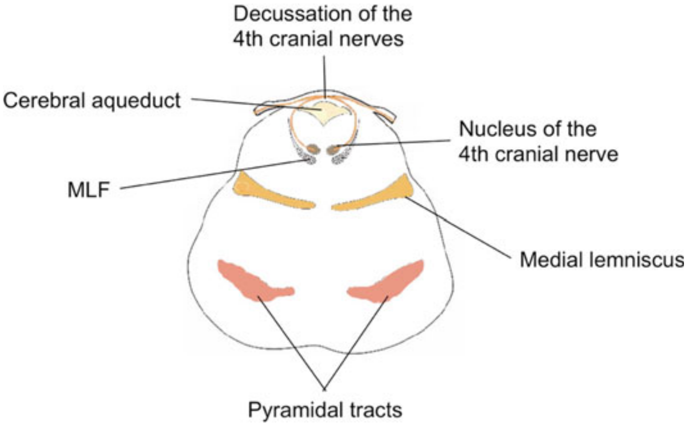


Fig. 17.4 Schematic representation of a cross section through the caudal midbrain, at the level of the fourth cranial nerve nuclei. *MLF* medial longitudinal fasciculus

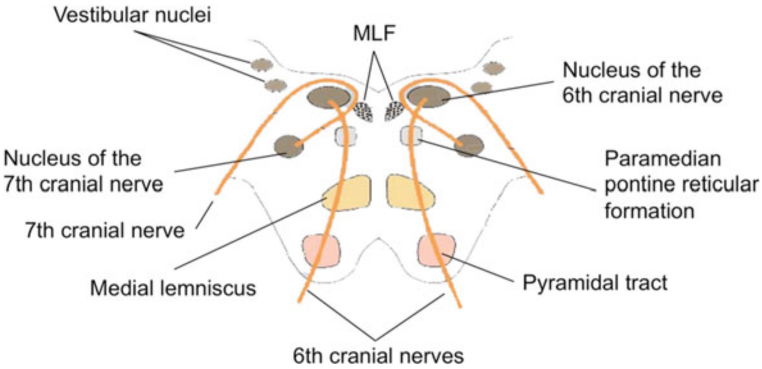


Fig. 17.5 Schematic representation of a cross section through the pons, at the level of the abducens nuclei. *MLF* medial longitudinal fasciculus

The clustered proximity of the sixth nerve nucleus and of its fascicle, and the anatomic neighbourhood to the PPRF, to the facial nerve nucleus and fascicle, to the (uncrossed) pyramidal tracts, and to the sensory and sympathetic long tracts, make collateral damage very common, and a series of eponymous syndromes exist (see Fig. 17.5). These classical brainstem syndromes, described in the nineteenth century, are seldom observed in isolation. The *Gasperini syndrome* (or pontine tegmental syndrome) is due to a lesion involving the PPRF, the seventh nerve, the medial lemniscus, and the spinothalamic tract. Clinically, an ipsilateral gaze palsy, and palsies of the fifth, seventh, and eighth nerve are observed. Moreover, sensory loss is found contralesionally, due to the damage of the medial lemniscus and of the spinothalamic tract.

The *Cestan Raymond syndrome* is due to a lesion of the cerebellar peduncle, of the corticospinal tract, and of the medial lemniscus. Ipsilateral to the lesion, cerebellar ataxia, contralateral hemiparesis, and hemihypesthesia (i.e., sensory loss) are observed.

The *Foville syndrome* is due to a lesion of the facial nerve nucleus, often combined with a lesion of the trigeminal neurons, of the ciliospinal tract and of the corticospinal tract. Ipsilesionally, a seventh nerve paresis and a Horner’s syndrome (ptosis, miosis, and enophthalmos) are found. Contralesionally, a hemiparesis is found.

The lesion leading to the *Millard Gubler syndrome* includes the facial nerve nucleus and the corticospinal tract. An ipsilesional facial palsy and a contralesional hemiparesis can be observed.

Internuclear ophthalmoplegia (INO; see also Box 1) may occur after a small paramedian stroke, and is due to a lesion of the medial longitudinal fasciculus (MLF; a heavily myelinated tract that allows conjugate eye movement, by connecting the abducens nucleus with the oculomotor nucleus on the contralateral side). After stroke, the syndrome is generally unilateral, with reduced adduction or reduced saccadic velocity of the eye on the affected side, and dissociated nystagmus in the abducting contralateral eye. Convergence is usually preserved, and sometimes a skew deviation (i.e., an abnormality in the vertical ocular alignment; *Hertwig-Magendie's syndrome*) is observed.

The *one-and-a-half syndrome*, found after more extensive paramedian strokes, is characterized by a conjugate horizontal gaze palsy to one side, and an INO on the other side. A lesion of the PPRF and of the ipsilateral MLF is at the origin of the syndrome. Alternatively, a lesion of the abducens nucleus, combined with an additional lesion of the ipsilateral MLF after its crossing of the midline, may also provoke a *one-and-a-half syndrome*.

The *locked-in syndrome* is a rare and very dramatic condition, in which patients are fully conscious, but cannot move or communicate verbally, due to a complete paralysis of nearly all voluntary muscles of the body, with exception of the eyes. Alexandre Dumas described a person suffering from locked-in syndrome in his novel "The Count of Monte Cristo" (1844). The patient in the novel is described as a "corpse with living eyes", who communicates with eye movements and expressions. In most cases, a bilateral horizontal gaze palsy is also present, and only lid movements and vertical eye movements are possible.

Finally, the *lateral pontine syndrome* (or *Marie-Foix syndrome*) consists of an ipsilesional one-and-a-half syndrome, ipsilesional limb ataxia (due to a lesion of the cerebellar tracts), and contralesional hemiplegia as well as sensory loss for pain and temperature (due to the lesion of the spinothalamic tract, which contains these fibres).

17.3.3 Medullary Syndromes

The *Wallenberg's syndrome* (or lateral medullary syndrome) is a constellation of neurological deficits due to a lesion of the lateral part of the medulla. It results from an occlusion of the posterior inferior cerebellar artery (PICA), or of one of its branches. Concerning saccades, hypermetria towards the affected side and hypometria towards the non-affected side are observed. Furthermore, Horner's syndrome and skew deviation are present, with the ipsilesional eye typically down. A rotatory (or horizonto-rotatory) primary position nystagmus, with the quick phase of the horizontal component beating contralesionally, is part of the syndrome. Ipsilesional ocular lateropulsion can also be observed. Ocular lateropulsion refers to an ocular motor bias towards the lesion side, without limitation of eye movements. It comprises steady-state ocular deviation, asymmetric horizontal saccades, and oblique misdirection of vertical saccades. Lateropulsion is due to damage of the climbing fibres running from the contralesional inferior olivary nucleus to the dorsal vermis of the cerebellum. Typically, a contralesional sensory loss in the trunk and the extremities (due to a lesion of the spinothalamic tract) and an ipsilesional sensory loss in the face (due to a lesion of the trigeminal nerve nucleus) are also found.

17.3.4 Oculomotor Disorders in Cerebellar Stroke

The part of the cerebellum involved in oculomotor control plays a central role in regulating all eye movements, including saccades, pursuit eye movements, eye vergence responses, and vestibulo-ocular responses. Different types of nystagmus, often associated with other cerebellar signs, are the most common oculomotor signs after cerebellar infarcts.

Horizontal nystagmus (after strokes in the territory of the posterior inferior cerebellar artery) is gaze-evoked, in both gaze directions, with larger ipsilesional amplitudes, and is accompanied by other prominent ipsilateral cerebellar signs. *Upbeat nystagmus* (after strokes in the territory of the superior cerebellar artery, rather uncommon) is a nystagmus beating in primary position, with an upward quick phase.

Downbeat nystagmus presents in primary position, with a downward quick phase (more prominent when the patient is looking down or laterally). It is often combined with smooth pursuit deficits, and an impairment of the optokinetic reflex. *Periodic alternating nystagmus* (after strokes near the midline of the cerebellum) presents in primary position. After a few minutes (the characteristic oscillation cycle is 2–4 min), this nystagmus stops for a few seconds, and starts beating again in the opposite direction. *Rebound nystagmus* (after floccular or parafloccular lesions) is observed after prolonged eccentric viewing. The eccentric viewing produces a directional bias in oculomotor control, in an attempt to counteract the centripetal tendency of the eyes to return to primary position. The nystagmus becomes evident when the eyes return to the primary position, and then show a tendency to return to the eccentric position. As a consequence, corrective saccades occur, with an opposite direction with respect to the initial eccentric position of gaze.

17.4 Eye Movement Disorders After Cortical Lesions

The cortical network subtending eye movement control in humans (Fig. 17.6) includes several regions (Pierrot-Deseilligny, Milea, & Müri, 2004). The cortical representation of oculomotor control is often redundant, and both frontal and parietal regions are involved. In contrast to eye movement disorders after brainstem lesions, eye movement disorders after cortical lesions are often not prominent, and the clinical examination is thus more complex. Therefore, quantitative oculography is often needed, using a test battery including different saccade paradigms, such as gap and overlap saccades, memory-guided saccades, and antisaccades. Furthermore, attention and perceptual factors have to be considered in the examination.

Oculomotor apraxia is characterized by marked difficulties in producing intentional (i.e., voluntary) saccades, both along the horizontal and the vertical dimension, whereas reflexive (i.e., externally triggered) saccades and the vestibulo-ocular reflex are largely preserved (Leigh & Zee, 2015). Patients may sometimes be able to produce voluntary gaze shifts by means of concurrent movements of the head and of

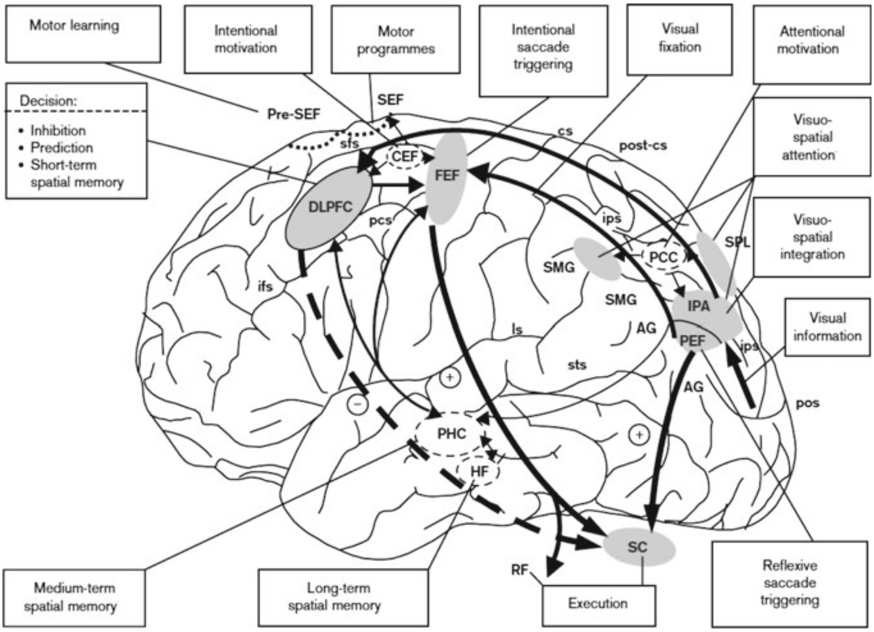


Fig. 17.6 Cortical areas and pathways involved in saccade control (Pierrot-Deseilligny, Milea, & Müri, 2004). *SEF* supplementary eye field; *sfs* superior frontal sulcus; *CEF* cingulate eye field; *cs* central sulcus; *DLPFC* dorsolateral prefrontal cortex; *pcs* precentral sulcus; *FEF* frontal eye field; *ips* intraparietal sulcus; *ifs* inferior frontal sulcus; *SMG* supramarginal gyrus; *PCC* posterior cingulate cortex; *SPL* superior parietal lobule; *IPA* intraparietal areas; *ls* lateral sulcus; *AG* angular gyrus; *PEF* parietal eye field; *sts* superior temporal sulcus; *pos* parieto-occipital sulcus; *PHC* parahippocampal cortex; *HF* hippocampal formation; *SC* superior colliculus; *RF* reticular formation—inhibitory pathway; + excitatory pathway

the eyes. Since the voluntary component of eye movements seems to be particularly affected, oculomotor apraxia has sometimes been referred to as “psychic paralysis of gaze”. Oculomotor apraxia most commonly arises from bilateral frontal, parietal, or frontal-parietal lesions, involving the FEF and/or the PEF (Leigh & Zee, 2015).

In *Balint’s syndrome*, oculomotor apraxia is part of a typical triad of symptoms, together with simultagnosia (defined as the impairment in perceiving more than one object at a time) and optic ataxia (defined as the impaired ability of reaching and grasping objects under visual guidance, which cannot be explained by elementary motor or sensory deficits). The main oculomotor difficulty encountered by patients suffering from Balint’s syndrome is in the visual guidance of saccades (Ptak & Müri, 2013). Accordingly, visual exploration patterns show impairments in disengaging fixation from the current location or object (sometimes referred to as ‘sticky fixation’) and in intentionally fixating new locations or objects (sometimes referred to as ‘wandering gaze’). The extent of the explored visual field is limited, and the scan path is usually characterized by small saccadic amplitudes. Balint’s syndrome typically arises after bilateral, parieto-occipital lesions (Ptak & Müri, 2013).

So-called *visual agnosia* can be defined as the impairment in the visual recognition of objects, which cannot be explained by an elementary sensory, linguistic, or memory-related deficit (Karnath & Thier, 2006). The basic parameters of visual fixations fall usually within normal range, but their spatial distribution during visual exploration is more heavily influenced by bottom-up (i.e., saliency-driven) elements than in healthy controls. Moreover, saccades with small amplitudes are proportionally over-represented. Taken together, these finding have been interpreted as evidence for an impairment of the cognitive processes integrating single features into a coherent object (Karnath & Thier, 2006). Visual agnosia is commonly due to bilateral lesions within the secondary visual areas. A special case is represented by *prosopagnosia*, typically following bilateral lesions of the fusiform gyrus, of the occipital face area, or of the anterior temporal cortex (Karnath & Thier, 2006). Patients suffering from prosopagnosia are unable to visually recognize faces, while visual recognition of objects is largely spared. During visual exploration of faces, the spatial distribution of fixations of these patients does not show the classic "triangular pattern" defined by eyes and mouth. In contrast, fixations are either directed to more peripheral features, or concentrate in one region (e.g., on specific facial features or paraphernalia), reflecting the aforementioned lack of integration (Karnath & Thier, 2006).

In *acquired dyslexia* patients present with a sub-total impairment of the reading ability, usually following a temporo-parietal lesion in the left hemisphere (Karnath & Thier, 2006). During reading tasks, their eye movements closely reflect their behavioural difficulties: in comparison to healthy controls, patients produce more regressive saccades (i.e., opposite to the reading direction), landing points of the saccades are biased towards the beginning of the words, and fixations durations are prolonged. During saccadic, non-reading tasks, the variability of saccadic latencies in patients with acquired dyslexia is sensibly greater than in healthy controls: patients show both anticipatory and express saccades in some trials, and prolonged saccadic reaction times in other trials.

Hemispatial neglect can be defined as the failure to detect, respond, or orient to relevant stimuli presented on the contralesional side of space. Hemispatial neglect is supramodal (i.e., can affect different sensory modalities), and cannot be solely explained by an elementary sensory or motor impairment. Neglect can arise after a lesion to an extended attentional network, most commonly in the right hemisphere, including cortical structures, subcortical structures, and connecting white matter fibre tracts (see Table 17.1). The most striking oculomotor finding in hemispatial neglect concerns the spatial distribution of visual fixations during visual exploration or visual search (although fixation durations per se are usually within normal range). The distribution of visual fixations is typically asymmetrical, and shifted towards the ipsilesional space, with frequent re-fixations (i.e., repeated fixations on the same locations). Neglect patients also show a bias of early attentional orientation towards the ipsilesional space, i.e., fist saccades are significantly more often oriented towards the ipsi-than the contralesional side in comparison to healthy controls. Saccades are commonly hypometric; this does not, however, follow any particular spatial pattern (i.e., saccades are shorter than in healthy controls, both towards the left and the right, both within the left and the right hemispace).

Table 17.1 Lesions sites (cortical structures, subcortical structures, and white matter fibre tracts) typically leading to hemispatial neglect

Cortical structures	Subcortical structures	White matter fibre tracts
Posterior and inferior parietal lobe (angular gyrus, supramarginal gyrus)	Pulvinar nucleus of the thalamus	Superior longitudinal fasciculus
Superior temporal lobe (superior temporal gyrus)	Putamen	Inferior occipito-frontal fasciculus
Temporo-parietal junction	Caudate nucleus	Superior occipito-frontal fasciculus
Inferior frontal lobe (middle frontal gyrus, inferior frontal gyrus)		

Note Lesions are typically located in the right hemisphere

17.5 Eye Movements in Neurodegenerative Diseases

17.5.1 Parkinson's Disease

Parkinson's disease (PD) is a neurodegenerative disorder characterized by slowness in initiation and performance of movements (i.e., hypokinesia, akinesia, and bradykinesia), with limb resistance to passive movement (called rigidity), and typically "pill-rolling tremor", usually more evident at rest, with a frequency of 4–6 Hz. The disease was first described by James Parkinson in 1817 as "paralysis agitans". PD is caused by the depletion of dopamine in the striatum, due to cell death of nigral dopaminergic neurones. A number of genetic loci have been associated with autosomal dominant or autosomal recessive PD. Most patients respond to dopamine replacement therapy. Along disease progression, response fluctuations become problematic. Dementia may also develop, with features similar to those observed in dementia with Lewy bodies.

PD is also associated with eye movement disturbances, which commonly vary according to the task and to the progression of the disease. Saccadic metrics and latencies show significant differences in comparison to healthy individuals. These differences are usually more pronounced in voluntary than in reflexive eye movements. However, mild impairments in optokinetic nystagmus, in the cancellation of the vestibulo-ocular reflex, and in smooth pursuit may also be observed. Saccades are mildly hypometric in simple tasks (such as visually guided saccades), and more substantially hypometric in complex tasks (such as memory-guided saccades and antisaccades). In complex oculomotor tasks, PD patients also commonly show an increased number of errors. Saccadic hypometria can be detected already early in the disease course. Concerning saccadic latencies, these are mildly prolonged in simple tasks, and more substantially prolonged in complex tasks. Prolonged saccadic latencies are detected later in the disease course, and are thought to be linked to cognitive

Table 17.2 Eye movement disturbances (according to quantitative eye movement analyses) in different neurodegenerative syndromes (based on Anderson & McAskill, 2013; Vidailhet et al. 1994)

	Visually guided saccades (horizontal)		Antisaccades	Smooth pursuit
	Latency	Gain	% of errors	Gain
PD	No	↓↑	No	no
PSP	No	↓	↑	↓
CBD	↑	No	No	↓
HD	↑	No	↑	↓
AD	↑	↓	↑	↓

PD Parkinson's disease; *PSP* Progressive supranuclear palsy; *CBD* Corticobasal degeneration. *HD* Huntington's disease; *AD* Alzheimer's disease; *No* normal; ↑ increased; ↓ decreased

impairment. Moreover, the prolongation in saccadic latencies has been reported to correlate with disease severity. Due to the above-described saccadic hypometria, and to the increased saccadic latencies, visual exploration in PD patients usually covers a significantly narrower area than in healthy individuals (Table 17.2).

17.5.2 Progressive Supranuclear Palsy

Progressive supranuclear palsy (PSP), or Steele-Richardson-Olszewski syndrome, is an akinetic rigid syndrome of unknown aetiology. Clinically, patients show early postural instability and falls, and the characteristic eye movement disturbances described below. Furthermore, cognitive decline may also be associated with PSP.

Oculomotor disturbances are generally more pronounced in PSP than in PD. In PSP, saccades show abnormalities in velocity, latency, and amplitude. Saccadic velocity is reduced, particularly for saccades performed in the vertical dimension (and, in some cases, particularly for downward saccades). This sign is already observable early in the disease course. The velocity of saccades performed in the horizontal dimension also decreases with the progression of the disease. However, the vertical slowing is usually more conspicuous in comparison to the horizontal one. Latencies of reflexive saccades are increased, especially on the vertical dimension. Saccadic hypometria can be observed, for both vertical and horizontal saccades. Moreover, when asked to perform antisaccadic tasks (i.e., saccades that are directed towards the opposite side of a visual target), PSP patients tend to commit more errors. Furthermore, fixation may be instable, being disrupted by small square wave jerks (i.e., saccadic intrusions that displace the gaze away from the fixation target). Vertical optokinetic nystagmus is also disrupted by horizontal saccadic intrusions. PSP patients may also present with lid-opening and lid-closing apraxia and reduced blink rate. With the progression of the disease, patients develop a supranuclear gaze palsy (more often upward than downward; with preserved vestibulo-ocular reflex) or complete ophthalmoplegia.

17.5.3 Multiple System Atrophy

Multiple system atrophy (MSA) is a sporadic neurodegenerative disorder, characterized by autonomic, parkinsonian, cerebellar, and pyramidal disturbance, which may occur in any combination. Thus, several subtypes have been defined, according to the predominant neurological symptoms:

- MSA-P: parkinsonism
- MSA-C: cerebellar
- MSA-M: mixed combination of neurological symptoms

Accordingly, the oculomotor disturbances observed in MSA may have common features with the presentation of PD and PSP, such as excessive square wave jerks, mild vertical supranuclear gaze palsy, mild or moderate saccadic hypometria, mild (or no) slowing of vertical saccades, impaired smooth pursuit, and reduced vestibulo-ocular reflex cancellation. However, gaze-evoked nystagmus and positionally-induced downbeat nystagmus may be distinctive features of MSA.

17.5.4 Huntington's Disease

Huntington's chorea (HD) is an autosomal dominant disease, causing motor, psychiatric, and cognitive dysfunction. HD is most often the result of a defect in the coding region of the gene encoding huntingtin (IT15), on the short arm of chromosome 4, namely an expansion of a CAG trinucleotide. Anticipation of symptoms is seen with increasing repeat length.

The most important clinical feature of HD is chorea (i.e., "dance-like", brief, semi-directed, and irregular movements), which is initially transient, then progresses to continuous, athetoid, and dystonic movements. In this stage of the disease, the patient is unable to feed, dress, or to use the toilet. Personality changes may also be observed, consisting of irritability, apathy, and depression. A subcortical type of dementia, leading over time to a global dementia, is also associated with HD.

At the oculomotor level, HD patients present with different abnormalities in saccadic parameters, such as latency, metric, and velocity. Saccades have increased latencies (particularly for voluntary saccades), and are often hypometric. Saccadic velocity is decreased, this sign being particularly evident in patients with an early onset of HD symptoms. Decreased saccadic velocities can already be detected in pre-symptomatic patients. HD patients have also difficulties in generating voluntary saccades, sometimes approaching the clinical picture of a full oculomotor apraxia. Accordingly, patients may thus make use of head turns or blinks to help initiating voluntary saccades. On the other hand, HD patients may also have difficulties in suppressing reflexive saccades (a phenomenon sometimes referred to as “distractibility of gaze”). Smooth pursuit may also be impaired by saccadic intrusions. The vestibulo-ocular reflex and steady fixation are generally preserved (although the latter may also be disrupted by saccadic intrusions).

Generally, the above-mentioned saccadic abnormalities are more evident on the vertical than on the horizontal dimension. Moreover, saccadic abnormalities severity has been shown to correlate with disease severity. Task demands tend also to modulate performance in HD, i.e., patients commit more errors than healthy controls in more demanding saccadic tasks (such as antisaccades and memory-guided saccades).

17.5.5 Corticobasal Degeneration

Corticobasal degeneration (CBD) is a rare, late-onset neurodegenerative disorder of unknown aetiology, which is characterized by a movement disorder, an asymmetric akinetic-rigid syndrome with marked dyspraxia, involuntary movements, and alien limb/hand behaviour, in combination with cognitive disorders resulting in cortical dementia.

Pathology shows nerve cell loss and gliosis in the cortex, the underlying white matter, the thalamus, the lentiform nucleus, the subthalamic nucleus, the red nucleus, the midbrain tegmentum, the substantia nigra, and the locus coeruleus. Cases of patients suffering from CBD, but presenting with clinical features of frontotemporal dementia, have also been reported, as well as cases of CBD patients with predominant parieto-occipital (Balint-like) cortical dysfunction.

Concerning eye movements, oculomotor apraxia is a characteristic sign of CBD. Saccadic latency—particularly for voluntary saccades—is increased, and this sign is more pronounced than in PD. However, saccadic velocity falls usually within normal range in CBD, and smooth pursuit eye movements are only moderately impaired. Moreover, CBD patients show increased error rates in complex saccadic tasks (such as in antisaccades). Up to one third of CBD patients develop supranuclear palsy during the course of the disease.

17.5.6 Alzheimer’s Disease (AD)

The causes of this type of dementia are not exactly known. However, so-called amyloid plaques (i.e., abnormal deposits of amyloid protein) and neurofibrillary tangles containing tau protein are found in the brain of Alzheimer patients. Furthermore, there is loss of acetylcholine. AD is a common neurodegenerative disease, first described by Alois Alzheimer in 1906. AD mainly affects the elderly population, and usually starts with episodic memory difficulties. Sporadic and familial cases have been reported, the latter tending to occur earlier in life. As the disease progresses, there is increasing difficulty with memory, language, and orientation, leading to a global impairment of cognitive functions within 5–10 years from symptoms onset.

The oculomotor deficits in AD affect both saccadic and fixational eye movements. Saccades are usually hypometric, their velocity is—at least mildly—reduced, particularly for vertical saccades, and their latency in increased, particularly for voluntary saccades. The vestibulo-ocular reflex is usually preserved. Steady fixation is disrupted by small-amplitude square wave jerks and large-amplitude saccadic intrusions. Some AD patients may be unable to suppress reflexive saccades towards expected targets, a phenomenon sometimes referred to as visual grasp reflex. Smooth pursuit is impaired by saccadic intrusions and by hypometric gains. In more complex saccadic paradigms, such as in antisaccades, AD patients show significantly higher error rates than healthy individuals, and errors usually remain uncorrected. This impairment has been shown to correlate with the severity of cognitive impairment. AD patients also present with impairments in visual exploration settings, with an exploration field that is narrower and less focused on given areas of interest.

17.5.7 Frontotemporal Dementia

The term frontotemporal dementia (FD) is used to describe a group of neurodegenerative disorders characterised by circumscribed atrophy of the frontal and/or of the temporal lobes. FD accounts for approximately 20% of cases of dementia in younger age groups (<60 years), and may be as common as AD. A number of sub-syndromes are characterised according to clinical symptoms, and depending on the anatomical distribution of pathology. With the progression of the disease, there is a tendency towards a merging of the clinical symptoms.

Oculomotor deficits in FD include reduced saccade velocities (particularly on the horizontal dimension), increased saccadic latencies, and poor performance in complex saccadic tasks, such as antisaccades: increased number of errors (but usually corrected), and a higher incidence of early saccades, in comparison to healthy individuals. Reduced saccadic velocity and deficient saccadic suppression seem to be largely dependent on the sub-syndrome type and on the anatomical distribution of the affected cerebral areas.

17.5.8 Dementia with Lewy-Bodies

The cause of dementia with Lewy body is unknown. The brains of such patients show abnormal clumps of a protein called alpha-synuclein. These clumps, called Lewy bodies, are found in nerve cells throughout the outer layer of the brain (the cerebral cortex) and deep inside the midbrain and brainstem.

Cognitive symptoms are usually the presenting features. DLB patients show a cortical dementia similar to AD, but with more severe impairments of visuospatial functions and visual memory, which start early in the disease course. Typically, cognitive performance fluctuates from day to day. Visual hallucinations are very common, usually non-threatening in nature, and the patients may recognise that these are not real. Parkinsonism is usually mild, and follows the cognitive symptoms.

The typical fluctuations in cognitive performance in DLB are also observed in oculomotor deficits: saccadic velocity and saccadic metric accuracy (i.e., gains) are usually reduced, but both parameters show a much greater variability than in healthy individuals. Saccadic latency is usually prolonged, and express saccades have a lower incidence. In complex saccadic tasks (e.g., antisaccades, target prediction), DLB patients show higher error rates than healthy individuals, compatibly with their cognitive impairments. Generally, vertical gaze seems to be more affected than the horizontal one, and upward gaze more than the downward one. In some cases, the oculomotor deficits can take the appearance of a vertical supranuclear gaze palsy.

17.6 Eye Movements in Multiple Sclerosis

Multiple sclerosis (MS) is a common inflammatory demyelinating disorder, which attacks the myelin of the white matter of the brain. MS starts typically in younger adults, and is twice as common in women than in men. Epidemiologic studies suggest greater risk for MS in people living further away from the equator. MS lesions have a widespread extension, and damage supra- and infratentorial structures. This explains why eye movement disorders are very common in MS, and may occur either as an early manifestation, or later during the course of the disease (Frohman, Zee, Mc Coll, & Galetta, 2005). An accurate examination of eye movements can thus help to diagnose and document the progression of the disease.

Patients with MS often complain about blurred vision, oscillopsia, diplopia, and a feeling of unsteadiness. During clinical examination, a variety of findings may be observed (Table 17.1).

Oculomotor manifestations of MS (after Frohman et al., 2005; Leigh & Zee, 2015)

Disorder	Characteristic signs	Localisation
Saccadic intrusions	Square wave jerks Macro-square waves Macrosaccadic oscillations Ocular flutter Opsoclonus Microsaccadic flutter	Cerebellum Brainstem

(continued)

(continued)

Disorder	Characteristic signs	Localisation
Internuclear ophthalmoplegia (INO)	Adduction slowing Adduction limitation Abduction nystagmus Low vertical pursuit gain Vertical nystagmus (in some cases) Skew deviation (in some cases) Vertical saccades preserved Vergence preserved (in most cases)	MLF
Wall-eyed and bilateral internuclear ophthalmoplegia	Wall-eyed and bilateral INO Loss of convergence	Rostral midbrain, involving bilateral MLF and vergence pathways or CN III pathways to medial rectus muscles
Saccadic dysmetria	Overshoots and undershoots Ipsipulsion Contrapulsion	Cerebellum Brainstem
Vestibulo-ocular reflex suppression	Can't suppress VOR during head/eye tracking	Cerebellar flocculus
Saccadic long latency	Slow initiation	Parietal lobe
Saccadic slowing	Reduced pulse (velocity)	PPRF
Cranial nerve palsies	Complete or partial Isolated In conjunction with other features	VI > III > IV
Horizontal gaze palsies (unilateral or bilateral)	Decreased velocity and amplitude of gaze to left, right, or both	PPRF Abducens (VI) nucleus VI + INO
One-and-a-half syndrome	Gaze palsy to one side INO to the other side Paralytic pontine exotropia	PPRF and/or VI nucleus and MLF on same side
Horizontal monocular failure	Absent or diminished adduction and abduction	VI fascicle + MLF on same side
Pursuit abnormalities	Low gain or saccadic High gain	Pursuit circuitry Attention/concentration
Skew deviation	Vertical misalignment Vertical diplopia Often with INO Head tilt-away from hyper Subjective vertical deviation Ocular counter-roll	Anywhere in vestibular system that can affect the linear otoliths (e.g., utriculus) Hyper on side of pontine lesionor opposite to medullary lesion

(continued)

(continued)

Disorder	Characteristic signs	Localisation
Nystagmus	Gaze evoked Multidirectional Upbeat and downbeat Vestibular Dysconjugate (with INO) Rebound (cerebellar) Pendular types Occult (with ophthalmoscopy) Periodic alternating	Cerebellum Brainstem Vestibular apparatus, either central or peripheral (as with BPPV)

BPPV: benign paroxysmal positioning vertigo; *CN* cranial nerve; *INO* internuclear ophthalmoplegia; *MLF* medial longitudinal fasciculus; *PPRF* paramediane pontine reticular formation; *VOR* vestibulo-ocular reflex

As shown in the table, most of the eye movement disturbances found in MS are due to lesions of the brainstem and the cerebellum. Indeed, these structures are well myelinated, and are therefore often a major target of the disease.

The exact prevalence of eye movement disorders in MS is not known, and varies from 32% (Jozefowicz-Korczynska, Łukomski, & Pajor, 2008) to 80% (Reulen, Sanders, & Hogenhuis, 1983). In a recent study (Servillo et al., 2014), a sample of 163 MS patients were assessed. Among them, 68% of the patients showed at least one abnormality of eye movements. Most frequent oculomotor dysfunctions were impaired smooth pursuit (42.3%), saccadic dysmetria (41.7%), unilateral INO (14.7%), slowing of saccades (14.7%), skew deviation (13.5%), and gaze evoked nystagmus (13.5%).

17.7 Conclusions

In the present chapter, we showed a wide spectrum of eye movement disorders found in neurological diseases. The pattern of eye movement disorders is dependent on the damage in the central nervous system. Lesions in the brainstem and the cerebellum provoke distinctive neurological deficits combined with eye movements disturbances, which were already described in the 19th century. Lesions in the cortex and the basal ganglia often lead to subtle eye movement disturbances and, in addition to the clinical examination, quantitative eye movement recordings are needed. We also presented the neural network controlling saccades and smooth pursuit, and illustrated the clinical consequences of isolated lesions within this network.

Finally, the role of eye movement disturbances as a marker in neurodegenerative diseases was illustrated. First results in this research area indicate that distinct patterns of eye movement disturbances may be used as biomarkers, already in the preclinical stages of neurodegenerative diseases. However, further research in this field is needed.

17.8 Suggested Readings

Schiefer U, Wilhelm H, Hart W. (Eds.). Neuro-Ophthalmology. A Practical Guide. 2007. Springer Berlin Heidelberg New York.

- *Well-illustrated clinical textbook, which also includes ophthalmological aspects.*

Leigh JR, Zee DS. (2015). The Neurology of Eye Movements. Contemporary Neurology Series. Oxford University Press. 5th edition.

- *Standard textbook, which includes basic, neurophysiological, and clinical aspects of eye movements and their diseases.*

17.9 Questions Students Should Be Able to Answer After Completing This Chapter

(1) Can you describe the important aspects of the clinical-neurological examination of the eye movements?
(2) Can you describe the anatomical organization of eye movement control?
(3) What is the difference between eye movement control in the brainstem and the cortex?
(4) What are the main findings related to eye movements in MS?
(5) What are the main findings related to eye movements in PD?

Bibliography

Anderson, T. J., & MacAskill, M. R. (2013). Eye movements in patients with neurodegenerative disorders. *Nature Reviews Neurology, 9,* 74–85.

Bogousslavsky, J., & Meienberg, O. (1987). Eye-movement disorders in brain-stem and cerebellar stroke. *Archives of Neurology, 44*, 141–148.

EROS Investigators. (2009). Incidence of stroke in Europe at the beginning of the 21st century. *Stroke, 40*, 1557–1563.

Frohman, E. M., Frohman, T. C., Zee, D. S., Mc Coll, R., & Galetta, S. (2005). The neuro-ophthalmology of multiple sclerosis. *The Lancet Neurology, 4,* 111–121.

Jozefowicz-Korczynska, M., Łukomski, M., & Pajor, A. (2008). Identification of internuclear ophthalmoplegia signs in multiple sclerosis patients: Saccade test analysis. *Journal of Neurology, 255,* 1006–1011.

Karnath, H.-O., & Thier, P. (Eds.). (2006). *Neuropsychologie*, 2nd edn. (p. 218). Berlin Heidelberg New York: Springer.

Mackay, J., & Mensah, G. (2004). *The Atlas of heart disease and stroke*. Geneva, Switzerland: World Health Organization. http://www.who.int/cardiovascular_diseases/en/cvd_atlas_15_burden_stroke.pdf.

Meienberg, O., & Müri, R. (1992). Nuclear and infranuclear disorders. *Baillieres Clinical Neurology, 1,* 417–434.

Pierrot-Deseilligny, C., Milea, D., & Müri, R. M. (2004). Eye movement control by the cerebral cortex. *Current Opinion in Neurology, 17,* 17–25.

Ptak, R., & Müri, R. M. (2004). The parietal cortex and saccade planning: lessons from human lesion studies. *Frontiers in Human Neuroscience, 7*, 254. https://doi.org/10.3389/fnhum.2013.00254.

Reulen, J. P. H., Sanders, E. A. C. M., & Hogenhuis, L. A. H. (1983). Eye movement disorders in multiple sclerosis and optic neuritis. *Brain, 106,* 121–140.

Servillo, G., Renard, D., Taieb, G., Labauge, P., Bastide, S., Zorzon, M., et al. (2014). Bedside tested ocular motor disorders in multiple sclerosis patients. *Multiple Sclerosis International, 2014,* 732329.

Vidailhet, M., Rivaud, S., Gouider-Khouja, N., Pillon, B., Bonnet, A. M., Gaymard, B., et al. (1994). Eye movements in parkinsonian syndromes. *Annals of Neurology, 35,* 420–426.

Chapter 18
Eye Movements as Biomarkers to Evaluate Pharmacological Effects on Brain Systems

Tatiana Karpouzian, Nadine Petrovsky, Ulrich Ettinger and James Reilly

Contents

Abstract Eye movements lend themselves to the study of pharmacological influences on sensorimotor and cognitive processes. In this chapter, we provide an introduction to the study of pharmacological influences on eye movements. We will first

---

T. Karpouzian (✉) · J. Reilly
Department of Psychiatry and Behavioral Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA
e-mail: t-karpouzian@northwestern.edu

N. Petrovsky · U. Ettinger
Department of Psychology, University of Bonn, Bonn, Germany

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_18

introduce key concepts and methods in pharmacological research before providing a selective but representative overview of studies that have evaluated pharmacological effects on oculomotor control. We will focus on benzodiazepines and other sedative compounds, antipsychotics, anticonvulsants, mood stabilizers, ketamine, nicotine, methylphenidate, and dextroamphetamine in healthy individuals, and we will survey treatment effects of antipsychotics in schizophrenia, stimulants in ADHD, and stimulants in Parkinson's disease. We will introduce the concept of pharmacological model systems and will explain how eye movements may be used in such designs.

18.1 Introduction and Learning Objectives

Eye movement paradigms have been widely used to probe the functional cortical and subcortical systems involved in sensorimotor and cognitive processes that are selectively intact or disturbed in various clinical conditions. Increasingly, these paradigms have been used as potential biomarkers to evaluate pharmacological effects on these functional brain systems. A biomarker is defined as a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a drug.

There are several reasons why eye movement paradigms are particularly useful tools for examining pharmacological effects on sensorimotor and cognitive systems. First, the neurophysiologic and neurochemical basis of oculomotor control has been well characterized both from single unit recording studies of non-human primates and functional imaging studies in humans. Second, advances in behavioral pharmacology have clarified effects of certain drugs on specific brain regions that are important for subserving oculomotor control. Third, oculomotor performance can be reliably measured and quantified, and paradigms can be experimentally manipulated to better understand underlying differences in performance changes. Fourth, eye movement tasks are relatively easy to perform with limited burden placed on patients and can be used across a range of ages and clinical severity levels. Finally, these paradigms may be used to examine dose dependent effects or to evaluate acute versus chronic exposure in clinical samples.

The focus of this chapter will be on the applied use of eye movements in clinical research settings as biomarkers for: (1) side effect profiles of drugs in healthy individuals; (2) identifying treatment targets in model systems of disease; (3) evaluating potential cognitive enhancement effects of drugs in healthy individuals; and (4) effects of pharmacological treatments on sensorimotor and cognitive systems in psychiatric conditions and neurologic conditions. We begin with an overview of methods and approaches in pharmacology to provide a context for the introduction to the literature examining pharmacological effects on eye movements in subsequent sections of the chapter.

A major learning objective of this chapter is to learn about fundamental methods in psychopharmacological research, both generally and specifically with regards to eye movements. Having read this chapter you should be able to design a pharmacological study of eye movements by making use of key recommendations for the conduct of such studies. You will also have acquired an overview of the effects of psychopharmacological compounds on eye movement control, both in healthy individuals and in psychiatric and neurological patients. On the basis of this introduction, you will be able to understand and critically evaluate current empirical studies in the field.

18.2 Historical Annotations

The study of pharmacological treatments on eye movements began as early as the 1940s. Initially, studies focused on the examination of agents or drugs on basic aspects of visual tracking, such as nystagmus (rapid involuntary eye movements) induced by chronic alcohol use (Bender & Brown, 1948) or barbiturate (Rashbass, 1959) or decreased vergence with barbiturate use (Westheimer & Rashbass, 1961). These findings were typically derived from single case studies or conducted using small samples of healthy individuals. Further, there was a movement in the field toward quantifying the sedating effects of drugs, such as alcohol, barbiturates, and anticonvulsants using eye movements in contrast to more subjective means such as an individual's rating of drowsiness. One of the earliest eye movement studies to evaluate the sedative effects of medication on eye movements was conducted by Aschoff (1968), in which reduced saccade velocity and increased reaction time was observed in response to intravenously administered diazepam. Several other studies demonstrated the effects of alcohol use on eye movements, including reduced velocity (Wilkinson, Kine, & Purnell, 1974), reduced amplitude, and overall decreased quality of saccade waveforms (Guedry, Gilson, Schroeder, & Collins, 1974).

In more recent years, research has focused on evaluation of pharmacological treatments on aspects of cognition in both healthy individuals and clinical populations, including pharmacogenetic effects, and use of eye movement measures as biomarkers for the evaluation of potential cognitive enhancers or and the identification of treatment targets in model systems of disease. These studies will be described in further detail in Sects. 18.4 and 18.5.

18.3 Methods and Approaches in Pharmacology

In this section, special emphasis is placed on methodological issues for designing, planning, and conducting a pharmacological study or clinical trial. Considerations for optimizing eye movement paradigms for pharmacological studies are presented.

18.3.1 What Is Pharmacology? What Is Psychopharmacology?

Initially, it is important to define some key terms. *Pharmacology* is the study of how therapeutic or non-therapeutic chemical agents (i.e., drugs) influence biological systems (at the cellular, physiologic, organ, and organism levels), and the mechanisms by which they do so. Pharmacology often divides a drug's interactions with the body into their pharmacokinetic and pharmacodynamics aspects (Tambour & Crabbe, 2010). *Pharmacokinetics*, sometimes described as what the body does to a drug, incorporates drug absorption, distribution, metabolism, and excretion. *Pharmacodynamics*, described as what a drug does to the body, involves receptor binding, post-receptor effects, and chemical interactions. A drug's pharmacokinetics and pharmacodynamics are both genetically and environmentally influenced (Tambour & Crabbe, 2010).

Psychopharmacology is the scientific field that utilizes drugs or other chemical agents to understand neural function, to prevent and treat mental illness and drug abuse, and to understand how nontherapeutic psychoactive drugs and natural substances alter human mood, cognition, motor activity, endocrine, and other centrally mediated functions (Stolerman, 2010). When using psychopharmacology as a tool, especially for psychologists and cognitive neuroscientists, it is of interest to understand how drugs influence cognition, perception, mood, and behavior through their actions on cellular receptors.

As mentioned above, pharmacokinetics and pharmacodynamics are both at least in part influenced by genetics. The field of *pharmacogenetics* studies how genetic variation influences the response to a drug. More specifically, it is the study of how polymorphisms in genes that encode transporters, metabolizing enzymes, receptors, and other drug targets are related with variations in responses to drugs, including toxic and therapeutic effects (Tambour & Crabbe, 2010). The terms *pharmacogenetics* and *pharmacogenomics* are often used interchangeably, however, we prefer to make the following distinction: while pharmacogenetics focus on single gene variants (i.e., single gene-drug interactions), pharmacogenomics refer to several gene variants across the whole genome (i.e., multiple gene-drug interactions).

Genes contain information that determines the structure of proteins. Any variations in the DNA sequence (mutations) may alter the expression or the structure of proteins. DNA mutations that occur at a frequency of 1% or greater are termed polymorphisms (Tambour & Crabbe, 2010). Polymorphisms in genes coding for a protein that carries a drug to its target cells or tissues may cripple the enzyme that activates a drug or aid its removal from the body, and thus may induce pharmacokinetic or pharmacodynamic variations leading to individual differences in the response to the drug. Examples of genetic variations that affect the response to pharmacological agents are single nucleotide polymorphisms (SNPs), repetitive sequences, deleted DNA sequences, and alterations in chromosome structure (Reilly, Lencer, Bishop, Keedy, & Sweeney, 2008). Common sources of genetic variation examined in pharmacogenetic studies include SNPs that are located in genes coding for sites of drug action (e.g. receptors or transporters), and drug metabolizing enzymes (e.g. cytochrome P450 enzyme variants) (Reilly et al., 2008) (see Box 1).

Box 1: Drug Metabolizing Enzymes

Cytochromes P450 are a good example of how drug metabolizing enzymes influence the response to a drug. Cytochromes P450 (CYPs) belong to a large superfamily of metabolizing enzymes. Within the CYP2 family, polymorphic CYP2D6 was one of the first and most important drug-metabolizing enzymes to be characterized at the genetic level (Tambour & Crabbe, 2010). It is estimated that CYP2D6 is responsible for the metabolism and elimination of approximately 25% of clinically used drugs (Wang et al. 2009). CYP2D6 is primarily expressed in the liver and there is substantial variation among people in the amount and the efficiency of CYP2D6 enzyme produced: four phenotypes have been described: "poor metabolizers", "Intermediate metabolizers", "rapid metabolizers", and "ultrarapid metabolizers". Ultrarapid metabolizers have multiple copies of the CYP2D6 gene expressed and greater-than-normal CYP2D6 activity. Therefore, ultrarapid metabolizers may not achieve therapeutic levels of a certain drug with usual doses and may require several doses to show a response. On the other hand, poor metabolizers are at increased risk of toxicity from CYP2D6 substrate drugs (Tambour & Crabbe, 2010).

What do these and other pharmacogenetic phenomena mean for psychologists, psychiatrists and cognitive neuroscientists? First, it is important to be aware that such phenomena as genetic differences in drug metabolizing exist. Cognitive effects of drugs in a sample of individuals may be blurred or may not be found at all due to the existence of inter-individual differences in genetic variation. Second, it might also be the case that "poor" and "good" responders to the drug under study may occur within the same sample—this might be due at least in part to differences in genetic make-up.

Accordingly, some researchers also take blood or saliva samples from their participants in order to perform DNA analyses to determine whether associations between gene variants and drug outcomes exist (see Fig. 18.1). Two scenarios are possible in this context: First, participants of a drug study may be genotyped after the study was conducted – this would be an a posteriori analysis strategy to find whether genetic polymorphisms influence drug outcome. Second, other researchers opt for an a priori strategy—that is having a specific hypothesis on the consequences of a genetic polymorphism beforehand.

An example of the latter approach is the evaluation of the relationship between working memory, variants in the gene coding for the enzyme COMT (the *COMT* gene) and the pharmacological substance tolcapone (a COMT inhibitor). COMT (catechol-*O*-methyltransferase) is an enzyme that breaks down catecholamines such as dopamine, epinephrine, and norepinephrine—all of which are neurotransmitters important for working memory and associated cognitive abilities. One already

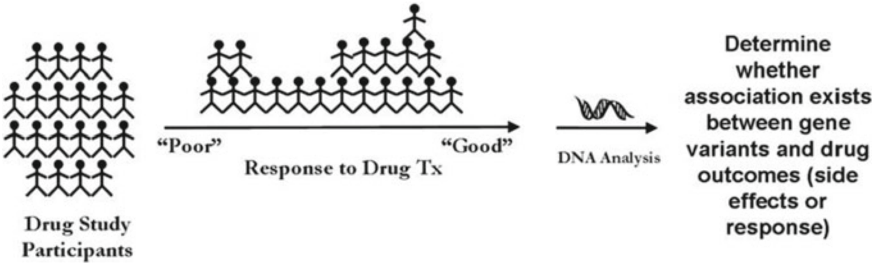


Fig. 18.1 Genetic variation may account for "poor" and "good" responders to a drug. Reprinted from Reilly et al. (2008), Copyright (2008), with permission from Elsevier

known and well-studied polymorphism in the *COMT* gene is the so-called Val158Met (rs4680) polymorphism. This is a functional SNP resulting in a valine (Val) to methionine (Met) substitute thereby altering the activity of the enzyme. The Met-allele is the low-activity allele; being homozygous (that is having two copies of the Met-allele) leads to a 3-4-fold reduction in enzymatic activity compared to being homozygous for the Val-allele (Lachman et al., 1996). As a consequence, carriers of the Met-allele degrade dopamine slower and thus are known to have higher tonic levels of dopamine in the frontal cortex (Bilder, Volavka, Lachman, & Grace, 2004).

Higher tonic levels of dopamine in prefrontal and frontal areas have been associated with better working memory performance (Bilder et al. 2004). A recent study by Farrell and colleagues (Farrell, Tunbridge, Braeutigam, & Harrison, 2012) found that males who are homozygous for the Met-allele outperformed those men who are homozygous for the Val-allele in the N-back task, a task measuring working memory performance. Furthermore, the COMT inhibitor tolcapone reversed these differences between the two genotype groups: it worsened N-back performance in MetMet participants but enhanced it in ValVal participants (Farrell et al., 2012).

18.3.2 *Methodological Issues in Pharmacological Studies of Eye Movements*

The following section addresses various topics on study design such as sample size, method of drug delivery, and the "window of opportunity" for evaluating drug effects. Furthermore, the use of eye movements to evaluate effects of acute versus long-term exposure to drugs as well as the reliability and sensitivity of eye movement measures is discussed.

Box 2: Is My Study a Clinical Trial?

Before conducting your research, you will need to check with the appropriate authorities in your country whether your study will be considered a clinical trial. For example, in the United States this is the Food and Drug Administration (FDA) which is responsible for drug approval, in the United Kingdom the Medicines and Healthcare Products Regulatory Agency (MHRA), and in Germany the Federal Institute for Drugs and Medical Devices (Bundesinstitut für Arzneimittel und Medizinprodukte, BfArM).

In medical research, a clinical trial refers to a study which provides measures of efficacy for new interventions, drugs, or devices (Van Spall, Toren, Kiss, & Fowler, 2007). Clinical trials are an essential part of the development and licensing of medicinal products. They are intended to prove the efficacy of new medicinal products and to demonstrate their safety. Clinical trials are performed before the product concerned is placed on the market and after licensing, for instance to provide important findings on long-term effects of treatment or to test whether the drug also works in other patient groups. An overview of the phases of clinical research is depicted in Table 18.1. It might be the case that the authority in your country will decide that your study might be a phase-IV clinical trial. In that case, it is advisable to check whether your university or research unit offers support for conducting a clinical trial. Some universities have specialized clinical trial units which offer support and guidance for the procedures relating to clinical trials such as registration and monitoring.

It is also helpful to familiarize oneself with the guidelines made by the International Conference on Harmonization (ICH). The ICH guidelines provide helpful documents on topics such as Good Clinical Practice (GCP) and the content of clinical study reports. Finally, for publication in most scientific journals it is often necessary to register your study with an eligible database such as http://www.clinicaltrials.gov. Another helpful link on the web is the checklist by the CONSORT (Consolidated Standards of Reporting Trials) group which gives recommendations for reporting clinical trials: http://www.consort-statement.org/consort-statement/.

18.3.2.1 How Do I Determine the Right Sample Size?

During the planning process of your study, you will need to generate an estimate of your sample size. This is important for several reasons. First, you will need a certain number of participants in order to be able to find a statistical effect of a certain size. Second, you will need to make a project schedule and get an idea of how long it will take to test all participants. Third, you also need the number of participants to calculate the costs of the study, e.g. for the reimbursement of the participants and for study personnel who will be in charge of running the study.

Table 18.1 Phases of clinical research

Phase	Question	Method/subjects
Pre-clinical	Is it worth testing?	In vitro: test tube or cell culture In vivo: animals
I	Dose ranging and safety—Is it safe?	Healthy humans
II	Efficacy—Does it work?	Patients (small sample)
III	Comparison—Does it work better?	Patients (large, multi-center samples)
IV	Comparison—What more can we learn?	After approval and marketing, e.g. other patient groups

See also https://www.nlm.nih.gov/services/ctphases.html or http://www.fda.gov/ForPatients/Approvals/Drugs/default.htm for further information

How do you determine the minimum number of participants for your pharmacological study? This is best done by performing an a priori power analysis. The power of a statistical test is the probability that its null hypothesis (H0) will be rejected given that it is in fact false. Obviously, significance tests that lack statistical power are of limited use because they cannot reliably discriminate between H0 and the alternative hypothesis (H1) of interest (Faul, Erdfelder, Lang, & Buchner, 2007).

In a priori power analyses, sample size N is computed as a function of the required power level $(1-\beta)$, the pre-specified significance level α, and the population effect size to be detected with probability $(1-\beta)$. For clarification, β is the so-called type II error (i.e. a “false negative”) when you reject H1 although in fact H1 is true. The type I error (i.e. α, the “false positive”) is the error you make when you reject H0 and decide that H1 is true when in fact H0 is true. Typically, your α error (i.e. your level of significance) is set at the 5% level ($p = 0.05$). It is often recommended that the power of your statistical test greater or equal to 80% ($p \geq 0.80$) (Faul et al., 2007). Sample size N and statistical power can be calculated for many different statistical tests such as t tests, F tests, χ^2 tests, z tests and some exact tests (e.g. using G*Power (Faul et al., 2007).

18.3.2.2 Study Population

One of the first considerations in choosing the study population should be whether you want to test male participants only or both males and females. Many pharmacological studies only include males in order to avoid possible confounding effects from fluctuations in hormone levels due to menstrual cycle. A disadvantage of the “males only” design is that you can only make generalizations about half of the population and you cannot analyze your data concerning any possible sex differences. Indeed, at least some cognitive functions vary during the phases of the menstrual cycle such as memory performance (Farage, Osborn, & MacLean, 2008; Sherwin, 2012). However, this issue has not been studied extensively.

If you test both males and females, it will be useful to document for females the first day of the last menstruation to be able to infer in which phase of the cycle (follicular or luteal) the cognitive testing took place. You might want to keep the menstrual phase constant for all female participants if you do not want additional variance in your data, or you might test females both in the follicular and in the luteal phase if you wish to distinguish between those and analyze your data accordingly. At the minimum, menstrual cycle should be well documented for your study to be able to account for this factor in data analysis.

Another issue concerns the decision to test healthy individuals only or to include a sample of patients. If you plan a pharmacological study that includes patients, then be aware that patients greatly vary intra- and inter-individually regarding psychopathology and medication history, even within a single diagnostic category. This means that drug effects cannot be studied without the confounding effects of psychopathology and medication. A good example of this issue is schizophrenia, an illness that is highly heterogeneous with significant variation in amount and kind of medication between patients. Again, it is crucial to document all medications the patients take and for some classes of medications there are conversions to standard units which make the different medications and doses of medications comparable between patients (e.g. a standard in schizophrenia research is to convert different antipsychotics to chlorpromazine equivalents). If you are interested in treatment efficacy of a substance and you test healthy individuals only, then treatment efficacy is difficult to assess: deficit correction cannot be determined in healthy participants unless deficits are first induced.

18.3.2.3 Study Design: Between- or Within-Subjects?

In a between-subjects design, different groups of participants are tested in parallel (sometimes called a parallel group design), with each group receiving a different substance (see Fig. 18.2). Note that in a between subjects design, each participant is only tested once.

In contrast, in a within-subjects design, all participants undergo each treatment condition, that is, each participant receives all substances (see Fig. 18.2) in what is referred to as a cross-over design. Note that in a within-subjects design, each participant is tested several times. In this instance randomization of treatment conditions across participants is crucial.

Importantly, it should be pointed out that randomization cannot be implemented in every study. Clinical studies of treatment effects in patients in particular are often limited in study planning, e.g. by self-assignment of patients rather than randomization (see Sect. 18.4 below).

What are the advantages and disadvantages of between- and within-subjects designs? In between-subjects designs, a large number of participants per group is required in order to have enough statistical power to detect treatment effects. By contrast, within-subjects designs offer higher statistical power and a smaller sample size is sufficient. The different treatment groups in a between-subjects design have to be

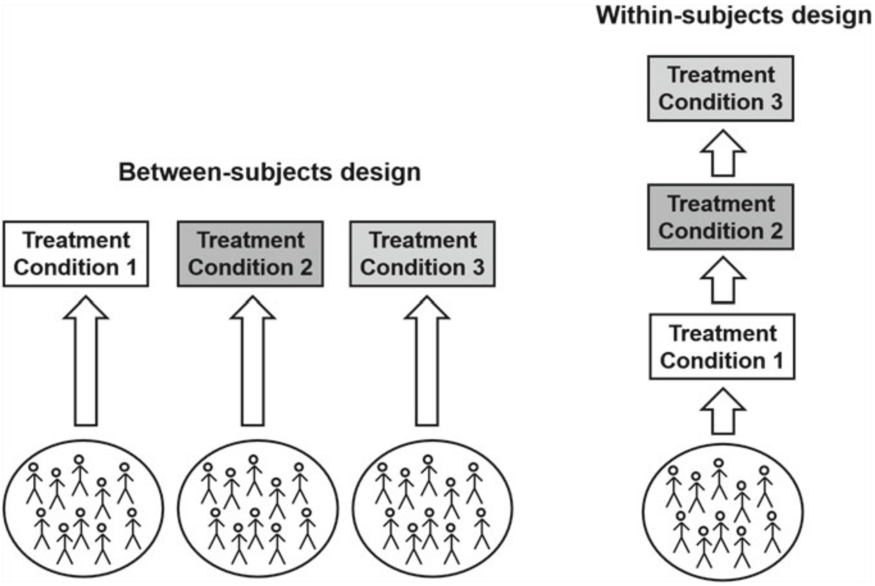


Fig. 18.2 Pharmacological studies can be divided in between-subjects design (i.e., group design) and within-subjects design (i.e., repeated-measures design)

carefully matched regarding demographic variables (and other possible confounding variables) such as level of education, age, and sex—otherwise the variance between the different treatment conditions is too high due to individual differences which arise from the participants. In within-subjects designs, variability between treatment conditions is generally low, as each participant undergoes all treatment conditions and by doing so serving as his/her own control (i.e. error variance associated with individual differences is low).

Despite the advantages of within-subjects designs, an important issue to consider in such studies concerns drug washout. The washout period refers to the time that is required for the administered compound to be eliminated from the body. Whilst this period varies between drugs, in experimental psychopharmacological studies with within-subjects designs, assessments often take place at one-week intervals. This is done not only to allow sufficient time for drug washout but also to control potential influences on behavior or performance that may vary throughout the week.

Another potential limitation of the sensitive within-subjects design concerns effects on performance of repeated exposure to the tasks under investigation (see also Sect. 18.3.3.2 below). Whilst oculomotor tasks tend to be trait-like and stable over time (Meyhöfer, Bertsch, Esser, & Ettinger, 2015), performance improvements have been observed following task repetition, likely reflecting effects of learning (e.g. Dyckman & McDowell 2005; Ettinger et al., 2003a). Such improvements may be problematic in within-subjects pharmacological studies as they may overlay or

Table 18.2 Latin square design for randomizing the order of treatment conditions in a within-subjects design (A = substance 1, B = substance 2, C = substance 3)

	Time 1	Time 2	Time 3
Group 1	A	B	C
Group 2	B	C	A
Group 3	C	A	B

counteract genuine drug effects (e.g. Ettinger et al. 2003b; Green, King, & Trimble, 2000).

As mentioned above, randomization is crucial in study design. When opting for a between-subjects design, you need to ensure that participants are randomly allocated to the different treatment groups. At the same time, you have to monitor the recruitment of participants so that all groups will be closely matched regarding individual variables of the participants. If you decide on a within-subjects design, you need to randomize the order of the treatment conditions across the participants. There are several ways of randomizing conditions; one established method is the Latin square method. A Latin square design is a blocking design with two orthogonal variables: the 2 blocking variables are divided into a tabular grid with the property that each row and each column receives each treatment exactly once. In a pharmacological study, these two blocking variables would be group of participants and time. If there are three different substances, then a Latin square design will lead to nine squares (3 points in time × 3 substances), see Table 18.2.

In summary, a pharmacological study should follow the “gold standard” – that is a randomized, controlled trial (RCT): randomized meaning randomization to treatments (see above) and controlled referring to comparison to a placebo condition. Furthermore, a pharmacological study should ideally be double-blind: both the participant and the experimenter (including individuals collecting, scoring, and analyzing data) are unaware which substance is administered in the respective treatment condition. This design controls for expectancy effects on the part of both parties.

18.3.2.4 The Window of Opportunity Is Crucial in Evaluating Drug Effects

Section 18.3.1 already introduced the concepts of pharmacokinetics and pharmacodynamics. Indeed, if you plan a pharmacological study, you need to familiarize yourself with some of the facts concerning your study drug, e.g. how quickly it is absorbed, when peak plasma concentrations are reached, how stable plasma concentrations are and how fast the drug is metabolized. Acquiring this information is important as you will only have a certain “window of opportunity” when plasma concentrations of your drug have reached a stable plateau in which you need to run your paradigms to be able to detect drug effects. Usually, you will find information on pharmacokinetics and pharmacodynamics in the summary of product characteristics (SPC) and in papers in pharmacology journals.

An important point to consider in this context is the route of administration for your drug. The route of administration will influence how fast the drug is absorbed and metabolized and this in turn will influence your window of opportunity for testing. An example is nicotine, which can be administered in several different ways, for instance via subcutaneous injection, a chewing gum or a lozenge, with a nasal spray or as a patch. While subcutaneous injection, chewing gum, lozenge and nasal spray deliver nicotine quickly (i.e. within minutes), nicotine from a patch is absorbed more slowly (a stable plateau is reached after approximately 2–4 h). However, the plasma nicotine concentrations will be more stable and longer lasting with a patch, thereby increasing the window of opportunity for testing.

With different routes of administration, different undesirable effects (side effects) arise. Thus, you will need to balance the way of administration against the severity of the side effects. For instance, it might seem appealing to deliver nicotine via nasal spray because it is quick, however the tingling sensations in the nose and a high incidence for nausea might be unfavorable. In contrast, nicotine delivery via patch might be time-consuming, but the side effects will be milder as nicotine is absorbed less rapidly.

18.3.2.5 Eye Movements Can Be Studied to Evaluate Both Acute and Long-Term Drug Effects

While most researchers opt for testing the acute effects of a certain drug on the sensorimotor and cognitive processes that underlie oculomotor behavior, eye movement paradigms also have the potential to provide information on the long-term effects of drugs on these processes.

A good example for the usefulness of eye movements in evaluating the long-term effects of medication is a study about the effects of two different benzodiazepines. As discussed below, benzodiazepines amplify the effect of the neurotransmitter gamma-aminobutyric acid (GABA), and thereby exerting sedative, sleep-inducing, and anxiolytic effects, and thus are often prescribed for sleep and anxiety disorders. Benzodiazepines are highly effective, but the risk of becoming addicted is also high. Therefore, they are usually only prescribed for a limited time period. However, there are a considerable number of patients who become chronic users. In a study by van Steveninck and colleagues (van Steveninck et al., 1997), the effects of the chronic use of the two benzodiazepines temazepam and lorazepam were evaluated. Chronic users of these medications (length of use was 1–20 years) were compared with control participants regarding plasma benzodiazepine concentrations, saccadic performance, visual analog scales, and other measures after having received a single dose of temazepam or lorazepam. Saccade performance was studied because the maximal velocity of saccadic eye movements is a highly sensitive parameter for sedative benzodiazepine effects (Hommer et al., 1986; van Steveninck et al., 1992), i.e. the sedative effect of benzodiazepines can be shown with saccadic slowing. Results showed that for temazepam, there were no significant differences between chronic users and controls regarding plasma concentration and saccadic velocity; both groups showed comparable saccadic slowing. However, for lorazepam, the chronic users showed reduced sensitivity to the effects of the drug as they exhibited less saccadic slowing than the controls. This effect on saccades demonstrates the reduced drug sensitivity in the chronic users. Nevertheless, despite this indication for tolerance in the chronic lorazepam users, a single dose of lorazepam still caused significant slowing effects on saccadic eye movements (van Steveninck et al., 1997).

This example shows that eye movement measures are sensitive not only to reveal acute drug effects, but also the effects of long-term exposure to drugs as discussed in Sects. 18.1 and 18.2 below.

18.3.3 Recommendations Towards Optimizing Eye Movement Paradigms for Pharmacological Studies

The following section provides some guidelines and considerations for optimizing eye movement paradigms for pharmacological studies. Topics include methods of standardization and addressing practice effects that occur with repeated testing (e.g., before and after treatment). Finally, we present some recommendations for task parameters in oculomotor paradigms.

18.3.4 Standardization: General Recommendations

Standardization of methods is crucial in order to be able to provide a safe environment for your participants and to obtain repeatable results. This section will focus on some general recommendations for the execution phase of a pharmacological study.

Prior to starting a research study, it is helpful to establish and document Standard Operating Procedures (SOPs). According to the ICH Guideline for Good Clinical Practice E6(R1), SOPs are *detailed, written instructions to achieve uniformity of the performance of a specific function*. More precisely, SOPs should include all the necessary steps the study personnel has to perform at each testing session. These steps may include, amongst others, how to administer the study drug, how to give the instructions to the participants, how to calibrate the eye tracker and how to run the eye movement paradigm with the respective computer software. It is important that SOPs are sufficiently detailed and easy to understand. It is also useful to have a time table for each testing session in which all the events of a testing session are documented, that means the experimenter has a list with each step that has to be performed (e.g. check vitals, hand out questionnaire, administer drug, etc.) and writes down when each step was performed. That way, timeliness is granted and no step is forgotten during the course of events.

Special emphasis should be given to the instructions for the participants. Written instructions are preferred; however, the experimenter should always ensure that the participant understands the task—a good way for doing this is to let the participant explain the task in her/his own words after she/he has read the text. When explanation of instructions is needed, it is important to use exactly the same words (and ideally stress the same syllables) because the way information is presented might influence the performance of the participant. For example, in an antisaccade task it may well make a difference whether participants are instructed "to look in the opposite direction" or "to look at the mirror image location" of the peripheral stimulus—the demands on the processes involved in the vector transformation are considerably greater in the latter instance (Hutton & Ettinger, 2006; Mosimann, Felblinger, Colloby, & Muri, 2004).

SOPs are also critical for data processing and analysis. Usually, physiological data such as eye movement data require visual inspection aside from automatic processing steps with computer software. Therefore, criteria should be defined for data analysis to which all raters adhere. Examples for such criteria include criteria for excluding artifacts in data and when trials need to be excluded. Moreover, it is important that the raters of the data are blind to experimental conditions, that is, they must not know whether it is data from the placebo or verum group conducting analyses—this will help to minimize expectancy effects.

18.3.5 Repeated Testing and Practice Effects

The section provides recommendations and considerations for addressing repeated testing and practice effects that arise in the context of a within subject design (see Sect. 18.3.2.3). A standard way of accounting for practice effects is to counterbalance experimental sessions (i.e. drug conditions) (see Sect. 18.3.2.3).

In addition to counterbalancing drug conditions, some researchers also introduce a so-called "baseline session" to their study design. A baseline session includes all procedures as the drug conditions (i.e. eye movement paradigms, questionnaires, visual analog scales, etc.) but without the substances (verum or placebo). Data from the baseline session can be taken into account when analyzing your data for drug effects: findings are interpreted relative to baseline performance and test whether there is improvement compared to baseline. Moreover, a baseline session has the advantage that participants will already be familiar with all the experimental procedures when they start with the randomized sessions. Practice effects from baseline to the first drug session (i.e. practice effects between first session and second session) might be the strongest but these practice effects might already decrease between the first and the second drug session (i.e. session two and three).

Another issue with repeated testing is that eye movement paradigms should exhibit high test-retest reliability (Ettinger et al., 2003a). Ideally, you consult published data on reliability or conduct your own reliability study before your pharmacological project to test whether performance on your eye movement paradigms stays stable over several sessions in the absence of pharmacological influences.

Finally, when statistically analyzing your repeated-measures data, it is advisable to include order of administration as a factor as it might play a role. It is thorough to look at effects of order in your data, however, most papers do not report on this issue when the "gold standard" of counterbalancing was used.

18.3.6 Recommendations for Designing an Eye Movement Paradigm

As aforementioned, one problem with eye movement research is the lack of standardization of task parameters across different research groups and studies. Therefore, it is difficult to compare results across studies. However, recently, researchers have made an effort to establish guidelines for particular parameters in eye movement paradigms, e.g. recommendations for stimulus size and trial length. These guidelines have been mainly designed for research in psychiatry and neurology; but these recommendations are transferable to the field of pharmacology. So far, recommendations for standardization have been addressed extensively by Smyrnis (2008) and Antoniades and colleagues (Antoniades et al., 2013). Although these recommendations might still be tentative, they provide a good starting point for working out the many details of a final synthesis of recommendations in an effort to achieve the standardization of oculomotor function test procedures and outcome measurements (Smyrnis, 2008). While the paper by Smyrnis gives advice on smooth pursuit and saccadic eye movement tasks, the paper by Antoniades and coworkers aimed at establishing an internationally standardized antisaccade protocol. For additional information on eye movement task recommendations, you can refer to the original publications.

18.4 Findings from Pharmacological Effects on Eye Movements in Healthy Individuals

In the following section, the main applications for studying the pharmacological effects on eye movements in healthy individuals are considered, specifically (1) studies designed to monitor side effects of drugs, (2) the use eye of movement tasks as treatment targets in model systems of disease, and (3) examination of eye movements as biomarkers for the study of cognitive enhancement.

18.4.1 Eye Movement Tasks Can Be Used to Monitor Side Effects of Drugs

Eye movement measures have been used as biomarkers to monitor sedative and other side effects across a range of drugs on sensorimotor and cognitive processes.

18.4.1.1 Benzodiazepines and Other Sedatives

Benzodiazepines are commonly used for the treatment of symptoms of anxiety. Benzodiazepines increase the effects of the neurotransmitter gamma-aminobutyric acid (GABA) and are mediated through agonism of GABA-benzodiazepine chloride receptor complex, such as $GABA_A$ 2 and $GABA_A$ 3 receptors (Reilly et al., 2008). The mechanism of this class of drugs results in anxiolytic, hypnotic, anticonvulsants, and muscle relaxant effects (Möhler, Fritschy, & Rudolph, 2002). The influence of benzodiazepines on eye movements is well established. Studies with nonhuman primates have shown that GABA-ergic projections from the substantia nigra pars reticulata to the superior colliculus are involved in the generation of saccades (Hikosaka & Wurtz, 1983; Reilly et al., 2008). A local injection of a GABA agonist into the superior colliculus reduces saccadic amplitude and increases latency of saccades (Hikosaka & Wurtz, 1985a, 1985b). Similarly, prolonged latency and fixation deficits following an injection of a GABA agonist into the frontal eye fields have been demonstrated (Dias, Kiesau, & Segraves, 1995).

In healthy control subjects, numerous studies have indicated a dose-dependent slowing of saccade peak velocity subsequent to benzodiazepines exposure, including diazepam, lorazepam, midazolam, and tenazepam (Ball et al., 1991; Rothenberg & Selkoe, 1981a). Additionally, there is evidence for log-linear correlations between peak saccadic velocity and serum benzodiazepine concentrations after exposure to temazepam, diazepam, and nitrazepam (Bittencourt, Wade, Smith, & Richens, 1981), as well as correlations between diazepam-induced changes in saccadic eye velocity and increasing plasma diazepam concentrations (Hommer et al., 1986). The most consistently reported effect of benzodiazepines, and one that has been shown to be more sensitive to the sedative effects of these medications than traditional neuropsychological measures, is decreased saccade velocity (Blom, Bartel, de Sommers, van der Meyden, & Becker, 1986; de Visser et al., 2003; Salonen, Aaltonen, Aantaa, & Kanto, 1986).

Benzodiazepines administered to healthy individuals can also influence other saccade parameters, including decreasing saccade acceleration and deceleration, and potentially increased saccade error (Ball et al., 1991). Masson and colleagues (2000) conducted a placebo-controlled double-blind study examining the effects of low dose lorazepam on prosaccade latency using a gap and overlap paradigm. This paradigm varies the temporal interval between central fixation offset and the peripheral target appearance. A gap condition refers to the temporal gap between the offset of the fixation target and peripheral target, while the overlap condition refers to the persistence of the fixation target after the appearance of the peripheral target (see Pierce et al., this volume). These manipulations result in a shortening of saccade latency in gap trials due to a release of the visual fixation system, or lengthening of latency in overlap trials due to persistent engagement of the fixation system after the peripheral target appearance. Compared to placebo, lorazepam significantly increased the latency of prosaccades under both conditions, though did not modify the gap/overlap effect; this finding suggests a generalized slowing of response time, without modification of the fixation release system. Notably, this effect on prosaccade latency may be drug specific, as it has not been observed by midazolam (Ball et al., 1991).

Similar to findings from saccade tasks, benzodiazepines have been shown to decrease smooth pursuit velocity (Padoan, Korttila, Magnusson, Pyykko, & Schalen, 1992). Further, reduced smooth pursuit velocity has also been shown to correlate with serum concentrations of benzodiazepines, such as temazepam and diazepam (Bittencourt, Wade, Smith, & Richens, 1983). In addition to effects on pursuit velocity, a placebo-controlled double blind study revealed that lorazepam resulted in increased smooth pursuit latency, reduced pursuit gain, and increased catch-up saccade activity to correct for reduced velocity (Masson et al., 2000), particularly at higher doses. Other studies have also observed dose-dependent reductions in smooth pursuit gain in healthy participants following diazepam exposure (Rothenberg & Selkoe, 1981b).

To conclude, the effects of benzodiazepines on eye movements have most consistently shown a dose-dependent slowing of peak saccade and smooth pursuit velocity. These findings have been shown to be more sensitive markers of the sedative effects of benzodiazepines compared to other measures of psychomotor speed and attention (Reilly et al., 2008). Furthermore, the association between serum concentrations of benzodiazepines and effects on velocity of saccadic and pursuit eye movements suggest that this is a sensitive and reliable measure of this class of drug's pharmacokinetics and pharmacodynamics.

18.4.1.2 Antipsychotics

Antipsychotic medications render their effectiveness, in part, through their ability to block central dopamine receptors, particularly D2 receptors. In addition to D2 antagonism, this class of medications also renders their benefit through 5HT-2 antagonism. Antipsychotics have been associated with the emergence of extrapyramidal side effects, such as involuntary movements. First-generation (or typical) antipsychotics refer to medications developed in the 1950s and were first used to treat psychosis, often at doses that resulted in sedation or adverse extrapyramidal effects (Reilly et al., 2008). Conversely, second-generation (or atypical) antipsychotics emerged in the 1990s, and are effective at typically lower doses and with less adverse side effects. First-generation antipsychotics block D2 receptors, while second-generation antipsychotics moderately block D2 receptors, but also block serotonin and other receptors.

As with benzodiazepines, the most consistent finding of the effects of first and second-generation antipsychotic medications is slowed saccadic peak velocity. In a study conducted by Lynch, King, Green, Byth, and Wilson-Davis (1997), healthy participants completed prosaccade (or visually guided saccade) and pursuit eye movement tasks before and 2, 4, and 6 h after administration of single doses of lorazepam (2.5 mg) or the first-generation antipsychotic haloperidol (2, 4, or 6 mg). Findings from this study indicated that similar to lorazepam, haloperidol exposure resulted in a dose-dependent decrease in peak saccade velocity on the visually guided saccade task at the higher doses (i.e., 4 and 6 mg). Further, there was no effect of haloperidol on saccade latency, and no effect of haloperidol on pursuit position error, velocity, or saccadic intrusions during the smooth pursuit task. These results parallel prior studies indicating that first-generation antipsychotics result in slowed peak saccade velocity, though do not adversely affect smooth pursuit (Holzman, Levy, Uhlenhuth, Proctor, & Freedman, 1975; King et al., 1995; Reilly et al., 2008).

Several studies have examined the effects of antipsychotics on cognitive or attentional control. In a randomized cross-over study completed by Green and King (1998) healthy subjects performed visual fixation (i.e., no saccade), prosaccade, and antisaccade tasks after receiving single doses of the benzodiazepine lorazepam (2 mg), the first-generation antipsychotic chlorpromazine (50, 75, and 100 mg), and placebo. The prosaccade task consisted of target displacements at 30 and 40°. The no-saccade task required the subject to fixate on a central location, and when the appearance of a peripheral target occurred, they were instructed to not look to the target, but remain fixated at the central location. The antisaccade paradigm, a measure of inhibitory behavioral control, requires a subject to fixate centrally to a target, and when a peripheral target appears, they are instructed to generate a saccade and to its mirrored location in the opposite direction. The results of the study indicated a dose-dependent decrease in peak saccade velocity in the antisaccade and prosaccade tasks after exposure to chlorpromazine. However, chlorpromazine exposure did not affect antisaccade latency or error rate (the percentage of trials where a saccade was incorrectly made to the target) on the antisaccade or no-saccade tasks. Chlorpromazine slowed prosaccade latency at the highest doses, while lorazepam increased both antisaccade latency, as well as errors during the antisaccade and no-saccade tasks.

In a parallel group placebo-controlled study comparing the effects of acute doses of the second-generation antipsychotics amisulpride (300 mg) and risperidone (3 mg), and the first-generation antipsychotic chlorpromazine (100 mg), healthy control subjects performed prosaccade and antisaccade eye movement tasks before and 3 h after drug exposure (Barrett, Bell, Watson, & King, 2004). Although amisulpride did not affect eye movement parameters, both risperidone and chlorpromazine resulted in a slowing of peak saccadic velocity and increased error rates on the antisaccade task. None of the antipsychotic medications had an effect on prosaccade or antisaccade latencies.

To summarize this section, antipsychotic medications have similar sedating effects to those reported in benzodiazepine use, particularly at higher doses, as both may result in a slowing of peak saccade velocity. However, the dose-dependent slowing of peak velocity has only been reported with use of first-generation antipsychotics, which may be suggestive of D2 antagonism. However, unlike benzodiazepines, antipsychotic medications do not result in significant changes in prosaccade or antisaccade latencies in healthy individuals, and first-generation antipsychotics do not seem to adversely affect pursuit velocity among healthy individuals (Reilly et al., 2008).

18.4.1.3 Anticonvulsants and Mood Stabilizers

Anticonvulsant and mood stabilizing medications are used for treatment of seizure disorders and as acute and maintenance treatments for bipolar disorder. These drugs have several mechanisms of action, with glutamatergic antagonism as a common mechanism among many of them.

Newer anticonvulsant medications, such as gabapentin, are associated with less serious side effects than those reported with use of older medications, such as carbamazepine. Noachtar, von Maydell, Fuhry, and Buttner (1998) compared the effects of gabapentin (600 mg) to carbamazepine (400 mg) and placebo in healthy individuals. Participants performed prosaccade and smooth pursuit tasks 2, 5, and 7 h after drug administration. Compared to placebo, both gabapentin and carbamazepine resulted in a reduction of peak saccade velocity; however, the effects of gabapentin on saccade velocity were present only at 2 h, whereas carbamazepine's effect was observed at 7 h after administration. Additionally, the duration of saccades was significantly longer with carbamazepine both at 2 and 7 h after intake, effects which were not observed with gabapentin. These findings illustrate the utility of eye movements for illustrating different pharmacokinetic properties of drugs, as it appears that the gabapentin may be cleared more readily than carbamazepine and with fewer adverse effects.

Another anticonvulsant that is commonly used as mood stabilizer in the treatment of affective disorders is lamotrigine, which is a glutamatergic antagonist and 5-HT agonist. In a double blind, crossover study conducted by Cohen and colleagues (1985), healthy volunteers completed prosaccade and smooth pursuit tasks after receiving 120 and 240 mg of lamotrigine, 10 mg of diazepam (a benzodiazepine known to cause sedation), 0.5 and 1.0 g of phenytoin (an anticonvulsant medication), and placebo. Diazepam resulted in a reduction in saccade peak velocity, which lasted 8 h after administration. Further, diazepam exposure increased saccade duration but not latency. These effects were not seen in lamotrigine or phenytoin. Lastly, both phenytoin and diazepam reduced smooth pursuit performance, while lamotrigine did not have this effect. The findings of this study are consistent with other studies that have shown no adverse effects on smooth pursuit tracking after lamotrigine administration (Peck, 1991), and which suggest that lamotrigine may have a more favorable side effect profile compared to other anticonvulsant medications.

Lithium is a mood stabilizing medication that is often used in the treatment of bipolar disorder, and early studies with this clinical population have suggested that adverse effects on smooth pursuit maybe associated with lithium exposure (Holzman, O'Brian, & Waternaux, 1991; Levy et al., 1985). To examine this hypothesis, Flechtner and colleagues (1992) examined smooth pursuit performance in healthy individuals after receiving either lithium carbonate injection or placebo. Subjects completed a smooth pursuit task at baseline and after 2 weeks of receiving treatment under randomized double blind conditions. No differences in the quality of smooth pursuit eye movements between the placebo and lithium groups were observed, suggesting that smooth pursuit impairment may be attributable to disease rather than lithium exposure effects in clinical populations.

18.4.2 Eye Movement Measures Are Studied as Treatment Targets in Model Systems of Disease

The following section will consider research that uses eye movements as surrogate treatment targets in pharmacological model systems of psychiatric and neurological diseases. A *pharmacological model system* refers to the administration of a drug with known effects (agonist/antagonist) on neurotransmitter systems to healthy humans or animals, resulting in changes that mimic symptom characteristics of a particular disease.

An example of such a model is administration of ketamine, an N-methyl-D-aspartate (NMDA) receptor antagonist that induces transient psychotomimetic effects in healthy individuals as well as temporary cognitive impairments similar to those observed in schizophrenia (Krystal et al., 2003). Individuals with schizophrenia demonstrate impairment in saccadic and pursuit eye movements (see chapter by Smyrnis et al. in this volume), leading Radant, Bowdle, Cowley, Kharasch, and Roy-Byrne (1998) to examine the effects of ketamine on eye movements in healthy individuals as a pharmacological model for this illness. Healthy individuals received either a placebo or ketamine infusion in a randomized single blind placebo controlled design with oculomotor performance measured using prosaccade, antisaccade, and smooth pursuit tasks. Subjects received progressively higher doses of ketamine with plasma concentrations of 50, 100, 150, and 200 ng/ml over a 2-h period, with oculomotor tasks administered after each infusion step. Compared to placebo ketamine induced a dose-dependent decrease in peak velocity prosaccade, increase in prosaccade latency, decrease in smooth pursuit gain, and increase in catch-up saccade frequency and amplitude during pursuit. Interestingly, ketamine did not adversely impact antisaccade performance. A recent study replicated the adverse effect of ketamine on smooth pursuit performance and additionally showed that these drug-induced impairments are accompanied by reductions in blood oxygen level dependent (BOLD) signal in a task-related network of primary visual cortex, area V5 and the right frontal eye field (FEF) (Steffens et al., 2016). Overall, these findings provide partial support for the effects of ketamine in inducing eye movement abnormalities observed in individuals with schizophrenia, such as impaired smooth pursuit, and illustrate the potential of NMDA receptor antagonism as a pharmacological model for schizophrenia (Radant et al., 1998).

Another application of a pharmacological model system is to evaluate the effects of antipsychotic medications among healthy individuals under the transient effects of ketamine and observe whether such medications can prevent or ameliorate any temporarily induced behavioral deficits. A recent study by Schmechtig et al. (2013) adopted this approach using a double-blind, randomized, placebo-control, parallel groups design in which healthy individuals performed prosaccade, antisaccade, and smooth pursuit tasks under one of four conditions: (1) placebo capsule and saline infusion, (2) placebo capsule and ketamine infusion, (3) risperidone capsule (2 mg) and saline infusion, or (4) risperidone capsule (2 mg) and ketamine infusion. As previously observed (Radant et al., 1998), ketamine was associated with impairment in smooth pursuit, reflected by increased saccadic frequency and decreased velocity gain, but was not observed to impact prosaccade or antisaccade performance. Risperidone administration resulted in a decreased gain and slower peak velocities in both the prosaccade and antisaccade tasks, and risperidone did not reverse any of the ketamine-induced oculomotor changes. These findings suggest that risperidone lacks cognitive enhancing effects on oculomotor biomarkers in the ketamine model system of schizophrenia (Schmechtig et al., 2013).

18.4.3 Eye Movements Are Used as Biomarkers for the Study of Cognitive Enhancement

Several drugs, both therapeutic and those of abuse, may alleviate cognitive deficits among clinical populations or enhance normative cognitive functions among healthy individuals. The most widely studied drug class for such cognitive enhancing effects is stimulants, which render their effects through augmentation of synaptic action of norepinephrine and dopamine neurotransmitter systems (and to a lesser extent, serotonin).

18.4.3.1 Nicotine

Nicotine is a cholinergic agonist that binds to nicotinic acetylcholine receptors and has stimulant properties. These receptors facilitate the release of other neurotransmitters, including dopamine, acetylcholine, and glutamate and are involved in regulating multiple cognitive functions. At certain doses, nicotine's beneficial effects include increased psychomotor speed, improved sustained attention, and greater performance on tasks of cognitive control (Reilly et al., 2008). Cholinergic inputs to structures in the brainstem, including the superior colliculi, facilitate motor outputs for the initiation of saccades (Kobayashi & Isa, 2002). For example, in single unit recordings of animals who are administered nicotine, there is a greater firing of cells in the substantia nigra, which resulted in shorter saccade latencies subsequent to increased inhibitory input to the fixation zone of the superior colliculus (Clarke, Hommer, Pert, & Skirboll, 1985).

In a study examining the effects of nicotine on antisaccade performance, a group of healthy individuals who smoked 10–20 cigarettes a day completed baseline and re-test antisaccade tasks on each of two test sessions separated between 2 and 7 days (Rycroft, Hutton, & Rusted, 2006). The antisaccade task employed in this study included a 200 and 500 ms gap condition in order to increase potential for observing the effects of nicotine on antisaccade performance. Sessions were counterbalanced, such that half of the participants smoked between baseline and retesting in their first session and did not smoke during the second session, and the other half of the sample did not smoke during the first session, but smoked between baseline and retesting in the second session. This study design enabled evaluation of whether acute nicotine exposure between baseline and retesting influence antisaccade performance. Results of the study indicated that nicotine exposure improved antisaccade performance reflected by a decrease in antisaccade error rate and antisaccade latency from baseline to retest. However, this effect was present only among subjects who smoked during the first session but not those who smoked during the second session. One explanation for these findings is that practice effects between session 1 and 2 were larger than the effect of nicotine. Thus, it is possible that the potential enhancing effects of nicotine are apparent when subjects have not been already exposed to the task. The findings of Rycroft et al., 2006 are consistent with other eye movements studies, which have demonstrated that nicotine administration results a reduction in antisaccade latencies and error rate (Depatie et al., 2002; Ettinger et al., 2009; Larrison, Briand, & Sereno, 2004; Powell, Dawkins, & Davis, 2002), possibly due to nicotine's contribution to improved attention and control in the inhibition of reflexive responses. Evidence from functional neuroimaging shows that antisaccade improvements with nicotine are accompanied by reduced, i.e. more efficient BOLD signal (Ettinger et al., 2009).

Less consistent findings of cognitive enhancing effects of nicotine on smooth pursuit performance have been reported (Kasparbauer et al., 2016). For example, in a study examining nicotine effects on smooth pursuit, Domino, Ni, and Zhang (1997) evaluated pursuit eye movements in healthy non-smokers and smokers before and after inhalation of a sham cigarette or a cigarette of their preferred choice. Smooth pursuit performance was measured at 5 min before and at 0, 3, 6, 10, 20, and 30 min after smoking a sham cigarette or tobacco cigarette. An increase in smooth pursuit velocity was observed when tracking a 15 degree per second velocity stimulus for both smokers and non-smokers, but not at slower tracking speeds (i.e., 6 degree per second velocity stimulus). Other studies, however, have reported contrasting findings of the effects of nicotine exposure on smooth pursuit performance. For example, Olincy, Ross, Young, Roath, and Freedman (1998) found that healthy smokers abstinent from cigarettes for several hours prior to testing did not show any change in pursuit gain 10–15 min after smoking.

There is therefore variable evidence for nicotine's potential enhancing effects on oculomotor measures, with perhaps some indication that it may improve antisaccade performance and smooth pursuit under certain conditions. The variability in findings may be influenced by methodological differences including subjects' familiarity with the task, baseline smoking status, and stimulus presentation conditions.

18.4.3.2 Methylphenidate

Methylphenidate is a commonly prescribed psychostimulant medication, most often to individuals with attention deficit hyperactivity disorder (ADHD), where its efficacy for treating clinical symptoms is well established and its beneficial effects on cognitive deficits in this clinical population is increasingly appreciated. Methylphenidate renders its clinical efficacy on ADHD symptoms and cognitive deficits through blockage of the dopamine transporter thereby increasing the availability of dopamine and noradrenaline. More recently, methylphenidate has been under consideration as a potential cognitive enhancer in healthy individual individuals, although there has been limited empirical support to support this use.

Allman, Ettinger, Joober, and O'Driscoll (2012) evaluated the effects of a single 20 mg dose of methylphenidate on oculomotor performance among healthy male volunteers in a double-blind placebo-controlled crossover design study. Subjects performed no gap prosaccade and antisaccade tasks, a predictive saccade task and a smooth pursuit task. Eye movement testing occurred at a baseline visit and then again at two subsequent visits separated by 1 week visits where subjects were randomized to receive methylphenidate or placebo (with the opposite condition assignment at the third visit). As noted above (Sect. 18.3.3.2), the baseline visit prior to randomization reduces the likelihood of practice effects on the medication trial. Compared to the placebo condition, methylphenidate administration resulted in significantly reduced prosaccade latency, significantly increased peak velocity and frequency of predictive saccades (particularly in conditions with predictable timing), and increased gain and reduced saccades during pursuit tracking. Antisaccade performance (latency or error rate) was unaffected by methylphenidate treatment in these healthy individuals. These findings of speeded prosaccade latency and improved smooth pursuit are generally consistent with those observed in studies of ADHD patient groups treated with methylphenidate, and may reflect enhancing effects in timing related behavioral functions. However, a more recent study did not find improvements in SPEM performance, albeit with a larger dose of 40 mg (Kasparbauer et al., 2016).

18.4.3.3 D-Amphetamine

Dextroamphetamine (D-amphetamine), is another psychostimulant medication that is prescribed for the treatment of ADHD and that is also under investigation as a potential cognitive enhancer among healthy individuals.

The effects of D-amphetamine on antisaccade and predictive saccade performance were evaluated by Allman et al. (2010) in a double-blind crossover design. Twenty-four healthy individuals completed an antisaccade task and a predictive saccade task at a baseline visit, and again at two subsequent visits where they were randomized to receive 0.3 mg/kg D-amphetamine or placebo (with the opposite condition assigned at the third visit). Unlike findings reported by Allman et al. (2012) no effect of drug was observed on the frequency of predictive saccades during the predictive saccade task. On the antisaccade task, however, error rate significantly decreased after D-amphetamine administration regardless of the baseline performance level whereas the effect of drug administration on antisaccade latency depended on baseline performance level. Among those individuals who had shorter antisaccade latencies at baseline (i.e., good performance), D-amphetamine resulted in a prolongation of antisaccade latencies suggesting an adverse effect for this set of individuals. In contrast, those individuals with longer antisaccade latencies at baseline had reduced latencies following drug administration, consistent with a beneficial effect. These findings are consistent with an inverted-U relationship between dopamine activity level and performance. Importantly, this study illustrates that the effects of a drug on oculomotor performance may depend on pre-exposure performance levels.

18.4.4 Benefits and Limitations to Studying Drug Effects in Healthy Individuals

We will conclude this part of our chapter by briefly evaluating the overall approach of studying pharmacological effects on eye movements in healthy individuals. There are both limitations and benefits to this approach.

A major benefit is that that the influence of confounds often present in studies with clinical samples are diminished. These confounds may include illness chronicity, co-morbid conditions, polypharmacy, and prior treatment exposure, all of which can obscure any effect of the drug or dose that is under investigation. Specifically, increasing efforts are now placed on including potential biomarkers of clinical endpoints earlier in the drug evaluation process. This can be complicated if patients are included, given heterogeneity with respect to disease severity and chronicity, and concomitant or prior treatment. However, most studies with healthy individuals are limited by use of acute rather than chronic treatment and the obvious absence of disease characteristics, which may be necessary when trying to evaluate efficacy (Table 18.3).

18.5 Findings from Pharmacological Effects on Eye Movements in Patient Groups

There are several reasons why there is utility in studying effects of clinical pharmacological treatments on eye movements in psychiatric and neurologic patient groups. First, eye movement tasks are relatively easy to perform, require relatively minimal engagement or cooperation from participants, and can be completed in patient groups across a range of ages and levels of clinical acuity. Second, the neurotransmitter systems and neural circuitry regulating eye movements are well characterized (Kobayashi & Isa, 2002; Leigh & Zee, 2006). Third, the increasing appreciation for cognitive deficits as underlying poor functional outcomes for patients, and how

Table 18.3 Summary of pharmacological effects on eye movements in healthy individuals (Reilly et al., 2008)

Pharmacological groups and agents	Presumed neurotransmitter system(s) influenced	Effects on saccades	Effects on pursuit
Sedatives			
Benzodiazepines	GABA agonist	Dose-dependent decrease in prosaccade peak velocity (Ball et al., 1991; Bittencourt et al., 1981)	Dose-dependent decrease in SPEM Velocity (Green et al., 2000; Bittencourt et al., 1983)
Diazepam		Decrease in prosaccade peak acceleration/deceleration (Ball et al., 1991)	Decrease in SPEM velocity (Pardoan et al., 1992)
Lorazepam		Increase in prosaccade latency (Masson et al., 2000); increase in antisaccade latency and error rate (Green & King, 1998)	Increase in SPEM latency and reduced SPEM gain (Masson et al., 2000)
Antipsychotics			
First Generation	DA antagonist	Dose-dependent decrease in prosaccade peak velocity (Lynch et al., 1997; King et al., 1995)	
Chlorpromazine		No effect on pro latency except at highest doses (Green & King, 1998); higher antisaccade error rate (Barrett et al., 2004)	
Haloperidol		Increase in antisaccade error (Barrett, 2004)	
Second Generation	DA antagonist	Decrease in prosaccade peak velocity (Barrett et al., 2004)	
Risperidone	5-HT antagonist	Increase in antisaccade error rate (Barrett et al., 2004)	

(continued)

Table 18.3 (continued)

Pharmacological groups and agents	Presumed neurotransmitter system(s) influenced	Effects on saccades	Effects on pursuit
Antidepressants			
Selective Serotonin Reuptake Inhibitors	5-HT agonist	No change in peak prosaccade velocity (Morrens et al., 2007)	Increase in SPEM velocity and gain (Friedman, Jesberger, & Meltzer, 1994; Gijsman et al., 2002)
Stimulants			
Dextroamphetamine		Decrease in antisaccade error rate (Allman et al., 2010)	
Methylphenidate	NE and DA reuptake inhibitor	Decrease in prosaccade latency, increase in prosaccade peak velocity (Allman et al., 2012)	Increase SPEM gain and decrease in saccades during tracking (Allman et al., 2012)
Nicotine	ACh agonist	Possible decrease in antisaccade latency and increase in error rate (Rycroft et al., 2006)	Increase in SPEM velocity to faster targets (Domino et al., 1997)
Anticonvulsants/mood stabilizers			
Carbamazepine	NE antagonist, DA and GABA agonist	Decrease in prosaccade peak velocity and increase in prosaccade duration (Noachtar et al., 1998)	
Gabapentin	Glutamate antagonist	Decrease in prosaccade peak velocity (Noachtar et al., 1998)	
Lamotrigine	Glutamate antagonist, possible weak 5-HT, DA, and NE antagonist	No reported effects on prosaccades (Cohen et al., 1985)	No reported effects on SPEM (Cohen et al., 1985; Peck, 1991)
Lithium	Modulation of glutamate, though mechanism still largely unknown		No reported effects on SPEM (Flechtner et al., 1992)

(continued)

Table 18.3 (continued)

Pharmacological groups and agents	Presumed neurotransmitter system(s) influenced	Effects on saccades	Effects on pursuit
Other			
Ketamine	NMDA antagonist	Dose-dependent decrease in peak prosaccade velocity and increase in prosaccade latency (Radant et al., 1998)	Dose-dependent decrease in SPEM gain and increase in catch-up saccade frequency and amplitude (Radant et al., 1998)

these deficits are impacted by existing treatments from a neural systems basis. Lastly, drugs targeting cognitive deficits in patient groups are, in the near term, likely to be adjunctive to existing treatments thereby increasing the importance for understanding the effect of such treatments on cognitive systems. In this section we provide exemplary discussions of three clinical disorders and how their corresponding medications impact eye movements.

18.5.1 Effects of Antipsychotic Medications in Schizophrenia

The effects of pharmacological treatments on eye movements in patients with schizophrenia are well documented. Schizophrenia is a chronic psychiatric illness that is characterized by positive and negative symptoms. Positive symptoms include hallucinations, delusions, and thought disorder, while negative symptoms include anhedonia, alogia, avolition and asociality (Tamminga, Buchanan & Gold, 1998). Cognitive deficits are also a characteristic of individuals with schizophrenia that appears independent of positive symptoms and are more closely related to negative symptoms. These cognitive deficits include attention, working memory, episodic memory and executive functioning, which is a set of cognitive processes that underlie reasoning, planning, problem solving, and mental flexibility. The general consensus is that schizophrenia is a complex disorder with a multifactorial etiology with multiple genes of small effect interacting with environmental insults leading to the development of the disorder (Siever & Davis, 2004).

In terms of the pathophysiology underlying schizophrenia, much research has supported the hypothesis that increased levels of striatal dopamine are related to positive symptomatology, and that such symptoms are reduced by antipsychotic medications through blockade of dopamine D2 receptors. Reduced function of the NMDA receptor leading to reductions in glutamate also likely plays a role and may further contribute to symptoms as well as associated cognitive deficits. This notion is also supported by studies that have shown that drugs such ketamine or phencyclidine, both of which are NMDA receptor antagonists, can induce the positive and cognitive symptoms characteristic of schizophrenia (as discussed above).

18.5.1.1 Effects of Antipsychotic Medications on Saccadic Eye Movements in Schizophrenia

Straube, Riedel, Eggert, and Muller (1999) and Muller, Riedel, Eggert, and Straube (1999) evaluated the effects of antipsychotic medications on eye movement performance in a group of first-episode schizophrenia patients who were either antipsychotic-naïve at the time of testing or had been antipsychotic free for at least four weeks prior to testing, as well as after antipsychotic treatment among a subset of these patients. Subjects completed gap and overlap visually guided saccade and antisaccade tasks and two different memory guided saccade tasks - one that involved memory of a single target location and another that involved memory for an ordered sequence of three target locations. Eye movement performance was compared between groups of medicated and unmedicated patients and healthy controls. First or second generation antipsychotic treatment resulted in a reduction in peak saccade velocity; this effect was larger for internally guided saccades (i.e. antisaccade and memory guided saccade) than for externally triggered saccades (i.e. visually guided saccade). There were no differences in saccadic velocity between the unmedicated patients and controls. Only mild and nonsignificant reductions in antisaccade latency and memory-guided saccade gain were observed in the medicated group compared to unmedicated group. Lastly, there were no significant treatment effects on antisaccade error rate, which is consistent with findings from studies with first-episode schizophrenia patients (Ettinger & Kumari, 2003; Harris, Reilly, Keshavan, & Sweeney, 2006; Hutton et al., 1998; Reilly et al., 2008).

Burke & Reveley (2002) examined prosaccade and antisaccade performance in schizophrenia patients in a within subject cross-over design that involved switching from a first-generation antipsychotic to the second-generation antipsychotic risperidone, or vice versa. A reduction in antisaccade error rate was observed in patients who switched from a first-generation antipsychotic to risperidone, while patients who switched from risperidone to a first-generation drug had increased antisaccade error rate. Thus, risperidone treatment was associated with improved antisaccade performance in this study.

Another study, which used a randomized treatment design, compared the effects of two second-generation antipsychotic medications, risperidone and olanzapine, on visually guided saccade, antisaccade and memory guided saccade tasks in first-episode patients with schizophrenia and healthy controls (Broerse, Crawford, & Den Boer, 2002). Compared to controls, patients made more errors on the antisaccade task, inhibition errors on the memory guided saccade task, and had reduced amplitudes of memory guided saccades; these effects were comparable between the two medication groups.

Several longitudinal (i.e., within group) studies of first-episode patients have examined antipsychotic treatment effects on saccadic eye movements over time after treatment initiation. A sample of antipsychotic naïve patients were evaluated before and after 6 weeks of treatment of the first-generation antipsychotic haloperidol or second-generation antipsychotic risperidone in comparison to healthy individuals followed over a similar time period (Harris, Wiseman, Reilly, Keshavan, & Sweeney, 2009; Harris et al., 2006; Reilly, Harris, Keshavan, & Sweeney, 2005; Reilly, Harris, Keshavan, & Sweeney, 2006). Prior to treatment initiation, patients demonstrated significantly faster visually guided saccade latencies compared to healthy controls. After 6 weeks of treatment with risperidone this atypical speeded response latency was not present among those taking risperidone but persisted among those taking haloperidol (Reilly et al., 2005). In addition, risperidone treatment was associated with a reduction in peak velocity, a modest decrease in prosaccade gain (Reilly et al., 2005), and a reduction in antisaccade latency (Harris et al., 2006).

Another study examined antipsychotic-naïve patients with schizophrenia performing an oculomotor delayed response (or memory guided saccade) task. Prior to treatment and early in the course of the illness, schizophrenia patients demonstrated an impairment in maintaining spatial location information in working memory (i.e. reduced memory guided saccade gain) at only the longest delay period duration (8 s) compared to controls (Reilly et al., 2006). After 6 weeks of risperidone treatment and clinical improvement, these deficits significantly worsened, such that patients demonstrated impaired gain across all delay period durations. Similar findings of reduced predictive saccade gain were also observed in these first-episode patients after treatment with risperidone (Harris et al., 2009), suggesting that accuracy of saccades made according to internal representation may be particularly susceptible to antipsychotic treatment effects.

18.5.1.2 Effects of Antipsychotic Medications on Smooth Pursuit in Schizophrenia

Several studies have also evaluated antipsychotic treatment effects in schizophrenia patients performing smooth pursuit tasks. Specifically, untreated patients have demonstrated comparable impairment in reduced smooth pursuit gain and more frequent catch up saccades to patients who were treated with first-generation antipsychotics (Ettinger & Kumari, 2003; Gooding, Iacono, & Beiser, 1994; Reilly et al., 2008; Sweeney et al., 1999; Thaker et al., 1999). For example, Campion and colleagues (1992) assessed smooth pursuit eye movement in healthy controls, and in drug-naive, chronic, and residual schizophrenia patients. Smooth pursuit gain was reduced across schizophrenia groups, and groups did not differ from each other. Taken together, these findings suggest that pursuit eye movements may not be impacted by first-generation antipsychotic treatment, but that smooth pursuit impairments may be related to extent of illness chronicity or may represent a trait marker of schizophrenia.

In another study examining smooth pursuit performance in schizophrenia, Hutton and colleagues (2001) conducted a study comparing smooth pursuit performance in groups of first-episode and chronic patients schizophrenia, as well as healthy controls. First-episode schizophrenia patients with less than 12 weeks of cumulative lifetime antipsychotic exposure and first-episode patients who were untreated at the time of testing, were compared on a smooth pursuit task to chronic schizophrenia patients who were either medicated with first-generation antipsychotics or were antipsychotic free for at least 6 months before testing. Chronic schizophrenic patients demonstrated significantly reduced velocity gain, longer latency to change in target direction, and increased catch-up saccades than first-episode patients and controls. There were no differences between antipsychotic-naive and treated first-episode patients. Antipsychotic-free chronic patients demonstrated less impairment in velocity gain than matched treated chronic patients. These findings suggest that impairment in pursuit performance may be worsened by chronic antipsychotic treatment.

The effects of second-generation antipsychotics on pursuit performance in schizophrenia patients is less known, although some studies have observed worsening pursuit performance in clozapine treated patients (Friedman, Jesberger, & Meltzer, 1992). To evaluate antipsychotic medication treatment effects on pursuit performance more directly, Lencer et al. (2008) evaluated smooth pursuit performance among antipsychotic naïve first-episode patients before and after 6 weeks of treatment with either risperidone or olanzapine and compared performance to controls studied in parallel. Before treatment latency of pursuit was shortened, pursuit gain was impaired (under conditions when tracking less predictable ramp targets requiring a relative high degree of sensorimotor processing), and catch-up saccade frequency was increased (when tracking predictable targets) compared to controls. After 6 weeks of treatment, pursuit gain decreased further to less predictable ramp targets while predictable pursuit performance did not change, suggesting that there may be a selective effect on second-generation antipsychotic treatment on tasks that require a greater extent of sensorimotor processing.

18.5.1.3 Pharmacogenetic Effects of Antipsychotics on Eye Movement in Schizophrenia

As discussed above (see Sect. 18.3.1), pharmacogenetic studies evaluate how genetic variation influences the response to a particular medication or class of medication. Recently, pharmacogenetic studies have demonstrated that the influence of antipsychotic medications on oculomotor measures in schizophrenia patients may be influenced by particular polymorphisms, indicating that genetic variation underlies much of the heterogeneity of treatment related response often observed in groups of patients. While studies have demonstrated an adverse response to risperidone on the accuracy of memory guided saccades after approximately 6 weeks of treatment (Reilly, Harris, Khine, Keshavan, & Sweeney, 2007, Reilly et al., 2006), considerable variability was observed in the extent and magnitude of this adverse effect across patients. In a subsequent study including some of these patients, Bishop et al. (2015) evaluated whether polymorphisms of the Type-3 metabotropic glutamate receptor gene (*GRM3*) and selected variants in candidate dopamine genes were associated with antipsychotic induced changes in memory guided saccade performance. The worsening of memory guided saccade accuracy observed after antipsychotic treatment was associated with variation in *GRM3* polymorphisms, such that those patients with the rs1468412_TT genotype exhibited a substantial worsening saccade accuracy compared to those with the rs1468412_AA genotype. While variants in candidate dopamine genes were associated with memory guided saccade performance, they were not associated with changes in performance following antipsychotic treatment. These findings suggest variation related to altered glutamate signaling exhibit increased sensitivity to the adverse effects of D2 antagonism from antipsychotic drugs on working memory.

In another pharmacogenetic study, Lencer et al. (2014) examined whether polymorphisms of two candidate genes, the dopamine receptor D2 gene (*DRD2*) and *GRM3*, were associated any antipsychotic medication related changes after approximately 6 weeks of treatment in first-episode schizophrenia patients. Variants of the −141C Del/Ins polymorphism in *DRD2* contributed to differences in the extent of slowed initial velocity gain during pursuit in patients, such that deletion carriers had significantly slower velocity gain compared to CC insertion carriers. While no effect of treatment on genotype differences was observed on initial velocity gain, treatment resulted in an increase in pursuit latency to a greater extent in CC insertion carriers compared to −141C deletion carriers. With regard to *GRM3* there was a significant association of rs274622 with maintenance velocity gain, such that rs_274622_CC carriers had poorer sustained pursuit maintenance than T-carriers. There was no differential treatment effect on sustained pursuit performance between genotypes.

18.5.1.4 General Comments Regarding Antipsychotic Medication Effects on Eye Movements in Schizophrenia

In sum, both first- and second- generation antipsychotic medications result in a reduction in peak saccade velocity and decreased saccade gain. Decreased gain is more apparent for internally driven saccades (i.e. memory-guided), than for externally driven (i.e. visually guided), suggesting that this effect may disproportionally affect accuracy of eye movements based on internal representations versus those driven by visual input. Additionally, antipsychotic treatment may result in lengthening of atypically shortened prosaccade latencies observed in untreated psychosis, suggesting some improvement in the regulation of visual attention. Generally, antisaccade error rate remains elevated in comparison to healthy controls despite antipsychotic treatment. Deficits on smooth pursuit tasks persist despite treatment and may be worse in chronically treated patients, suggesting potential cumulative adverse medication effects on pursuit systems (Reilly et al., 2008). Finally, more recent studies have suggested that pharmacogenomic effects may account for some of the observed changes in saccade and pursuit performance among patients after starting antipsychotic medication treatment, and that genetic variation may account for heterogeneity of treatment response in patient groups.

18.5.2 Effects of Stimulants in Attention Deficit Hyperactivity Disorder (ADHD)

Attention deficit hyperactivity disorder (ADHD) is a neurodevelopmental disorder characterized by symptoms of inattentiveness and hyperactivity/impulsivity. Inattentive symptoms include difficulty maintaining focus, becoming easily distracted, and difficulties with organization. Hyperactive/impulsive symptoms include excessive talking, fidgeting, impatience, and interrupting others. If these symptoms are present in several different settings (i.e. work, school, etc.), are present before the age of 12, and interfere significantly with quality of daily functioning, an ADHD diagnosis may be likely. Under current diagnostic nomenclatures (DSM-V, American Psychiatric Association, 2013), subtypes of ADHD include a predominantly inattentive type, predominantly hyperactive type, and a combined type with features of both inattention and hyperactivity. The pathology of ADHD may be due to deficits in the frontal-striatal pathway, the neural pathway underlying executive functions that links the frontal lobe to the basal ganglia. Stimulants such as methylphenidate are an effective treatment for ADHD, as these medications increase the amount of dopamine and norepinephrine available at the synapse, which in turn modulates neural circuits including the frontal lobe and basal ganglia (Reilly et al., 2008).

Klein, Fischer, Fischer, and Hartnegg (2002) studied prosaccade and antisaccade performance in patients (ages 10–15 years) on or off methylphenidate in two separate eye movement sessions. In order to limit any confound of practice effects, pediatric patients were randomized to a different testing order, either on-off medication or off-on mediation. Subjects completed a prosaccade overlap condition and antisaccade gap condition. Methylphenidate resulted in faster latencies on both the prosaccade and antisaccade tasks, reduced antisaccade error rate, and improved error corrections. Collectively, these findings suggested that under treatment with methylphenidate, individuals with ADHD demonstrate a greater control of the voluntary attention systems.

These findings of improved voluntary attentional control were supported by another study that examined the effects of methylphenidate on executive control and response inhibition in children with ADHD compared to controls (O’Driscoll et al., 2005). In addition to prosaccade, antisaccade and predictive saccade tasks, this study also used a task switching paradigm which required subjects to perform either a prosaccade or antisaccade depending on the color of the central fixation. Children with ADHD, inattentive and combined types, were compared to controls under a baseline condition, and were then evaluated in a double blind crossover trial of a single dose of methylphenidate or placebo, with the order of drug administration counterbalanced. When compared to placebo, methylphenidate administration resulted in decreased prosaccade and antisaccade latency, as well as a decrease in antisaccade error rate comparably in both subtypes, and did not impact performance on task switching. These findings suggest that improved executive control is evident during methylphenidate treatment for either clinical subtype.

In a study assessing smooth pursuit performance in individuals with ADHD, 20 children with ADHD were compared to 20 age-matched healthy controls on a smooth pursuit eye movement task (Bylsma & Pivik, 1989). Individuals with ADHD were assessed both on and off medication. Smooth pursuit performance was characterized by velocity arrests, or periods where pursuit velocity reached below 2 degrees per second, and root mean square errors, or eye position error relative to target error. Greater velocity arrest scores were observed in unmedicated children with ADHD compared to healthy controls; however, there were no differences in smooth pursuit error. Furthermore, there were no changes in smooth pursuit performance when comparing patients evaluated on or off methylphenidate.

18.5.3 Effects of Dopamine Agonists in Parkinson's Disease

Parkinson's disease (PD) is a progressive, neurodegenerative disorder of the central nervous system. PD is characterized by motor abnormalities including difficulty initiating or continuing (eye) movements, problems with balance and walking, and tremors (see chapter by Müri et al. in this volume). Symptoms of PD are caused by dopaminergic disruption in the basal ganglia, which plays a central role in (oculo)motor control. The pathology of PD is attributed abnormalities in the nigrostriatal pathway, a dopaminergic pathway that connects the substantia nigra to the striatum. Specifically, levels of dopamine in individuals with PD are greatly reduced. The basal ganglia also has connections to regions responsible for executive control and cognitive processing, and therefore individuals with PD may develop a range of cognitive difficulties, including executive impairment and memory loss, as well as oculomotor abnormalities (Hood et al., 2007). Levodopa (L-Dopa) is pharmacological treatment for Parkinson's disease that is a considered a pre-cursor to dopamine. Once it crosses the blood brain barrier it is converted to dopamine by dopamine-decarboxylase, thereby increasing concentrations of dopamine in the brain.

Prior studies examining pharmacological effects on eye movements in PD have been somewhat mixed. A study conducted by Hood and colleagues (2007) examined the effects of levodopa on individuals with PD. Patients with PD completed eye movement testing both on and off their optimal dose of levodopa, and were also compared to a healthy control group. All subjects completed a prosaccade and antisaccade task, with a gap overlap manipulation. Patients had significantly higher error rate on the antisaccade task (both on/off levodopa) when compared to healthy controls. However, when on levodopa, patients demonstrated had fewer antisaccade errors when compared to their off-levodopa baseline. Patients had slower prosaccade latencies when under on verses off conditions, findings that have been reported by others (Michell et al., 2006). Thus, levodopa may improve voluntary cognitive control, and may slow reflexive responses. Improvements in prosaccade accuracy, which are hypometric in PD, have also been reported in patients on levodopa (Gibson, Pimlott, & Kennard, 1987; Rascol et al., 1989), although not consistently (Nakamura et al., 1991).

Several studies have shown that patients with PD demonstrate reduced maintenance gain on smooth pursuit tasks (Bares et al., 2003; Waterston, Barnes, Grealy, & Collins, 1996). In a study of untreated patients with PD, clinical improvement with dopaminergic treatment was associated with an improvement in smooth pursuit gain (Gibson et al., 1987), suggesting that oculomotor changes in PD may be related to dopamine levels in the basal ganglia. However, in a study completed by Sharpe, Fletcher, Lang, & Zackon (1987), smooth pursuit gain did not improve after Levodopa. Specifically, patients with PD completed a smooth pursuit task during predictable off-periods and during on-periods, with no apparent change in performance.

18.6 Conclusions

There is increasing evidence that eye movements are sensitive biomarkers of drug effects on discrete sensorimotor and cognitive processes both for examining side effects in early phase studies with healthy individuals as well as in studies of potential cognitive enhancers or therapeutics in clinical settings.

In many examples reviewed above, eye movement paradigms have shown greater sensitivity to pharmacological effects than neuropsychological measures or subjective ratings, with strong dose–response effects that potentially could be used to individualize drug dosing for patients.

In addition, use of eye movements in studies involving pharmacogenetics and genetics remains highly promising and may lead to improved understanding of heterogeneity in drug response as well as in drug selection. With the need for biomarkers to support the development of new drugs and assess their efficacy, eye movement studies offer particular methodological benefits including the translational linkages to discrete neurotransmitter and neural systems, the rational manner to evaluate pharmacological effects for proof of concept/target engagement, and for understanding individual differences in drug response.

Despite these advantages, challenges for the practical implementation of eye movement biomarkers remain such as variability in laboratory set-up and paradigm characteristics, the degree of technical sophistication to data collection and analysis (which may limit transportability), and the challenge of substantiating the linkage of eye movement measurements, and drug effects upon them, to functional outcomes in patients.

18.7 Suggested Readings

Ansari A, Osser DN (2015). Psychopharmacology: A Concise Overview for Students and Clinicians (2nd Edition). CreateSpace Independent Publishing Platform. North Charleston, South Carolina, USA

– *An overview of the most important pharmacological therapies currently used in psychiatry.*

de Visser SJ, van der Post JP, de Waal PP, Cornet F, Cohen AF, van Gerven JM (2003). Biomarkers for the effects of benzodiazepines in healthy volunteers. Br J Clin Pharmacol 55(1):39–50.

– *A review underscoring the role of saccadic tasks in the assessment of sedative effects.*

Golan DE, Armstrong EJ, Armstrong AW (2017) Principles of Pharmacology: The Pathophysiologic Basis of Drug Therapy, 4th Edition. Wolters Kluwer Health. Philadelphia, PA, USA.

– *A comprehensive textbook of the principles of pharmacology.*

Reilly JL, Lencer R, Bishop JR, Keedy S, Sweeney JA (2008). Pharmacological treatment effects on eye movement control. Brain Cogn 68(3):415–35.

– *A comprehensive review of drug effects on eye movements.*

Thaker GK (2007). Schizophrenia endophenotypes as treatment targets. Expert Opin Ther Targets 11(9):1189–206.

– *A review of the overlap between oculomotor endophenotypes and biomarkers and their application to pharmacological studies.*

18.8 Questions Students Should Be Able to Answer

(i) Define the terms pharmacokinetics and pharmacodynamics
(ii) What are the advantages and disadvantages of a repeated-measures design in pharmacological studies?
(iii) What are the most consistent effects of benzodiazepines on eye movements in healthy individuals?
(iv) What are the advantages and disadvantages of studying pharmacological influences on eye movements in healthy individuals?
(v) What is the pattern of effects of antipsychotics on eye movements in patients with schizophrenia?
(vi) What are some of the challenges or limitations of studying pharmacological effects on eye movements in clinical populations?

Bibliography

Allman, A., Benkelfat, C., Durand, F., Sibon, I., Dagher, A., Leyton, M., et al. (2010). Effect of d-amphetamine on inhibition and motor planning as a function of baseline performance. *Psychopharmacology, 211*(4), 423–433.

Allman, A., Ettinger, U., Joober, R., & O'Driscoll, G. A. (2012). Effects of methylphenidate on basic higher-order oculomotor functions. *Journal of Psychopharmacology, 26*(11), 1471–1479.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Antoniades, C., Ettinger, U., Gaymard, B., Gilchrist, I., Kristjansson, A., Kennard, C. … Carpenter, R. H. (2013). An internationally standardised antisaccade protocol. *Vision Research, 84,* 1–5.

Antoniades, C. A., Xu, Z., Mason, S. L., Carpenter, R. H., & Barker, R. A. (2010). Huntington's disease: Changes in saccades and hand-tapping over 3 years. *Journal of Neurology, 257*(11), 1890–1898.

Aschoff, J. C. (1968). The effect of diazepam (Valium) on the saccadic eye movements in man. *Archiv fur Psychiatrie und Nervenkrankheiten, 211*(4), 325–332.

Ball, D. M., Glue, P., Wilson, S., & Nutt, D. J. (1991). Pharmacology of saccadic eye movements in man. 1. Effects of the benzodiazepine receptor ligands midazolam and flumazenil. *Psychopharmacology (Berl), 105*(3), 361–367.

Bares, M., Brazdil, M., Kanovsky, P., Jurak, P., Daniel, P., Kukleta, M., & Rektor, I. (2003). The effect of apomorphine administration on smooth pursuit ocular movements in early Parkinsonian patients. *Parkinsonism & Related Disorders, 9*(3), 139–144.

Barrett, S. L., Bell, R., Watson, D., & King, D. J. (2004). Effects of amisulpride, risperidone and chlorpromazine on auditory and visual latent inhibition, prepulse inhibition, executive function and eye movements in healthy volunteers. *Journal of Psychopharmacology, 18*(2), 156–172. https://doi.org/10.1177/0269881104042614.

Bender, M. B., & Brown, C. A. (1948). Character of Nystagmus Induced by Amytal in Chronic Alcoholics. *American Journal of Ophthalmology, 31,* 825–828.

Bilder, R. M., Volavka, J., Lachman, H. M., & Grace, A. A. (2004). The catechol-O-methyltransferase polymorphism: Relations to the tonic-phasic dopamine hypothesis and neuropsychiatric phenotypes. *Neuropsychopharmacology, 29*(11), 1943–1961.

Bishop, J. R., Reilly, J. L., Harris, M. S., Patel, S. R., Kittles, R., Badner, J. A. … Sweeney, J. A. (2015). Pharmacogenetic associations of the type-3 metabotropic glutamate receptor (GRM3) gene with working memory and clinical symptom response to antipsychotics in first-episode schizophrenia. *Psychopharmacology (Berl), 232*(1), 145–154.

Bittencourt, P. R., Wade, P., Smith, A. T., & Richens, A. (1981). The relationship between peak velocity of saccadic eye movements and serum benzodiazepine concentration. *British Journal of Clinical Pharmacology, 12*(4), 523–533.

Bittencourt, P. R., Wade, P., Smith, A. T., & Richens, A. (1983). Benzodiazepines impair smooth pursuit eye movements. *British Journal of Clinical Pharmacology, 15*(2), 259–262.

Blom, M. W., Bartel, P. R., de Sommers, K., Van der Meyden, C. H., & Becker, P. J. (1986). The effects of alprazolam, quazepam, and diazepam on saccadic eye movements, parameters of psychomotor function and the EEG. *Fundamental & Clinical Pharmacology, 4*(6), 653–661.

Broerse, A., Crawford, T. J., & den Boer, J. A. (2002). Differential effects of olanzapine and risperidone on cognition in schizophrenia? A saccadic eye movement study. *Journal of Neuropsychiatry and Clinical Neurosciences, 14*(4), 454–460.

Burke, J. G., & Reveley, M. A. (2002). Improved antisaccade performance with risperidone in schizophrenia. *Journal of Neurology, Neurosurgery and Psychiatry, 72*(4), 449–454.

Bylsma, F. W., & Pivik, R. T. (1989). The effects of background illumination and stimulant medication on smooth pursuit eye movements of hyperactive children. *Journal of Abnormal Child Psychology, 17*(1), 73–90.

Campion, D., Thibaut, F., Denise, P., Courtin, P., Pottier, M., & Levillain, D. (1992). SPEM impairment in drug-naive schizophrenic patients: Evidence for a trait marker. *Biological Psychiatry, 32*(10), 891–902.

Clarke, P. B., Hommer, D. W., Pert, A., & Skirboll, L. R. (1985). Electrophysiological actions of nicotine on substantia nigra single units. *British Journal of Pharmacology, 85*(4), 827–835.

Cohen, A. F., Ashby, L., Crowley, D., Land, G., Peck, A. W., & Miller, A. A. (1985). Lamotrigine (BW430C), a potential anticonvulsant. Effects on the central nervous system in comparison with phenytoin and diazepam. *British Journal of Clinical Pharmacology, 20*(6), 619–629.

de Visser, S. J., van der Post, J. P., de Waal, P. P., Cornet, F., Cohen, A. F., & van Gerven, J. M. (2003). Biomarkers for the effects of benzodiazepines in healthy volunteers. *British Journal of Clinical Pharmacology, 55*(1), 39–50.

Depatie, L., O'Driscoll, G. A., Holahan, A. L., Atkinson, V., Thavundayil, J. X., Kin, N. N., & Lal, S. (2002). Nicotine and behavioral markers of risk for schizophrenia: a double-blind, placebo-controlled, cross-over study. *Neuropsychopharmacology, 27*(6), 1056–1070.

Dias, E. C., Kiesau, M., & Segraves, M. A. (1995). Acute activation and inactivation of macaque frontal eye field with GABA-related drugs. *Journal of Neurophysiology, 74*(6), 2744–2748.

Domino, E. F., Ni, L. S., & Zhang, H. (1997). Effects of tobacco smoking on human ocular smooth pursuit. *Clinical Pharmacology and Therapeutics, 61*(3), 349–359. https://doi.org/10.1016/S0009-9236(97)90168-5.

Dyckman, K. A., & McDowell, J. E. (2005). Behavioral plasticity of antisaccade performance following daily practice. *Experimental Brain Research, 162*(1), 63–69.

Ettinger, U., Antonova, E., Crawford, T. J., Mitterschiffthaler, M. T., Goswani, S., Sharma, T., et al. (2005). Structural neural correlates of prosaccade and antisaccade eye movements in healthy humans. *Neuroimage, 24*(2), 487–494.

Ettinger, U., & Kumari, V. (2003). Pharmacological studies of smooth pursuit and antisaccade eye movements in schizophrenia: Current status and directions for future research. *Current Neuropharmacology, 1*(4), 285–300.

Ettinger, U., Kumari, V., Crawford, T. J., Davis, R. E., Sharma, T., & Corr, P. J. (2003a). Reliability of smooth pursuit, fixation, and saccadic eye movements. *Psychophysiology, 40*(4), 620–628.

Ettinger, U., Kumari, V., Zachariah, E., Galea, A., Crawford, T. J., Corr, P. J. … Sharma, T. (2003b). Effects of procyclidine on eye movements in schizophrenia. *Neuropsychopharmacology, 28*(12), 2199–2208.

Ettinger, U., Williams , S. C., Patel, D., Michel, T. M., Nwaigwe, A., Caceres, A., … & Kumari, V. (2009). Effects of acute nicotine on brain function in healthy smokers and non-smokers: Estimation of inter-individual response heterogeneity. *Neuroimage, 45*(2), 549–561.

Farage, M. A., Osborn, T. W., & MacLean, A. B. (2008). Cognitive, sensory, and emotional changes associated with the menstrual cycle: A review. *Archives of Gynecology and Obstetrics, 278*(4), 299–307.

Farrell, S. M., Tunbridge, E. M., Braeutigam, S., & Harrison, P. J. (2012). COMT Val(158)Met genotype determines the direction of cognitive effects produced by catechol-O-methyltransferase inhibition. *Biological Psychiatry, 71*(6), 538–544.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Flechtner, K. M., Mackert, A., Thies, K., Frick, K., & Muller-Oerlinghausen, B. (1992). Lithium effect on smooth pursuit eye movements of healthy volunteers. *Biological Psychiatry, 32*(10), 932–938.

Friedman, L., Jesberger, J. A., & Meltzer, H. Y. (1992). Effect of typical antipsychotic medications and clozapine on smooth pursuit performance in patients with schizophrenia. *Psychiatry Research, 41*(1), 25–36.

Friedman, L., Jesberger, J. A., & Meltzer, H. Y. (1994). The effect of apomorphine, MK-212 (6-chloro-2-[1-piperazinyl]-pyrazine) and placebo on smooth pursuit gain and corrective saccades in normal subjects. *Neuropsychopharmacology, 11*(1), 49–62.

Gibson, J. M., Pimlott, R., & Kennard, C. (1987). Ocular motor and manual tracking in Parkinson's disease and the effect of treatment. *Journal of Neurology, Neurosurgery and Psychiatry, 50*(7), 853–860.

Gijsman, H. J., van Gerven, J., Verkes, R. J., Schoemaker, R. C., Pieters, M. S., Pennings, E. J. … Cohen, A. F. (2002). Saccadic peak velocity and EEG as end-points for a serotonergic challenge test. *Human Psychopharmacology: Clinical and Experimental, 17*(2), 83–89.

Gooding, D. C., Iacono, W. G., & Beiser, M. (1994). Temporal stability of smooth-pursuit eye tracking in first-episode psychosis. *Psychophysiology, 31*(1), 62–67.

Green, J. F., & King, D. J. (1998). The effects of chlorpromazine and lorazepam on abnormal antisaccade and no-saccade distractibility. *Biological Psychiatry, 44*(8), 709–715.

Green, J. F., King, D. J., & Trimble, K. M. (2000). Antisaccade and smooth pursuit eye movements in healthy subjects receiving sertraline and lorazepam. *Journal of Psychopharmacology, 14*(1), 30–36.

Griffiths, A. N., Tedeschi, G., & Richens, A. (1986). The effects of repeated doses of temazepam and nitrazepam on several measures of human performance. *Acta Psychiatrica Scandinavica, 74*(S332), 119–126.

Guedry, F. E., Jr., Gilson, R. D., Schroeder, D. J., & Collins, W. E. (1974). *Some effects of alcohol on various aspects of oculomotor control* (No. NAMRL-1206). Naval Aerospace Medical Research Lab, Pensacola, FL.

Halliday, J., & Carpenter, R. H. S. (2010). The effect of cognitive distraction on saccadic latency. *Perception, 39*(1), 41–50.

Harris, M. S., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2006). Longitudinal studies of antisaccades in antipsychotic-naive first-episode schizophrenia. *Psychological Medicine, 36*(4), 485–494.

Harris, M. S., Wiseman, C. L., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2009). Effects of risperidone on procedural learning in antipsychotic-naive first-episode schizophrenia. *Neuropsychopharmacology, 34*(2), 468–476.

Hikosaka, O., & Wurtz, R. H. (1983). Effects on eye movements of a GABA agonist and antagonist injected into monkey superior colliculus. *Brain Research, 272*(2), 368–372.

Hikosaka, O., & Wurtz, R. H. (1985a). Modification of saccadic eye movements by GABA-related substances. I. Effect of muscimol and bicuculline in monkey superior colliculus. *Journal of Neurophysiology, 53*(1), 266–291.

Hikosaka, O., & Wurtz, R. H. (1985b). Modification of saccadic eye movements by GABA-related substances. II. Effects of muscimol in monkey substantia nigra pars reticulata. *Journal of Neurophysiology, 53*(1), 292–308.

Hill, S. K., Reilly, J. L., Harris, M. S., Khine, T., & Sweeney, J. A. (2008). Oculomotor and neuropsychological effects of antipsychotic treatment for schizophrenia. *Schizophrenia Bulletin, 34*(3), 494–506.

Holzman, P. S., Levy, D. L., Uhlenhuth, E. H., Proctor, L. R., & Freedman, D. X. (1975). Smooth-pursuit eye movements, and diazepam, CPZ, and secobarbital. *Psychopharmacologia, 44*(2), 112–115.

Holzman, P. S., O'Brian, C., & Waternaux, C. (1991). Effects of lithium treatment on eye movements. *Biological Psychiatry*, *29*(10), 1001–1015.

Hommer, D. W., Matsuo, V., Wolkowitz, O., Chrousos, G., Greenblatt, D. J., Weingartner, H., & Paul, S. M. (1986). Benzodiazepine sensitivity in normal human subjects. *Archives of General Psychiatry, 43*(6), 542–551.

Hood, A. J., Amador, S. C., Cain, A. E., Briand, K. A., Al-Refai, A. H., Schiess, M. C., & Sereno, A. B. (2007). Levodopa slows prosaccades and improves antisaccades: An eye movement study in Parkinson's disease. *Journal of Neurology, Neurosurgery and Psychiatry, 78*(6), 565–570. https://doi.org/10.1136/jnnp.2006.099754.

Hutton, S. B., Crawford, T. J., Gibbins, H., Cuthbert, I., Barnes, T. R., Kennard, C., & Joyce, E. M. (2001). Short and long term effects of antipsychotic medication on smooth pursuit eye tracking in schizophrenia. *Psychopharmacology (Berl), 157*(3), 284–291.

Hutton, S. B., Crawford, T. J., Puri, B. K., Duncan, L. J., Chapman, M., Kennard, C. … Joyce, E. M. (1998). Smooth pursuit and saccadic abnormalities in first-episode schizophrenia. *Psychological Medicine, 28*(3), 685–692.

Hutton, S. B., & Ettinger, U. (2006). The antisaccade task as a research tool in psychopathology: A critical review. *Psychophysiology, 43*(3), 302–313.

Janset, A. A. I., Verbaten, M. N., & Slangen, J. L. (1988). Acute effects of bromazepam on signal detection performance, digit symbol substitution test and smooth pursuit eye movements. *Neuropsychobiology, 20,* 91–95.

Kasparbauer, A. M., Meyhöfer, I., Steffens, M., Weber, B., Aydin, M., Kumari, V., … & Ettinger, U. (2016). Neural effects of methylphenidate and nicotine during smooth pursuit eye movements. *NeuroImage, 141,* 52–59.

King, D. J., Best, P., Lynch, G., Mannion, M. F., Montgomery, R. C., Tiplady, B., & Yisak, W. (1995). The effects of remoxipride and chlorpromazine on eye movements and psychomotor performance in healthy volunteers. *Journal of Psychopharmacology, 9*(2), 143–149.

Klein, C., & Fischer, B. (2005). Instrumental and test-retest reliability of saccadic measures. *Biological Psychology, 68*(3), 201–213.

Klein, C., Jr., Fischer, B., Fischer, B., & Hartnegg, K. (2002). Effects of methylphenidate on saccadic responses in patients with ADHD. *Experimental Brain Research, 145*(1), 121–125.

Kobayashi, Y., & Isa, T. (2002). Sensory-motor gating and cognitive control by the brainstem cholinergic system. *Neural Networks, 15*(4–6), 731–741.

Konradi, C., & Heckers, S. (2003). Molecular aspects of glutamate dysregulation: Implications for schizophrenia and its treatment. *Pharmacology & Therapeutics, 97*(2), 153–179.

Krystal, J., d'Souza, D. C., Mathalon, D. M., Perry, E., Belger, A., & Hoffman, R. (2003). NMDA receptor antagonist effects, cortical glutamatergic function, and schizophrenia: Toward a paradigm shift in medication development. *Psychopharmacology, 169*(3–4), 215–233.

Krystal, J. H., Karper, L. P., Seibyl, J. P., Freeman, G. K., Delaney, R., Bremner, J. D. … Charney, D. S. (1994). Subanesthetic effects of the noncompetitive NMDA antagonist, ketamine, in humans. Psychotomimetic, perceptual, cognitive, and neuroendocrine responses. *Archives of General Psychiatry, 51*(3), 199–214.

Lachman, H. M., et al. (1996). Association of codon 108/158 catechol-O-methyltransferase gene polymorphism with the psychiatric manifestations of velo-cardio-facial syndrome. *American Journal of Medical Genetics, 67*(5), 468–472.

Larrison, A. L., Briand, K. A., & Sereno, A. B. (2004). Nicotine improves antisaccade task performance without affecting prosaccades. *Human Psychopharmacology, 19*(6), 409–419. https://doi.org/10.1002/hup.604.

Leigh, R. J., & Zee, D. S. (2006). *The neurology of eye movements* (4th ed.). New York, NY: Oxford University Press.

Lencer, R., Bishop, J. R., Harris, M. S., Reilly, J. L., Patel, S., Kittles, R. … Sweeney, J. A. (2014). Association of variants in DRD2 and GRM3 with motor and cognitive function in first-episode psychosis. *European Archives of Psychiatry and Clinical Neuroscience, 264*(4), 345–355.

Lencer, R., Sprenger, A., Harris, M. S., Reilly, J. L., Keshavan, M. S., & Sweeney, J. A. (2008). Effects of second-generation antipsychotic medication on smooth pursuit performance in antipsychotic-naive schizophrenia. *Archives of General Psychiatry, 65*(10), 1146–1154.

Lencer, R., & Trillenberg, P. (2008). Neurophysiology and neuroanatomy of smooth pursuit in humans. *Brain and Cognition, 68*(3), 219–228.

Levy, D. L., Dorus, E., Shaughnessy, R., Yasillo, N. J., Pandey, G., Janicak, P. G., et al. (1985). Pharmacologic evidence for specificity of pursuit dysfunction to schizophrenia: Lithium carbonate associated with abnormal pursuit. *Archives of General Psychiatry, 42*(4), 335–341.

Litman, R. E., Hommer, D. W., Radant, A., Clem, T., & Pickar, D. (1994). Quantitative effects of typical and atypical neuroleptics on smooth pursuit eye tracking in schizophrenia. *Schizophrenia Research, 12*(2), 107–120.

Lynch, G., King, D. J., Green, J. F., Byth, W., & Wilson-Davis, K. (1997). The effects of haloperidol on visual search, eye movements and psychomotor performance. *Psychopharmacology (Berl), 133*(3), 233–239.

Macdonald, R. L., McLean, M. J., & Skerritt, J. H. (1985). Anticonvulsant drug mechanisms of action. In *Federation Proceedings* (Vol. 44, issue 10, pp. 2634–2639).

Masson, G. S., Mestre, D. R., Martineau, F., Soubrouillard, C., Brefel, C., Rascol, O., & Blin, O. (2000). Lorazepam-induced modifications of saccadic and smooth-pursuit eye movements in humans: attentional and motor factors. *Behavioural Brain Research, 108*(2), 169–180.

McDowell, J. E., Dyckman, K. A., Austin, B. P., & Clementz, B. A. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades: Evidence from studies of humans. *Brain and Cognition, 68*(3), 255–270.

Meyhöfer, I., Bertsch, K., Esser, M., & Ettinger, U. (2015). Variance in saccadic eye movements reflects stable traits. *Psychophysiology, 53,* 566–578.

Michell, A. W., Xu, Z., Fritz, D., Lewis, S. J., Foltynie, T., Williams-Gray, C. H. … Barker, R. A. (2006). Saccadic latency distributions in Parkinson's disease and the effects of L-dopa. *Experimental Brain Research, 174*(1), 7–18.

Möhler, H., Fritschy, J. M., & Rudolph, U. (2002). A new benzodiazepine pharmacology. *Journal of Pharmacology and Experimental Therapeutics, 300*(1), 2–8.

Morrens, M., Wezenberg, E., Verkes, R. J., Hulstijn, W., Ruigt, G. S., & Sabbe, B. G. (2007). Psychomotor and memory effects of haloperidol, olanzapine, and paroxetine in healthy subjects after short-term administration. *Journal of Clinical Psychopharmacology, 27*(1), 15–21.

Mosimann, U. P., Felblinger, J., Colloby, S. J., & Muri, R. M. (2004). Verbal instructions and top-down saccade control. *Experimental Brain Research, 159*(2), 263–267.

Muller, N., Riedel, M., Eggert, T., & Straube, A. (1999). Internally and externally guided voluntary saccades in unmedicated and medicated schizophrenic patients. Part II. Saccadic latency, gain, and fixation suppression errors. *European Archives of Psychiatry and Clinical Neuroscience, 249*(1), 7–14.

Müri, R. M., & Nyffeler, T. (2008). Neurophysiology and neuroanatomy of reflexive and volitional saccades as revealed by lesion studies with neurological patients and transcranial magnetic stimulation (TMS). *Brain and Cognition, 68*(3), 284–292.

Nakamura, T., Kanayama, R., Sano, R., Ohki, M., Kimura, Y., Aoyagi, M., & Koike, Y. (1991). Quantitative analysis of ocular movements in Parkinson's disease. *Acta Oto-Laryngologica. Supplement, 481,* 559–562.

Nawrot, M., Nordenstrom, B., & Olson, A. (2004). Disruption of eye movements by ethanol intoxication affects perception of depth from motion parallax. *Psychological Science, 15*(12), 858–865.

Noachtar, S., von Maydell, B., Fuhry, L., & Buttner, U. (1998). Gabapentin and carbamazepine affect eye movements and posture control differently: A placebo-controlled investigation of acute CNS side effects in healthy volunteers. *Epilepsy Research, 31*(1), 47–57.

O'Driscoll, G. A., Depatie, L., Holahan, A. L., Savion-Lemieux, T., Barr, R. G., Jolicoeur, C., & Douglas, V. I. (2005). Executive functions and methylphenidate response in subtypes of attention-deficit/hyperactivity disorder. *Biological Psychiatry, 57*(11), 1452–1460.

Olincy, A., Ross, R. G., Young, D. A., Roath, M., & Freedman, R. (1998). Improvement in smooth pursuit eye movements after cigarette smoking in schizophrenic patients. *Neuropsychopharmacology, 18*(3), 175–185.

Padoan, S., Korttila, K., Magnusson, M., Pyykko, I., & Schalen, L. (1992). Effects of intravenous diazepam and thiopental on voluntary saccades and pursuit eye movements. *Acta Oto-Laryngologica, 112*(4), 579–588.

Peck, A. W. (1991). Clinical pharmacology of lamotrigine. *Epilepsia, 32*(s2), S9–S12.

Pierrot-Deseilligny, C., Muri, R. M., Ploner, C. J., Gaymard, B., & Rivaud-Pechoux, S. (2003). Cortical control of ocular saccades in humans: A model for motricity. *Progress in Brain Research, 142,* 3–17.

Powell, J., Dawkins, L., & Davis, R. E. (2002). Smoking, reward responsiveness, and response inhibition: Tests of an incentive motivational model. *Biological Psychiatry, 51*(2), 151–163.

Radant, A. D., Bowdle, T. A., Cowley, D. S., Kharasch, E. D., & Roy-Byrne, P. P. (1998). Does ketamine-mediated N-methyl-D-aspartate receptor antagonism cause schizophrenia-like oculomotor abnormalities? *Neuropsychopharmacology, 19*(5), 434–444.

Rascol, O., Clanet, M., Montastruc, J. L., Simonetta, M., Soulier-Esteve, M. J., Doyon, B., & Rascol, A. (1989). Abnormal ocular movements in Parkinson's disease. Evidence for involvement of dopaminergic systems. *Brain, 112*(Pt 5), 1193–1214.

Rashbass, C. (1959). Barbiturate nystagmus and the mechanisms of visual fixation. *Nature, 183,* 897.

Reilly, J. L., Harris, M. S., Keshavan, M. S., & Sweeney, J. A. (2005). Abnormalities in visually guided saccades suggest corticofugal dysregulation in never-treated schizophrenia. *Biological Psychiatry, 57*(2), 145–154.

Reilly, J. L., Harris, M. S., Keshavan, M. S., & Sweeney, J. A. (2006). Adverse effects of risperidone on spatial working memory in first-episode schizophrenia. *Archives of General Psychiatry*, *63*(11), 1189–1197.

Reilly, J. L., Harris, M. S., Khine, T. T., Keshavan, M. S., & Sweeney, J. A. (2007). Antipsychotic drugs exacerbate impairment on a working memory task in first-episode schizophrenia. *Biological Psychiatry, 62*(7), 818–821.

Reilly, J. L., Lencer, R., Bishop, J. R., Keedy, S., & Sweeney, J. A. (2008). Pharmacological treatment effects on eye movement control. *Brain and Cognition, 68*(3), 415–435.

Remler, B. F., Leigh, R. J., Osorio, I., & Tomsak, R. L. (1990). The characteristics and mechanisms of visual disturbance associated with anticonvulsant therapy. *Neurology, 40*(5), 791–796.

Roche, D. J., & King, A. C. (2010). Alcohol impairment of saccadic and smooth pursuit eye movements: impact of risk factors for alcohol dependence. *Psychopharmacology (Berl), 212*(1), 33–44.

Rothenberg, S. J., & Selkoe, D. (1981a). Specific oculomotor deficit after diazepam. I. Saccadic eye movements. *Psychopharmacology (Berl), 74*(3), 232–236.

Rothenberg, S. J., & Selkoe, D. (1981b). Specific oculomotor deficit after diazepam. II. Smooth pursuit eye movements. *Psychopharmacology (Berl), 74*(3), 237–240.

Roy-Byrne, P., Cowley, D. S., Radant, A., Hommer, D., & Greenblatt, D. J. (1993). Benzodiazepine pharmacodynamics: Utility of eye movement measures. *Psychopharamcology, 110*(1–2), 85–91.

Rycroft, N., Hutton, S. B., & Rusted, J. M. (2006). The antisaccade task as an index of sustained goal activation in working memory: Modulation by nicotine. *Psychopharmacology (Berl), 188*(4), 521–529.

Salonen, M., Aaltonen, L., Aantaa, E., & Kanto, J. (1986). Saccadic eye movements in determination of the residual effects of benzodiazepines. *International Journal of Clinical Pharmacology, Therapy, and Toxicology, 24*(5), 227–231.

Schmechtig, A., Lees, J., Perkins, A., Altavilla, A., Craig, K. J., Dawson, G. R. … Ettinger, U. (2013). The effects of ketamine and risperidone on eye movement control in healthy volunteers. *Translation Psychiatry, 3,* e334.

Sharpe, J. A. (2008). Neurophysiology and neuroanatomy of smooth pursuit: Lesion studies. *Brain and Cognition, 68*(3), 241–254.

Sharpe, J. A., Fletcher, W. A., Lang, A. E., & Zackon, D. H. (1987). Smooth pursuit during dose-related on-off fluctuations in Parkinson's disease. *Neurology*, *37*(8), 1389–1389.

Sherwin, B. B. (2012). Estrogen and cognitive functioning in women: Lessons we have learned. *Behavioral Neuroscience, 126*(1), 123–127.

Siever, L. J., & Davis, K. L. (2004). The pathophysiology of schizophrenia disorders: Perspectives from the spectrum. *American Journal of Psychiatry, 161*(3), 398–413.

Smyrnis, N. (2008). Metric issues in the study of eye movements in psychiatry. *Brain and Cognition, 68*(3), 341–358.

Steffens, M., Becker, B., Neumann, C., Kasparbauer, A. M., Meyhöfer, I., Weber, B., … & Ettinger, U. (2016). Effects of ketamine on brain function during smooth pursuit eye movements. *Human Brain Mapping, 37*(11), 4047–4060.

Stolerman, I. P. (2010). *Encyclopedia of psychopharmacology*. New York: Springer.

Straube, A., Riedel, M., Eggert, T., & Muller, N. (1999). Internally and externally guided voluntary saccades in unmedicated and medicated schizophrenic patients. Part I. Saccadic velocity. *European Archives of Psychiatry and Clinical Neuroscience, 249*(1), 1–6.

Sweeney, J. A., Haas, G. L., Li, S., & Weiden, P. J. (1994). Selective effects of antipsychotic medications on eye-tracking performance in schizophrenia. *Psychiatry Research, 54*(2), 185–198.

Sweeney, J. A., Luna, B., Haas, G. L., Keshavan, M. S., Mann, J. J., & Thase, M. E. (1999). Pursuit tracking impairments in schizophrenia and mood disorders: Step-ramp studies with unmedicated patients. *Biological Psychiatry, 46*(5), 671–680.

Tambour, S., & Crabbe, J. C. (2010). Pharmacogenetics. In I. P. Stolerman (Ed.), *Encyclopedia of psychopharmacology* (pp. 998–1001). New York: Springer.

Tamminga, C. A., Buchanan, R. W., & Gold, J. M. (1998). The role of negative symptoms and cognitive dysfunction in schizophrenia outcome. *International Clinical Psychopharmacology, 13*(Suppl 3), S21–S26.

Tedeschi, G., Casucci, G., Allocca, S., Riva, R., Di Costanzo, A., Quattrone, A. … Bonavita, V. (1989). Computer analysis of saccadic eye movements: Assessment of two different carbamazepine formulations. *European Journal of Clinical Pharmacology, 37*(5), 513–516.

Tedeschi, G., Griffiths, A. N., Smith, A. T., & Richens, A. (1985). The effect of repeated doses of temazepam and nitrazepam on human psychomotor performance. *British Journal of Clinical Pharmacology, 20*(4), 361–367.

Thaker, G. K., Ross, D. E., Buchanan, R. W., Adami, H. M., & Medoff, D. R. (1999). Smooth pursuit eye movements to extra-retinal motion signals: Deficits in patients with schizophrenia. *Psychiatry Research, 88*(3), 209–219.

Van Spall, H. G., Toren, A., Kiss, A., & Fowler, R. A. (2007). Eligibility criteria of randomized controlled trials published in high-impact general medical journals: A systematic sampling review. *JAMA, 297*(11), 1233–1240.

van Steveninck, A. L., Verver, S., Schoemaker, H. C., Pieters, M. S., Kroon, R., Breimer, D. D., & Cohen, A. F. (1992). Effects of temazepam on saccadic eye movements: concentration-effect relationships in individual volunteers. *Clinical Pharmacology and Therapeutics, 52*(4), 402–408.

van Steveninck, A. L., Wallnofer, A. E., Schoemaker, R. C., Pieters, M. S., Danhof, M., van Gerven, J. M., & Cohen, A. F. (1997). A study of the effects of long-term use on individual sensitivity to temazepam and lorazepam in a clinical population. *British Journal of Clinical Pharmacology, 44*(3), 267–275.

Wang, B., Yang, L. P., Zhang, X. Z., Huang, S. Q., Bartlam, M., & Zhou, S. F. (2009). New insights into the structural characteristics and functional relevance of the human cytochrome P450 2D6 enzyme. *Drug Metabolism Reviews, 41*(4), 573–643.

Waterston, J. A., Barnes, G. R., Grealy, M. A., & Collins, S. (1996). Abnormalities of smooth eye and head movement control in Parkinson's disease. *Annals of Neurology, 39*(6), 749–760.

Westheimer, G., & Rashbass, C. (1961). Barbiturates and eye vergence. *Nature, 191,* 833–834.

Wilkinson, I. M. S., Kime, R., & Purnell, M. (1974). Alcohol and human eye movement. *Brain, 97*(4), 785–795.

Part IV
Applications in Industrial Settings

Chapter 19
Neuromarketing

Matthias Rothensee and Philipp Reiter

Contents

Abstract The following chapter will illustrate how eye movement research is applied to neuromarketing research across various consumer touchpoints. After offering a definition and justification of neuromarketing research, we review the history and current state of neuromarketing. We continue to describe specific methodological innovations and tools used to embrace the special focus on efficiency that is typical

---

M. Rothensee (✉) · P. Reiter
eye square GmbH, Berlin, Germany
e-mail: rothensee@eye-square.com

P. Reiter
e-mail: reiter@eye-square.com

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_19

of marketing research projects. The main body of this chapter consists of a number of example cases in two domains: testing marketing communication both in advertising and in-store communication. The examples demonstrate that eye movement studies provide advertisers and manufacturers with deep insights into the perception and behavior of humans as consumers. Eye movement insights help to optimize communication with consumers and improve our understanding of what drives consumer behavior on the shop floor. We hope that this chapter provides comprehensive insights into how visual attention research adds value to the study of consumer attitudes.

19.1 Introduction and Learning Objectives

Neuromarketing is a subdivision of traditional consumer marketing, with a particular focus on research. Marketing is often described as the alignment of all company activities to the circumstances of its market (Springer Gabler Verlag, 2016). Marketing uses products, prices, placement, and promotion strategies to ensure business success in the market. Regarding advertising, understanding the consumer is important for creating better communication. Understanding consumers and speaking to them with the right arguments plays a pivotal role in the marketing agenda of any company. Marketing research uses various methods to understand consumers, their preferences and their needs. It provides insight on how to act on customer insights with communication at various touchpoints (e.g., advertising, digital interfaces, and the shop floor).

The term "neuromarketing" is used to indicate a subdivision in marketing research that applies methods from the neurosciences to the study of consumer behavior. Neuromarketing extends the classical repertoire of marketing research and complements it by opening a new, behavioral perspective on the consumer. Neuromarketing is defined as the application of methods from the neurosciences to the field of marketing (Zurawicki, 2010). Scheier (Scheier & Held, 2012) argues that neuromarketing uses insights from such diverse disciplines as neuropsychology, psychophysiology, artificial intelligence, cultural studies, and developmental psychology. Neuromarketing leverages these methods to optimize the effects of marketing communication.

A distinction should be made between *consumer neuroscience* and *neuromarketing*: consumer neuroscience is focused on academic research using brain imaging tools only, whereas neuromarketing is "aiming at practitioners and commercial interest in neurophysiological tools to conduct company-specific market research" (Plassmann, Ramsoy, & Milosavljevic, 2012).

The justification for neuromarketing techniques in the consumer research domain is often based on three arguments, which are elaborated in the following section. First, however we will consider the learning objectives of this chapter.

After reading this chapter, the student will know how eye movement research is used in a neuromarketing context and be able to answer the following questions:

- How and when did neuromarketing develop, and which limitations of traditional methods paved the way for its sustained success in the domain of marketing research?
- What are the special needs of clients who buy commercial eye movement research, and how does technical innovation in the field resonate with them?
- Which techniques are used to evaluate the performance of static and video ads (e.g., in the field of out-of-home advertising)?
- Which key performance indicators (KPIs) tell if a poster is the winner when comparing two versions?
- How does eye movement research inform the optimization of product packaging on the shelf?

19.1.1 Three Arguments for Neuromarketing

The rise of neuromarketing as a discipline was born out of dissatisfaction with the results of traditional marketing research, as Zaltman (2003) describes: "The world has changed, but our methods for understanding consumers have not. We keep relying on familiar but ineffective research techniques and consequently misread consumers' actions and thoughts. The products we create based on those techniques, simply are not connecting with consumers." In the next section, an explication of how this critique can be structured into three distinct arguments is presented.

The evolution argument

The first argument takes an evolutionary stance in arguing that "the customer's brain is 100,000 years old" (Pradeep, 2010). This metaphorical statement is used to underline that even though our environment, society and living conditions have dramatically changed throughout the past millennia, on a molecular level the structures and functional mechanisms of our brains have remained largely the same. This leads to the conclusion that applying methods to understand how the human brain receives information and processes it to produce outcomes will also inform modern forms of communication such as marketing communication. In contemporary psychology, human evolution is understood as important for explaining the development of human emotions and perception patterns (Tooby & Cosmides, 2015).

The filter argument

The second justification for neuromarketing is the notion that our environment contains far more stimuli than humans can process. Scheier argues that our sensory systems have limited bandwidth. The amount of information that we are able to process each second is small compared to the amount of available information (Scheier & Held, 2012). More precisely, Wilson (2002) claims that humans are exposed to 11 million bits of information each second but can only process 50 bits of that information, filtering out most sensory input. The so-called "cocktail party" effect demonstrates that in the auditory domain, humans have the capacity to filter out irrelevant information (Colin, 1953). Studies have shown that such filter mechanisms also exist in the visual domain (Itti, Koch, & Niebur, 1998; Wolfe & Horowitz, 2004).

Recent changes in the media landscape add momentum to the notion of filtering: scholars such as Sorensen (2009) argue that the sheer number of products and amount of marketing communication is continually rising. A study by Media Dynamics showed that in 2014, American audiences were exposed to 362 ads across TV, radio, the internet, newspapers and magazines (Media Dynamics Inc., 2016). Of those, only 153 were actually noted by consumers (i.e., attracted the audiences' full attention for a few seconds or more). Adding subtle brand messages (e.g., logos on products, labels, and sponsorships on bus stops) increases this figure to approximately 5000 messages every day (Munzinger & Musiol, 2008). The part of consumers' daily lives with the highest density of brand messages and logos is brick and mortar supermarkets, which stock approximately 30,000–50,000 products (US data) (Sorensen, 2009).

Whether or not to attend to each of those brand messages or to examine whether to buy all those products is far beyond the number of conscious decisions a human may be able to make. This is why it is important to illustrate that understanding how people filter—or pay attention to some things and not others—is of crucial importance to anyone wanting to sell customers products. We suggest that this *filter argument* is one of the core justifications of eye movement research in marketing and its enormous success in this field. The eye is one of the primary gatekeepers to consumers' minds, and this is where many brands want to be present.

The willing-and-able problem

The third justification for eye movement research in neuromarketing is that consumers are often neither willing nor able to completely report on their perceptions. This is likely due to two reasons:

1. As consumers are overwhelmed by advertising in all media channels, they tend to have a negative evaluation of advertising in general as a communication medium. Already in 1999, consumers answered in a survey that only 20% of the ads communicate trustworthy information about products or services (Mayer & Illmann, 1999). This shows that advertising is a controversial issue from consumers' perspectives. Therefore, generating valid, neutral information from merely asking consumers about their perceptions and the effects of advertising on their shopping behaviors is unlikely.
2. Consumer contact with marketing messages is often very brief and mundane. Looking at a banner on a website can range from a fraction of a second to a few seconds, as we will demonstrate in a later section. These very short contacts go unnoticed or are very quickly forgotten by consumers. In the application section "media consumption and multiscreen," we will include statistics on how much consumers underestimate their amount of media contact.

Taken together, evolution, filtering and the willing-and-able problem are the most important reasons to introduce eye movement research in neuromarketing. The next section broadens this focus. It lists other techniques that fall under the umbrella of neuromarketing and explains why eye movement research is an especially important complementary method for extending the focus of the classical neuromarketing toolset.

19.1.2 How Eye Movement Research Relates to Other Neuromarketing Disciplines

Neuromarketing itself is not a technique but rather a label for a set of measurement technologies and approaches to consumer understanding. The most prominent group of methods in the neuromarketing domain are brain imaging technologies. Perrachione and Perrachione (2008) list the most relevant methods in neuromarketing research. They list nine different brain imaging techniques and three complimentary methods. Two groups of neuroimaging techniques can be distinguished: 3D representations of metabolic processes in the brain, such as fMRI imaging, show which areas in the brain are active when consumers process, for example, brand messages (McClure et al., 2004). These techniques have high spatial resolution in their three-dimensional representations of consumers' brains but are very complex and resource-demanding in regard to applications to consumer research. Furthermore, the temporal resolution of these methods is poor. On the other side of the spectrum, EEG-based methods measure electrical activity on the scalp's surface. These techniques are fast but have low spatial resolution because they only reflect brain activity in 2D regions on the scalp surface and have little chance to distinguish lower-level activity from surface activity in, for example, the neocortex. EEG and fMRI and variations of these techniques compose the group of brain imaging techniques.

Three complimentary methods extend the neuromarketing toolset beyond brain imaging (Perrachione & Perrachione, 2008). The additional methods do not measure brain signals directly but rather focus on automatic, mental or implicit measures that are usually not consciously accessible to the consumer (Zurawicki, 2010). One of these methods is eye movement research (the other two concern voice pitch analysis and galvanic skin response, which are omitted here for the sake of brevity).

Eye movement research is often regarded as one of the most important complementary methods in the neuromarketing toolset (de Oliveira Joaquim dos Santos, Caldeira de Oliveira, Rocha, & Giraldi, 2015). This is because eye movements and, consequently, visual intake are temporal preconditions of most subsequent neural processes. This is why visual attention is regarded a universal prerequisite of the effectiveness of all marketing efforts. Visual attention is a necessity for most of the further information processing in the brain. We are primarily visual beings and live in a highly visual environment (Horsley, 2014). Furthermore, perception is not solely defined as the passive information processing of our surroundings but rather seen equivalent and directly linked with our actions (Tatler, 2014). Popular eye movement research wisdom from the point of sale says that "unseen is unsold" (Chandon, Hutchinson, & Young, 2016). The notion of attention being a scarce resource in our cluttered media environment led to the term of "attention economy" more than ten years ago (Davenport & Beck, 2002), and today's market researchers realize how important it is to acquire the consumer's attention.

There are two additional, more practical reasons for the sustained success of eye movement research in marketing. First, the method is comparably inexpensive and easy to apply for non-technical researchers outside of laboratories. This issue will be further elaborated in the technology section (Sect. 19.3.1) of this chapter. Second, as will be demonstrated in the practice cases below, insights from eye movement studies are often clearly related to visual properties of the investigated stimulus. The fact that consumers do not pay attention to a brand logo, for instance, clarifies the importance of emphasizing this element of a print ad, and its effectiveness can be ascertained in a follow-up study. How insights from eye movement research are applied to optimize marketing communication will be the core issue of Chap. 5.

In a nutshell, eye movement research is a complimentary method used to gain a holistic understanding of the consumer. While EEG and fMRI methods address processes in the brain when stimuli are evaluated, eye movement research informs the marketer about which information is viewed in the first place (e.g., when a customer has contact with an advertising message). Now that we have defined and reviewed the field of neuromarketing, the remainder of this chapter is dedicated exclusively to eye movement research, leaving all other neuromarketing methods aside.

19.2 Historical Annotations

Marketing has a long tradition in economics. The term has been used in the academic world since the beginning of the twentieth century. The first studies observing consumers' eye movements appeared as early as 1924 but were more anecdotal and limited by shortcomings of the technical equipment of the times (Nixon, 1924). As markets became increasingly saturated and new products were developed, the need grew for a consumer-oriented strategic activity to ensure market success. Methods such as market segmentation studies and conjoint analyses flourished. However, as scholars became increasingly aware of the bounded rationality under which consumers often make their purchase decisions (Simon, 1959), the need for a more consumer-centric approach to marketing appeared.

In the 1990s, the foundations were laid for the advent of advanced eye movement research due to advances in recording technology and other technological innovations (e.g., increasing processing speed and high-resolution cameras). At the end of the 1990s, studies appeared that demonstrated the value of eye movement research for understanding marketing communication (e.g., yellow pages ads) (Lohse, 1997).

The 2000s then saw a substantial increase in eye movement research applications and neuromarketing in general. Figure 19.1 illustrates how the number of Google hits, published articles and dedicated neuromarketing research companies increased after 2003 (Plassmann, Ramsoy, & Milosavljevic, 2012).

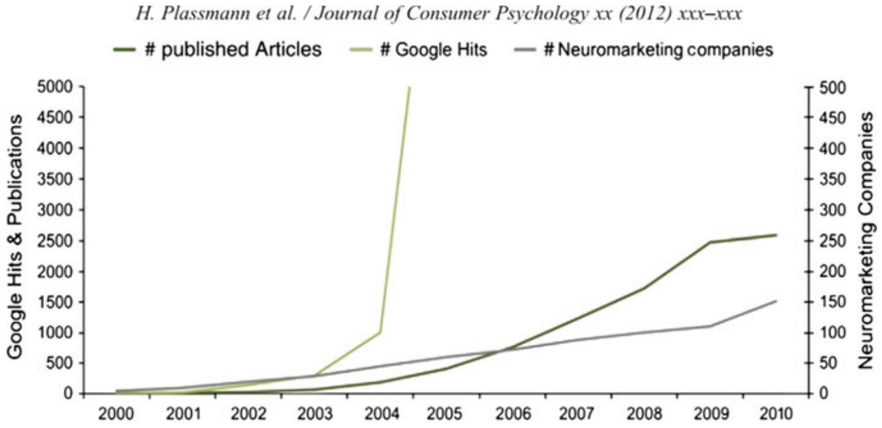


Fig. 19.1 Neuromarketing over time

In academics, research focused on model-based analyses of attention patterns and their memory effects. Wedel and Pieters (2000) for example, used Bayesian models to predict the memory effects of fixations on various elements of print magazine advertisements. They also investigated how brand salience on the shelf influences product purchase decisions (van der Lans, Pieters, & Wedel, 2008).

Box 1: Wedel and Pieters' Theory of Visual Attention

In the late 2000s, Wedel and Pieters also established a theory of visual attention, stating that visual behavior is influenced by consumer characteristics (endogenous factors) and marketing stimuli (exogenous factors) as well as the research situation (Wedel & Pieters, 2008a, b). Figure 19.2 shows these influences in a simplified way.

Automatic (bottom-up) processes are not under conscious control and occur in the first seconds of the attentional process. Contrast, shapes and saliency are processed before motivations come into play. Goal-oriented factors (motivation, product involvement, brand familiarity, etc.) are controlled by individuals and set the focus in cognitive information tasks. They typically appear later than bottom-up processes. This dual process distinction mirrors the Elaboration Likelihood Model that distinguishes high elaboration and low elaboration ad contacts and describes their differential effects on attitude formation (Petty & Cacioppo, 1986). Wedel and Pieters (2008a, b) added a third influence: the exposure situation. The setting and instructions will affect the visual pattern depending on the level of control that a consumer is given (forced exposure or free browsing) and the instruction or task. Yarbus (1967) describes that attention is "determined by the nature of the object and the problem facing the observer at the moment of perception."

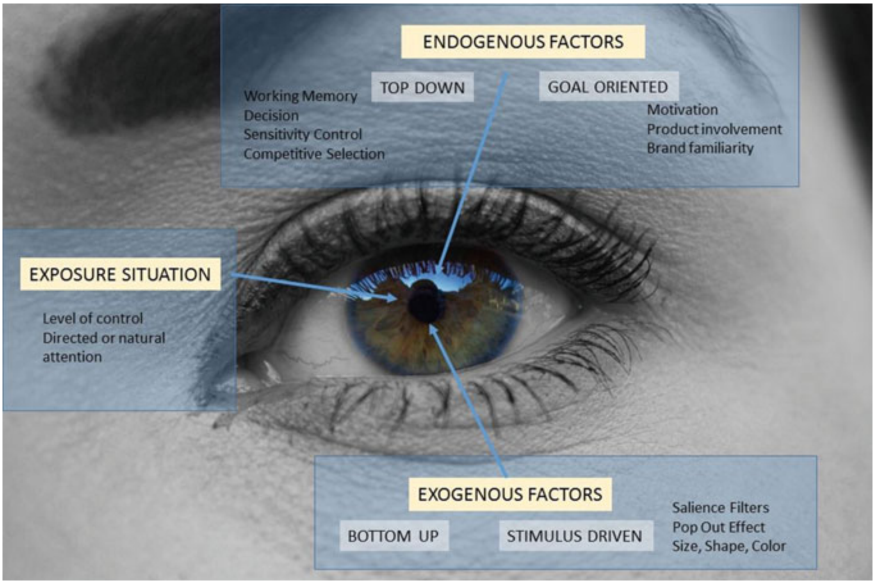


Fig. 19.2 Three factors influence consumer attention, according to Wedel and Pieters (2008a, b)

In the commercial domain, the early hype of the field was created because neuromarketing promised to introduce neutral, objective methods into market research and, moreover, to do so at a micro level that is neither conscious nor deliberately reflected upon. The hope of the 2000s was to improve communication at the individual consumer level and create a whole new field of marketing communication that rather speaks to consumers' subconscious brain areas (Scheier & Held, 2012). The often cited "buy-button" in consumers' minds is one of the most iconic rhetoric examples of this phase (Lindström, 2008).

However, this hope has not come true (Satel & Lilienfeld, 2013). As a result, a disillusionment set in with many marketers questioning the value of neuromarketing techniques for commercial purposes. Wiedmann, Lischka, and Schiessl (2017) claim that the term "*neuro*" was heavily overused in the contemporary science and business world. This critique, however, mainly concerned brain imaging based neuromarketing methods but not so much eye movement research.

Six signs of increasing maturity of the field can be observed from 2005 until today:

1. *Consolidation of the institute landscape.* Larger traditional research institutions acquired some commercial suppliers of brain imaging techniques. For example, NeuroFocus, a company applying fMRI and other methodologies to customer understanding, was purchased by the Nielsen Group in 2011.
2. *Standardization of measurement methods.* Commissioning eye movement research from institutes repeatedly over time created the challenge to standardize measurement methodology (e.g., how attention is defined, how a fixation is detected, etc.). Using standardized measurement methodology, research insti-

tutes in the field enabled building up benchmark databases to compare results of one study with past studies or, more importantly, the competition.

3. *Change of results presentation from the visual to the numeric.* In the early days, the heatmap was a fascinating, almost ubiquitous visual representation to demonstrate where consumers are looking, (e.g., when surfing on a website). However, with repeated exposure to heatmaps, clients' interest shifted focus increasingly to numerical and comparable results. This also demonstrates that the purchasers of eye movement research are now much more informed on how to interpret its results than at the end of the 1990s.
4. *Academic curricula.* Eye movement research has made its way into academic curricula in marketing research studies at universities (e.g., the European Summer School on Eye Movements).
5. *Establishment of dedicated special interest groups.* The international Neuromarketing Science and Business Association (NMSBA) was established in 2012 to promote the development and dissemination of insights from this field into marketing.
6. *Purchasing standards.* Certain standards or quality criteria were developed (ESOMAR, 2012). These provide assistance regarding quality assurance when commissioning neuromarketing research including questions about the company, technical equipment, tools and metrics, portability, and sample size.

All in all, in 2016 we see a field that has matured and consolidated with much understanding and trust from clients when commissioning eye movement research. Through consolidation, the total volume of eye movement studies has decreased to a healthy amount and concentrates on the most important domains such as analysis and optimization of advertising messages and point-of-sale consumer communication.

19.3 Getting Practical: The Eye Movement Research Ecosystem

In the first section of this chapter, the origins of eye movement research have been laid out for the reader, its relationship to other methods from the neuromarketing domain have been explicated, and the history of the method has been briefly sketched. As a next step, we turn our attention to more practical aspects of the application of eye movement research in marketing research practice. We first explain the technical requirements of the method in commercial marketing research applications. The particular focus on the effectiveness of clients in marketing research emphasizes the development of innovative techniques that increase speed and scalability of attention measurement. Second, we briefly describe the stakeholder relationships in a typical commercial project. Third, we provide an overview of industrial applications in the neuromarketing domain and give an outlook for the selected two areas of applications that will be discussed in more detail in subsequent chapters.

19.3.1 Technical Innovations Unique to Neuromarketing

The next section presents an admittedly subjective enumeration of clients' interests in commercial eye movement research that we experience in our daily business. Our eye movement studies in marketing often have a stronger focus on effectiveness than traditional academic applications. From our perspective, clients are highly demanding concerning the optimization of the efficiency of the method. This effectiveness scope has four components: budget, speed, required sample sizes, and ecological validity.

1. *Budget.* Even though large global companies frequently have monetary resources that exceed university budgets for this kind of research, they are highly interested in spending their funds in the most efficient way possible. This introduces budget pressure on eye movement studies in the marketing domain. In the 1990s, a number of institutes specialized on eye movement studies were established (eye square, Package InSight, Sticky, and Tobii Pro). They now form a strong competitive field. Furthermore, the purchasers of such research are tightly controlled by financial departments and have a high pressure to justify the budget for proper eye movement research.
2. *Speed.* Timing is a critical issue in commercial eye movement research, primarily because marketing decisions are made at a fast pace and require insights in a short time to inform these decisions. Many markets such as the e-commerce domain change in a matter of months, not years. Launching an advertising campaign for a product typically takes a few months' time. Having created a TV commercial, the time frame to optimize it using "eye of the customer" data is very short. Typically, optimization timing is a matter of weeks, if even that. If the marketing research is not able to deliver optimization recommendations in a few weeks from the commission of the study its value is close to zero.
3. *Sample size.* It is of crucial importance for companies to ensure that results of an eye movement analysis generalize to the consumer group that is targeted by the enterprise. When an academic study produces results on viewing patterns based on research with students, the question of representativeness of the sample is often secondary. In practical marketing research, it becomes the primary focus. As a very simple example, eye movement research is used to identify the more effective print advertising version between two candidate ads. As we will demonstrate later, this is possible for determining, for example, pop-out effects that one ad has but the other does not). In reporting the results to the client, the responsibility of the agency is to ensure that the more intensive perception of one ad over the other will also manifest in the general population. The consequence is that the neuromarketing studies using eye tracking often operate on higher, statistically sound sample sizes and clearly defined target population definitions than academic research does.
4. *Ecological validity.* Related to sample size is the demand for a high ecological validity of the results of eye movement studies. Many clients request that research be based on real-world stimuli, such that internet advertising must be conducted

> in the homes of participants because they doubt that results found in the lab generalize to the real daily lives of people. That is why, from the beginning of eye movement research, there was a high demand for mobile and online solutions that carry measurements of perception into the real world.

In the following paragraphs, we present three technical innovations that are, in part, answers to these demands from clients. These solutions are not exclusive to commercial eye movement research but from our perspective have a higher proliferation in the marketing research world than in the academics. These three innovations are mobile eye tracking, restricted-focus viewers, and webcam eye tracking.

Mobile eye trackers (see also chapters by Hutton and Foulsham). The contemporary standard in this domain of research are eye trackers that can appear similar to a pair of glasses. The eye tracker records most parts of the visual field of the subject.

As an example, the open source PUPIL eye tracking system (pupil-labs.com) uses a dedicated set of pattern recognition algorithms that allow for the detection of the pupil using a small webcam and project pupil movement after calibration on the video stream recorded by a second webcam facing the viewing focus of the participant. These systems are lightweight, record on a tablet PC connected to the glasses or via mobile-phone. They weigh less than 100 g and can record several hours of interaction in 1080p HD video. The frame is a custom-made 3D-print and is produced in high quantities to meet the increasing demand for large-scale mobile eye movement studies. They are used on a global scale and can provide conclusive results in very short time spans. Mobile eye trackers are most often used in stores because store orientation can be challenging to measure in laboratory settings.

An important technical innovation related to mobile eye tracking is visual based marker tracking that automatize analysis. Visual markers circumvent manual coding of the video material. The analysis software automatically detects the little black and white markers in the recorded video and projects gaze onto areas that are defined in size and orientation. In the near future, even more advanced and automated analysis frameworks will appear. They make use of neural networks to recognize objects in the visual field as it is currently explored in the German Research Center for Artificial Intelligence (Barz, 2016).

The next two dedicated innovations in eye movement studies in marketing are online measurement techniques of viewing patterns. In market research in general, online sampling has become one of the primary sources of data, especially for questionnaire-based research (see Box Comparison "Online vs. offline sampling"). There are a number of online access panels that offer extensive customer sample pools to invite to market research studies directly. Participants receive incentives for participation. Many demographic background variables are stored in the panel databases so that specific target groups can be easily recruited with efficiency regarding budgets and speed.

Restricted-focus viewer. The first online measurement solution was developed in the early 2000s by several companies in different variants. It is often referred to as "online restricted focus viewer," "spot viewer" or similar. Berger explains the process: "Participants keep the highest resolution area aligned with their foveal vision using manual cursor movements" (Berger, 2012). At its simplest level, participants are asked to move the mouse in the places where they are looking. A more advanced technique is restricting the field of clear focus to an area around the mouse pointer and blurring or masking the remaining part of the stimulus. This method simulates the focused viewing field of humans and forces the viewer to move the mouse wherever he or she wants to see the stimulus clearly. These approaches have the advantage that they are technically simple and do not require further hardware or software on the participants' computers. However, since eye movement is much faster than mouse movement, it comes as no surprise that the resulting data is of lower accuracy in the first few seconds of attention. It is doubtful whether these approaches can exactly mimic viewing behavior. However, academic studies (Berger, 2012) do show a correlation (e.g., ranking perception duration of certain areas on an advertising for both approaches and mouse tracking and real hardware-based eye tracking). This technique is still used in some marketing research applications, especially when the main focus is on speed and budget.

Webcam eye tracking. A more sophisticated innovation in online attention measurement is webcam-based eye tracking (Sewell & Komogortsev, 2010). Most consumer laptops and desktop PCs are equipped with built-in webcams. Several companies and university spin-offs (e.g., visagetechnologies.com and sightcorp.com) develop solutions for recording a video stream of the viewer sitting in front of his PC watching a stimulus (e.g., a website or digital print magazine). Pattern recognition algorithms for face, head pose and eye detection are combined to be able to distil pupil movement from this recorded video stream and to synchronously map this pupil movement data on the stimulus also being recorded. The analysis proceeds analogous to standard hardware based eye tracking data. This technology becomes especially important in attention research and the neuromarketing industry because it circumvents the cost of purchase and maintenance for hardware eye trackers and the lab-based logistics costs that come with central-location testing.

However, at this point, there are three limitations of the webcam-based approach:

1. *Technical limitations.* Technically, the resolution of webcams and the bandwidth for transferring HD video to the analysis servers still limits the accuracy of the analysis. In proprietary validation studies, a spatial resolution under 75 px on a standard monitor is rarely achieved. Furthermore, a significant amount of data drop-outs occurs because of lighting conditions, lack of focus on the viewers' pupils and other environmental factors that produce insufficient video quality for analysis.
2. *Privacy concerns.* Webcam eye tracking studies face some privacy concerns from participants. Approximately 50% of the participants being invited to webcam eye tracking studies refuse access to their webcam for this purpose (own statistics). They fear that the webcam recording could extend over the sole aim of the individual study that they are recruited for. This reservation can even lead some individuals to mask their webcams with tape to make any recordings impossible at any time. To increase participants' acceptance of webcam recordings, we developed a solution that processes the eye tracking data within the participant's

browser. The fact that there is neither installation necessary nor transmission of any video or picture to our servers has increased the acceptance of such studies in online panels. From a methodological perspective, this development is especially important because limiting the panel only to very "privacy-relaxed" individuals also introduces a sampling bias, a situation that must be avoided to be able to make general statements about viewing patterns rather than producing biased results.

3. *Mobile phone webcam recording.* Recordings of front-facing cameras on smartphones and tablets are challenging because of angle, light and movement artifacts. Correcting for hand movements remains problematic. Furthermore, smartphone screens are much smaller and have a higher resolution than laptop screens. Consequently, as of now, it is difficult to conduct eye movement studies on participants' own tablets and smartphones.

The major streams of development try to alleviate the shortcomings of webcam eye tracking recording regarding the drawbacks mentioned above. From our perspective, when these drawbacks are mitigated, webcam-based recording of viewing patterns is clearly the most promising avenue for future development in eye movement research.

Box 2: Comparison of Offline Versus Online Sampling

An example comparison of two projects illustrates the benefit of online sampling to a study's budget and timing: in 2012, a study was conducted on viewing patterns on several formats of online banners. This study was commissioned by the Arbeitskreis Online Forschung e.V. (AGOF), a major industry cooperative that measures the reach of the various website networks in Germany. We researched numerous ad formats with 700 participants in a lab-based setup. Participants visited the lab and were asked to surf several websites on which we integrated standard banners. We then compared the viewing duration and viewer reach of the various formats (AGOF e.V., 2012). Overall the project lasted for approximately nine months. In 2014, we conducted a similar project for an individual marketer of online banners. The experimental study had a similar focus and required a sample of 1000 participants because of the design and representativeness requirements. This time online sampling was used to invite participants who again surfed various websites with banners that were controlled by our browser plugin. The analysis yielded comparable yet even more detailed results. This project lasted no longer than one month and cost about a quarter of the budget of the first study. This comparison demonstrates tremendous advantages concerning budget and timing are possible to expect when switching from central-location testing to online testing.

19.3.2 The Client and Supplier Landscape

Most of the purchasers of eye movement research in neuromarketing are market researchers from large national and multi-national brands. Running eye movement research remains reserved to market researchers in larger companies. The market researcher typically receives internal requests for a study or to answer certain questions about consumer perception and, after consultation, on the right combination of method and setup and the size of the research commission for the project. In the domain of in-store communication, we often work with trade and category managers. In advertising, we work with media agencies, advertising agencies and TV stations. A special case in these domains is industry associations (such as the association of the digital industry of Germany), who frequently conduct studies on media usage and consumption. These studies can focus, for example, on the strengths of a certain media channel because the associations often have a more-focused interest on demonstrating why their domain is especially relevant for customers.

Finally, clients in the government and other institutions of the state conduct eye tracking based research. As an example, eye square conducts research for the Ministry of Transport of the United Kingdom. Eye movement research is used to discover viewing patterns on traffic signs and roadside advertisement for a safer driving style. These approaches have a similar focus as standard marketing studies but results are used to optimize communication with drivers on public roads to lower the number of traffic accidents coming from an overly dangerous driving style. The number of these governmental requests, however, is much lower than the traditional marketing departments of big companies.

Usability studies are another example of institutionalized eye tracking when eye tracking is required by law before a certain product such as a navigational system in a car can be used. This is the case for German ISO 15007, which requires the automotive industry to address the “measurement of driver visual behavior with respect to transport information and control systems.”

19.3.3 Applications and Typical Use Cases in Commercial Research

The landscape of touchpoints of brands with consumers is becoming increasingly more complex. Paid, owned and earned media; digital mobile interfaces; and street and shopper marketing work together to form a holistic brand experience (Rothensee, 2015). This chapter focuses on two consumer touchpoints: advertising and in-store communication. (User experience is another important application of eye movement research. However, since this subject is being treated in a dedicated chapter by Thoma and Dodd in this textbook, we deliberately omit it from our chapter.)

Advertising contains all means of direct communications of manufacturers of products (physical and informational). Advertising has a clear purpose: to inform the customer (my brand does exist and offers products in the domain X) and persuade, with the eventual aim of strengthening a product's attractiveness for a future purchase. This includes brand likeability messages such as "this brand does exist" and "this brand is likable, capable of producing high-value goods." The forms of persuasion can be very subtle and long-term oriented. Primary channels of advertising are defined by the various media that consumers use daily: TV advertising, print magazines, out-of-home billboards, banners on websites, Google AdWords, and social media. We will present research from those domains in the advertising section of this chapter. Findings will be presented that relate to the interplay and possible competition between these channels in contemporary everyday media usage.

In-store communication reaches consumers in brick and mortar stores. As we will clarify later, we differentiate three levels of analysis: the product itself (package design), the shelf (category management) and the physical store (in-store orientation). All three contribute to a holistic buying experience. In the past years, the digital shopper, the convergence between brick-and-mortar stores with digital channels, so-called omni-channel marketing has increasingly become reality (McGoldrick & Collins, 2007). Partly, our work is dedicated to harmonizing these different approaches to form an attractive brand world for customers (i.e., the cross-touchpoint customer experience) (Rothensee, 2015).

19.4 Applications in Marketing Communication

Pretesting marketing communication is one of the major applications of eye movement research in consumer research. Marketing communication has many forms, most important are advertisements in print and online magazines, TV commercials, and out-of-home campaigns. Every consumer has contact with one or more forms of such classical consumer marketing formats every day. As already mentioned, the amount of communication that consumers are exposed to by far over exceed the cognitive capacities of consumers. Eye movement analysis is mainly requested when questions about visibility, the optimization of creative executions of advertising messages, and media-usage trends need to be answered.

19.4.1 Measuring Ad Contact—Countering Ad Fraud and Delivering Ad Sales Arguments

The first goal is to measure if communication is actually seen. This rather mundane task is of high importance, given that there is much more information available to see but mainly not attended to or not processed thoroughly (Wilson, 2002). A typical print magazine these days contains 30–100 advertisements. A high number of these ads is glanced at only very superficially if at all (Wedel & Pieters, 2000). The same is true for other ad placements. Out-of-home marketing studies have been conducted that prove that passersby and vehicle drivers look at roadside billboards to justify pricing for these ad placements. In one of these studies, the attention of 1000 drivers in a driving simulator driving through a simulated city was analyzed. The results highlight gaze durations on various out-of-home media formats (smaller and larger billboards) (Ströer, 2016).

Online banners. The main application of visibility measurement relates to online advertising. Online display advertising is typically integrated in the form of banners into the content of the website. Here, they interrupt reading and attract the attention of the viewer, at least theoretically. Whether that actually happens, however, depends on the creative execution itself (Bundesverband Digitale Wirtschaft, 2014). The duration of viewing and the reach, the percentage of website visitors who look at the banner at all are important statistics for several actors in the field. Domain managers want to ensure that website visitors look at the banner positions to justify pricing on placing banner on these places. In this way, eye movement study results are used in pricing negotiations as a strong argument for a higher price for more attractive banner positions.

Ad visibility. Another important metric works on an even more basic level than gaze durations: validating ad impressions in digital advertising by means of ad visibility measurement (RealVu, 2010). Digital publishers that host websites, create content and sell ad spaces for a given price are increasingly under pressure to prove that digital ad placements are seen in the first place. The discussion on viewable ad impressions surfaced because companies claimed that only a fraction of the ad placements that media agencies buy from the publishers is seen at all (Google, 2014). The reason being that many websites are much longer vertically than the area seen when opening the website and rely heavily on hyperlinks as connections to other pages of the domain. Only by scrolling down can the audience see the content and ad placements further down the website, but this does not always happen. Many visitors scroll very little or not at all before finding an interesting article, clicking it and leaving the page to open a new one (the one with the article itself). In the past, media agencies bought ad space in lower regions of the website, often to a lower price at the risk that a fraction of visitors never scroll down to the place where the ad is. Ad placements were billed on the basis of page impressions but not viewable impressions. The result was that a significant amount of ad space was sold that was never visible to the audience. Increasingly, media agencies requested billing only impressions when they became visible to the visitor. This led to the development of ad visibility measurement. Ad visibility reach and duration are calculated by analyzing scrolling on the website. This results in a fine-grained statistic of how long and by what fraction an ad was in the visible space that appeared on the screen, the so-called browser viewport. A common definition of the IAB (2015) states that an ad is only visible when at least 50% of its size is present on the screen for at least one second. This measurement is currently used widely as a criterion for billing viewable impressions on websites. The first major publisher to introduce this for all sold ad spaces was Google, in 2013 (Serafin, 2013). It shows how high the importance of visibility for digital publishers is, even though technical visibility does not mean that the ad has been looked at.

It only relates to the *possibility* of its having been seen. A study for the German publisher United Internet Media demonstrated that in fact the correlation between viewer share in ad visibility and viewer share measured with eye tracking is moderate and depends on ad format (Klar, 2014). Some formats such as the sticky skyscraper, which stays in place when the website is scrolled down, have a very low correlation with eye tracking gaze contacts. Despite impressive statistics about viewing duration in ad visibility measurement, they are very infrequently looked at.

Banner blindness. Banner blindness describes the phenomenon in which the duration of viewing an ad when it is present decreases continuously due to conscious ad avoidance (Benway & Lane, 1998). In the first years of the internet, online banners were interesting to the audience because advertising in the digital space was new. However, with the passing of years, people got accustomed to the presence of banners and started avoiding them (Drèze & Hussherr, 2003). Today, banner blindness is most strongly related to ad formats on the right regions next to a website (so-called skyscrapers). The visitors of websites have learned that ads are typically placed in this space and avoid looking there. However, there are newer formats that compensate this banner blindness with size: the dynamic sidebar is a format that expands with larger widescreen computer displays. This leads to ad messages becoming very large on the right side of the screen, and since larger ads attract more attention than smaller ones (Google, 2014), they receive more attention than do traditional skyscrapers. With most other online banner formats and mobile ads, banner blindness is not a problem because these ads integrate into the content of the website and are less predictable in position.

Search Engine Advertising. Google also uses eye movement analysis to underline the effectiveness of paid advertising positions. Google AdWords, the search results that are sold on the top positions of the search engine results page (SERP), deliver clear performance-based and controlled-effect advertising. Advertisers can purchase these campaigns via auctions visitor by visitor and receive a detailed report on ad impressions. Viewability per se is not a problem for most of the placements on the standard SERP because they can be found on the top position of the results list and are therefore visible upon page load. However, Google wants to show that these AdWords also generate a branding effect (i.e., brand lift even without a click on the AdWords) when used. To this end, Google repeatedly conducts eye tracking experiments and publishes the results of them (Google, 2016). In one of these experiments, 101 participants visited the test studio and fulfilled a number of standard search tasks, such as “please search for an insurance company that you could contract for car insurance.” Their eye movements were recorded with a Tobii X1 standard eye tracker. Figure 19.3 shows the fixation distribution on the SERP after a standard search (task based). As seen, a high number of fixations spread in the upper left section of the SERP. People scan the results page using a combination of horizontal and vertical movement. The first paid search results gain much attention as do the first two or three organic (i.e., non-paid) results in the SERP, which are also of central interest for the viewers and are carefully scanned and read. Since the pattern evolved in other studies as well, it was named as the “golden triangle” or “F-scheme” of attention on the Google SERP. These analyses of attention distribution lead Google to the

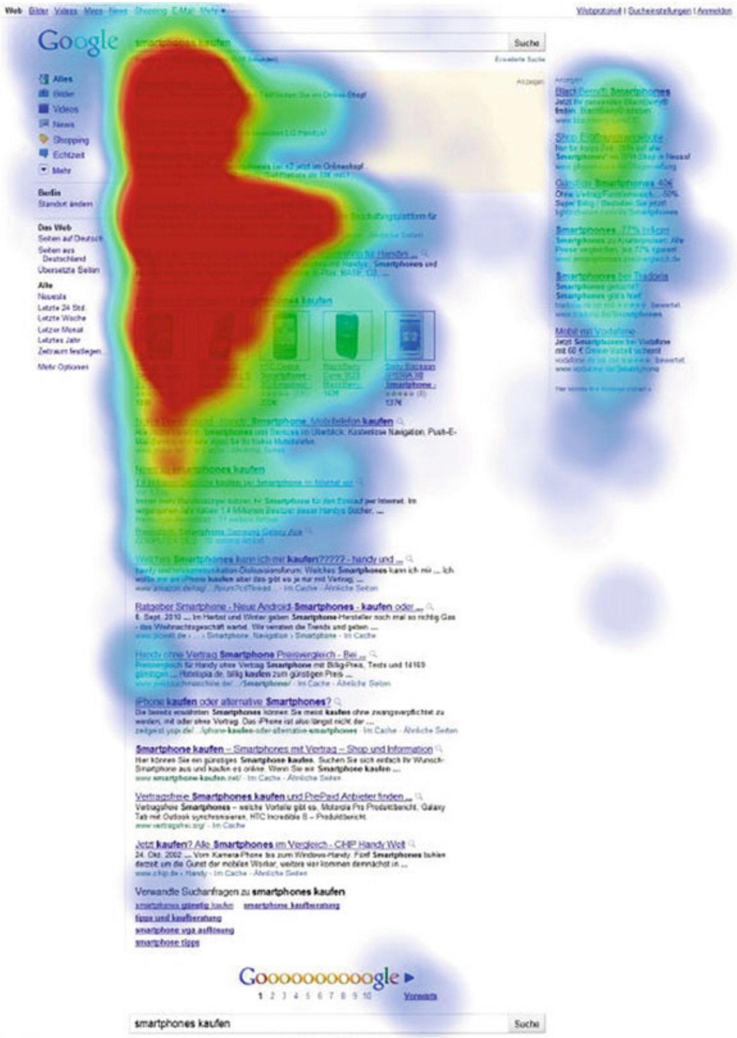


Fig. 19.3 Attention distribution on a Google SERP (Google, 2016)

conclusion that AdWords can increase awareness for the brand even if the audience does not click on the AdWords at all. This can be an added value the advertiser gains when using search engine advertising on Google.

More recent eye tracking research on Google SERP shows that the scanning behavior of the Google SERP has changed significantly between 2005 and 2014. The distinctive triangle shape center of attention on the top left of the SERP has evolved into a more vertically stretched pattern in which the page is increasingly scanned downward. This is considered to be a habituation effect coming from mobile usage. Since the SERP on mobile screens is longer viewers learn to scroll more and have a closer look at elements on lower positions in the search list. Attention also distributes more horizontally because Google introduces additional informational elements on the SERP such as the knowledge graph to keep visitors on the SERP (Mediative, 2014).

19.4.2 Optimization of Creative Executions—The Ad Itself

The second important case is the optimization of the creative execution of the advertising message. Online advertising research has shown that a huge gap exists between banners of poor creative quality and those of high creative quality (Bundesverband Digitale Wirtschaft, 2014). This optimization gap likely also exists for ads in the other media, such as print ads: poor ads receive very little attention, while powerful, highly salient executions attract much more initial attention and successfully retain viewers' attention longer.

There are best practices how to draw attention on those marketing stimuli, one of which is presented in the text box (Case "Vampire Diaries"). The optimization before flight of newly created creative executions is one of the most frequently requested applications of eye movement studies in marketing communication.

Box 3: Case "Vampire Diaries"

In 2012, German TV station ProSiebenSat.1 GmbH wanted to determine which campaign to use to promote the new series "Vampire Diaries" in their female content channel SIXX. The mascot of the channel is a chicken. The program is about a group of young people who act as vampires in the night. ProSieben tested several versions of print ads and out-of-home posters. The task of the study was to determine the most effective variant among a six candidates. Two classes can be reliably distinguished. One was a classical, testimonial-based approach. The billboard as well as the print ad showed the actors together on the ad, looking at the audience. The ads contained a rather provocative slogan, the logo of the TV channel SIXX, as well as detailed information on when the program is being aired, the starting date and time. The second group of ads contained the mascot of the channel, the chicken. It was designed to look similar to a bat with vampire teeth hanging from the "ceiling" (i.e., the top of the ad).

The study was carried out with 120 women in two German cities. Participants came to the lab and received two tasks: first have a walk through a "virtual city." This was a series of images from a city in which several billboards were visible. In one of these images, we integrated the different versions of the billboards. Then, participants were asked to flip through a digital version of a general interest lifestyle magazine. Likewise, we integrated the print ad at a specific position approximately in the middle of the magazine.

Reading/visiting sessions for both tasks lasted 20 min and were followed by a standard questionnaire.

Figure 19.4 shows the heat maps depicting the fixation hot spots on the out-of-home billboards.

We determined that the ads receive comparable amounts of attention, with a 95% viewer rate and approximately 6 s of viewing time. These statistics are higher than our benchmarks for billboard ads, so we concluded that these specific motives are stronger in catching the attention of passersby in the city than other campaigns. Furthermore, we noticed that the viewing times of the billboards were comparable with others. Since the chicken's motives are much less information-rich than the actors' motives, we concluded that there must be a certain point of fascination with the chicken's motive itself, together with a curiosity to determine what this symbolic depiction means. We also asked participants afterward which motive they like most. We determined that the traditional actor motive was viewed longer but less-liked, than the chicken motive. Combining both data, we could conclude that in this specific case the longer the actor's motive is viewed does not mean fascination but rather irritation because the motive was visually difficult to decode, especially from a distance.

Figure 19.5 shows the gaze transitions on the advertising. It shows that the major transitions on the ad occur between the several actors. There are no significant connections to the small chicken in the upper right corner, nor are there many transitions downward to the detailed information in the lower section of the ad. This is a typical pattern called the "vampire effect" (Kilian, 2009). It appears when there are many intensive visual stimuli (here: actors, attractive people) present and draw attention away from important written information or other content. This pattern is also problematic because it is known from psychological research in memory that our short-term memory can hold only approximately 7 pieces of information (Miller, 1956). Counting the number of elements on the ad, it becomes clear that the actors and the chicken, brand logo, program info, etc., are so many information pieces that the billboard is difficult to "digest" for the viewer and is therefore not optimal.

Figure 19.6 shows an illustration that was produced to underline the most likely scan path on the advertising. We continuously find, as in other studies, that people tend to "read" ads as they read text on a page. They start scanning in the upper left corner and work their way down the ad (cf., the Google SERP heatmap in the previous section). This notion, together with the aforementioned vampire effect, explains why very few people look at the small chicken in the upper right corner in this ad: it does not stand in the natural reading direction of the audience and therefore has a low probability of being seen.

What is more, the ad is very dark and has a large amount of black in the lower section, which also renders it less attractive for the viewer (Fig. 19.7).

The gaze transitions between areas of attention in the chicken motive are more straightforward. The central visual element catches most of the attention. After having looked at the chicken, most viewers' gazes travel down to the logo of the program and further on to the detailed information. The poster contains fewer elements and is thus easier to decode than the actor motive, even though it is very non-traditional. The few elements do not overtax viewers' short-term memory. Furthermore, it has larger areas in bright green than black and therefore has a stronger pop out effect when placed in the streets. Bright green as a color code was especially effective in this case because the campaign was launched in February when German streets are rather dark and grey. The bright green is a strong contrast and thus sticks out of the street and can thus attract more attention from the viewer than a darker, essentially black motive.

Based on this visual analysis, the additional analysis of the print ads and other measures, such as ad recall, brand evaluations, viewing intention, etc. we recommended the more innovative motive "chicken" to the TV channel. The motive was used in a large combined campaign of out-of-home billboards and print ads in magazines and created a huge success for the station SIXX. Until then, no other SIXX program had such high viewer rates in the German TV population as "Vampire Diaries."

Optimizing creatives and optimizing digital banner formats follow the same logic: to have an effect, visual ads in the various media have to gain attention from the viewer and hold their attention. To this end, a deep analysis of advertising can result in a) selection of the most effective ad from a series of different versions or b) recommendations from the eye movement experts can help to improve a given campaign so that it will attract even more attention than it did in the test. Both strategies are often combined in marketing testing and lead to more attractive advertisements that

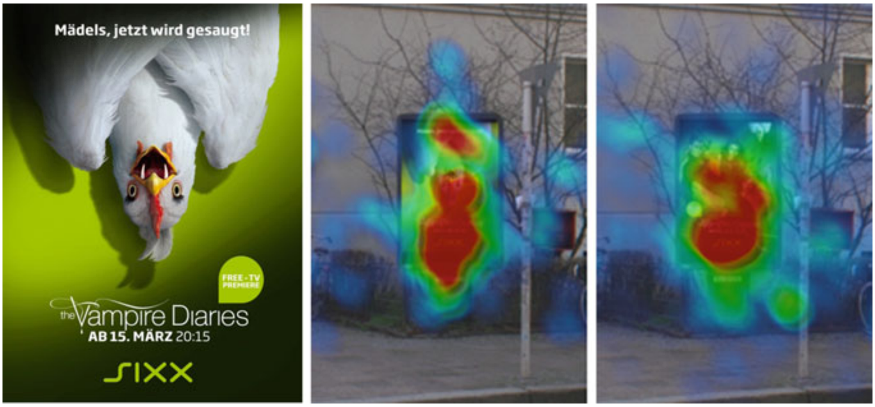
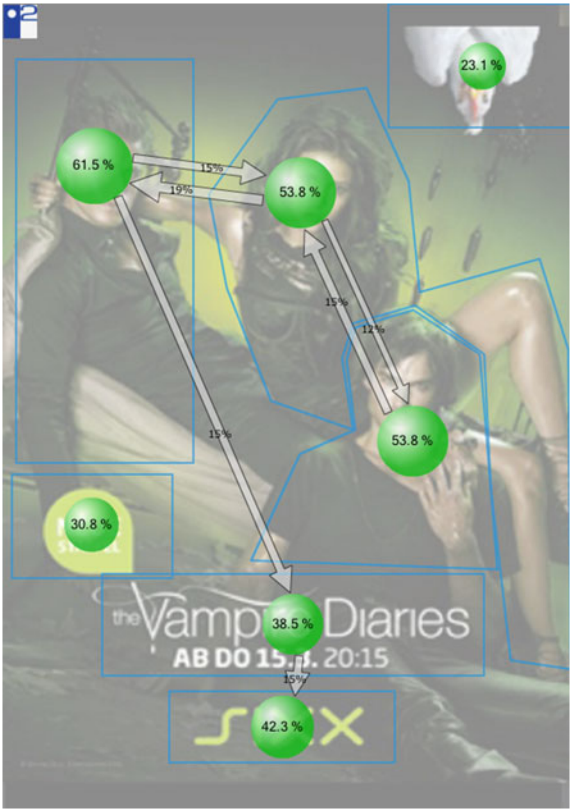


Fig. 19.4 Heatmap attention distributions on two versions of the poster

Fig. 19.5 Gaze transitions on the advertising



receive more attention from the audience. More attention is considered important because, first, the longer viewers look at an ad, the more likely is that they will receive the message that was communicated to them correctly. Second, relating to dynamic messages, considering that digital advertising is not static but tries to communicate a story it is important to follow the story, having gaze contact not only once but repeatedly throughout the entire flow of the animation of a banner for instance is also vital for understanding the marketing message. If consumers look at banner animation cycles in the "wrong" moments only, it is possible that a marketing message is understood (e.g., "it is important to have an insurance when you travel abroad") without learning the associated branding message ("our brand offers good travel insurance at reasonable costs"). Third, the longer the ad contact is, the deeper the viewers' immersion into a branded world of the ad, where sustainable branding effects become more likely to reach the goal of the campaign at hand (i.e., marketing campaigns often act not only to communicate availability and attractiveness of a certain product but also to strengthen the attractiveness of a brand for other products or product categories).

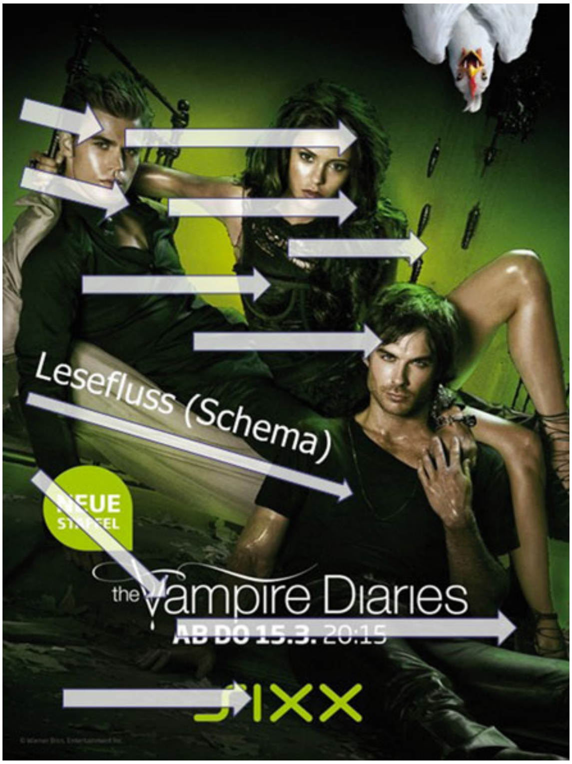


Fig. 19.6 Estimated reading flow on the ad for illustration to the client

The selection of the winner from an array of possible variations of an ad and the specific recommendations for optimization are also combined with expertise about the average viewing times of ads in the various media channels. This helps the creative agencies to adapt to the attention reality that exists in the various media. Some years ago, we noticed abundance of digital ads that operate on 15–20 s animation cycles and tried to tell a story from a TV commercial. A benchmark analysis showed that typical online banners, as the medium rectangle, a 250*300 px ad placement found on most popular websites are viewed for no more than 800 ms on average (Schiessl, 2013). These figures help creative agencies to start new campaigns with realistic expectations on the "attention canvas" that they have in typical media consumption situations. Thereby, they can avoid producing banner ads that cannot work at all because the branding message appears only after a 15 s story in a banner. Table 19.1 lists the average viewing times of ads in various media channels (Statista, 2016). They are computed on the basis of 25,000 data points collected between 2008 and 2013 and can help to draw a broad picture of the attention that ads or commercials typically receive. However, these statistics must be interpreted with caution: they stem from lab-based research and the digital banners, especially, are represented by a small format, the medium rectangle. Since there is a wealth of formats in the digital ad space, this number must not be taken as representative for online banners

Fig. 19.7 Gaze transitions on the alternative motive

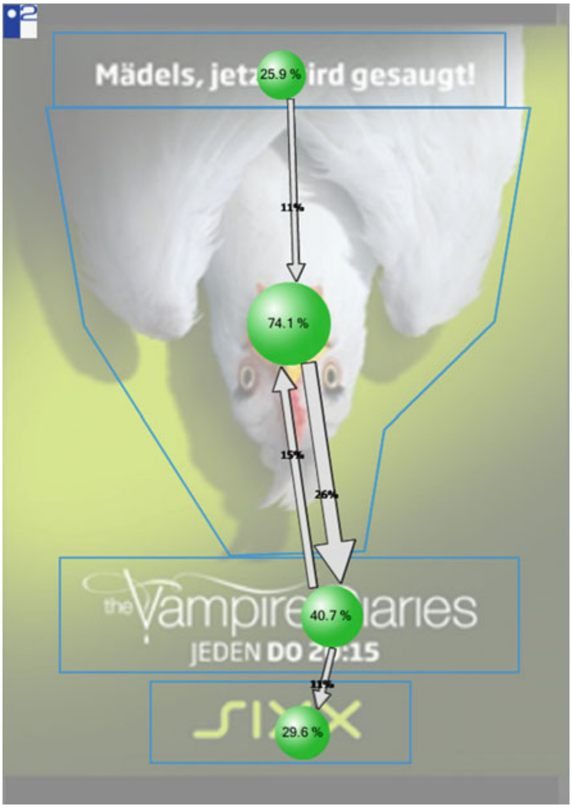


Table 19.1 Average viewing times in various media channels, own data, published on Statista (2016)

Format/media channel	Average viewing time (s)
Online banner (medium rectangle)	0.8
Print ad (whole page in a magazine)	3.5
Online video ad (e.g., YouTube preroll)	9.3
TV commercial	19

in general. Eye tracking studies also demonstrate that larger, more effective online banner formats can generate much more attention than the medium rectangle.

Ultimately, having prolonged visual contact with the audience can help to render traditional advertising more effective, thereby increasing cost-effectiveness of mass media advertising for the advertiser. This is especially important because we notice that media consumption in general becomes increasingly intensive (SevenOne Media, 2014): as people spend more time with different media, they do so in a combined, mixed, parallel fashion so that situations in which there are concurrent branding messages available become ever more likely. This change in media consumption will be the focus of the next section.

19.4.3 Understanding Changing Media-Usage Patterns (Multiscreen)

An important field of application for eye tracking in neuromarketing is the analysis of the change of media consumption patterns. Eye tracking can help to accurately determine how the attention of media consumers' changes and which channel gets how much attention in the so-called attention economy (Davenport & Beck, 2002). As an example of this field of application we take a closer look at the notion of multiscreen media usage in the interplay between TV and digital channels.

Since 2013, there is the notion of multiscreen media usage. It is now popular to use multiple devices at a time; for example, when you watch a TV show you use your smartphones and tablets in parallel. Survey-based research has yielded inconsistent results regarding the average multiscreen consumption time budgets of consumers. Fifty-nine percent of the internet users in Germany say they regularly use multiple devices when watching TV (Universal McCann, 2014). Another study reports that only 28 percent of internet users confirm this activity (TNS, 2013). This created doubt as to whether people can accurately report on their media usage at all. As one possible solution to this dilemma, eye movement analysis with mobile eye trackers is used to quantify the amount of time dedicated to the various media channels when in front of the TV.

In an ethnographic study commissioned by United Internet Media in Germany in 2014, 20 households were visited with mobile eye trackers for an evening of free TV-watching. Two hours of media consumption per household were selected and analyzed. A representative sample of households of multiscreen users was invited to the study. Participants used a PUPIL mobile eye tracker to record their attention while watching TV. It returned the following insights about multiscreen media usage (United Internet Media, 2014):

- while the TV is running, attention is distributed evenly between TV and digital devices, with 49% of the time spent watching TV and 47% spent viewing digital devices such as smartphones, tablets, and laptops
- when there is a commercial break on TV, viewers do not switch to another program but rather shift their attention to the digital devices, paying more attention to smartphones, tablets, etc., and less to TV
- the average audience switches attention from TV to digital and back 112 times, approximately 2 "channel switches" (gaze transitions) per minute.

From a methodological perspective, the last figure is particularly interesting: the high number of switches from one device to the other makes clear that these switches occur often and are short and mundane in nature. Checking a Facebook profile, checking for new mail or a short chat are quick, mundane usage situations that participants have a hard time counting and remembering the frequency. This can explain why self-reports on multiscreen usage vary so strongly between different surveys: people cannot remember their media usage behaviors. This is a strong case for the use of very accurate and detailed analyses for a better understanding of media usage in many situations.

The economic background to the multiscreen discussion is who gets the attention of the viewer when multiple advertising messages (e.g., on the TV and simultaneously on the smartphone) are present. Which advertising is more effective in such direct competition situations? Does a TV commercial still have an effect if the audience is surfing on their Facebook profiles and only hear but do not see the advertising message? Research is being conducted to produce best practices for optimizing TV commercials to have an effect even without active viewing (Teichmann, 2014), and there is much potential for synchronized campaigns running on the TV and on smartphones in parallel.

19.5 Applications at the Point of Sale

The final section focuses on applications on the point of sale. We will present research methodologies and case examples of three levels of analysis. The point of sale is typically investigated on the level of the packaging, the shelf and the store as a whole. On each of these levels of investigation, eye tracking can help optimize products, placements and the store environment. Contrary to the advertising field, in the store another component is the target of analysis: orientation. The greatest number of studies are primarily concerned with marketing purposes, with the eventual aim of increasing sales at the point of sale. However, retailers are often interested in store redesigns that make the store more attractive without sacrificing orientation in the store. In this way, the study's objectives parallel what is the focus in the user experience chapter of this book (Thoma & Dodd): ability to find products, speed of orientation, identifying shopper expectations and analyzing success rates.

Eye movement analysis on the point of sale is justified because retail stores often offer a plethora of marketing messages and available brands in densely organized spaces. A typical larger retail store can easily integrate 50,000 products (Sorensen, 2009) on a few 100 square meters. It comes as no surprise that the orientation and visibility of individual products are primarily visual matters that are driven by underlying orientation principles such as colors, contrasts, and shapes. As we have learned in the previous chapter (multiscreen), consumers are often neither willing nor able to accurately report on their perception. The store, with its high number of visual elements, is a pivotal case for detail analysis and making optimizations about which shoppers would not have reported in the first place.

Retailers strategically follow a paradoxical, double-bind optimization logic: rendering the store attractive for customers, making them feel comfortable, allowing for a barrier-free, hedonic shopping experience will increase brand perception, consistency and store loyalty in the long-term, but retailers try to maximize the total sales volume of the store by attracting as many customers as possible and increasing the sales volume of every individual in the store. Customers therefore have to optimize the potential contrary aims to clean up the store environment (i.e., for an easy, intuitive shopping experience) and at the same time integrate in-store marketing materials as effectively as possible. Additionally, integrating digital and brick-and-mortar stores into a holistic brand proposition becomes increasingly important as an additional strategy.

19.5.1 Packaging Design—Optimizing Perception of Product Features and the Brand

Introducing new products in a saturated market is a difficult task. The advertising agency network BBDO estimated that in Germany, 910 products enter the consumer market every week (Batten & Company, 2001). As convergent figures across many markets, industries and brands show, typically more than 90% of product introductions in a market in a year fail (i.e., they leave the market within the same year that they are introduced). One reason for that is a mismatch in consumer demand and product availability (i.e., many products are introduced that are not interesting for consumers). Another reason, however, is the visibility and distinctiveness of the product. Therefore, one target of packaging design is catching the attention of the shopper and creating a desire to purchase the product. Furthermore, branding has the aim of rendering the product distinct from other brands in the same category, so that shoppers can find the exact product that is shown in advertising. Brand parity is a problem in that many brands, in the minds of customers, are not distinguished from each other. In 2009, on the average 64 percent of shoppers said that brands in a category are not distinguishable from each other (BBDO Consulting, 2009). The packaging is a very important feature of almost every modern product in the consumer market. It communicates product features and the brand proposition. It is often the first point of contact of a customer with a brand (i.e., when it is from a brand with little or no classical advertising volume) and therefore has to tell a story about its content.

Box 4: Werther's Originals Optimizes Packaging with Eye Tracking Research [Update]

In 2013, Storck—a German candy manufacturer—started a study to investigate possible optimizations to the product packaging of one major product in the caramel sweets category: Werther's Original. The product is sold in packs of 250 g and has a long tradition in the German and other markets. The goal of the redesign was to strengthen the brand positioning in the various markets, rendering the packaging more attractive and visible on the shelves together with the competition and letting them tell a compelling story about the origins of the Werther's Original brand. Werther's Originals image is traditional and emphasizes the quality and success of the sweets. As many other products in the fast moving consumer good category Werther's Originals uses the country of origin effect to add an image component to the value proposition of the candy, namely that the shopper not only buys a good-tasting candy but also buys a piece of German candy-maker tradition.

Research was conducted in Germany, the UK and France to test a variety of packaging, including the original and a number of redesigns. The products were positioned in virtual shelves together with the typical set of competitors in the category. Participants in the study had several tasks, including a typical finding task ("please locate Werther's Originals on the following shelf as quickly as possible") and a shopping task ("please shop virtually for candy on the following shelf"). When shoppers had selected a product on the shelf by clicking it, the product was shown in zoomed view so that shoppers could take a closer look at all the products' details. After selecting a product, shoppers answered a questionnaire on product attractiveness, features, brand evaluation, purchase intention and other parameters.

Figure 19.8 shows the results of the eye tracking analysis of one particular packaging. As seen, attention is spread evenly across the packaging, and the majority of the viewers read the product name and brand. The value proposition and product depiction receive attention.

The results of the study were successful for Storck. The optimal package design was identified for every market. Interestingly, in the UK, a different design won the competition than in Germany. This shows that there are cultural and competition effects in the various markets. As a result, Storck changed the packaging design of the Werther's Originals in the different markets.

The case of the Tropicana packaging redesign (Visible Logic, 2009) underlines the importance and effect of packaging redesigns: Tropicana, one of the major brands for orange juice in the US decided to change the packaging from a traditional product image to a more modern, abstract product depiction on the pack. With the introduction of the new package, the market share of Tropicana dropped by 20% (Fett, 2013). Within a very short time, the redesign of the packaging was reversed and consequently Tropicana market shares recovered.

The Werther's study and the Tropicana case show the importance of creating packaging that stands out on the shelf and tells a story about the manufacturer and its content. Eye movement research can help improve packaging in real-life contexts at the point of sale.

Fig. 19.8 Heatmap attention distribution on packaging

19.5.2 *On the Shelf—Finding a Product and Adapting to the Mental Model of the Shopper*

The shelf layout is typically organized by retailers' category managers. The category manager tries to optimize the layout so that the brand gets a fair share of interest of the shopper. Simultaneously, brands can often intervene in the shelf architecture to appear in the optimal positions. Shelf layout can take roughly two different forms. The first type is functional organization, the division of shelf regions in functional zones and sorting of various brands according to these functional criteria. As an example, a hair shampoo shelf can be divided into female and male haircare sections, special sections for colored hair or special dandruff shampoos for men. Within each of these functional sections, brands are sorted so that the shopper can make a primary distinction on what type of shampoo he or she wants and afterwards choose from a number of brands. The second type is the branded shelf architecture. In this version, the shelf is sorted by brand, so that the dandruff shampoo from brand A will be found in the brand A section whereas the dandruff shampoo of brand B will appear in the brand B section. These branded sections are often further distinguished by additional branded material such as shelf displays, logos, etc. Depending on the retailer, the category and the shoppers mission the functional or branded layout of a shelf has advantages concerning findability of products and generated sales volume (Sorensen, 2009).

Eye movement research on the shelf has demonstrated that the visual search for a product can be very fast when shoppers have a loyalty to a certain brand as the most common brands serve as key brands in visual macro-scanning processes (Deherder, 2011). Scanning a shelf searching for a certain product is therefore strongly influenced by brand preference. In this situation, newly introduced brands have a hard time catching the attention of the shopper since the category complexity can lead to "visual blindness" (Sorensen, 2009). Brand salience can be positively influenced by increasing the number of facings on the shelf, which in turn has a positive effect on visual attention (Chandon, Hutchinson, Bradlow, & Young, 2009). It is also known from eye movement studies (Horsley, 2014) that the viewing patterns of customers can accurately predict where the attention will go next and even which product is most likely to be purchased (Wedel & Pieters, 2008a, b). The primary metrics in this prediction are the time to first contact (e.g., when shoppers look at a product for the first time, the earlier the product catches their attention, the higher the probability of a purchase) and the number of repeated visual contacts (products that are looked at more than once have a much higher purchase probability than products with only one eye contact).

That is why manufacturers pay considerable attention to create highly distinctive products on the shelf. Color, shape of the packaging, logo integration are important factors in this matter. This is the reason why studies of packaging redesign often include a shelf findability task.

Box 5: Predicting Product Choice on the Shelf

A study conducted in 2012 demonstrated the value of on-the-shelf eye tracking analysis for predicting choice in a subsequent purchase task (Schiessl, 2012). One hundred twenty women were invited to a studio to watch TV commercials for a variety of shampoo and soft drink brands. After having seen the commercials, they were presented with a virtual shelf for each category projected in real size onto a wall using a beamer. Their eye movements were recorded with a mobile eye tracker. Figure 19.9 shows a typical virtual shelf with a large number of shampoo brands. Shoppers were asked to select a shampoo for purchase on this shelf.

Regarding how to predict the (virtual) purchase of a shampoo on the shelf, it was shown that an analysis of implicit indicators on the shelf such as eye tracking increased prediction accuracy by 10% over the accuracy of an analysis of the effect of the TV commercials. Thus, in this particular study, watching the TV commercials and analyzing customer brand loyalty predetermined 63% of the actual product choice. The shelf added another 10% to the prediction. This study, however, is biased toward the TV commercial effect since the situation watching a TV commercial and directly after that looking at a shelf is somewhat artificial. Therefore, it is highly likely that in natural situations on the shelf where the customer might have seen a TV commercial sometime in the past, a week ago or not at all, the influence of TV commercial contact is probably much lower, giving room for the product presentation on the shelf itself. The results for predicting soft drink choice were consistent, with an 11% surplus in prediction accuracy from eye tracking indicators on the shelf.

The study emphasizes that eye tracking on the shelf can help predict the purchase probability of a product. This method can be used to pretest several versions of a

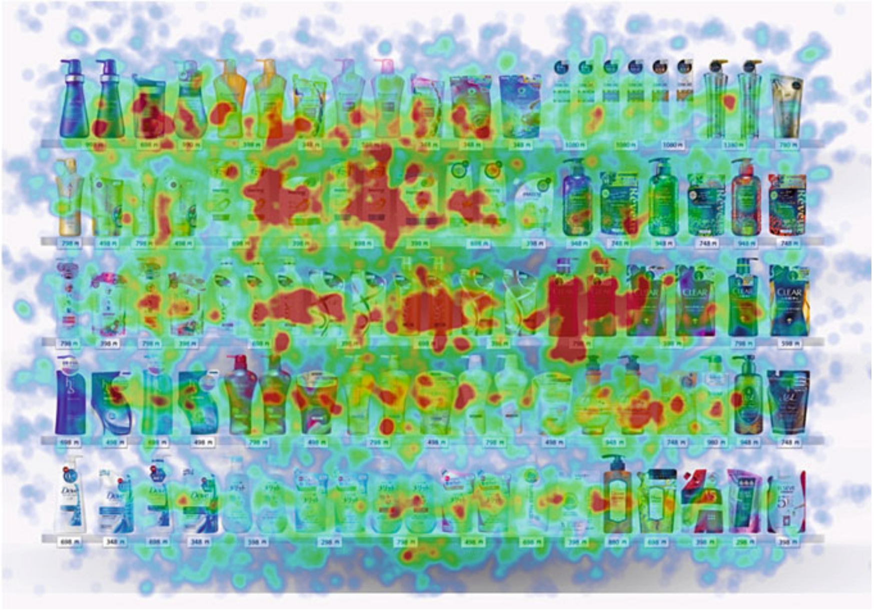


Fig. 19.9 Virtual shampoo shelf with attention distribution (own data)

package (re)design. The result of these optimization studies is typically a recommendation for selection of a particular product packaging to be introduced. In the redesign of packaging for a long-established brand, there especially lies a certain risk that the product will not be recognized by its loyal customers and consequently will lose market share against competitors. Therefore, a high number of studies is conducted in package design for established brands. Anecdotally, we can report on the global redesign of a major cosmetics brand in which a series of shelf findability tests were conducted: the packaging redesign was embedded in a brand rejuvenation strategy that was planned over 3 years and the redesign itself proceeded in several evolutions over the course of 1 year to accustom the shoppers to the new brand design, logo presentation and packaging design. This shows that brand relaunches in packaging design are often an effortful, risky task with a substantial investment in accompanying market research.

19.5.3 The Store—Orientation and Point of Purchase Materials

Optimizing the store architecture and introducing point of purchase materials helps to create a certain store atmosphere. Modern shoppers expect from a store that it offers a unique, hedonic, and seamless shopping experience. In this way the retailer can build a strong loyalty. Shopping in a market of 50,000 products is a difficult selection task. Research shows that approximately 80% of the time in the shop is not actually spent with direct purchase related activity but instead walking around and searching but not looking at products (Sorensen, 2009). Therefore, any amendments that render the store "easier to use" or more understandable from the customers' perspective are highly welcome optimizations from the perspective of the shopper. Eye movement studies help to identify the best positions for store guidance materials, (e.g., on the floor, hanging on the ceiling, or at the ends of the shelves). Typically, several redesign iterations of a retail store are compared against the status quo. Analyses of reach and reader frequency of the orientation materials, together with performance metrics regarding shopping (speed, error rates, finding products, etc.) help to identify whether a store redesign offers an improved orientation for the shopper or not.

In a recent study, Harwood and Jones investigated orientation in a UK department store (Harwood & Jones, 2014). Their in-depth coding revealed among other findings that when in the department store, shoppers dedicate 72% of their eye movements to the actual products, 6% to their shopping partner and 6% to product labels and prices. In the apparel and garden departments, more attention is paid to the products and less to shopping partners. They show in detail how sight lines and contrasts influence the visual saliency of components of the store.

The introduction of effective point of purchase material can further improve the shopping experience. It can also create moments of inspiration for shoppers that distract them from functional and fast shopping and seduce them into making unplanned impulse purchases. Experts estimate that approximately 55% of purchases in a market are not well-planned but impulsive purchases elicited, among other reasons, by effective point-of-purchase advertising (Baun, 2013; Neue Verpackung, 2002).

Eye movement research helps to identify zones in which the attention of shoppers is particularly high and movement speed is low. These zones can be used to present the shopper with additional promotional material. The material itself is often tested in several variations to identify the most effective material. Success criteria for such point of purchase promotions are typically the purchase frequency of the advertised products.

19.6 Conclusions

The application of eye movement analysis in neuromarketing is a very active field with a number of important research institutions and many projects commissioned per year. The majority of all global blue chip brands conduct eye movement analysis in the fields of advertising or shopper marketing on a regular basis. The last sections showed, in a number of examples the most prominent use cases for eye movement in the marketing world. Generally, eye movement analysis adds to our understanding of customer behavior and preference. It makes it possible to better understand the critical moments when consumers make contact with a brand, whether in the form of advertising or at the physical point of sale. Brand managers are aware that to create successful marketing, moment-by-moment analysis of consumer behavior provides crucial insights and helps to stay ahead of competition.

The enumeration of use cases is not meant to represent the full bandwidth of applications of eye movement research in neuromarketing. This would be a tedious task since there is a great variation of use cases today. It is also biased toward applications in Europe and Germany in particular.

However, we hope that by presenting the most prominent examples from brand and media as well as shopper research we could demonstrate how substantial the added value of eye movement research in the field of consumer marketing is. It is undoubtedly true that in the future technical innovation and the change of the media landscape will create even more demand for dedicated eye movement analyses and enlarge the play field for our profession in the future. A significant effort is made toward further automation and standardization of eye movement research so that it constantly adds to the growing body of data for brands to make informed decisions on marketing strategies in their markets.

19.7 Questions Students Should Be Able to Answer

1. For which market research questions is eye-tracking most helpful?
2. Which key performance indicators (KPIs) tell if a poster is the winner when comparing two versions?
3. How does eye movement research inform the optimization of product packaging on the shelf?
4. Why is it important to test products or ads in natural contexts like shelves and magazines when using eye-tracking?

19.8 Suggested Readings

Ariely, D., & Berns, G. (2010, April). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience*, pp. 284–292.

- *Gives a good overview of market research and neuromarketing tools. Highlights the benefits of these measures in early stages of product development.*

Beck, D. &. (2002). *The Attention Economy: Understanding the Currency of Business.* Boston: Harvard Business School Press.

- *The authors explain why we live in an information-saturated world. An example of how to get a customers attention is to customize a product.*

Kahneman, D. (2012). *Thinking, fast and slow.* New York: Penguin.

- *Explaining the Implicit and Explicit Systems 1 and 2 with various examples this is a behavioral economics classic.*

ESOMAR. (2012). 36 questions to help commission neuroscience research. Amsterdam, The Netherlands.

- *This paper helps buyers of neuromarketing to select the right tools and ask the companies that offer the service the right questions.*

Perrachione, T. K., & Perrachione, J. R. (2008). Brains and brands: Developing mutually informative research in neuroscience and marketing. *Journal of Consumer Behaviour*, 303–318.

- *The author explain the motivation and use neuroimaging techniques used in neuromarketing as well as in traditional neuroscience.*

Plassmann, e. a. (2012). *Branding the brain: A critical review and outlook.* Journal of Consumer Psychology: Elsevier.

- *Good overview of neuroscience in combination with consumer psychology as well as list with studies about branding using neuromarketing methods.*

Zurawicki, L. (2010). *Neuromarketing, Exploring the Brain of the Consumer (Pg 42–53).* Berlin Heidelberg: Springer Verlag.

- *A complete textbook about the field of Neuromarketing with an extensive bibliography and information about other methods like segmenting consumer groups based on neuromarketing insights.*

Bibliography

AGOF e.V. (2012). AGOF untersucht Werbewahrnehmung im Web. Frankfurt/Main. Abgerufen am 2. 9 2015 von http://www.agof.de/pressemitteilung-2012-09-12/.

Ariely, D., & Berns, G. (2010). Neuromarketing: the hope and hype of neuroimaging in business. *Nature Reviews Neuroscience, 11*(4), 284–292.

Barz, S. (2016). Gaze-guided object classification using deep neural networks for attention-based computing. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 253–256.

Batten & Company. (2001, December 1). *Brand Equity Excellence*. Düsseldorf. Abgerufen am 2. 9 2015 von http://www.batten-company.com/uploads/media/2001_12_Brand_Equity_Excellence_1_Review.pdf.

Baun, D. (2013). *Impulsives Kaufverhalten am Point of Sale*. Heidelberg: Springer.

BBDO Consulting. (2009). *Brand Parity Studie 2009*. Düsseldorf. Abgerufen am 2. 9 2015 von http://www.batten-company.com/uploads/media/20090219_Brand_Parity_Studie_2009_01.pdf.

Beck, D. (2002). *The attention economy: understanding the currency of business*. Boston: Harvard Business School Press.

Benway, J., & Lane, D. (1998). *Banner blindness: Web searchers often miss 'Obvious' links*. Houston, TX. Abgerufen am 2. 9 2015 von http://www.ruf.rice.edu/~lane/papers/banner_blindness.pdf.

Berger, W. S. (2012). Assessing advertising effectiveness: The potential of goal-directed. *Psychology & Marketing, 29*(6), 411–421.

Bundesverband Digitale Wirtschaft. (2014). The Power of Creation. *Results of the Online-Vermarkterkreis (OVK) Study 2013*. Düsseldorf. Abgerufen am 2. 9 2015 von http://www.bvdw.org/mybvdw/media/download/bvdw-ovk-power-of-creation-2013-powerofcreation-english.pdf?file=3101.

Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing, 73,* 1–17.

Chandon, P., Hutchinson, J., & Young, S. H. (2016, November 15). *Unseen is unsold: Assessing visual equity with commercial eye-tracking data.* Von https://faculty.wharton.upenn.edu/wp-content/uploads/2012/04/Pcpc.pdf abgerufen.

Colin, C. E. (1953). Some experiments on the recognition of speech, with one and with two ears. *Journal of the Acoustical Society of America, 25*(5), 975–979.

Davenport, T., & Beck, J. (2002). *The attention economy: Understanding the new currency of business*. Bosto, MA: Harvard Business Review Press.

de Oliveira Joaquim dos Santos, R., Caldeira de Oliveira, J., Rocha, J., & Giraldi, J. (2015). Eye tracking in neuromarketing: A research agenda for marketing studies. *International Journal of Psychological Studies, 7*(1):32.

Deherder. (2011). *Shopper intimacy.* New Jersey.

Drèze, X., & Hussherr, F.-X. (2003). Internet advertising: Is anybody watching? *Journal of Interactive Marketing, 17*(4), 8–23.

Duchowski, A. (2007). *Eyetracking methodology*. London: Springer.

Egner, I. S. (2000). Implicit measures of consumer cognition: A review. *Psychology & Marketing, 27,* 921–937.

ESOMAR. (2012). *36 questions to help commission neuroscience research*. Amsterdam, The Netherlands. Abgerufen am 2. 9 2015 von https://www.esomar.org/uploads/public/knowledge-and-standards/codes-and-guidelines/ESOMAR_36-Questions-to-help-commission-neuroscience-research.pdf.

Fett, E. (2013, September). *Tropicana sales drop after changing logo on orange juice bottle*. Abgerufen am 2. 9 2015 von http://www.worldwideweirdnews.com/2013/09/w7700.html.

Google. (2011). Eye tracking study: Wahrnehmung der Suchmaschinen- und Displaywerbung Mobil vs. Desktop. Mountain View, CA. Abgerufen am 2. 9 2015 von (http://google.inxshare.de/09_Agency/2011/August/Eye%20Tracking%20Results_deutsch.pdf).

Google. (2014). *The importance of being seen*. Mountain View, CA. Abgerufen am 2. 9 2015 von https://think.storage.googleapis.com/docs/the-importance-of-being-seen_study.pdf.

Google. (2016, December 2). *Eye-Tracking-Studie vergleicht Mobilgeräte und Desktop*. Von www.thinkwithgoogle.com: https://www.thinkwithgoogle.com/intl/de-de/research-study/eye-tracking-studie-vergleicht-mobilgerate-und-desktop/ abgerufen.

Harwood, T., & Jones, M. (2014). Mobile eye-tracking in retail research. In M. Horsley, M. Eliot, B. A. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research* (pp. 183–199). Cham: Springer International Publishing.

Horsley, E. K. (2014). *Current trends in eye tracking research*. London: Springer.

IAB. (2015). http://www.iab.net/media/file/2015_State_of_Viewability_Transaction.pdf. Abgerufen am 2. 9 2015 von http://www.iab.net/media/file/2015_State_of_Viewability_Transaction.pdf.

Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 20,* 1254–1259.

Kahneman, D. (2012). *Thinking, fast and slow*. New York: Penguin.

Kilian, K. (2009, August). Was ist ein Vampireffekt? *Absatzwirtschaft*, p. 72.

Klar, J. (2014). Wie Werbetreibende Online-Kampagnen richtig bewerten. Karlsruhe. Abgerufen am 2. 9 2015 von http://united-internet-media.de/fileadmin/uim/media/home/Studien/Gemeinschaftsstudie_UIM_Meetrics.pdf.

Kroeber-Riel, W., & Gröppel-Klein, A. (2013). *Konsumentenverhalten*. München: Verlag Franz Vahlen.

Lindström, M. (2008). *Buyology: the new science of why we buy.* Currency Doubleday.

Lohse, G. L. (1997). Consumer eye movement patterns on yellow pages advertising. *Journal of Advertising, 26*(1), 61–73.

Mayer, H., & Illmann, T. (1999). *Markt- und Werbepsychologie*. Stuttgart: Schäffer Pöschel.

McClure, S., Li, J., Tomlin, D., Cypert, K. S., Montague, L. M., & Montague, P. R. (2004). Neural correlates of behavioral preference for culturally familiar drinks. *Neuron, 44*(2), 379–387.

McGoldrick, P., & Collins, N. (2007). Multichannel retailing: Profiling the multichannel shopper. *International Review of Retail, Distribution and Consumer Research, 17*(2), 139–158.

Media Dynamics Inc. (2016, October 17). Von http://www.mediadynamicsinc.com/uploads/files/PR092214-Note-only-150-Ads-2mk.pdf abgerufen.

Mediative. (2014). *The evolution of Google's search results pages & effects on user behavio*ur. Abgerufen am 2. 9 2015 von http://www.mediative.com/whitepaper-the-evolution-of-googles-search-results-pages-effects-on-user-behaviour/.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *The Psychological Review, 63,* 81–97.

Munzinger, U., & Musiol, K. G. (2008). *Markenkommunikation: Wie Marken Zielgruppen erreichen und Begehren auslösen*. Landsberg/Lech: Verlag Moderne Industrie.

Neue Verpackung. (2002). Trend zum Impulskauf. Heidelberg. Abgerufen am 2. 9 2015 von http://www.neue-verpackung.de/wp-content/uploads/migrated/docs/173_6588.pdf.

Nixon, H. (1924). Attention and interest in advertising. *Archives of Psychology, 72,* 5–67.

Perrachione, T. K., & Perrachione, J. R. (2008). Brains and brands: Developing mutually informative research in neuroscience and marketing. *Journal of Consumer Behaviour, 7*(4–5), 303–318.

Petty, R. E., & Cacioppo, J. T. (1986). The elaboration likelihood model of persuasion. In *Advances in experimental social psychology*, pp. 123–193.

Plassmann, H., Ramsoy, T., & Milosavljevic, M. (2012). Branding the brain: A critical review and outlook. *Journal of Consumer Psychology, 22*(1), 18–36.

Pradeep, A. K. (2010). *The buying brain: Secrets for selling to the subconscious mind*. Hoboken, NJ: Wiley.

RealVu. (2010). *RealVu revolutionizes internet advertising measurement with the viewable impression*. Abgerufen am 2. 9 2015 von http://www.marketwired.com/press-release/realvu-revolutionizes-internet-advertising-measurement-with-the-viewable-impression-1271884.htm.

Rothensee, M. (2015). Brand Experience messen und steuern. Quantitative Verfahren im Überblick. In A. Baetzgen (Ed.), *Brand experience: An jedem Touchpoint auf den Punkt begeistern* (pp. 40–56). Stuttgart: Schaeffer Poeschel.

Satel, S., & Lilienfeld, S. (2013). *Brainwashed: the seductive appeal of mindless neuroscience*. Jackson: Basic Books.

Scheier, C., & Held, D. (2012). *Wie Werbung wirkt. Erkenntnisse aus dem Neuromarketing*. Freiburg: Haufe.

Schiessl, M. (2012). *Foodregio.* Abgerufen am 2. 9 2015 von http://www.foodregio.de/file/vortrag-michael-schiessl-foodregiomarketingtag2012.pdf.

Schiessl, M. (2013). Nutzer schauen nur Millisekunden hin. *Horizont* (32). Abgerufen am 2. 9 2015 von http://www.horizont.net/marketing/nachrichten/Onlinewerbung-Nutzer-schauen-nur-Millisekunden-hin-116690.

Schmidt, S., Wiedmann, K.-P., Reiter, P., & Kurlbaum, C. (2016). Anticipating the effects of marketing communication: a neuroeconomic framework for marketing purposes. In T. Langner, F.-R. Esch, & M. Bruhn (Eds.), *Handbuch Sozialtechniken der Kommunikation*. Wiesbaden: Springer Fachmedien.

Serafin, L. (2013, December 18). *Google's move to viewable impression based pricing*. Abgerufen am 2. 9 2015 von http://www.gearylsf.com/googles-move-to-viewable-impression-based-pricing/.

SevenOne Media. (2014). Navigator Mediennutzung 2012. München. Abgerufen am 2. 9 2015 von http://viewer.zmags.com/publication/b9c9d2fc.

Sewell, W., & Komogortsev, O. (2010, April 10–15). Real-time eye gaze tracking with an unmodified commodity webcam employing a neural network. In *CHI'10 Extended abstracts on human factors in computing systems*, pp. 3739–3744.

Simon, H. (1959). Theories of decision making in economics and behavioural science. *American Economic Review*, 253–283.

Sorensen, H. (2009). *Inside the mind of the shopper*. Hudson, NY: Pearson Education.

Springer Gabler Verlag. (2016). Gabler Wirtschaftslexikon. online: Springer Gabler Verlag. Von http://wirtschaftslexikon.gabler.de/Archiv/1286/marketing-v9.html abgerufen.

Statista. (2016, December 2). *Statista*. Von https://de.statista.com/statistik/daten/studie/271067/umfrage/betrachtungsdauer-von-werbung-in-verschiedenen-medien/ abgerufen.

Ströer. (2016, november 18). *ATLAS—Das Out-of-Home Wirkungsbarometer*. Von ATLAS—Das Out-of-Home Wirkungsbarometer: http://www.stroeer.de/wissen-inspiration/knowledge-base/wirkstudien/atlas.html abgerufen.

Tatler. (2014). *The active eye: Perspectives on eye movement research.* In M. Horsley, M. Eliot, B. Knight, & R. Reilly (Eds.), *Current trends in eye tracking research* (pp. 3–16). Berlin Heidelberg: Springer.

Teichmann, J. (2014). Werbewirkung bei Parallelnutzung. München. Abgerufen am 2. 9 2015 von https://www.sevenonemedia.de/c/document_library/get_file?uuid=d01f4ac3-81d7-4b1c-bd29-6a28d6aa72b9&groupId=10143.

TNS. (2013, August 27). TNS CONVERGENCE MONITOR: TV und Internet profitieren voneinander. Bielefeld. Abgerufen am 2. 9 2015 von https://www.tns-infratest.com/presse/presseinformation.asp?prID=3253.

Tooby, J., & Cosmides, L. (2015). The theoretical foundations of evolutionary psychology. In D. M. Buss (Ed.), *The handbook of evolutionary psychology. Volume 1. Foundations* (pp. 3–87). Hoboken, NJ: Wiley.

United Internet Media. (2014). *A peek into the living room*. Karlsruhe. Abgerufen am 2. 9 2015 von http://united-internet-media.de/en/studies/category-studies/a-peek-into-the-living-room/.

Universal McCann. (2014). *Wave 7—cracking the social code*. Abgerufen am 2. 9 2015 von http://wave.umww.com/assets/pdf/wave_7-cracking-the-social-code.pdf.

van der Lans, R., Pieters, R., & Wedel, M. (2008). Competitive brand salience. *Marketing Science, 27*(5), 922–931.

Visible Logic. (2009, February 25). *Good design vs. good branding: Tropicana case study*. Abgerufen am 2. 9 2015 von http://www.visiblelogic.com/blog/2009/02/good-design-vs-good-branding-tropicana-case-study/.

Wedel, M., & Pieters, R. (2000). Eye fixations on advertisements and memory for brands: A model and findings. *Marketing Science, 19*(4), 297–312.

Wedel, M., & Pieters, R. (2008a). Eye tracking for visual marketing. *Foundations and Trends in Marketing, 1*(4), 231–320.

Wedel, M., & Pieters, R. (2008b). *Visual marketing.* Taylor & Francis Group, LLC.

Wiedmann K. P., Lischka G., & Schiessl M. (2017). Applied neuroeconomics: Science meets business practice—profound insights or witchcraft? (Abstract). In M. Stieler (Eds.), Creating marketing magic and innovative future marketing trends. Developments in marketing science: Proceedings of the academy of marketing science. Cham: Springer.

Wilson, T. D. (2002). *Strangers to ourselves: Discovering the adaptive unconscious*. Cambridge, MA: Belknap Press of Harvard University Press.

Wolfe, J., & Horowitz, T. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5,* 495–501.

Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

Zajonc, R. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology, 2,* 1–27.

Zaltman, G. (2003). *How customers think*. Boston: Harvard Business School Press.

Zurawicki, L. (2010). *Neuromarketing, exploring the brain of the consumer* (pp. 42–53). Berlin Heidelberg: Springer Verlag.

Chapter 20
Neuroeconomics

Susann Fiedler, Ulrich Ettinger and Bernd Weber

Contents

Abstract In this chapter, we will describe how the analysis of eye movements can be used in the study of the cognitive and neural mechanisms underlying economic decision making. We will first introduce the area of neuroeconomics by describing relevant theories and methods. We will then demonstrate how oculographic methods have been used profitably in this field. While initially, considerable effort and energy had to be applied to find a common ground and language for the different disciplines involved in neuroeconomics, in recent years, testable hypotheses and models have developed within a common framework. These models have led to intriguing insights into the cognitive and neural mechanisms leading up to economic decisions. We

---

S. Fiedler (✉)
Max Planck Institute for Research on Collective Goods, Bonn, Germany
e-mail: fiedler@coll.mpg.de

U. Ettinger
Department of Psychology, University of Bonn, Bonn, Germany

B. Weber
Institute of Experimental Epileptology and Cognition Research, University Hospital
Bonn, Bonn, Germany

Centre for Economics and Neuroscience, University of Bonn, Bonn, Germany

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_20

conclude that the field of neuroeconomics is a flourishing and fruitful endeavour and that especially the application of knowledge about the oculomotor system can make important contributions.

20.1 Introduction and Learning Objectives

Which coffee brand to buy? Which university to apply for? Buy risky assets or rather make a deposit in the savings account? We make some decisions automatically and without much cognitive effort, whereas others take time and careful deliberation. While we can observe the final choice of each of these decisions, the underlying thought processes are unknown to the observer. Understanding human decision making is an endeavour that has been approached by a variety of disciplines, including psychology, economics and neuroscience (Newell & Bröder, 2008). Improved knowledge about the processes underlying choices will not only allow for better models to predict behaviour, but also to create interventions which may help people in the decision process. These implications range from health-related decisions, like food choices, to financial decisions, like retirement savings.

The eye has been argued to offer privileged access to the study of decision making and, more generally, the mind, i.e. to human and animal cognition, motivation, emotion, ability and personality as well as psychopathology (Stern & Dunham, 1990). The eye offers a wealth of information about ongoing cognitive and emotional processes, first and foremost about where the person is directing her overt visual attention, i.e. how people are searching for information in a decision context. Additionally, the measurement of eye blinks as well as the characterisation of the more subtle features of the eye, such as the size of the pupil, can be used to make inferences about state variations in the viewer's arousal levels as well as demands on information processing, to name a few.

In this chapter, we will describe how recordings of the eye and its movements can be used in the study of the cognitive and neural mechanisms of economic decision making. We will first introduce the area of neuroeconomics generally, and then demonstrate how oculomotor methods in particular have been used successfully in this field.

Learning Objectives

- In this chapter, you will acquaint yourself with important theories and paradigms of neuroeconomics.
- This knowledge will enable you to gain an understanding of the principles underlying current empirical studies in this field, especially those employing oculographic methods. To achieve this objective, the chapter draws upon knowledge from basic oculomotor work (see relevant chapters of this textbook).
- Having read this chapter, you will be able to understand the role that the study of eye movements has played in the accumulation of empirical evidence and theory testing in neuroeconomics, both in human and in animal work.

20.2 Introduction to Neuroeconomics: Historical Annotations and Current State

Contemporary economics has contributed fundamentally to our understanding of decision making by analysing how costs and benefits connected to a particular choice are weighted and opportunity costs of decisions are identified. Both from a rational choice as well as a behavioural economics perspective, theories have been put forward to explain and predict economic decision making and inform policies and institutions. By drawing on insights gained from analysing final choices only, most motives and processes at play stay hidden and the application of new methods to record and understand the underlying mechanisms of choice seem warranted. However, the assessment of cognitive processes in economic decision making is a complex challenge (Newell & Bröder, 2008). Directly asking people why they behave in specific ways is often not insightful. The answers may not be veridical or valid and might be distorted by social desirability, self-deception, or demand effects, i.e. the participants replying in a way they expect the experimenter wants them to reply. More importantly, individuals often simply lack adequate introspection into their mental processes, driving forces, and decision criteria involved in their judgments and decisions (Nisbett & Wilson, 1977). In order to understand the mechanisms guiding economic choices, researchers commonly use experimental manipulations to test their effects on choice behaviour and use data to make inferences about the unobservable. This methodological approach has been extended since the foundation of neuroeconomics. Insights from neuroscience have been utilized in this interdisciplinary scientific endeavour that aims to understand human decision making, particularly in the context of economic choices. The goal of neuroeconomics is to integrate physiological and behavioural data to improve our understanding of the deliberation process and motivation behind a given decision with the ultimate goal of improving models of human choice behaviour (Reuter & Montag, 2016).

Neuroeconomics makes use of multiple methods and provides the possibility to measure the neural correlates of reasoning underlying the decision process, thereby building the foundation for competitive model testing. The methods are drawn from its constituent scientific fields, namely neurobiology, behavioural and experimental economics as well as cognitive, social and evolutionary psychology. Many models of economic decision making provide similar predictions for choice behaviour, but state different assumptions about the motives and processes leading up to a particular choice. Hence, the ability to test these process assumptions of existing theories furthers our understanding by dismissing inaccurate models of the decision making process and allowing the development of more psychologically and biologically plausible models of the process. Therefore, pairing standard experimental paradigms and neuroscience methods has a unique value for understanding the driving factors of behaviour and the effects of interventions. In the following, we will introduce the reader to the set of most frequently used economic paradigms and neuroscientific methods.

20.2.1 Standard Research Paradigms to Capture Economic Decision Making

Economic games aim to model real world decision situations with and without social interactions within the controlled environment of a lab. As such, economic games have provided insights into simple choices and valuation processes, from two- or n-party social interactions to complex market interactions. By creating tightly controlled situations in the lab, control can be gained over situations which are normally influenced by many known and unknown factors in the real world, so that researchers in the lab can separate the various influences from each other and identify their causal effect on economic choices. The targets of such investigations are often factors like risk, ambiguity, time, and social considerations. The games are mostly presented via a computer interface, but are fully incentivized to introduce real world (mostly monetary) consequences of behaviour. In the following, we will briefly discuss a selection of the most commonly used paradigms to elicit individuals' value estimations, preferences and strategic considerations.

Preference Elicitation

For a wide range of economic decision situations, a multitude of information has to be considered in order to form a subjective evaluation that serves as the basis for the subsequent decision. Many of these economic decisions have to be performed every day and consists of simple, **value based choices**, be it food or other product choices. To investigate the cognitive processes underlying simple choices and evaluations, a basic design has been established and used in many different ways. Participants are presented with different products for which their preference may vary. Subsequently, they are asked to evaluate the product and enter their subjective willingness-to-pay or preference for each product (see Fig. 20.1). Here behaviour is merely used as a measure of value, to be correlated with neural processes. This category of experiments has provided important insights into the processes of valuation and comparison of choice options (see a recent review by Bartra, McGuire, & Kable, 2013).

Another very important domain are evaluations of risk. Research in the area of behavioural economics as well as judgment and decision making traditionally has a strong focus on research employing the **risky choice paradigm**. Within this

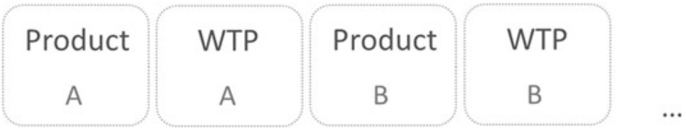


Fig. 20.1 Example of a design to investigate simple, value based choices. In this paradigm, subjects see a series of products, which are followed by a subjective valuation, i.e. they enter their willingness-to-pay (WTP). The WTP is the maximum price the participant would be willing to pay for the product. At the end of the experiment, a product and price are drawn at random and if the price is below the participant's WTP, the purchase is made

	Lottery A	Lottery B
Probability 1	28%	77%
Value 1	€ 46.6	€ 46.6
Probability 2	72%	23%
Value 2	€ 5.6	€ 14.4

Fig. 20.2 Example of the risky choice paradigm: Decision between two gambles (Lottery A&B) with two outcomes ("Values") and their corresponding probabilities [example adapted from Fiedler and Glöckner (2012)]

paradigm, real world risky choices are reduced to simple decisions between gambles (lotteries) with monetary outcomes, assuming that economic risk can be represented by probability differences of potential payoffs. The typical experimental setup is the presentation of the potential payoffs, the respective probabilities of their realization and the request to choose one of the presented alternatives (see Fig. 20.2).

Besides economic decisions that involve no consequences for others, a large strand of literature in behavioural economics focuses on interdependent choices. The simplest experimental paradigm in this context is the **Dictator Game**. Participants are asked to divide a given set of resources (money) between themselves and another, often unknown person. The first published version of this paradigm was used by Kahneman, Knetsch, and Thaler (1986) to refute the then prevailing theory of a purely money maximizing economic actor. Looking at the evidence accumulated since then, Kahneman and colleagues' historical results still hold true. More than 63% of all participants pass some positive amount to the other person within these decision tasks (Engel, 2011). In the past, Dictator Games have been used in a wide set of contexts (e.g. in loss or gain contexts) and forms (e.g. mini dictator games reducing the originally free allocation decision to a decision between two well-defined alternatives varying in the payoffs received by the dictator and the receiver (see e.g., Fig. 20.3).

	Option A	Option B
Own payoff	€ 20	€ 12
Other persons payoff	€ 9.3	€ 2.2

Fig. 20.3 Decision between two outcome distributions [decomposed Dictator Game adapted from Liebrand and McClintock (1988)]

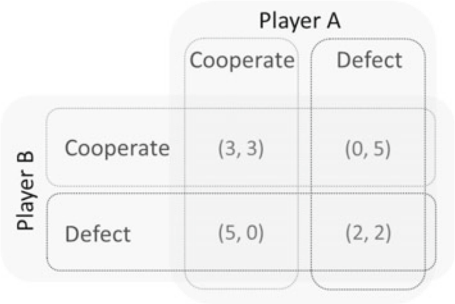


Fig. 20.4 Example of a Prisoner's Dilemma. The first number in brackets refers to Player A's payoff, while the second number refers to Player B's payoff. Both players choose without knowing the action of the other player

Based on the Dictator Game Paradigm, various extensions and alterations have been developed and used to test for example norm-enforcing punishment and compensation [Three-person Game; Fehr and Fischbacher (2003)] or ultimatum bargaining [Ultimatum Game; Güth, Schmittberger, and Schwarze (1982)].

Eliciting Strategic Considerations

One of the most important paradigms to capture strategic components of decision making is the **Prisoner's Dilemma** (Rapoport & Chammah, 1965). The paradigm uses the conflict of interest between the social incentives of cooperation and the individual incentives of defection to recreate real world situations like price-competition, bankruptcy and public good problems in the lab. Within the game two players (A and B) have to decide between cooperation and defection. The incentive structure creates a situation, in which each player can always increase individual payoffs by defecting in a single interaction (see Fig. 20.4), but results from experimental studies show that people vary in their decisions and that joint cooperation sometimes occurs even if the two players interact only once.

The **Public Good paradigm**, structurally very similar to the Prisoner's Dilemma, was developed for n-player situations. All players receive an endowment of €20 (see a schematic example in Fig. 20.5) that they can invest into a common pool, knowing that the money within the pool will be doubled and distributed equally among all players irrespective of their contributions in the game. However, players can also decide to keep the initial endowment and follow their temptation to free-ride on the other players' contributions to the common pool (see for example player 4 in Fig. 20.5).

Many adaptations of the above introduced paradigms as well as other paradigms testing for example time discounting (e.g. intertemporal choice paradigm) or sophistication in reasoning (e.g. beauty contest game) are used in neuroeconomics. To learn about them in more detail, see Binmore (2007) and Camerer (2003).

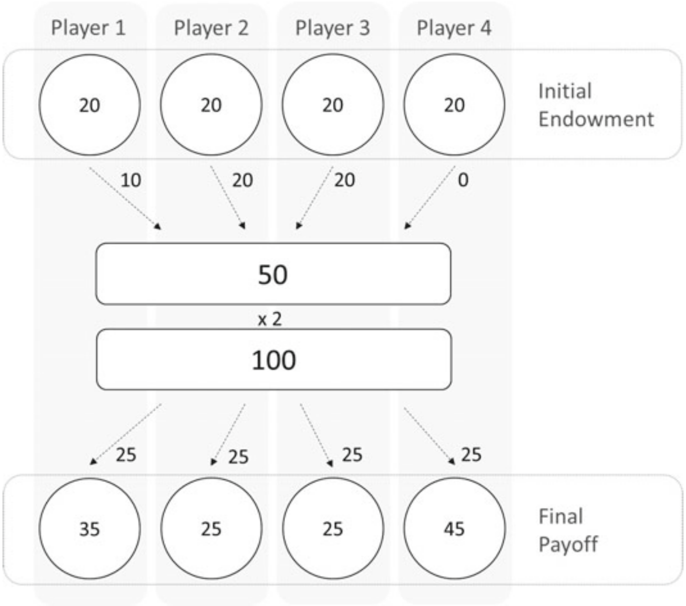


Fig. 20.5 Example of a 4-Person Public Good Game

20.2.2 *From Rational Choice to Process Models of Decision Making*

In order to account for the range of observed decision behaviours, a wealth of models have been developed and tested. In this process, a tremendous amount of empirical results has been presented as evidence for violations of standard economic rationality assumptions, which held that one should always pick the one alternative that has the highest expected utility.[1] Today, it is recognized that individuals vary in their decision patterns. Studies show that these decision patterns are sensitive to relatively small manipulations of context factors (e.g. framing, anchoring, and endowment status), unrelated to the underlying payoff structures, resulted in changes in decision patterns (e.g., Tversky & Kahneman, 1981). Reactions to these violations have been twofold.

On the one hand, extensions and refinements of the original rational choice model have been developed (e.g., Birnbaum, 2008; Kahneman & Tversky, 1979; Tversky & Kahneman, 1992). These new models integrated the multitude of empirically established behavioural effects. A large strand of new outcome-based and intention-based theories have been developed in the context of independent and interdependent decision making (e.g., Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999). In contrast

---

[1] The expected value model is a model integrating values and their corresponding probabilities of their materialization by multiplication. As an example: the expected value of a 50:50 gamble for €1 is 0.5 x €1 = €0.5. Rationality requires people to assign subjective values (i.e. “utilities”) to outcomes, which they then try to maximize.

to the standard assumption of classic economics viewing the individual as purely selfish, money-maximizing agent social preferences models assume that many individuals do care about consequences (e.g., material outcomes) for others (see e.g., Balliet, Parks, & Joireman, 2009; Fehr & Schmidt, 1999; Van Lange, 1999). As individuals are heterogeneous, they diverge in their social preferences (e.g., cooperators, competitors, individualists, or altruists). Social preference models aim to integrate behavioural assumptions to improve psychological realism of the predictions for situations involving others, but still lack predictions about the cognitive processes involved. Since the underlying processes are not accounted for in these theories, statements about them remain at an "as-if" level, meaning that theories of this class state that people behave "as if" they are computing some kind of expected utility.

On the other hand, there is a new line of work aimed at developing process models of choice. These models focus not solely on the relation between input and output, but rather on the decision process itself by describing how a decision occurs from an initial state of information gathering through a sequence of evaluations to a particular decision. Two important subsets of models have been identified: (1) heuristics, assuming that decision makers avoid integrating all information and decide via simple rules of thumb (e.g., Simon, 1955) and (2) automatic integration models (e.g., Busemeyer & Townsend, 1993; Glöckner & Betsch, 2008), which strive to understand and predict the cognitive processes leading up to a choice as well as the choice itself. According to one main representative subset of this model class, the Evidence Accumulation models, which include for example the Decision Field Theory (Busemeyer & Townsend, 1993), decision processes are based on a stochastic sampling process of available outcomes during every given moment of the decision. For each potential outcome, an affective reaction is generated and accumulated over time within a noisy process until a given threshold for the decision is reached (see a schematic presentation of the sampling process in Fig. 20.6). While only the potential outcomes are accumulated, the probabilities of their realization are the driving forces for the sampling process. Another important specification of these kinds of models is the speed of accumulation (drift rate) which represents the quality of the sampled evidence (Liu & Pleskac, 2011; Milosavljevic, Malmaud, Huth, Koch, & Rangel, 2010; Ratcliff, 1978). These models aim at understanding the underlying internal processes of decision making that cannot be inferred from basic analyses of reaction time and choice data.

Box 1

An influential framework to investigate value-based decisions has been proposed by Rangel, Camerer, and Montague in (2008). In their neural network model, they divide the process of value-based choices into separate stages: First, a representation of the decision problem is created, which incorporates important internal and external states, like hunger and the environment as well as the available choice options. In a second step, values are assigned to the identified actions. In a third step, these identified action values are compared

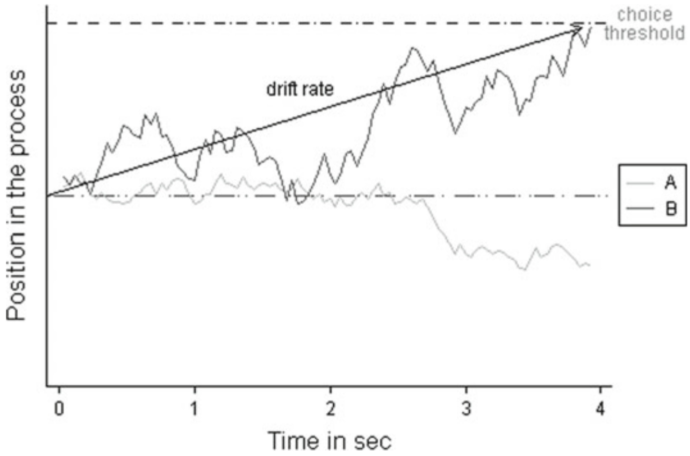


Fig. 20.6 Schematic representation of the evidence accumulation process for 2 presented options ("A" and "B") as proposed by decision field theory

and an action is selected. In a fourth step, the actual outcome is evaluated and finally compared to the previous expectations, so that learning can take place and expectations for future similar choices can be updated. Even though, as in any scientific domain, there still is some controversy about this model, there is also strong consensus about the neural networks involved in value-based choices. Hence, as a recent review by Bartra et al. (2013) shows, the proposed framework has been used successfully in various studies.

20.2.3 Methods and Directions in Neuroeconomics

In contrast to purely "choice-observing studies", neuroeconomic studies additionally include oculography, physiological measures, such as heart rate and skin conductance, electroencephalography (EEG), magnetoencephalography (MEG), positron emission tomography (PET), functional magnetic resonance imaging (fMRI), molecular genetics or pharmacological manipulations and even single unit recordings in humans. In addition, animal studies have been performed in different species, involving single-neuron recordings.

The role of oculography in neuroeconomic research was first uncovered in the seminal paper by Platt and Glimcher in the journal Nature in 1999, in which they presented compelling evidence that neurons in the parietal cortex of monkeys coded important decision variables, such as the expected value of choice options. In this study, the authors studied eye movements in monkeys to investigate choice behaviour.

Since that time, hundreds of papers have been published investigating various aspects of the neural basis of decision making in humans and other species. Bringing together researchers with different backgrounds from neuroscience to economics is a profound strength of this growing research field. This approach is corroborated by a recent meta-analysis that stresses the importance of different scientific backgrounds being involved in a research group as the variety of scientific backgrounds increases the group's impact on the scientific community (Levallois, Clithero, Wouters, Smidts, & Huettel, 2012).

Besides simple value-based choices, several studies have investigated the neural basis of social influences on human decision making (see also relevant chapters for a review in Neuroeconomics: Decision Making & the Brain, 2013), investigating phenomena like altruism (Harbaugh, Mayr, & Burghart, 2007), social comparison processes (Fliessbach et al., 2007), reciprocity (Sanfey, Rilling, Aronson, Nystrom, & Cohen, 2003) or trust (King-Casas et al., 2005). These studies are much more complex compared to simple value-based choices, but also develop a converging view on the integration of social norms into subjective valuations. What these studies nicely showed is how the brain integrates the social context in choice, allowing the better understanding of norm-defections but also how pathological brain changes may lead to disturbances in social interactions.

Recently, neuroeconomic studies have aimed at answering more applied questions, like understanding how people make food choices in order to identify factors leading to obesity. Their purpose is to better understand the involved computational processes, which may improve interventions (Rangel, 2013). Working towards the same goal within a quite different context, two studies on risky decision making in the financial market showed how the brain's valuation regions track the values of stocks in a market situation and how they react during the formation of market bubbles (De Martino, Fleming, Garrett, & Dolan, 2013; Smith, Lohrenz, King, Montague, & Camerer, 2014). Even though this more applied research is still rather young, it may help to design new interventions and provide ideas for more efficient regulations.

Despite these applications of neuroeconomic methodologies to interesting and important economic and social phenomena, there are only few models of decision making that integrate knowledge stemming from the neurosciences. Webb and colleagues recently presented a theoretical paper and empirical evidence on a model integrating biological constraints on neural information processing, i.e. divisive normalization (Webb, Glimcher, & Louie, 2014). In that paper, the authors showed that the limited capacity of neurons to code information requires neurons to adapt their firing rate to a choice set. Their model predicts and explains different disturbances in simple choice behaviour like the influence of irrelevant alternatives on choices. This is one of the first—if not the first—study trying to integrate basic neurobiological knowledge into models of individual choice. This new perspective on the questions of how choices emerge and how context influences the resulting behaviour highlights the need for a new framework of thinking. Of particular value will be the integration of the assumption that the brain codes information in the most *efficient* way, given its limited processing resources.

Neuroeconomics has proven over the—albeit small number of—years of its existence to be a very fruitful endeavour. While in the beginning a lot of effort and energy had to be applied to find a common ground and language for the different disciplines involved, in recent years, testable models and hypotheses have developed within a common framework. Now that a firm ground has been developed, the studies become more applied, addressing important phenomena and providing even first models integrating biological insights into models of human decision making.

20.3 What Can Oculography Tell Us About Economic Decision Making?

Since its beginning, economics has accommodated a wealth of theories and now presents itself with a history of careful empirical testing. Traditional theories often contain "as-if" assumptions about underlying processes by proposing that people behave as if they are maximizing their utility. Specifically, that means that choices are modelled as if the decision maker transforms each objective payoff into a subjective value and then selects the alternative with the highest subjective value. These theoretical prepositions are a result of the fact that for the longest time, empirical studies in the field of economic decision making were based on observable behaviour only. Often only final choices were observed and the underlying decision making process was reconstructed based on the recorded choice patterns. With the emergence of neuroeconomics and, in particular, the utilization of oculography, the information search and integration process became observable and opened up a new level of understanding in economics. Methodological developments of oculography (detailed in the chapter by Hutton in this volume) as well as the advanced understanding of the association between attention processes and eye movements and their neural mechanisms (see chapter by Pierce et al. for saccades and chapter by Alexander & Martinez-Conde for fixations) have provided a fruitful basis for a more thorough investigation of underlying mechanisms of decisions. Making the transformation process visible through oculography is a new layer and an important step to better understand economic behaviour and develop psychological plausible models of economic choice.

Oculographic recordings in the study of economic decision making were first made in the 1970s (Russo & Rosen, 1975) and have since been used profitably to elucidate the underlying processes of economic decision. After a productive phase of early work, oculography was in decline in decision research for many years (but see e.g., Lohse & Johnson, 1996, and Russo & Leclerc, 1994, for exceptions). Only recently, several psychologists and economists have taken up the challenge again to use oculography for investigating decision processes (e.g., Ashby, Dickert, & Glöckner, 2012; Innocenti, Rufa, & Semmoloni, 2010; Krajbich & Rangel, 2011; Reutskaja, Nagel, Camerer, & Rangel, 2011).

A decisive advantage of oculography is its unobtrusive nature. Unlike previously employed methods (e.g. where participants are forced to open one piece of information at a time, or to describe what they are currently thinking) that reduce the possible inferences that could be drawn from the collected data, oculography does not alter the information search by forcing participants into a serial information search (see for a discussion Glöckner & Betsch, 2008). Instead, oculography quietly records the gaze of decision makers who engage in their natural decision processes. As a consequence, it yields more realistic measures of decision formation, and with that more valid inferences on decision processes. Oculography cannot only shed light on the importance of the various kinds of information retrieved in a decision making process, but also on the underlying arousal state and cognitive load under which a decision is made, especially when inner features such as pupil size are additionally considered (see also the chapter by Laeng & Alnaes in this volume). The recording of fixations and saccades allows for the investigation of attention and information search in a very fine-grained manner at a high temporal and spatial resolution.

The basic assumption underlying all oculography experiments is that people tend to look at the information that they are currently processing. Hence, one way to investigate the computational processes involved in decision making is to record the information attended leading up to a decision point. Attention allocated to particular outcomes can be interpreted as indicators of the corresponding value weights in the underlying utility function (e.g. Fiedler & Glöckner, 2012; Fiedler, Glöckner, Nicklisch, & Dickert, 2013). Also, the processes underlying a subjective evaluation of alternatives can be further described. This is particularly interesting in the context of preference formation. Multiple models within the context of economic choice make different process assumptions, but provide similar predictions. Oculography can help to discriminate and comprehensively test these theories by recording the temporal dynamics and distribution of attention during each decision making process.

Process Models Utilizing Process Measures

Economic research has largely progressed through the use of models that describe and predict observable behaviour. In the past this often meant models based on choice outcomes only, but great progress has been made in developing new models taking a process perspective on decision making. One of the most recent and commonly used models is the **Attentional Drift Diffusion Model**. Outlined by Krajbich, Armel, and Rangel (2010) it builds on evidence accumulation models, more specifically, the drift diffusion model, by including eye-tracking data (see also Krajbich & Smith, 2015). In the standard drift diffusion model, one assumes that the accumulation speed is driven by the value difference between the two stimuli displayed, and that decisions are made when a certain decision threshold is reached. Krajbich and colleagues extended the model by including fixations, which they hypothesized are involved in the value integration process. More specifically, they assumed that the slope of the accumulation process (i.e., the drift rate) is biased toward the item that is fixated. This in turn leads to a choice bias for the more fixated item. As the authors point out, their assumptions about visual attention are quite simplistic, but increase the predictive and descriptive power relative to models dismissing gaze behaviour as input.

Another model connecting the evidence from neurobiology and game theory is the **Economic Model of Brain Processes**, in which Brocas and Carrillo (2012, also see Brocas, 2012) developed an evidence accumulation account for strategic decision making. They proposed a decision threshold model, assuming that the sensory system encodes information by means of cell-firing. The level of cell-firing depends stochastically on the information obtained. Information is thereby accumulated until sufficient information supporting one voluntary action is gathered and a certain threshold is reached. Within the model, neuronal activity is driven by changes in payoffs and beliefs, since it is assumed that neurons automatically compute the approximate expected values of the presented alternatives and update their priors through Bayesian inferences conditional on the accumulated information.

The above introduced process models are starting points of a more general development into the direction of descriptive models utilizing new methodological innovations and the need for insights into underlying mechanisms of decision making.

20.4 Illustrative Examples: Behavioural Studies

In the following, we will provide a short overview of published work in neuroeconomics that has used oculographic methods. We will give selected examples to illustrate the richness of oculographic approaches and the way this research answered some of the fundamental questions in neuroeconomics. Many of the experiments employed the economic paradigms introduced earlier in the chapter and the remaining part is therefore arranged into two subsections on (1) preference elicitation and (2) elicitation of strategic considerations. Interested readers are invited to additionally read the overview article on individual decision making and attention by Orquin and Mueller Loose (2013) as well as to read further relevant articles (e.g., Costa-Gomes, Crawford, & Broseta, 2001; Glöckner & Herbold, 2011; Lohse & Johnson, 1996; Middlebrooks & Sommer, 2011; Stritzke, Trommershäuser, & Gegenfurtner, 2009) for a more comprehensive overview.

Oculographic experiments in the area of neuroeconomics target a wide range of research questions. The literature contributes to our understanding about specific effects as well as complex strategy use. Research aims to understand the *influence of task and personality factors* that alter choice behaviour (e.g. through framing, complexity or differences in individuals' social preferences see Fiedler et al., 2013; Fiedler & Glöckner, 2012; Glöckner & Herbold, 2011; Krajbich et al., 2010; Krajbich, Lu, Camerer, & Rangel, 2012; Krajbich & Rangel, 2011). A further goal is to identify properties of the dynamic search algorithm of decision makers when instructions and context conditions aim to *manipulate the decision process* itself (e.g., through time pressure or motivation, see Ettlin & Bröder, 2015; Milosavljevic, Navalpakkam, Koch, & Rangel, 2012; Pieters & Warlop, 1999; Reutskaja et al., 2011). Additionally, oculography has recently been used in economic experiments to distinguish between different possible decision processes leading to otherwise similar behavioural results

(Arieli, Ben-Ami, & Rubinstein, 2011; Knoepfle, Tao-yi Wang, & Camerer, 2009; Reutskaja et al., 2011; Wang, Spezio, & Camerer, 2010).

20.4.1 Preference Elicitation

While a number of companies have already postulated for quite some time that oculographic methods might help in predicting consumer behaviour, only in recent years, **value based choices** have also been more intensively investigated in the academic domain. Oculography allows insights into the mechanisms involved when consumers integrate different attributes of products (like brands or labels) into the choice process. For example, Krajbich and colleagues (2012) investigated simple purchasing decisions using oculography. First, they used a simple Becker-DeGroot-Marschak auction to reveal the subjective values of different products for each participant. Within this procedure, participants make a bid for a certain product, which is then compared to a price determined by a random draw. In case the participants bid is higher than the drawn price, the participant pays the price and receives the product. In the event of the bid being lower than the drawn price the participant will not receive the product, but will also not pay any price. In the second part of the study, participants were presented with each item and a corresponding price. This part was performed in front of an eye-tracker and fixations on the prices and the products were measured. The authors show how the attentional drift-diffusion model extends to purchasing decisions involving products and prices (Krajbich et al., 2012).

In one other recent example, Arieli et al. (2011) recorded eye movements in a **Risky Choice task** where participants chose between two lottery scenarios with differing potential outcomes and their respective probabilities. They aimed to compare two possible strategies in the decision making process, namely the holistic and the component strategy. The holistic procedure assumes that participants would treat the two alternatives as a whole, for example by evaluating the certainty equivalent of each alternative and selecting the one with the higher certainty equivalence, or by computing the expectation of each of the two alternatives and selecting the one with the higher expectation. In the component procedure, it is assumed that participants would compare the outcomes and the probabilities as separate components. Using eye movement recordings, Arieli et al. (2011) found that more difficult problems were associated with saccades between the same components of the two alternatives, suggesting a component procedure, whereas in easier problems, saccades between elements within the same alternative were more frequent, suggesting a holistic processing approach.

Similarly, Fiedler and Glöckner (2012) used oculography to understand decision making while participants decided between two lotteries (see Fig. 20.7 for an example of the observed gaze behaviour). Even though on average 93% of the information was visually attended to, the authors find brief decision times as well as short to medium-long fixations. The extent of information search was influenced by the similarity of the presented options and the stakes that were at play. The authors concluded that

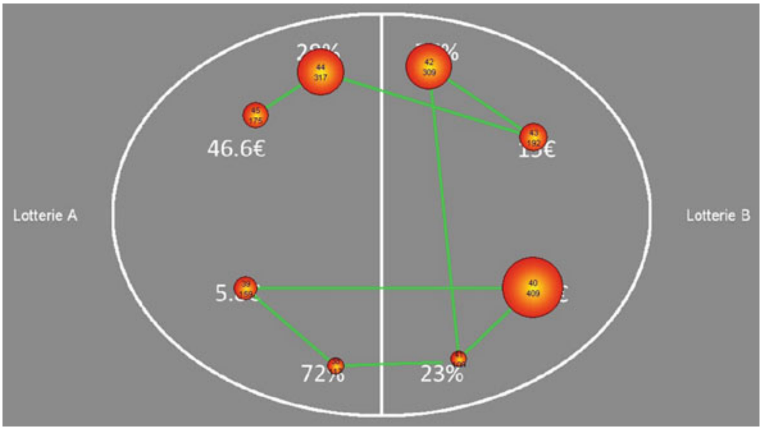


Fig. 20.7 Example taken from Fiedler and Glöckner (2012). Participants have to choose one of two lotteries (“A” on the left and “B” on the right) consisting of two possible payoffs and the respective probabilities of their realization which later on was played out and determined the payoffs of the experiment, see also Fig. 20.1. The orange dots illustrate the last sequence of fixations whereby the size of the individual dot represents the length of the particular fixation

process models postulating evidence accumulation and coherence construction seem to better account for risky choices, compared to simple heuristic models and models assuming complex deliberate calculation.

Together, these studies demonstrate how the investigation of eye movements can provide valuable information on the likely cognitive processes taking place during decision making involving risky prospects.

Oculography has also contributed significantly to our understanding of decision making and economic behaviour in the context of social preferences. First studies presented evidence that the selective distribution of attention serves as an indicator of the decision strategies applied in the context of decomposed **Dictator Games** (Fiedler et al., 2013; Funaki, Jiang, & Potters, 2010). These studies showed that differences in social preferences were accompanied by consistent differences in information search and processing. Most importantly, by using combinations of experimental manipulations, oculography and computational modelling, they were able to directly test process theories of judgment and decision making. The data by Fiedler et al. for example indicated that in the Dictator Game neither the common assumption of individuals being of certain types (e.g., cooperators, competitors, altruists) and using corresponding stable strategies nor the assumption of persons stochastically mixing between distinct strategies receive empirical support. Rather, evidence accumulation models seemed to account best for the overall pattern of data. Furthermore, inter-individual differences in social values/preferences (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999; McClintock, 1972; Van Lange, 1999) seemed to be reflected in consistent and time-stable differences in cognitive processes underlying decision making in other-regarding choice.

20.4.2 Eliciting Strategic Considerations

Eye-tracking studies of information acquisition can also inform our understanding of complex strategic behaviour. Traditional game theory yields elaborate descriptions about how we should behave when our outcomes do not only depend on our own decisions but are additionally dependent on the choices of other people and vice versa. Evidence gathered in the lab as well as observations of our daily lives show that the prediction of purely money-maximizing behaviour made by this traditional approach is often not observed in the collected empirical data. Hence, more recently theories have been developed, which aim to provide a description of the decisions people actually make in strategic settings in order to explain why people sometimes deviate from the rational predictions of game theory (see Camerer, Ho, & Chong, 2004; Costa-Gomes & Crawford, 2006; Nagel, 1995; Stahl & Wilson, 1995).

Camerer and colleagues (e.g., Bhatt & Camerer, 2005; Johnson et al., 2002) have used oculographic methods repeatedly to distinguish between different neuroeconomic theories of strategic sophistication that make similar predictions at the choice level. As these theories cannot be distinguished purely in terms of the participants' behavioural decisions, the recording of eye movements provided a necessary level of analysis that helped ruling out alternative explanations. Since people re-fixate rather than remember payoffs, gaze behaviour was used repeatedly in the past to investigate strategic reasoning (Costa-Gomes et al., 2001). The number of fixations has thereby been used as an indicator of attention and possibly inattention throughout multiple studies (Costa-Gomes et al., 2001; Devetag, Di Guida, & Polonio, 2016; Knoepfle et al., 2009) and helped to disentangle sets of process assumptions. The collective evidence shows that individuals rarely engage in backward induction, which is one of the core underlying mechanisms to solve strategic games assumed in many theories.

In line with this observation, using the **Prisoner's Dilemma** game, Stewart et al. (2015) showed that the observed information search is not compatible with level-k or cognitive hierarchy models. Instead, they show that the processing data is well described by evidence accumulation models. They presented evidence that the patterns of eye-movements are stable throughout the decision making process with the exception that participants present an attention bias towards the later chosen option. The results presented first evidence that individuals may search for information in a sensible fashion, but their information integration process follows very simple rules. Specifically, they suggested that with every fixation on a specific alternative (i.e., cooperation or defection) the likelihood of this action to be finally taken is increasing.

Also in the context of a **sender-receiver game**, a paradigm developed to focus on trust as one specific aspect of strategic interactions, a similar misfit between k-level models and recorded eye-movements has been reported. In this paradigm two players are matched and a privately informed player (sender) sends a message to another player (receiver) who will then take a payoff-relevant action. In the experiment, the sender was incentivized to deceive the receiver about his best action. Wang et al. (2010) showed that before sending their message senders attended to the payoff information of the true state of the world more than to the payoffs connected to the information that they were sending. This pattern provides evidence that the senders fail to take the perspective of the receivers who have no knowledge of the true state of the world when making their decision. Put differently, senders appear to not think strategically enough, but rather present a truth bias (see also Bhatt & Camerer, 2005) and overweigh their own payoffs in comparison to their matched partner. Using both senders' messages as well as their attention allocation towards the potential payoffs the authors show that the true state can be predicted successfully and would increase receivers' payoffs substantially. Wang and his colleagues also utilized the measurement of pupil size over the course of the decision and present evidence that the senders' pupil dilation increased right before as well as after the message was sent to the receiver. Recorded arousal was thereby particularly higher when the magnitude of the lie was large. Taking this evidence together, much is learned that would have been inaccessible through simple recordings of the final decision. Observing the overweighting of own payoffs, the tendency to allocate most of the attention towards the true state of the world and the experienced arousal gives us insights on the limitations of the underlying strategic reasoning and might be used for predicting truth telling.

Expanding the view from simple two-player interactions the literature also presents first evidence for processing and decision strategies in the context of a **Public Good game**. Fiedler et al. (2013) showed that social preferences are not only predictive in dictator games, but rather guide information search and processing also in more complex strategic consideration. Specifically, individuals with more prosocial preferences present a more thorough information search attending not only reported payoffs for all players but also the contributions made by each player, while individualistic participants focus mainly on the payoff information. These results point to personal as well as situational factors driving attention.

Besides giving great insights into the underlying information search and processing, attention distribution has additionally been claimed to be an important factor influencing decision making. The most robust finding concerning the influential role of attention on final choices is the gaze-cascade effect. First introduced by Shimojo, Simion, Shimojo, and Scheier (2003), it describes the observation that individuals increasingly focus visually on the option they tend to prefer over time. The authors discuss how the gaze bias informs decision making and thereby leads to observable choice biases. Replicating the gaze-cascade effect in two studies within the context of risky choices, Fiedler and Glöckner (2012) presented evidence that attentional biases mainly appear in the last part of the decision process. The data showed that the risky asset, which will be chosen later on, receives more visual attention than the unchosen alternative. Going beyond the mere observation of this attention-choice link recent publications present evidence of a causal link between attention and the subsequent choices by manipulating the location of the last fixation or overall fixation exposure externally. The evidence suggests that the last fixated item is more likely to be chosen (Armel, Beaumel, & Rangel, 2008; Milosavljevic et al., 2012; Pärnamets et al., 2015).

Much work has been done investigating cognitive processes underlying judgment and decision making, addressing individual, inter-individual and strategic decisions. The evidence gathered by oculomotor studies has contributed tremendously to our understanding of the cognitive processes of decision making and has contributed to the development of more fine grained models of decision making, which take the sampling process of information into account. Yet, compared with the established psychological knowledge about cognition, these approaches are still simplistic. Processes involved in decision making are a complex interplay of bottom up and top down processes. Since most studies are designed to investigate natural information search and processing strategies, experiments have to be carefully designed to avoid unwarranted conclusions about top down or bottom up processes that are merely induced by specific instructions and saliency of particular information. Overall, the research reviewed here has clearly shown the additional value of studying oculomotor data such as saccades, fixations and measures of pupil size.

20.5 Illustrative Examples: Neurophysiological Studies

Eye movements have been studied in the field of neuroeconomics not only to provide an additional level of data to enrich behavioural experiments, but also to shed light on the neural mechanisms of economic decision making (Ettinger & Klein, 2016). In these studies, economic choices are often indicated by saccadic eye movements to specific locations, instead of button presses or verbal responses, which may also be used to indicate decisions.

The value of using eye movements to study the neural mechanisms mediating economic decisions lies at least in part in the already well-characterised neural system that supports eye movements. Knowledge concerning the neural mechanisms of eye movements stems from multiple lines of inquiry, including human lesion and neuroimaging studies (see chapters by Pierce et al. for saccades and Alexander & Martinez-Conde for fixations) as well as intracerebral recording studies in non-human primates (see chapter by Johnston & Everling).

Much of the eye movement work on neuroeconomics has adopted simple saccadic paradigms (see chapter by Pierce et al.) and has been rooted in the broad framework that perceptual decisions are enabled by sensorimotor transformation processes along the fronto-parietal dorsal pathway including the lateral intraparietal area (LIP) in monkeys and the intraparietal sulcus (IPS) in humans (Goodale & Milner, 2013). The dorsal pathway has been described as the “where” pathway or the “vision-for-action” system, enabling the transformation of a visual signal into a motor response (see Fig. 20.8). Of particular importance in relation to saccadic responses is area LIP in monkeys, an area known to play an important role in the neural circuitry supporting saccadic eye movements (Johnston & Everling, 2008).

In the context of neuroeconomics, sensorimotor neurons along this pathway, in particular in posterior parietal cortex, have been argued to represent a decision value. This model thus considers decision making processes to be integrated within the

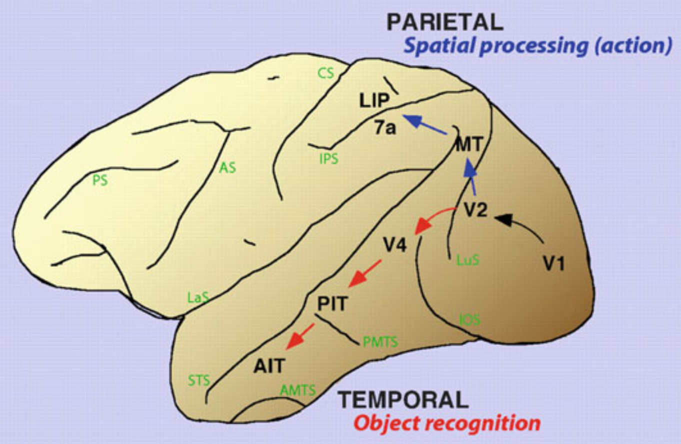


Fig. 20.8 Illustration showing primate area LIP. The figure shows the dorsal (blue) and ventral (red) visual pathways in the macaque brain. LIP: lateral intraparietal area; AIT: anterior inferotemporal cortex; STS: superior temporal sulcus; AMTS: anterior medial temporal sulcus; PMTS: posterior medial temporal sulcus; LuS: lunate sulcus; LaS: lateral sulcus; CS: central sulcus; AS: arcuate sulcus; PS: principal sulcus; MT: medial temporal area. Adapted from Lehky and Sereno (2007)

sensorimotor network (Tosoni et al., 2014). Experiments following this framework and drawing upon our knowledge about the implementation of eye movements in the brain have thereby begun to study the neural mechanisms of economic decision making in both humans and non-human primates.

A seminal study in this line of work was carried out by Platt and Glimcher (1999). These authors observed that when monkeys made saccades to a specific target, activity in LIP neurons early in the trial was modulated both by the magnitude of the expected reward and by its probability. These are, of course, the two bits of information that are needed in order to calculate the expected value of an action to be performed according to expected utility theory (Edwards, 1954; Savage, 1954; von Neumann & Morgenstern, 1944). However, this first finding may not necessarily provide evidence of the mechanisms of economic decision making, as in that experiment monkeys were not given a choice as to where to look. Therefore, in a second experiment monkeys were given the choice which target to make a saccade to. Now, the expected value of the saccade correlated both with the frequency of the performed response and with LIP neuron activity prior to the movement. Therefore, it was concluded that LIP neurons represent the expected value of an action in form of a map of the relative expected subjective utility (Trommershäuser, Glimcher, & Gegenfurtner, 2009).

Completing the process of sensorimotor transformations and decision making, the output of such an LIP map is passed forward to the frontal eye fields (FEF) and the superior colliculus. Neurons in the FEF, in turn, have been shown to display topographic firing over the stimulus position in a “winner-takes-all” fashion that encodes the direction of the saccade; when the level of activity at that location exceeds a certain threshold, a saccade is executed (Hanes & Schall, 1996; Schall & Thompson, 1999).

Other experiments have shown that information on the expected value of a saccade comes from a number of areas in the oculomotor network, e.g. from neurons in caudate, dorsolateral prefrontal cortex, supplementary eye fields, substantia nigra pars reticulata and superior colliculus (McCoy & Platt, 2005). Information on the expected value of a saccade thus appears to be represented throughout the oculomotor network. Such distributed representations are thought to bias saccade selection such that the expected reward is maximised. This neurophysiological evidence is impressive and has been immensely influential on our current thinking about saccadic decision making.

However, there is also evidence that challenges the specificity of these findings. This evidence will briefly be considered in the following paragraphs.

First, there is evidence of value-based modulation of neural activity in areas outside the oculomotor decision network (Lim, O'Doherty, & Rangel, 2011). For example, Serences (2008) observed in a human fMRI study that the value indicated by a cue influenced the blood oxygen level dependent (BOLD) signal in primary visual cortex (V1), even when no saccade was performed.

Second, there are studies showing that LIP/IPS activation in saccadic choice tasks is modulated not only by the expected value but also by other factors. For example, in a memory task in fMRI, the authors (Sestieri, Corbetta, Spadone, Romani, & Shulman, 2014) observed that the BOLD signal in the middle IPS and other cortical areas closely reflected perceived memory strength, with older items associated with greater BOLD signal change compared to newer items. Activation in the IPS has also been found to relate to demands on spatial programming of pro- and antisaccades (Herweg et al., 2014), a finding that likely reflects purely visuo-motor transformations without any influence of economic evaluations. Most critically from a neuroeconomics perspective, however, a recent study of monkeys observed that the firing of LIP neurons was increased not only by reward cues but also by cues signalling punishment (Leathers & Olson, 2012). Leathers and Olson (2012) concluded from these data that LIP neurons respond not specifically to signals of reward but more generally indicate the cues' motivational salience. Their conclusion thus contradicted both the previous hypothesis that LIP neurons represent action value and the view that LIP neurons code value-based decisions [however, see rebuttal by Newsome, Glimcher, Gottlieb, Lee, and Platt (2013)].

To summarise this section, the knowledge gained from neurophysiological, lesion and neuroimaging studies of the non-human and human primate saccadic system has been applied fruitfully in the development of a truly translational model system of economic decision making. Saccadic tasks can be implemented in both monkeys and humans, and data on neural activity can be concurrently collected. Saccadic responses and their underlying physiology can be reliably modulated by the expected value of a decision. A central conclusion from this work is that activation in the oculomotor network, including monkey area LIP and its human homologue IPS, scales with the expected value of the to-be-performed saccade.

20.6 Conclusions

The present chapter has provided an introduction to the study of eye movements in neuroeconomic research. We have presented an overview of key methods and concepts in neuroeconomics and have discussed exemplary evidence from oculographic studies in this field. We conclude that the study of eye movements provides a rich body of behavioural and neuroscientific evidence that can be profitably incorporated into neuroeconomic research. We argue that neuroeconomic research may benefit from the systematic integration of eye movement recordings in diverse research settings, e.g. in the behavioural lab, the MRI scanner and in animal work. Such recordings provide critical data on the process of value computations in the decision process and, more generally, an additional level of analysis with which it may be possible to separate competing accounts of economic decision making. The availability of cross-species models of economic decision making provides a highly innovative model system that builds on animal neurophysiology, behavioural oculomotor research and human neuroeconomics. Work utilizing eye-tracking data provides an important foundation for extending research in the domain of utility theory and neuroeconomic decision making to animal behaviour and neurophysiology.

An important challenge will be to further combine lab and field data to show that including eye-tracking data into models of choice actually improves the prediction about real-world behaviour. While there are some studies that already suggest this with respect to simple value-based choices, more evidence is needed for real-life settings.

20.7 Suggested Readings

Bartra O, McGuire JT, Kable JW (2013). The valuation system: a coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. NeuroImage 76:412–27.

– *A meta-analysis of studies investigating the neural basis of subjective valuation.*

Camerer CF (2003). Behavioral Game Theory: Experiments in Strategic Interaction. Princeton University Press.

– *A great compilation of experimental evidence in games and theoretical contributions that helps the interested reader to understand strategic thinking in depth, drawing from the fields of psychology and economics.*

Glimcher PW, Fehr E (2013). Neuroeconomics: Decision Making and the Brain. Academic Press.

– *A comprehensive textbook providing a basis as well as in-depth knowledge about the field, its topics and current state of knowledge.*

Platt ML, Glimcher PW (1999). Neural correlates of decision variables in parietal cortex. Nature 400(6741):233–8.

- *The seminal paper in neuroeconomics showing that single neurons code important decision variables.*

Rangel A, Camerer C, Montague PR (2008). A framework for studying the neurobiology of value-based decision making. Nat Rev Neurosci 9(7):545–56.

- *Provides a framework for the study of simple value based choices.*

Reuter M, Montag C (2016). Neuroeconomics. Springer Verlag.

- *An edited book providing a comprehensive survey of methods, findings and theories in the field of neuroeconomics.*

20.8 Questions to Students

(i) Why do you think we need neuroscientific methods to understand human decision-making instead of just observing behaviour?
(ii) An argument often raised in neuroeconomics is that an improved understanding of the neuroscience of decision-making will enable more effective interventions. Do you agree?
(iii) Which role does the monkey area LIP play in saccadic decision making?

References

Arieli, A., Ben-Ami, Y., & Rubinstein, A. (2011). Tracking decision makers under uncertaint. *American Economic Journal: Microeconomics, 3*(4), 68–76.

Armel, K. C., Beaumel, A., & Rangel, A. (2008). Biasing simple choices by manipulating relative visual attention. *Judgment and Decision Making, 3*(5), 396–403.

Ashby, N. J. S., Dickert, S., & Glöckner, A. (2012). Focusing on what you own: Biased information uptake due to ownership. *Judgment and Decision Making, 7*(3), 254–267.

Balliet, D., Parks, C., & Joireman, J. (2009). Social value orientation and cooperation in social dilemmas: A meta-analysis. *Group Processes & Intergroup Relations, 12,* 533–547.

Bartra, O., McGuire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *NeuroImage, 76,* 412–427.

Bhatt, M., & Camerer, C. F. (2005). Self-referential thinking and equilibrium as states of mind in games: fMRI evidence. *Games and Economic Behavior, 52*(2), 424–459.

Binmore, K. (2007). *Playing for real: A text on game theory*. Oxford: Oxford University Press.

Birnbaum, M. H. (2008). New paradoxes of risky decision making. *Psychological Review, 115*(2), 463.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American Economic Review, 90*(1), 166–193.

Brocas, I. (2012). Information processing and decision-making: Evidence from the brain sciences and implications for economics. *Journal of Economic Behavior & Organization, 83*(3), 292–310.

Brocas, I., & Carrillo, J. D. (2012). From perception to action: An economic model of brain processes. *Games and Economic Behavior, 75*(1), 81–103.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision field theory: A dynamic-cognitive approach to decision making in an uncertain environment. *Psychological Review, 100*(3), 432.

Camerer, C. (2003). *Behavioral game theory: Experiments in strategic interaction*. Princeton, NJ: Princeton University Press.

Camerer, C. F., Ho, T.-H., & Chong, J.-K. (2004). A cognitive hierarchy model of games. *The Quarterly Journal of Economics, 119*(3), 861–898.

Costa-Gomes, M. A., & Crawford, V. P. (2006). Cognition and behavior in two-person guessing games: An experimental study. *The American Economic Review, 96*(5), 1737–1768.

Costa-Gomes, M., Crawford, V. P., & Broseta, B. (2001). Cognition and behavior in normal-form games: An experimental study. *Econometrica, 69*(5), 1193–1235.

De Martino, B., Fleming, S. M., Garrett, N., & Dolan, R. J. (2013). Confidence in value-based choice. *Nature Neuroscience, 16*(1), 105–110.

Devetag, G., Di Guida, S., & Polonio, L. (2016). An eye-tracking study of feature-based choice in one-shot games. *Experimental Economics, 19*(1), 177–201.

Edwards, W. (1954). The theory of decision making. *Psychological Bulletin, 51,* 380.

Engel, C. (2011). Dictator games: A meta study. *Experimental Economics, 14*(4), 583–610.

Ettinger, U., & Klein, C. (2016). Eye movements. In M. Reuter & C. Montag (Eds.), *Neuroeconomics*. Berlin: Springer.

Ettlin, F., & Bröder, A. (2015). Perceptual grouping does not affect multi-attribute decision making if no processing costs are involved. *Acta Psychologica, 157,* 30–43.

Fehr, E., & Fischbacher, U. (2003). The nature of human altruism. *Nature, 425*(6960), 785–791.

Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition, and cooperation. *Quarterly Journal of Economics, 114*(3), 817–868.

Fiedler, S., & Glöckner, A. (2012). The dynamics of decision making in risky choice: An eye-tracking analysis. *Frontiers in Psychology, 3,* 335. https://doi.org/10.3389/fpsyg.2012.00335.

Fiedler, S., Glöckner, A., Nicklisch, A., & Dickert, S. (2013). Social Value Orientation and information search in social dilemmas: An eye-tracking analysis. *Organizational Behavior and Human Decision Processes, 120*(2), 272–284.

Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., & Falk, A. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science, 318*(5854), 1305–1308.

Funaki, Y., Jiang, T., & Potters, J. (2010). Eye tracking social preferences. Working Paper.

Glöckner, A., & Betsch, T. (2008). Multiple-reason decision making based on automatic processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 34*(5), 1055.

Glöckner, A., & Herbold, A. K. (2011). An eye-tracking study on information processing in risky decisions: Evidence for compensatory strategies based on automatic processes. *Journal of Behavioral Decision Making, 24*(1), 71–98.

Goodale, M., & Milner, D. (2013). *Sight unseen: An exploration of conscious and unconscious vision*. Oxford: Oxford University Press.

Güth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behavior & Organization, 3*(4), 367–388.

Hanes, D. P., & Schall, J. D. (1996). Neural control of voluntary movement initiation. *Science, 274*(5286), 427–430.

Harbaugh, W. T., Mayr, U., & Burghart, D. R. (2007). Neural responses to taxation and voluntary giving reveal motives for charitable donations. *Science, 316*(5831), 1622–1625.

Herweg, N. A., Weber, B., Kasparbauer, A., Meyhöfer, I., Steffens, M., Smyrnis, N., & Ettinger, U. (2014). Functional magnetic resonance imaging of sensorimotor transformations in saccades and antisaccades. *NeuroImage, 102,* 848–860.

Innocenti, A., Rufa, A., & Semmoloni, J. (2010). Overconfident behavior in informational cascades: An eye-tracking study. *Journal of Neuroscience, Psychology, and Economics, 3*(2), 74–82.

Johnston, K., & Everling, S. (2008). Neurophysiology and neuroanatomy of reflexive and voluntary saccades in non-human primates. *Brain and Cognition, 68,* 271–283.

Johnson, E. J., Camerer, C., Sen, S., & Rymon, T. (2002). Detecting failures of backward induction: Monitoring information search in sequential bargaining. *Journal of Economic Theory*, *104*(1), 16–47. https://doi.org/10.1006/jeth.2001.2850.

Kahneman, D., Knetsch, J. L., & Thaler, R. H. (1986). Fairness and the assumptions of economics. *Journal of business*, S285–S300.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society, 47,* 263–291.

King-Casas, B., Tomlin, D., Anen, C., Camerer, C. F., Quartz, S. R., & Montague, P. R. (2005). Getting to know you: Reputation and trust in a two-person economic exchange. *Science, 308*(5718), 78–83.

Knoepfle, D. T., Tao-yi Wang, J., & Camerer, C. F. (2009). Studying learning in games using eye-tracking. *Journal of the European Economic Association, 7*(2–3), 388–398.

Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience, 13*(10), 1292–1298.

Krajbich, I., Lu, D., Camerer, C., & Rangel, A. (2012). The attentional drift-diffusion model extends to simple purchasing decisions. *Frontiers in Psychology, 3,* 193.

Krajbich, I., & Smith, S. M. (2015). Modeling eye movements and response times in consumer choice. *Journal of Agricultural & Food Industrial Organization, 13*(1), 55–72.

Krajbich, I., & Rangel, A. (2011). Multialternative drift-diffusion model predicts the relationship between visual fixations and choice in value-based decisions. *Proceedings of the National Academy of Sciences, 108*(33), 13852–13857.

Leathers, M. L., & Olson, C. R. (2012). In monkeys making value-based decisions, LIP neurons encode cue salience and not action value. *Science, 338*(6103), 132–135.

Lehky, S. R., & Sereno, A. B. (2007). Comparison of shape encoding in primate dorsal and ventral visual pathways. *Journal of Neurophysiology, 97*(1), 307–319.

Levallois, C., Clithero, J. A., Wouters, P., Smidts, A., & Huettel, S. A. (2012). Translating upwards: Linking the neural and social sciences via neuroeconomics. *Nature Reviews Neuroscience, 13*(11), 789–797.

Liebrand, W., & McClintock, C. (1988). The ring measure of social values: A computerized procedure for assessing individual differences in information processing and social value orientation. *European Journal of Personality, 2*(3), 217–230.

Lim, S.-L., O'Doherty, J. P., & Rangel, A. (2011). The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *Journal of Neuroscience, 31*(37), 13214–13223.

Liu, T., & Pleskac, T. J. (2011). Neural correlates of evidence accumulation in a perceptual decision task. *Journal of Neurophysiology, 106*(5), 2383–2398.

Lohse, G. L., & Johnson, E. J. (1996). A comparison of two process tracing methods for choice tasks. *Organizational Behavior and Human Decision Processes, 68*(1), 28–43.

McClintock, C. (1972). Social motivation: A set of propositions. *Behavioral Science, 17*(5), 438–455.

McCoy, A. N., & Platt, M. L. (2005). Risk-sensitive neurons in macaque posterior cingulate cortex. *Nature Neuroscience, 8*(9), 1220–1227.

Middlebrooks, P. G., & Sommer, M. A. (2011). Metacognition in monkeys during an oculomotor task. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(2), 325.

Milosavljevic, M., Malmaud, J., Huth, A., Koch, C., & Rangel, A. (2010). The drift diffusion model can account for the accuracy and reaction time of value-based choices under high and low time pressure. *Judgment and Decision Making, 5*(6), 437–449.

Milosavljevic, M., Navalpakkam, V., Koch, C., & Rangel, A. (2012). Relative visual saliency differences induce sizable bias in consumer choice. *Journal of Consumer Psychology, 22*(1), 67–74.

Nagel, R. (1995). Unraveling in guessing games: An experimental study. *The American Economic Review, 85*(5), 1313–1326.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Newell, B. R., & Bröder, A. (2008). Cognitive processes, models and metaphors in decision research. *Judgment and Decision Making, 3*(3), 195–204.

Newsome, W. T., Glimcher, P. W., Gottlieb, J., Lee, D., & Platt, M. L. (2013). Comment on "In monkeys making value-based decisions, LIP neurons encode cue salience and not action value". *Science, 340*(6131), 430–430.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231.

Orquin, J. L., & Mueller Loose, S. (2013). Attention and choice: A review on eye movements in decision making. *Acta Psychologica, 144*(1), 190–206.

Pärnamets, P., Johansson, P., Hall, L., Balkenius, C., Spivey, M. J., & Richardson, D. C. (2015). Biasing moral decisions by exploiting the dynamics of eye gaze. *Proceedings of the National Academy of Sciences, 112*(13), 4170–4175.

Pieters, R., & Warlop, L. (1999). Visual attention during brand choice: The impact of time pressure and task motivation. *International Journal of Research in Marketing, 16*(1), 1–16.

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*(6741), 233–238.

Rangel, A. (2013). Regulation of dietary choice by the decision-making circuitry. *Nature Neuroscience, 16*(12), 1717–1724.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9*(7), 545–556.

Rapoport, A., & Chammah, A. M. (1965). *Prisoner's dilemma: A study in conflict and cooperation*. Ann Arbor: University of Michigan Press.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85*(2), 59.

Reuter, M., & Montag, C. (2016). *Neuroeconomics*. Berlin/Heidelberg: Springer.

Reutskaja, E., Nagel, R., Camerer, C. F., & Rangel, A. (2011). Search dynamics in consumer choice under time pressure: An eye-tracking study. *The American Economic Review, 101*(2), 900–926.

Russo, J., & Leclerc, F. (1994). An eye-fixation analysis of choice processes for consumer nondurables. *Journal of Consumer Research, 21*(2), 274–290.

Russo, J. E., & Rosen, L. D. (1975). An eye fixation analysis of multialternative choice. *Memory and Cognition, 3*(3), 267–276.

Sanfey, A. G., Rilling, J. K., Aronson, J. A., Nystrom, L. E., & Cohen, J. D. (2003). The neural basis of economic decision-making in the ultimatum game. *Science, 300*(5626), 1755.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Schall, J. D., & Thompson, K. G. (1999). Neural selection and control of visually guided eye movements. *Annual Review of Neuroscience, 22*(1), 241–259.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron, 60*(6), 1169–1181.

Sestieri, C., Corbetta, M., Spadone, S., Romani, G. L., & Shulman, G. L. (2014). Domain-general signals in the cingulo-opercular network for visuospatial attention and episodic memory. *Journal of Cognitive Neuroscience, 26*(3), 551–568.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317–1322.

Simon, H. A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics, 69*(1), 99–118.

Smith, A., Lohrenz, T., King, J., Montague, P. R., & Camerer, C. F. (2014). Irrational exuberance and neural crash warning signals during endogenous experimental market bubbles. *Proceedings of the National Academy of Sciences, 111*(29), 10503–10508.

Stahl, D. O., & Wilson, P. W. (1995). On players' models of other players: Theory and experimental evidence. *Games and Economic Behavior, 10*(1), 218–254.

Stern, J. A., & Dunham, D. N. (1990). The ocular system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Principles of psychophysiology: Physical, social and inferential elements* (pp. 513–553). Cambridge: Cambridge University Press.

Stewart, N., Gächter, S., Noguchi, T., & Mullett, T. L. (2015). Eye movements in strategic choice. *Journal of Behavioral Decision Making, 29,* 137–156.

Stritzke, M., Trommershäuser, J., & Gegenfurtner, K. R. (2009). Effects of salience and reward information during saccadic decisions under risk. *JOSA A, 26*(11), B1–B13.

Tosoni, A., Corbetta, M., Calluso, C., Committeri, G., Pezzulo, G., Romani, G., & Galati, G. (2014). Decision and action planning signals in human posterior parietal cortex during delayed perceptual choices. *European Journal of Neuroscience, 39*(8), 1370–1383.

Trommershäuser, J., Glimcher, P. W., & Gegenfurtner, K. R. (2009). Visual processing, learning and feedback in the primate eye movement system. *Trends in Neurosciences, 32*(11), 583–590.

Tversky, A., & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science, 211*(4481), 453–458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*(4), 297–323.

Van Lange, P. A. M. (1999). The pursuit of joint outcomes and equality in outcomes: An integrative model of social value orientation. *Journal of Personality and Social Psychology, 77*(2), 337.

Wang, J., Spezio, M., & Camerer, C. (2010). Pinocchio's pupil: Using eyetracking and pupil dilation to understand truth telling and deception in sender-receiver games. *The American Economic Review, 100*(3), 984–1007.

Webb, R., Glimcher, P. W., & Louie, K. (2014). Rationalizing context-dependent preferences: Divisive normalization and neurobiological constraints on decision-making. *Under Review*.

Chapter 21
Web Usability and Eyetracking

Volker Thoma and Jon Dodd

Contents

Abstract Eyetracking is one of several methodological tools researchers can employ when they need to evaluate in how far an interface or environment is amenable to a users' tasks or action goals. This chapter provides an introduction to the field of web usability, what user research looks like, and how it is embedded in an applied research

---

V. Thoma (✉)
Psychology, University of East London, London, UK
e-mail: v.thoma@uel.ac.uk

J. Dodd
Service Design and User Experience, Bunnyfoot Ltd., London EC1V 4JL, UK
e-mail: jon@bunnyfoot.com

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_21

and design setting. We will then describe and discuss when and how eyetracking is useful as a methodology for measuring aspects of web usability, and what it can tell us about the interplay between users and interface. Different eyetracking measures are explained in the context of how they can help the research team to make inferences about users' mental states and experience, and how these may be used to improve the interface design. This includes a discussion of some "myths" or common misconceptions about eyetracking in usability studies. A few brief case studies are given to illustrate eyetracking in applied web usability research.

21.1 Introduction and Learning Objectives

Every time a person wants to access an online service, perhaps for entertainment, to use social media, or because they want to buy something, they need to interact with an interface. The effectiveness, efficiency and satisfaction with which a person can deal with the interface is termed its 'usability'. Interfaces that have higher usability are easier and more pleasurable to use.

In recent years the term 'user experience' (or UX) has become popular. UX is related to usability but places more emphasis on the satisfaction component of usability, often considering aspects such as the emotional effect of the overall experience.

In principle, an interface can be anything that can be perceived and operated on, and can as such utilise any sensory modality. Currently, visual and touch (or click) interfaces predominate, although advances in auditory, haptic and voice controlled interfaces (amongst others) are also gaining popularity.

In this chapter we will focus on digital visual interfaces and how eyetracking may be used to enhance the usability and user experience they provide. Such interfaces may be delivered via webpages, smartphones or tablet apps, kiosks, smart TVs, games consoles etc. In recent years eyetracking has been increasingly combined with traditional usability tests to allow a richer and more objective picture of users' experience with an interface.

After reading this chapter, you should be able to answer the following questions:

1. What is usability, in particular, web usability?
2. Why is usability research necessary and useful?
3. What are common methods and practices in web usability research (e.g., user-centred design process)?
4. When is eyetracking a useful method in web usability testing? When not?
5. What are common misconceptions (or 'myths') when people consider eyetracking results in user interface research?
6. What are useful eyetracking measures and visualisations in web usability research?
7. What are common issues and potential pitfalls in using and interpreting eyetracking results in web usability projects?

21.2 Historical Annotations

The concept of usability originated in studies on man-machine-interaction and human factors research, with documented research starting just before World War 1 in human efficiency studies. In 1957 the Human Factors and Ergonomics Society was formed (www.hfes.org). John Bennett published an article that is considered one of the first (academic) papers on usability in 1979, and dedicated labs—such as at IBM—began with usability tests (although reports indicate earlier user test labs, e.g., AT&T Bell from 1970). Important publications followed, such as the “Psychology of Human Computer Interaction” (Card, Newell, & Moran, 1983), and then, after the emergence of the internet, Jacob Nielsen published the influential book “Designing web usability” in 2000. Shortly earlier, in 1998, usability was enshrined as an ISO standard (ISO 9241, pt. 11). Eyetracking in usability studies was used in the late 1940s, famously with studies in pilots’ cockpits (Fitts, Jones, & Milton, 1950). By 1986 there were scientific articles describing how eyetracking was used in evaluating magazine advertisement research. Eyetracking in usability and then web usability research was becoming more common from the late 1990s, when manufacturers launched tracking kits that were less expensive and easier to handle for researchers in usability test settings than previous equipment. Publications on eyetracking in usability and then web usability followed soon after, notably recent books by Nielsen and Pernice (2009) and Bojko (2013).

21.3 What Is (Web) Usability?

21.3.1 Definitions and Aspects of Web Usability?

Usability is about making interfaces for products or systems easier to use, and matching them closely to the needs and requirements of the people who will actually interact with them (the ‘users’). There are a number of ways to define usability but one of the best known definitions comes from the International Standards Organisation (ISO).

The international standard ISO 9241-11 (Guidance on Usability 1998) defines usability as:

> the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use. (“ISO 9241-11,” 1998)

The above definition is robust and still highly relevant even though technology has evolved greatly since 1998. The important components are:

- **Effectiveness**: the degree to which the users complete their goals (or tasks) successfully. Often measured as a percentage of completed tasks.
- **Efficiency**: the amount of effort users must expend in completing their goals. Often measured in time or number of clicks/taps.

- **Satisfaction**: what the users actually think about the experience of using the product/system. Often measured with a subjective questionnaire.

Importantly all the above three components are measureable (albeit they are not mutually exclusive measures) and are affected by the:

- **Users**: This concerns the people using the product, their needs and specific characteristics. For example, users may be experts or novices, and they use it daily or only once.
- **Goals**: This addresses the tasks/outcomes that users need to achieve in using the product/system, e.g., finding and successfully purchasing a specific product.
- **Context of use:** This could mean a range of situations, e.g., where, when and on what device the product/system is being used—in the office, on a desktop computer using high speed broadband, on the train using a smartphone with patchy signal, with a specific browser, etc.

Other definitions of usability do exist (for an overview, see e.g., Petrie & Bevan, 2009) but tend to cover the above components from the ISO definition and add aspects such as:

- **Learnability**: How easy it is for users to perform tasks the first time they encounter the product?
- **Memorability**: After a period of non-use how easy it is for users to get back to their previous performance level?
- **Error tolerance**: How many errors the users make during a task? How easy it is to recover from errors made?

User experience (UX) as a term has grown popular in the last few years. There are many definitions of UX but they generally encompass usability (usability is a key component to delivering a good UX) and in addition stress more subjective, often emotional, perceptions of the interaction and emphasize longer term holistic relationships ("What Is User Experience Design?," 2010). Here is an example from Kuniavsky (2010):

> 'The user experience is the totality of end-users' perceptions as they interact with a product or service. These perceptions include effectiveness (how good is the result?), efficiency (how fast or cheap is it?), emotional satisfaction (how good does it feel?), and the quality of the relationship with the entity that created the product or service (what expectations does it create for subsequent interactions?).

While the User Experience Professionals Association (UXPA) defines user experience as ("Definitions of User Experience and Usability," n.d.):

> Every aspect of the user's interaction with a product, service, or company that make up the user's perceptions of the whole. User experience design as a discipline is concerned with all the elements that together make up that interface, including layout, visual design, text, brand, sound, and interaction. User Experience works to coordinate these elements to allow for the best possible interaction by users.

A number of principles for interface and interaction design have been produced, many derived from cognitive psychology and Human Computer Interaction research (e.g., Nielsen & Molich, 1990; Norman, 1988; Shneiderman, 1986; Shneiderman & Plaisant, 2005; Stone, Jarrett, Woodroffe, & Minocha, 2005).

Principles such as these (along with other points of reference such as design patterns and strict technology or operating system guidelines/specifications) are used to guide the design and the development of interfaces, but also to assist with evaluating existing or evolving designs.

Although tried and tested principles and guidelines such as those above exist and are highly useful, they can only elevate the design of an interface or experience so far. Input and evidence is needed from the end users themselves in order to ensure designs match their wants, needs, expectations, capabilities and contexts of use.

Such evidence may be gained through the many forms of user experience research with different techniques being appropriate for different stages of the design lifecycle. An important method is usability or user experience testing, which in many circumstances may be enhanced by the addition of eyetracking technology.

21.3.2 Why User Experience Research Is Useful?

It's hardly surprising that if an interface—like any other tool or service—is too difficult, then people will fail to complete their goals, work inefficiently, or simply walk away. For example, if people try to purchase goods on the Internet, but cannot find them easily, are unable to choose their preferred delivery methods, or cannot easily find the right payment option, they will leave and seek alternative services that offer a better experience. Once a better alternative is found they are more likely to return to that service and highly unlikely to return to the poor service.

An interface with good usability (and user experience) increases user satisfaction and productivity, heightens customer loyalty and advocacy, and also is typically cheaper to produce and maintain (Bias & Mayhew, 2005). Ultimately user experience research leads to increased profitability—it is essential to the bottom line.

A meta-study of 43 web usability projects (Nielsen & Gilutz, 2003) found an average of 135% improvement in core business performance indicators, such as sales numbers, visitor traffic, user performance (productivity), and use of specific target features on the website. Another survey by e-consultancy (2007) found between 35 and 55% improvements after usability research and consequent redesign on variables such as brand perception, customer loyalty and retention, increased traffic, increased search rankings, and increased customer advocacy.

The UX consultancy Bunnyfoot (where one of the authors works) have published figures from some more recent (2013/14) iterative design projects that involved usability testing and eyetracking ("Clients - What our clients say about us and our work," n.d.). The data show:

- 38.5% increased online booking conversion (Great Rail Journeys)

- 37% increase in landing page conversion rate and a 15% increase in donations (Oxfam)
- 59% quicker completion of online checkout, which was preferred by 80% of users and caused 6 times fewer abandonments (Boden)

A general rule of thumb often cited (originally from Gilb, 1998) which still appears to be broadly relevant today is that once a system is in development it costs 10 times as much to fix a problem than it would have done to fix that same problem earlier at the design stage. In addition, once the product is released it costs 100 times as much to fix the problem compared to fixing it earlier in design. This is commonly stated as: the cost benefit ratio for usability is £1:£10:£100 (or $1:$10:$100).

21.4 User Experience Research and Eyetracking

As described in other chapters (e.g., Hutton, see this volume) in more detail eyetracking is a technique to measure the movements of the eyes of an observer in order to track and record where they are looking. Measurements can be made of the observers' eye movements (such as when they first look at something, how long they look at something, how many times they look at something). In addition, eyetracking can potentially provide clues as to *why* they are looking at something (by looking at their gaze behavior and correlating with other data such as spontaneous comments).

Until relatively recently, eyetracking was of limited use for usability and user experience research because older technology required intrusive interventions such as the requirement to keep the head absolutely still. These intrusions imposed by the old technology caused unnatural behavior (people do not normally keep their head absolutely still) and a likely 'white coat effect' (the test conditions changing the results, e.g., because users may behave perhaps according to some assumed expectations of the experimenters). Modern eye trackers used for usability [see chapter Hutton, in this volume] and user experience research generally nowadays only require a short and simple calibration process, allow natural movement of the head, and impose little, if any, other requirements or intrusions for the observer. Thus, modern eyetrackers are good tools for acquiring extra insight into behavior without much threat of changing it.

For usability and UX research there are two main ways eyetracking is used:

1. **To provide qualitative insight** to enhance observations of behavior during exploratory research, and to detect and help explain usability and user experience issues (so called 'formative' testing).
2. **To provide quantitative data** to measure and compare performance of alternative designs (so called 'summative' testing). These data are used to support the optimisation of the design and layout of elements within a screen/page.

The ultimate goal of any research and testing is to produce actionable insight. The particular type of research, and indeed, the choice of whether eyetracking can usefully support it, is determined somewhat by the stage in the overall user experience design process that it takes place. This is discussed below with reference to a user centred design (UCD) approach to web development.

21.4.1 User Centred Design and Eyetracking

Eyetracking during usability or user experience research is just one of many tools and activities that may be employed in the process of designing interactive systems such as websites, apps, or software.

A commonly used, and successful approach is to adopt a User Centred Design (UCD) methodology (see Fig. 21.1). UCD places the user of the eventual product or service at the heart of the design effort. The user is considered, explicitly referenced, and researched at each stage, thus ensuring that design is driven forward based on evidence rather than assumptions or speculation.

Executing a successful UCD program of work requires a variety of different skills in a team, and typically includes roles such as: designers, researchers, information architects, marketing experts, business analysts, programmers, and project managers. Sometimes an individual may fulfil multiple roles.

UCD, together with iterative usability/UX testing, fits with traditional 'waterfall' approaches to development and also with more modern and increasing popular 'agile'

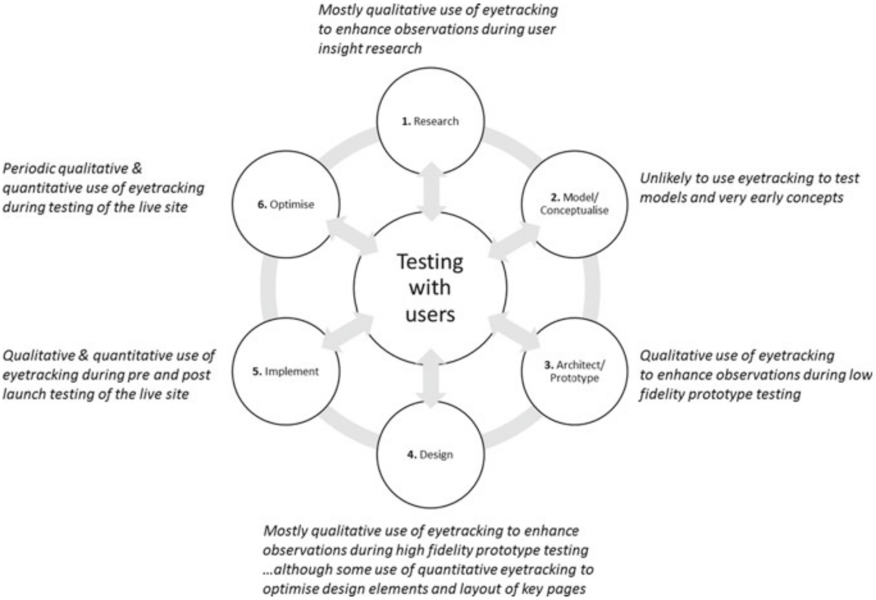


Fig. 21.1 A User Centred Design (UCD) cycle for a web design project. Eyetracking can be usefully employed at different stages

approaches (see Jones & Thoma, 2019). In an iterative development process, research, design, prototyping, and testing may be revisited and refined a number of times (the iterations) before a final product is released.

There are many different models and representations of UCD but all follow roughly the same stages as in Fig. 21.1. The potential use of eyetracking within each of the stages is discussed below:

1. **Research**: Activities at this stage include exploring the business opportunity and setting objectives, performing competitive evaluations and importantly performing user insight research aimed at understanding user needs and the context of use.
   Eyetracking is mostly used at this stage to enhance observations during qualitative research activities such as:

   - *Observing people use the current site/system*
   - *Observing people use competitor sites/systems*
   - *Observing people use early concepts or prototypes*
   - *Observing people in their own environment behaving naturally (potentially using wearable eyetracking to aid this contextual research)*

   Outputs contribute towards requirements analyses, and towards gaining a deeper understanding of user needs and priorities which may be modeled in the next stage.
2. **Model/conceptualize**: Activities in this stage distil the research into useful design tools such as personas (descriptions of who the users are, their goals and priorities), mental models (how the users might approach using the system) and scenarios (descriptions of different contexts and uses of the system). These design tools are used to drive the next stage of the process, and may be useful reference points throughout the whole process.
   Whilst testing and validation of models can occur within this stage eyetracking is not the ideal tool to probe insight into models and scenarios, compared to more qualitative methods probing users' reasons.
3. **Architect/prototype**: Activities at this stage include detailing the overall structure of the system (information or system architecture) and producing prototypes that simulate how the system will be used (interaction design). The prototypes can be of varying degrees of fidelity from hand drawn early prototypes, through to black and white 'wireframe prototypes', through to dynamic functional prototypes.
   Eyetracking at this stage is mostly used to help inform qualitative observations during iterative usability testing—essentially trying to find and understand problems/issues within the design to overcome them.
4. **Design:** This is where the application of branding and fine scale visual design takes place, resulting in the precise look and feel that the final product will have.
   Eyetracking at this stage can be used to inform qualitative observations during iterative usability testing to ensure that the visual design enhances the experience. Also more quantitative eyetracking may be used to optimize layout of elements on

key pages (e.g., assessing salience of key items such as calls to action—leading to perhaps changes in location or visual form).

5. **Implementation**. This is where the coding occurs leading ultimately to the launch of the product. In a more agile approach the implementation may be incorporated tightly with the other stages in short bursts (or sprints) leading to viable products that incrementally improve or extend features.

   Eyetracking testing (qualitative and quantitative) may occur pre-launch (or within Agile sprints) to check usability and attempt to measure and optimize the experience.
6. **Optimisation**. Often post launch performance will be assessed and changes made to the design to make improvements or address deficiencies. Performance might be assessed using a whole range of methods including website analytics. Live tests of the effects of tweaks or alternative designs may be performed using split testing (often A/B testing) or Multivariate testing techniques.

   Periodic eyetracking testing is often employed to gain qualitative insight into the effects of changes and/or to seek to optimize performance (perhaps through quantitative techniques similar to those performed in the design stage).

21.4.2 Different Testing Approaches

There are a number of ways that usability and UX testing may be performed. Table 21.1 shows a summary of the different characteristics of these different testing approaches. Currently (in 2017) eyetracking is only really available in face-to-face research (either lab based or intercept testing), whereas in the future it is likely to be available for remote methods using technology, such as webcams, built into common consumer devices (Fig. 21.2).

21.5 The Role of Eyetracking in Usability/UX Testing

21.5.1 Advantages, Caveats, and Myths About Eyetracking in UX Research

Ever since reliable, non-intrusive eyetracking became available to UX researchers and commercial agencies much debate has occurred about the appropriateness of its use. Often people appear to be firmly and vociferously in a pro-eyetracking or anti-eyetracking camp. These opposing parties often make strong claims and counter claims: Some see eyetracking as the silver bullet for everything and every project, whilst others claim it is ‘research snake oil’ and has little value. Unsurprisingly as authors of a chapter on eyetracking and UX we are generally pro eyetracking and

Table 21.1 A spectrum of usability testing methods—often design projects would use more than just one method (depending on the stage and the needs of the research)

	Traditional lab based	Remote moderated	Intercept testing	Remote un-moderated	Live site analytics/split testing MVT
Location of testing	Constrained: Lab based	Unconstrained (users own equipment)	Constrained: where users can be found	Unconstrained	Unconstrained
Source of test users	Pre-recruited and scheduled	Pre-recruited and scheduled	Opportunistic interception	Invited from-databases or web link	Live traffic
Approach	1 to 1 face to face	1 to 1 via screen sharing	1 to 1 face to face	User completes alone	User unaware of test (performing real task)
Test session length	Typically 30–60 min	Typically 30–60 min	Typically up to 15 min	Depends on the number of tasks	Depends on the real life task the user is trying to complete
Number of users	Typically low (due to cost)	Typically low (due to cost)	Medium	Medium-high	High
Cost per user	High	High	Medium	Medium-low	Low (real traffic)
Level of qualitative insight	High	Medium	High-medium	Medium-low	Low
Eyetracking possible?	YES	Potentially in the future	YES	Potentially in the future	NO/unlikely?

positive about its use, but only when it is used for appropriate studies, when objectives are clearly defined and when the goal is to create actionable insight.

We recognize that remote claims and inappropriate use may have damaged the reputation of eyetracking as a tool to some, and equally some negative myths might also obscure its real benefits to others. Below we examine and discuss some common potential myths and misconceptions:

1. **‘Eyetracking is a silver bullet for UX research’**
   Some enthusiastic proponents of eyetracking seek to use it on every project without discrimination, they produce data visualisations to include in reports without reflection, because they deem it generally useful just to ‘know where people looked’.
   Instead, eyetracking should be seen as a potentially helpful tool to be used in conjunction with other tools and research methods when and only when appropriate and beneficial to gaining insight. Importantly, eyetracking is the only real

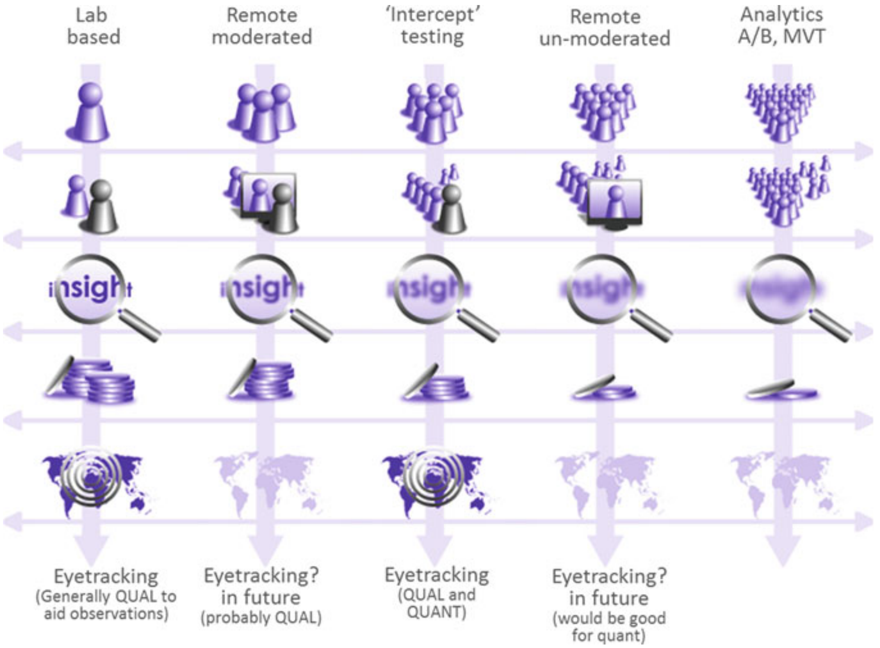


Fig. 21.2 A continuum of usability testing methods in relation to numbers of users, clarity of insight, cost per participant, and location (local vs. global)

physiological measure of attention readily available to UX research. The precise use and role of eyetracking along with any analyses should be planned specifically in advance for each research project (see below). Post hoc unplanned analyses, and in particular speculative viewing of eyetracking visualisations can lead to misinterpretation and misinformation. Often clients/stakeholders can get obsessed with the outputs of eyetracking because they are so striking and visual, to the point that they start requesting 'eyetracking' of their product rather than UX/usability testing of it (e.g., they may demand heat-maps or 'hotspots' of a page when only 4 or 5 people have viewed it in a formative usability test—and start making inappropriate conclusions from the data).

Much valuable research and testing can and should be performed without eyetracking, for example, contextual interviews, user observation, testing of lo-fidelity paper prototype testing, etc.

2. **'Eyetrackers are like mind readers'**

   Just because an eyetracker recorded a certain fixation does not necessarily mean that the user actually registered the contents of the location. The user may have been briefly attracted to the location by some random feature (or just the color) without reading or recognizing the visual element. Alternatively, it could just be that the observer briefly "rested" their eye while thinking of something else.

Another reason why eyetracking data should not automatically be equated with what people are processing is the fact that an area has not been fixated does not necessarily mean that information contained in this location will not be processed. Research has shown that a face or an object shown in the periphery of a display can still be recognized even it was irrelevant to the main task, and even if it wasn't fixated (Thoma & Henson, 2011; Thoma & Lavie, 2013).

Eyetracking data alone does not give any information at all about *why* someone looked at something or their interpretation of what they looked at—this insight must be gained in combination of other means (see Table 21.2).

3. **'Knowing where people looked isn't useful'**

   The eye-mind hypotheses, which is supported by a substantial amount of research (Hoffman, 1998; Just & Carpenter, 1976) states that there is a close relationship between what the eyes are gazing at and what the mind is engaged with. Furthermore, this is most likely to hold when a person is engaged in a task that requires information presented in a visual modality (such as most tasks on a web site during a usability test). Therefore, what a person is looking at often indicates the thought "on top of the stack" of cognitive processes, essentially what they are attending to and thinking about.

   Whilst there are limitations in that eyetracking is not a mind reading tool (see above) having access to gaze data (principally the location, duration and order that people look at a scene) in the context of what they were trying to achieve at the time, can give extra valuable information regarding their cognitive processes. For example, certain eye-movement patterns in a particular task may hint at possible confusion of participants when seeing a particular visual element. Eyetracking therefore assists other observation and helps interpreting of what occurred (see

Table 21.2 Some common eyetracking patterns/metrics and their interpretation

Eye movement metric/pattern	Potential interpretation
Time to first fixation on target or AOI	Target is salient (attention grabbing) or in expected location if short, potentially not salient enough if long
Duration of fixation on target or AOI	Can indicate difficulty in extracting information (confusion) or that the item is interesting
Number of fixations on AOI divided by the total number of fixations	Depending on tasks: May indicate low search efficiency or higher level of interest, or complexity
Time from first fixation on target AOI to making an action (e.g., clicking a button)	If long then unclear call to action (doesn't look clickable, label not clear), or too many potential distractors
Saccades of a user changing significantly in direction on page	User's goals have changed, or interface layout does not match/support expectations
Returning to/rechecking target content	Difficulty with understanding, or highly engaging—or indicates high interest

e.g., the classic study by Yarbus (1967), showing that eyetracking patterns for identical scenes can differ depending on the task observers were given).

At its most basic, eyetracking produces real-time output that observers can view whilst test participants are engaged in tasks. This can help motivate non-research team members to attend test sessions and helps them engage in the session and empathize with the participants, therefore asking questions of the design that would not have occurred otherwise. More importantly, the real-time eyetracking output assists trained researchers with interpretation of findings or usability issues, and can be used to guide further probing as to what occurred.

Posttest examination of data, when planned in advance, can assist with both qualitative insights and quantitative comparisons or performance measures. Data visualisations, such as hotspots, and gazeplots, whilst they should not be abused (i.e. taken at face-value, without reflection about the task and consideration of other data, see Sect. 21.6), can help illustrate test findings.

In short knowing where people look can be useful—if the eyetracking methods are used properly.

4. **'Eyetracking is too expensive to add value to usability/UX research'**

   In the past it was very expensive to buy eyetracking equipment—in particular for lone researchers or small organisations. But as of 2017 high-end machines can be purchased for under £10,000 (and is likely to drop still further). Extra costs though should be factored in for things like staff expertise (in use of software and importantly in test design and interpretation of data) and renewal of software updates.

   When considering all the other costs it takes to produce a website or piece of software of high UX quality, the cost of using eyetracking, whilst not insignificant, is unlikely to be a major impediment. Alternatives are to hire the equipment, or hire a company that has the expertise and equipment.

5. **'Eyetracking research is off-putting to users in the tests'**

   The goal of UX research is to observe and potentially measure real and representative behavior. Consequently, it is good research practice to seek to minimize factors that might cause behaviors to become non-natural (the 'white-coat' effect). Imposing any unnecessary burden or conditions to a user in a test should be avoided, and so if eyetracking does not have a specific and defined benefit then it follows that its use should be avoided.

   Where there is a defined benefit though, and where the context of the test is appropriate (e.g., a test of a website in a usability lab using modern non-imposing eyetracking equipment) the inclusion of a short eyetracking calibration exercise (often less than 15 s) does not add undue extra burden or imposition on test users. Informal observations by the authors during many hundreds of usability studies using eyetracking also suggest that most users forget about the eyetracker and being eyetracked when they are engaged in the usability test tasks—it would be useful to confirm this anecdotal observation in more formal studies.

   For some types of test the use of eyetracking can actually increase the behavioral validity of the test. This is when a retrospective protocol (RP), facilitated by the use of eyetracking replay, is used instead of the concurrent think aloud protocol (where users are required to continuously and unnaturally verbalize their thought processes while performing a task)—However, some researchers question the validity of this approach (see later sections for more about different test approaches, and Boijko, 2013, for a discussion).

6. **'Eyetracking is an extra burden to the researcher'**

   This is true to the extent that, as with any technique or tool, a researcher will need to first learn how to operate the equipment properly. Even more importantly they must invest time and effort learning how to identify and design appropriate studies and how to analyze and interpret data correctly.

   Physically setting up an eyetracking study often requires little extra effort to that required in testing/lab setups where (audio/video) recording of the session would take place in any case. Indeed, most eyetracking software (e.g., Tobii Studio, SMI Experiment Suite) for commercial usability testing has the same or similar features for set up, running and analysis of tests as non-eyetracking testing software.

   At its most basic use, if eyetracking just improves the real-time observation of natural behavior in formative usability studies, no extra analysis is required, and it has given a benefit without any appreciable extra burden.

7. **'An algorithm could be used to predict how people look at a scene so why bother with eyetracking'**

   Gaze behavior is determined by both bottom up stimulus driven shifts of attention (for example to a particularly salient visual feature) and by top down goal driven shifts of attention (e.g., actively looking for a brand or specific product in an online catalogue).

   In principle if attention was only directed by bottom up factors then it should be possible to create an algorithm that could predict with relatively good accuracy where and when people would look at the features in any given scene. Indeed this has been attempted and can give some interesting basic results, but of course fails to reflect reality in most cases because of top down influences. Thus, actual gaze behavior is dependent on both the features of the visual scene and factors such as the observer's goals, knowledge, experience and expectations. This 'true' behavior is what is measured in eyetracking UX/usability research.

8. **'Mousepointer–data can provide equally valid information as eyetracking.'**

   Some researchers claim a correlation between mouse movements and gaze data, or mouse movements and attention (Chen, Anderson, & Sohn, 2001)—thus opening up the possibility that the more economical method of tracking the mouse pointer could be used as a substitute for insights gained from eyetracking. However these studies are mostly from web pages with low complexity or restricted function (such as search engine results pages) and recent studies (e.g., Huang, White, & Buscher, 2012) have shown that gaze-to-cursor correlation is situational as it depends on the time spent on the page, personal browsing habits, and on the user's current cursor mode of behavior (inactive, examining, reading, action). In addition, where correlations occur it appears to lag gaze position by about 700 ms. A further restriction is that mousepointer data is not available for touch systems like kiosks or tablets, or devices like interactive TV.

21.5.2 Uses of Eyetracking Data in UX/Usability Testing

Many usability/UX tests typically deliver examples of where people had problems, their qualitative verbalizations and measures such as the percentage of completed tasks (effectiveness), speed of completion (efficiency) and subjective satisfaction with the experience. Eyetracking can assist with qualitative insight into why something happened, as well as support quantitative measures and comparisons between different designs.

21.5.2.1 Qualitative Insight

An important role of eyetracking is to assist with the understanding of what occurred during a test session. Eyetracking can augment interpretation concerning cognitive processes or usability issues that occur, but critically this must be combined with other information available such as knowledge of the context of the test, the tasks/goals themselves, verbalizations made, knowledge of the participants etc. For example: eyetracking may show that a number of test participants fail to look at a key action button (such as an 'add to basket' button) indicating that it may not be salient enough, or in the correct position (or both). Whereas if participants were found to look at the button but not select it then this may indicate issues with the button label, or its visual form (perhaps it does not look clickable enough).

21.5.2.2 Quantitative Comparison

The goal of using quantitative eyetracking metrics is to compare different alternative designs, or compare performance based on benchmarks such as competitor designs. The eyetracking data can be used to support and augment non-eyetracking performance comparisons (such as time to complete task) and/or compare factors concerning visual attention and attraction and potentially even emotional engagement (e.g., including motional face recognition software). In practice, relatively low samples sizes (i.e., number of participants) for most UX research projects mean that the statistical power is usually limited and therefore findings cannot be readily generalized to the larger population.

21.5.2.3 Eyetracking Measures, Patterns and Interpretation

Foundation of Common Measures: Time, Gaze Position, Fixations, Pupil Size, Blink Rate

Most eyetracking technology produces raw data that shows a time stamp (e.g., a data entry every 20 ms), and the associated position of convergence of the eyes (sometimes position is reported for each eye separately, sometimes computed as X, Y, Z coordinates), and some eyetrackers also record pupil size and derived measures/events such as fixations, eye-movements ('saccades') and the occurrence of blinks. In the following section we discuss the data capture possibilities that are most useful for UX studies (for a comprehensive overview see also Ehmke & Wilson, 2007).

Most eyetracking software permits the definition of Areas of Interest or AOIs. AOIs are specific regions of the visual interface or scene that the researcher may be interested in. Measures can be made with reference to target AOIs, and often eyetracking software will automatically produce different standard charts and data such as the order of fixation of defined AOIs, time to first fixation on AOIs, duration of gaze on AOIs, number of fixations on AOIs etc. It is also possible to extract and compute other measures.

To many people fixations are the easiest understandable measure in eyetracking research. Fixations typically indicate situations when a user is taking in information (according to the eye-mind hypothesis—see above) hence why they are interesting. An algorithm must be used in order to detect fixations from the raw gaze data. Typically the minimum duration of 70–100 ms is set to define a fixation along with either a setting for the maximum position deviation (usually 0.5°–1.0° visual angle), or maximum velocity of movements between samples (typically 20–30 deg/s)—depending on the type of algorithm used.

Many (though not all) eyetrackers are able to monitor the pupil size as well as the blink rate (the number of times per minute somebody blinks). Blink rate and pupil size can be used as an index of cognitive workload (see Chap. 11 by Laeng & Alnaes, in this volume). A lower blink rate is assumed to indicate a higher workload, because users are focussed on a goal or tasks to be solved. In contrast, a higher blink rate may indicate fatigue (Brookings, Wilson, & Swain, 1996; Bruneau, Sasse, & McCarthy, 2002). When participants' pupils get larger ('dilate') this may also indicate more cognitive effort or increased attention or emotional arousal (Marshall, 2002; Pomplun & Sunkara, 2003). However, pupil size and blink rate can depend on many other factors (Goldberg & Wichansky, 2003). Indeed, there is evidence from studies that pupil dilation does not vary with mental effort in a web usability study (Józsa, 2011). Consequently, pupil size and blink rate are not often used in UX eyetracking research. However, if researchers are interested in detecting fatigue and strain ('user cost'), a combination of blink rate and other physiological measures (e.g., heart rate, or galvanic skin response) can be employed (e.g., Bruneau et al., 2002).

Interpretations of eyetracking measures

According to the eye-mind hypothesis longer than expected fixations may mean that a user is mentally engaged in processing the visual element at that location (Just & Carpenter, 1976). However, depending on the user's task and the type of stimulus presented to them, interpretations about fixation data may vary. For example, long fixation durations may mean that people are really interested in what they are seeing. Alternatively, long fixations may indicate that users may be really confused and are trying to figure out what they are looking at, or simply think about the visual information longer. The interpretation of fixation durations therefore needs to take into account what tasks users are performing and what they are reporting during or after a session. In addition, longer fixations may be relative independent of task and stimulus, but may mean that the observer is less interested or slower than others to process the information.

Similar limitations apply to the interpretation of number of fixations. A high number of overall fixations may indicate less efficient visual search, and therefore indicate that the spatial layout needs a redesign (Goldberg & Kotval, 1999). At the same time, a high number of fixations on a particular location or area of interest may also indicate increased interest for the related content (Poole & Ball, 2006). This is why eyetracking data alone cannot give the UX team conclusive answers—researchers need to know the users (their background, such as their experience with a particular technology, their goals, their motivations) to be able to interpret the data sensibly.

Saccades are defined as rapid eye movements between fixations. As with fixation measures, the interpretation of saccade measures is not straightforward, and will depend on the context of the particular research, and on the interpretation of other data. A high number of saccades across the visual display may indicate that users have to (or want to) search for more information. Larger saccades (i.e., a greater distance between two temporal adjacent fixations) may indicate more attention pull or capture from the information displayed at the second location (Goldberg, Stimson, Lewenstein, Scott, & Wichansky, 2002).

During saccades it is unlikely that there is any encoding (that is, any meaningful intake of information) possible (see chapter by Pierce, in this volume). Therefore, saccades are less apt to tell us much about any potential issues with user interface elements themselves. However, if saccades indicate that users seem to change direction (“backtrack”) repeatedly (so-called regressive saccades) this may indicate some processing difficulty during encoding or information processing. For example, there is considerable amount of research on word-reading that suggests that larger (phrase-length) regressions can mean that readers are confused by the content of a text (Mitzner, Touron, Rogers, & Hertzog, 2010; Rayner & Pollatsek, 1989; Reichle, Pollatsek, Fisher, & Rayner, 1998). Poole and Ball (2006) also claim that regressions could be used as a measure of recognition value—the more salient or recognisable a text is, the fewer regressions should be observed. Very large saccades could mean that a user’s task or goal has changed or that there is a considerable mismatch between the information layout and the user’s expectation (Cowen, Ball, & Delin, 2002).

The interpretation of fixations and saccades needs also to take into account the context in which the areas of interests (AOIs) appear. For example, a relatively high number of saccade regressions in a text area may indicate that a word or phrase was unfamiliar. However, a similar result of regressive saccades found for the same text in a navigation AOI may indicate that—despite the text or label being very familiar—users may be confused where they are in the navigation process, rather than by the text itself.

Scanpaths (see chapter by Foulsham, in this volume) show the sequence and duration of fixations that a person made during a certain period of time (often whilst engaged in a particular activity). An optimal or ideal scan path (e.g., in a search task or to buy an item and ‘go’ to checkout) is viewed as being a straight line from the initial point to a desired target, with comparatively short durations of fixations at the target (Goldberg & Kotval, 1999). If scanpaths are relatively long, and with a number of changes in direction, this may suggest inefficient searching (perhaps due to poor visual organisation or design; Goldberg et al., 2002). The interpretation of scanpaths can be complex, and Goldberg and Kotval (1999) provide an overview of different measures and their implications. There are also measures that are derived (calculated from) the scanpath data, such as the ‘transition matrix’ and ‘transition probability’ which indicate the degree to which users switch their gaze between AOIs (see also Jacob & Karn, 2003) and may indicate the efficiency of the overall layout of visual elements.

Box 1: Fixation does not always mean attentional processing

The main assumption behind much applied (and scientific) eyetracking research is that what information is processed is determined by the location the eye is fixating on (the ‘mind’s eye’ hypothesis)—observers process the information at the ROI. But everyone who has held a book lying in bed knows that this is not always the case-we can read perfectly without taking much in, because we are thinking of something else or already half-asleep. Related to this is the phenomenon of “attentional blindness”, which is the finding that people do not perceive a stimulus they are looking at. You may have seen a famous video (it is available online) in which people are told to watch two teams, one in black shirts and one in white shirts, each passing a basketball between them. Observers are asked to count how many passes the team in white manages, ignoring the team in black. A distracting item in black also passes through the scene and even remains highly visible in the center—a co-experimenter dressed in a black (!) gorilla costume. Nevertheless, about 70% of observers completely miss the “gorilla” when asked afterwards—because their task forced them focus on the “white” information (Simons & Chabris, 1999). This is of course one reason why it so important to know about the task (or tasks) that people have in mind when doing eyetracking, as discussed earlier in this section.

On the other hand, not looking at an item does not necessarily mean that it is not being perceived. One of the authors of this chapter (VT) has studied how objects and faces outside the focus of attention are being processed. First, there is ‘covert attention’, meaning the ability that even when you look for example at the center of a screen, you can perfectly “pay attention” to say the right of the screen. Thoma, Hummel, and Davidoff (2004) showed that when people did not have enough time to move their eyes from the central fixation point, they were highly accurate and fast to name an object shown 4° to the left or right of fixation (and greater distances are possible). But even more interesting is the fact that completely ignored stimuli on the opposite side of where attention was directed to can still be recognized (this is measured in so-called priming paradigms, where objects that were to be ignored in one trial would be shown in a later block, and response performance would show that they are recognized faster and more accurately than previously unseen objects). However, the ability to process the identity of ignored (spatially unattended) objects depends on many factors, such as the object's orientation (Thoma, Davidoff, & Hummel, 2007) and whether there are many other objects in the center of the display (Lavie, Lin, Zokaei, & Thoma, 2009). Similar results are found for faces, although ignored faces seem even more likely to be processed when ignored compared to objects, as longs as they are familiar to us (Thoma & Lavie, 2013). Therefore, even if observers in an eyetracking study seem not to be looking at, say, an advertisement banner, this may not necessarily mean they have not processed some aspects of it.

Typical Eyetracking Measures for UX

Experienced researchers have learned to identify typical eye-movement metrics or gaze pattern signatures. Some of these common eyetracking data patterns and their potential interpretations are shown in Table 21.1 (for a more thorough summary see Ehmke & Wilson, 2007).

Holmquist et al. (2011) documented over 100 different eyetracking measures and categorized these into the following four categories:

- Movement measures
- Position measures
- Numerosity measures
- Latency and distance measures

In the book 'Eyetracking the user experience' Aga Bojko (Bojko, 2013) identifies key measures suitable for UX practitioners and classifies them according to two main types of UX question that can be answered:

- Questions related to attraction of elements (often during non-goal specific browsing behavior). For example: Where is the best placement of a promotion, advertisement or key messaging in order to be noticed easily? What product picture retains the customer's interest longer?
- Questions related to performance during goal directed behavior. For example: Which listing layout makes it easiest to find and comprehend key information in order to support a choice? Where will the buy now button be found most easily?

One can see from Table 21.3 that measures of attraction and performance are often shared and so it is vital to understand the context of the behavior in order to determine what is being measured (e.g., passive attraction versus active search/location, emotional arousal vs. workload). In addition, it often makes sense to report multiple

Table 21.3 Adapted from Bojko (2013) shows a practical taxonomy of eyetracking measures often used in UX

Questions related to:	Potentially relevant eye movement measures
Attraction: how easy an area is to notice	• % of participants who fixated AOI—higher % = more noticeable • Number of fixations before first fixation on AOI—higher number = less noticeable • Time until first fixation on AOI—longer time = less noticeable
Attraction: how interesting an area is	• Number of fixations on an AOI—higher number = of more interest • Duration of fixations (or dwell time) on AOI—longer time = of more interest • Percentage of total time on AOI—higher % = of more interest
Attraction: how emotionally arousing an area is	• Pupil diameter—larger = more aroused
Performance: how easy a target was to find	• % of participants who fixated target AOI (higher % = more noticeable) • Number of fixations before first fixation on target AOI—higher number = less noticeable • Time until first fixation on target AOI—longer time = less noticeable
Performance: effect on mental workload	• Pupil diameter—larger = higher workload/processing
Performance: effect on cognitive processing	• Average fixation duration—longer duration = more processing
Performance: How easy a target was to recognize	• Number of gaze visits to target prior to selection—lower number = easier to recognize • Time from first fixation on target to selection—shorter time = easier to recognize

different measures together (e.g., the percentage of participants who fixated an AOI, together with time until first fixation) as only a single measure may give a misleading picture (e.g., a short duration until first fixation can indicate high ease of finding a target, but this would not be true if only a small percentage of overall people actually fixated on the target in the first place.

The number of fixation measures are often correlated highly with time (e.g., number of fixations before fixation correlated with time before fixation, number of fixations correlated with total duration of fixations), hence it can be useful to report both figures, or sometimes just the duration figure as stakeholders find it often easier to understand the concept of ‘looking time’ rather than number of fixations (as these depend more on parameter settings). Total fixation duration (sum of all fixations in an area) can indicate attraction but a higher average fixation duration also indicates increased cognitive processing. When comparing two areas or two designs for attraction it is necessary also to determine that the average fixation duration do not differ significantly.

As discussed above although pupil dilation is a potential source of data and insight for the UX researcher, difficulties with controlling light conditions, and interpreting the cause of dilation often preclude its use in a practical perspective.

21.6 Performing a Usability/UX Test

There are many factors that need to be considered in order to execute an effective usability or UX test, and indeed a number of good books have been written on the subject (e.g., Bojko, 2013; Nielsen & Pernice, 2009). In this section, we introduce and discuss only the major considerations for lab based tests and how eyetracking fits into these.

Box 2: UX/usability tests compared to other behavioral research

A UX/usability test is an assessment/evaluation of the experience that a product/service provides by testing it with representative users (test participants). Typically test participants will attempt to complete relevant and representative tasks using the product/service while researchers observe and measure their behaviour. Although UX/usability tests share similarities with other behavioural research, there are differences:

- The focus is on observation of participant behaviour *not* opinion, prediction of future behaviours, or description of past behaviours (e.g., unlike many surveys, or interviews): Though researchers need to listen to their participants, the data need to be qualified and interpreted in the context of behaviours exhibited. Critically, what they do is far more important than what they think they might do, or their opinion about the design (and researchers should discount any design solutions offered by participants).
- The context of the observation/measurement of behaviour is known *rather than* inferred (e.g., unlike in web site traffic analytics): In UX/usability sessions information such as the participants' background knowledge, their expertise, and importantly their specific goals are known and/or controlled such that data can be evaluated with reference to a specific context (which of course includes the environment).
- Tests should feature controlled and structured tasks—*not* any task (e.g., unlike in ethnography): In a typical lab test, specific tasks are set for participants, which can give the testing a very similar structure across participants and therefore allows for comparative testing/accurate measurement of performance etc. This type of structured testing is necessary for meaningful quantitative results from eyetracking data.

21.6.1 Roles of People Involved with Executing a Usability/UX Test

The following roles are often involved in the execution of a usability/UX test. Sometimes one person may play more than one role (e.g., sometimes a single person may recruit, moderate, and be the lone observer for the test).

- **Product/service stakeholder**—the 'owner' of the product or service under test—they should be the ultimate beneficiaries of the improvements gained from acting on insights obtained during the testing. They will often provide domain knowledge and assist with setting the research and business objectives for the test.
- **Test moderator/facilitator**—a trained and experienced researcher who will interact with the test participant and ensure the test runs smoothly, as well as plan and analyze the sessions. The behaviour of the moderator towards the test participant is of critical importance to the success of the test. They must support and encourage the test participant to perform the tasks required, whilst also remaining objective and non-leading.
- **Test participant**—a person or persons who are representative of the real users of the system. Test participants have to meet certain criteria such as demographics, level of experience with the product or similar systems etc.
- **Participant recruiter**—a person or an agency tasked with finding, screening, and scheduling suitable test participants to take part in tests.
- **Design team**—people who will act on the results of the test; often it is their designs that are specifically under test. It is generally a good idea for members of the design team to attend and observe testing, but they should not moderate the testing of their own designs.
- **Observers/researchers/analysts**—trained researchers in addition to the moderator/facilitator who will assist with planning, observation, interpretation and analysis of results. They may also play a role in explaining events to stakeholder and design team observers in order to avoid possible mis-interpretation of events.

21.6.2 The Usability Lab

Many UX/usability tests are performed in specialist usability labs or studios. These allow observation of test participants' behaviours under controlled conditions. A usability lab typically consists of two rooms; a testing room where the test participant and facilitator/moderator execute the test, and a separate viewing room, which permits a number of stakeholders or researchers to unobtrusively observe the test in progress. Traditionally the two rooms were adjacent and separated by a large one-way mirror, although nowadays many labs forego the potentially imposing one-way mirror and instead use audio-visual equipment to transmit the activity in the test

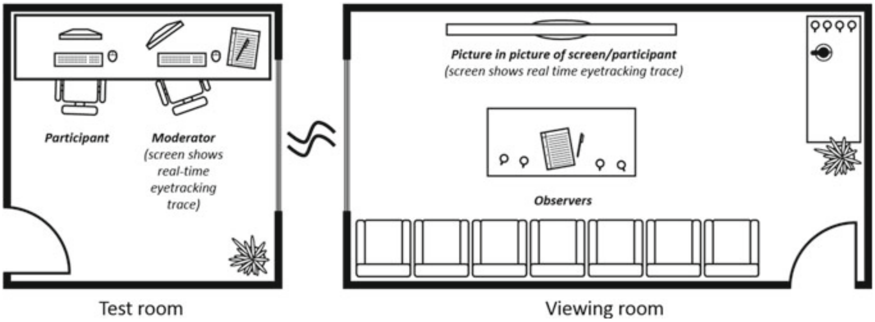


Fig. 21.3 Schematic of a usability lab. The rooms may be adjacent (with or without one-way glass) or remote. The moderators screen will typically show real-time eyetracking trace which is also transmitted to observers

room to observers who may be in an adjacent or remote room, or even in several different locations.

Although UX/usability research is certainly not constrained to the lab, often eyetracking equipment functions best in a more controlled lab-environment where factors such as lighting, and the participants seating position can be controlled. Ideally the lab should be set up to resemble the environment that the test participant would be in if they were to use the product under test in real life. For example, many labs have ‘office’ or ‘domestic’ configurations, some are set up to resemble living rooms. Figure 21.3 shows a schematic of a typical usability lab set-up, Figs. 21.4 and 21.5 are photos from real test rooms and observation rooms respectively.

21.6.3 Pre-test Considerations/Activities/Planning

The key to success in any research is planning. Good planning should ensure that overall the combination of tests deliver actionable insight on appropriate research goals. Good planning ensures that each test itself is executed smoothly, appropriate data are collected, and therefore analysis and reporting is as efficient as possible.

Below we discuss several considerations that would need to be considered as they determine how a test can actually be performed and indeed whether eyetracking should play a role (see also Sect. 21.5).

21.6.3.1 Research Goals and Research Questions

It is essential that there is a specified overall goal or objective for each test before data collection begins, and that then this goal is broken down into a number of appropriate research questions. The goal and research questions will drive the planning for how the test should be executed and analyzed.

Fig. 21.4 Test room. The moderator (left) sees a picture in picture representation of screen and participant (right). The eyetracker is the small box on the lower bezel of the monitor

Typically research goals are about either (1) detecting issues or areas for improvement (formative research), or (2) comparing alternative designs to see which performs best (summative research). Sometimes it can be a combination of both.

Example 1

Overall objective: To identify issues that could lead to improvements in the overall experience that an ecommerce site supports.

Type of study: Formative study.

Examples of research questions:

- Does the landing page convey the purpose of the site?
- On the landing pages and product pages do the users notice key persuasive messages? (e.g., depending on the page these might be social proof, trust indicators, scarcity messaging etc.).
- On product pages do users interact with the page in the appropriate order (e.g., this might be select size, select color, add to basket), if not why not?

Example 2

Overall objective: Compare a new product listing design to the current design.

Type of study: Summative study (with also some formative components).

Fig. 21.5 Viewing room. This is an example of a remote viewing room. Observers view the screen and a separate feed of the participant's facial expressions

Examples of research questions:

- What are the key information items that people seek in order to differentiate between products?
- Does the new design make it easier to find critical information?
- Does the new design make it easier to understand and differentiate between critical information?
- Does the new design make it quicker to navigate to product details screens?
- Which design do potential customers prefer, and why?

One can see that some of the above research questions can be informed by eyetracking (those to do with noticing information, interacting with and understanding elements) but the observations must be supported by other observations available during the test session (e.g., actions, spontaneous verbal utterances, discussion at the end of tasks) to get a complete picture. In addition, many questions will not be informed by eyetracking. Those questions that can be informed by eyetracking and the appropriate observations and analyses should be identified in advance. For example, stakeholders and the research team may define the areas of interest or expected scanpaths, and make predictions about expected eyetracking patterns and measures.

21.6.3.2 Number of People to Test with

Despite usability testing being established for many years the appropriate number of people to test with is still a matter of relatively hot debate. Several books have whole chapters devoted to the subject (Bojko, 2013; Rubin, 2010; Sauro & Lewis, 2012), and go into great detail, elaborating theoretical and empirical arguments. They also present formulae that can be used to calculate the number of participants required to support the likelihood of detecting a certain percentage of issues, or deliver metrics with certain degrees of certainty. We cannot go into such detail here but encourage readers to investigate this issue more in depth.

From a practical perspective the major distinction that determines the number of participants is whether the study is formative or summative. For formative studies a usual practice is to test between three and six people per specified user group (sometimes a persona type, that is based on a category of typical user e.g., see Thoma & Williams, 2009) with the number of people per user group generally decreasing as the number of user groups increases (e.g., 6 people if only 1 user group tested, 2 × 5 people for 2 groups, 3 × 4 people for 3 groups etc.)

For summative studies additional considerations apply including:

- Task complexity
- Complexity of interface, and differences between the alternatives
- The reason for the measurement (absolute measure or to compare to alternatives)
- Number of alternatives to be compared
- Whether there is a need for within subjects or between subjects design (for cost reasons most studies are performed as within subjects designs, where people interact with multiple alternatives of the design within the same test session—this requires balancing or pseudo-randomizing of the order of designs that different participants see in order to attempt to control for possible effects due to the order that the participants see the alternative designs)
- Degree of certainty or statistical significance required.

In practice—despite theoretical ideals—most summative studies test between 20 and 80 participants. If possible you should use formulae to predict the number of participants required for a study, often though you will need to work within constraints of available resources (e.g., time and money) which will determine the number of participants you can possibly work with and use the theory to set your expectations for the testing (see e.g., Sauro & Lewis, 2012 for useful tables and formulae). Needless to say, if your study does not require summative testing then you should avoid it and rely on lower number formative studies (ideally as part of an iterative test-improve-test cycle). It is important to involve stakeholders in these deliberations as it will have repercussions for later interpretations (and permissible generalizations) of results.

21.6.3.3 Material/Stimulus that Can Be Tested

The material that should be tested is somewhat determined by the test stage and the research aims—but in principle anything can be tested from paper prototype concept sketches through to implemented finalized software/sites. Often eyetracking is more useful the higher fidelity (nearer finished) a design is, but this does not preclude testing things like black and white wireframes, or wireframe prototypes and getting useful qualitative results from eyetracking. Quantitative (summative) studies are only generally performed on finished or near finished high fidelity designs.

21.6.3.4 Concurrent Thinking Aloud Versus Retrospective Protocol

The majority of usability/UX tests (especially formative tests) are performed using a concurrent think aloud (CTA) protocol (Dumas & Redish, 1999; Ericsson & Simon, 1998). During CTA the participant is asked to verbalize their thought processes as they complete the various tasks required. CTA is used because it allows observers access to the participant's underlying thought processes, and thus provides richer qualitative information which might be useful in explaining why the participant behaved in the way that they did. For example, the verbalizations may reveal misinterpretations that lead to usability issues, or underlying mental models that drive the way a participant interacts with an interface. The problem with the use of CTA is that it may require the participant to behave unnaturally in that most people when they interact with interfaces do not introspect and verbalize (e.g., Klein et al., 2014) so using CTA changes the way people interact with an interface (e.g., they may be more cautious or hesitant, some people report feeling 'silly'), and therefore also changes measurements of behavior such as time on task (generally longer), or eyetracking measures (people tend to look at what they are talking about).

An alternative to the use of CTA is to use a retrospective protocol (RP, also called retrospective thinking aloud, RTA), which specifically does not ask participants to verbalize their thought processes concurrently but just lets them get on with tasks as they normally would. Sometimes a moderator might be present, sometimes they may leave the room entirely. The advantage of this approach is that performing the tasks is more natural and therefore less likely to be changed by the condition of the test compared to CTA (Eger, Ball, Stevens, & Dodd, 2007; Guan, Lee, Cuddihy, & Ramey, 2006) but the disadvantage is that it loses some of the potentially rich qualitative information that may explain behavior.

To gain this information with RP, at the end of completing tasks users are replayed their performance (this is often a screen recording) and asked to provide a commentary on what they did and why. Including the participant's eyetracking trace with the replay (sometimes called the post experience eyetracked protocol or PEEP) can enhance the commentary as it provides an additional cue to help guide recall (e.g., Eger et al., 2007; Elling, Lentz, & de Jong, 2011) but there is an ongoing controversy (e.g., Alshammari, Alhadreti, & Mayhew, 2015; Willis & McDonald, 2016) about the effectiveness and validity of CTA versus RP/RTA. Sometimes a hybrid of the two techniques are used in that participants are not asked to think aloud during normal completion of tasks but a skilled moderator will view real-time eyetracking and use that and other observations to identify points of interest to probe during follow up questioning or replay. The major disadvantage of using RP or the hybrid technique is that test sessions are generally significantly longer, and require skilled facilitators.

21.6.3.5 Other Measures Questionnaires, User Test Scores

In addition to the use of eyetracking and recording user observations and comments, the UX team will usually gather quantitative data relating to the users' experience in form of questionnaires. Depending on the stage of the product being tested (e.g., early prototype, or almost finished product) the type of questionnaires may vary. From the observations (or video recording) of user behaviours the team may quantify data such as spontaneous positive versus negative comments, types of facial expressions, hesitations to click on certain elements, etc. (for a good overview as wells an in-depth description of usability metrics, see Tullis & Albert, 2013).

As a rule of thumb, the team may also include simple measures of perceived effectiveness (to what degree did the user think they achieved set tasks or goal completion), perceived efficiency as well as objective efficiency (was the task or goal achieved in a reasonable amount of time or effort), and users' satisfaction. This can be done by using simple Likert-type rating scales (participants have to rate on a subjective scale from 1 to 5, e.g., “How efficient could you achieve task completion using this interface”). Usually it is recommended to use some freely available (or paid for) standard questionnaires that were carefully developed to measure these dimensions of psychological experiences reliably (e.g., SUS, QUIS, WAMMI; and its relationship with the Net Promoter Score, NPS; see Sauro & Lewis, 2012).

21.6.4 Testing

There are certain steps in preparing—as described above—preparing, executing, and analyzing a usability test, and these are summarized in Fig. 21.6.

21.6.4.1 Pre-test Planning

The research team meet with stakeholders to confirm the purpose of the testing, and gather any pertinent background information such as the business goals, information on the target audience, previous research or analytics, technical or brand constraints. Research goals are then formulated, specific research questions listed, sample size and the characteristics of the target audience groups for the test identified. The approach for the test will also be decided e.g., summative vs. formative, concur-

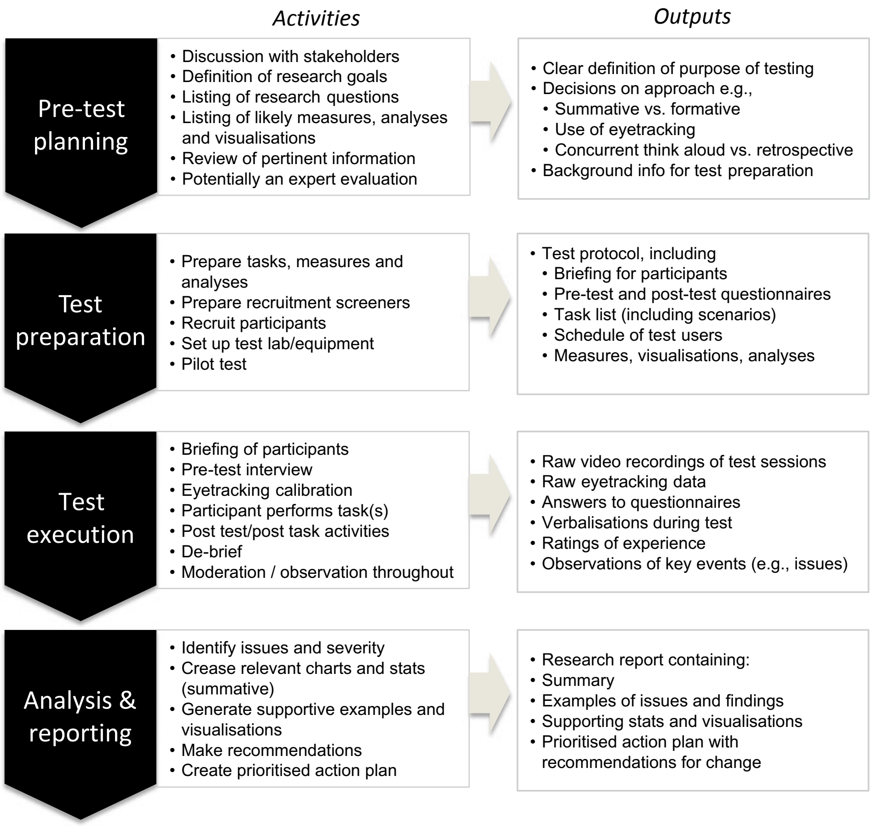


Fig. 21.6 A brief overview of the activities involved in carrying out a usability test

rent think aloud vs. retrospective protocol, and the use and purpose of eyetracking, as well as the target measures and their analyses.

An expert evaluation of the site or system may also be performed, in which a UX specialist performs an assessment using established usability and UX principles (e.g., Nielsen & Molich, 1990). The expert evaluation may identify issues that can potentially be fixed before testing takes place, and also helps give the researchers in-depth familiarity of the system which assists with planning and executing the test.

21.6.4.2 Testing Execution

Test protocol

Test preparation often involves producing a detailed test protocol which specifies all aspects of the test and guides the moderator/facilitator and test observers. During preparation of the protocol the research objectives and specific questions are reviewed and confirmed, tasks for the participants are formulated, and all other aspects of the test are detailed. A protocol often includes the following:

- Background and objectives
- Specification of materials to be tested (e.g., URLs of live sites, version of prototypes etc.)
- Profiles of participants to be recruited for the testing (subsequently modified to show the profiles of the actual people taking part in the testing)
- Participant briefing script (used so that each participant receives exactly the same information to avoid potential bias)
- Pre-test and post-test questions
- List of tasks and scenarios
- Notes specifying exactly how the test will be run (e.g., task order, procedural aspects such as operation of the eyetracking equipment, when to calibrate etc.)
- Notes to guide the moderator and observers relating to the research questions and tasks
- Participant de-brief information.

Other materials such as non-disclosure and consent forms, observation/logging sheets, and participant recruitment screeners are also produced.

For tests involving eyetracking it is important to specify the role of eyetracking (formative vs. summative, e.g., uncovering usability problems or ignored interaction areas, vs. benchmark testing or optimization) and also to identify the post-test eyetracking analyses (if any) that will be performed. It is important to re-iterate with the team and stakeholders that these need to be mapped to the research goals and research questions.

Setting tasks

Much has been written in other texts (Dumas & Redish, 1999; Rogers, Sharp, & Preece, 2011; Rubin, 2010) about how to set appropriate tasks for usability and UX testing, we will only summarize some of the main considerations here. The nature and details of the tasks will be informed by other research stages, such as a task analysis or other requirements engineering processes (see, e.g., https://www.usability.gov/how-to-and-tools/methods/index.html), and in discussions with stakeholders. It is important that the tasks will produce behavior that will inform the research objectives. For quantitative (summative) eyetracking studies this often means ensuring exactly the same task conditions (e.g., starting position, scenario, task goal, task end point) for each participant to ensure the data are collected over the same pages/views. For qualitative (formative) studies tasks can be more varied, sometimes informed by each particular participant's personal circumstances, for example rather than asking everyone to '[starting from the home page] book a flight from London to Barcelona for yourself next Tuesday' the moderator might discuss with the participant their particular upcoming travel needs and ask them to complete the following task: 'You mentioned that your next trip will be for business and is from Birmingham to Brussels on the 12th of next month—please try to you use the site to complete a booking that will meet your needs'. Generally tasks should:

- Be as natural and relevant to the participant as possible
- Be appropriate for the site/system under test
- Be provided with appropriate context (sometimes this is provided with a scenario)
- Not lead or prime the participant (e.g., they should avoid guiding them to a specific feature, or give clues about the steps involved)
- Have a defined end point
- Be achievable (although this is not absolutely necessary, having many tasks that will definitely result in failure can have a negative effect on the participant)
- Produce interaction/behavior with the site/system and not seek opinion.

Participant recruitment

Tests should be performed with people who are representative of the real user population of the site or system (so not just who is available in the office). During planning and preparation, the profiles of the appropriate user groups will be identified, and from this a screener which lists required characteristics for people to take part in the testing will be produced. Screeners typically define requirements for participants which could be related to their behavior (e.g., active users of social media), their job role (e.g., project managers), their demographics (age, gender etc.) and potentially many other factors such as attitude, experience, geographical location, use of competing products etc. Often there are clear primary characteristics that define individual participant groups (e.g., job role) and other characteristics might be balanced within groups (e.g., gender if that is not a defining property of the group). Specialist fieldwork recruitment agencies exist that can assist with recruiting test participants from databases of volunteers. Care must be taken not to recruit people that frequently take part in research (‘professional usability test participants’), or people that have a stake or role in the project (e.g., your known colleagues or stakeholders)—test participants should be real and representative.

For eyetracking studies, it is important to ask people about their vision and use of eyeglasses. Some eyetrackers have difficulties with certain types of spectacles or coatings on lenses, and so people wearing these should be avoided. In recent years, modern eyetracking technology has become much more robust, such that only a small proportion of people cannot be tracked (for reasons of eye color, shape of eyes, use of spectacles etc.; see chapter by Hutton chapter, in this volume).

Pilot testing

It is essential to include pilot testing as part of the preparation. This ensures that the flow of the test is appropriate, the tasks are understandable and appropriate, the technical set-up works, and the outputs are appropriate for analysis etc. Ideally the pilot test should be with an appropriate representative participant, but often for time and budget reasons might be someone naïve to the project but from a neighboring office. The results of the pilot test are used to make amendments and improvements to the protocol.

Test sessions

A user test schedule involving eyetracking has this general structure (which of course may be adapted depending on the project, see also Fig. 21.5):

Just before the testing session:

1. Ensure technology set-up is appropriate and working, and all conditions are re-set to a common base level (e.g., reset browser history).
2. Greet participant and show to testing room, introduce to test facilitator.

At the start of the session:

3. Participants are briefed about general goals and nature of tasks and data recording, they are invited to ask questions.
4. Participants are put at ease regarding their contribution and are asked to read and sign consent forms (and sometimes non-disclosure agreements). It is usually stressed that the test is not a test of them in any way and that they should try to behave as naturally as possible, they are free to stop the test at any time. The skill of the facilitator is not to involuntarily induce any behavior by the participants that they would not do by themselves.
5. A pre-test questionnaire/interview is conducted to gain insight into the participant's background, their attitudes and experiences, and any other relevant information that may have a bearing on the test or their use of the site/system. Often information gained at this stage can be used to help define the specifics of appropriate tasks in order to make them as relevant and natural to the individual as possible.

Execution of testing:

6. The eyetracker is calibrated, sometimes a trial task may follow, which may not be part of the formal data collection but rather designed to get the percipient used to the form of the testing and put them at ease. For example, this might be a simple task, such as the use of a search engine, that the user will find easy, and during which they will 'practice' aspects such as thinking aloud (if that is the protocol selected for the test).
7. Participants perform tasks, while their eye movements are tracked (as well as video and audio recorded). Sometimes instructions for each task might be given verbally, sometimes they may be given on screen, sometimes on cards. A disadvantage of giving instructions on cards especially for eyetracking tasks is that some participants may frequently look at the task card for reassurance and so break their gaze (and it also potentially adds an unnatural condition to the behavior) so unless necessary it is often better to avoid tasks on cards that will be present throughout the task. Whatever form of task instruction is given care should be taken to ensure the participant has understood the task before they commence, and that the instructions are not unduly influence the otherwise 'natural' looking behavior.

8. For some tests there may be rating, questionnaire or discussion following each task. These are designed to gain further insight and potentially measures of effectiveness, efficiency and satisfaction. When using the retrospective protocol (RP) the participant may be invited to review their actions using the eyetracking trace as an additional memory prompt.

On completion of test tasks:

9. After all the tasks have been completed a post testing interview is often conducted to help clarify the users' experiences, in addition users may be asked to fill out post-task questionnaires. If a retrospective protocol has been used and this was not performed following each individual task then this will be performed on completion of all tasks.
10. Participants are debriefed about the research project.
11. Participants are thanked and reminded of any non-disclosure agreements and given their reward/incentive for participating.

Box 3: Tips when using Eyetracking

Even if you have already some experience with using an eyetracker or running usability tests, combining both methods creates new challenges. Here are some important points to keep in mind:

- Although the eyetracker needs to be calibrated and it is good practice to explain what it is for, drawing more attention than is absolutely necessary to the eyetracker should be avoided. The goal is to make the conditions of the test as natural as possible and avoid the white coat effect as much as possible.
- During eyetracking testing sessions the aim will be to try to keep the participant eyetracked throughout. This can be influenced by aspects such as the chair that the participant sits in (avoid chairs with wheels or that rotate, use adjustable desks), where the moderator is sat in relation to the observer (usually a short way away and slightly behind not in the immediate vision), how tasks are given (visible task cards can cause disruption), and several other factors including the natural behavior of the participant—some participants sit still whilst some are very animated and may lean in and out and point at the screen etc. A skilled facilitator learns to control these factors (now often assisted by the eyetracking software) as much as possible and to subtly encourage the participant to regain tracking if they drift out without drawing too much attention to the eyetracking equipment itself.
- If during a task eyetracking is lost then the facilitator has to decide whether to intervene in order to restore the eyetracking. Obviously, this is more vital on quantitative eyetracking studies, but may be less of a concern if eyetracking is only used as an aid to observation and interpretation. Most eyetracking

software now offers tools that helps monitoring of tracking accuracy. As for everything the research goals determine the appropriate course of action.

21.6.4.3 Analysis and Visualizing Eyetracking

General Considerations

For formative studies, observations and notes on pertinent events and incidents are made in real time by observers (and sometimes also the facilitator). Sometimes events are coded electronically using specialised logging software which includes time-stamping of events, some observers prefer to use paper, or spreadsheets or annotate screenshots of the design under test. When there are many simultaneous observers post-it notes are often used so that individual observations can be collated and analysed together. The choice of how to log observations is somewhat dependent on the test conditions and sometimes also personal preference. Playback of video is often used just to check specific observations or to produce compelling illustrations of particular issues or events. Whilst the whole session recording can be played back for post-session observation it is time consuming and often of little additional benefit.

When all test sessions are complete the observations from each session are combined and collated and analysed. There are a number of techniques for performing this (such as affinity-diagramming using post-it notes, see Kuniavsky, Goodman, & Moed, 2012) with the aim to identify common themes, behaviours, issues, positive aspects etc. Observations such as usability issues are assessed for severity and impact, ultimately feeding into recommendations leading to a prioritised action plan for change.

Visualisations (see below, next subsection) may be produced to illustrate certain findings but care should be taken not to misinterpret or invoke misinterpretation of visualisations such as hot spots particularly when they are based on low numbers of participants. Further to consider is that if tests sessions have different durations this can skew data so it is important to potentially use relative duration instead.

In summative studies often the same type of real-time observation will occur but also post-test analysis of the quantitative data will produce comparative measures of behaviour such as task success, satisfaction scores and relevant gaze data (such as number of fixations on a certain Area of Interest—see Sect. 2.5). Additional metrics (questionnaire data, counting of errors etc.) also are analysed in this type of study (see Tullis & Albert, 2013 on how to use, analyse and present metrics).

Reporting

Heat Maps

Heat maps (e.g., Fig. 21.7) are colourful overlays on the study material (e.g., a webpage) to indicate the average overall amount of fixations that participants made (or for how long they fixated). It's the most often used and a well-known visualisation technique, because it is fairly easy to understand: Red areas (traditionally, you can change these of course in the settings) indicate where people looked most, and variations in shading relate the degree to which people did that (the darker the red, the more fixations); blue areas (or when no colours are overlaid) often mean that people looked there least (or not at all). The advantage of heat maps is that they are a straightforward method to communicate results to clients. A disadvantage is that differences between individuals are smoothed over, but it is in the discretion of the researcher to correct for this, e.g., by producing different heat maps for different task scenarios or stimuli. Other considerations are important when producing and interpreting heatmaps: For example, because of certain looking biases most people look first and hence more often in the centre of a display, or the setting of correct parameters (such as adjusting kernel size for heatmaps depending on users' distance from screen and screen size).

Fig. 21.7 Gaze opacity and heatmap ('hotspot') visualisations from 5 example participants from a single target test group, these illustrate that some key content areas were being engaged with in the way intended. The final figure shows 2 representative gaze plots

Opacity or Focus Maps

These are variants of heat maps (Fig. 21.7), with the main difference being that they use a transparency gradient (rather than colour) to indicate the main area of eye fixations. A dark area indicates low levels of measures, whereas opaque/visible areas indicate high levels of eye fixations.

Gaze Plots

Gaze plots are a visual representation of fixations (like heat maps) but also indicates saccades (eye movements). Fixations are usually depicted as points or circles over the visual environment, and saccades (in idealised form) are represented as lines connecting them. The fixation circles are often numbered indicating the point in sequence they were visited, and the number of circles and circle size can be used to indicate the number and/or duration of fixations. Gaze plots are therefore more appropriately used for representing an individual's looking behaviour in space and time (and therefore show the 'journey' the eyes have taken). Gaze plots may also help you to identify outlier behaviour (e.g., some fixations of an individual that are very untypical). The disadvantages are that this type of visualisation is harder to use for groups of participants, and that it may give a false impression for certain circumstances, e.g., repeat visits of a webpage may yield different results.

Bee swarm analysis

The bee swarm visualization is a replay of gaze behavior of more individuals. Like in the gaze plots small disks indicate where people looked in the display, and the density indicates the frequency of looking. However, there is no information on sequence order, but indicates consistency in looking behavior (as long as all participants saw the same stimulus), with different colored disks indicating different participants (or participant groups, e.g., regular vs. irregular users).

Mixed visualisations

A useful tool for summative studies is to provide a type of visualization such as hotspots overlaid with statistics (e.g., percentages of fixations) from quantitative analyses of gaze data. Visualisations such as hotspots can also be useful to augment graphs and charts of gaze data to provide some visual context concerning the measures for the different designs under test. It is also worth mentioning here that different eyetracking software solutions will have different or additional variations of visualizing eyetracking measures.

21.7 Case Studies—Eyetracking in Digital Marketing

In the following section we illustrate some points made in this chapter based on research done at the UX agency Bunnyfoot (UK).

21.7.1 Eyetracking as an Insight Tool

For this retail client, eyetracking was used to test their brochure and lead generation-site during key stages of its design cycle, namely early low fidelity prototype stage and final visual design. During early prototype testing the principle use of eyetracking was to assist with observations and interpretation of issues, but it also helped confirm key regional (international) differences with respect to the type of product information customers in those different regions were looking for. For example Dutch and English customers when viewing the same design (albeit translated into the appropriate language) and performing the same task, focused their spontaneous visual search on different product attributes, and Dutch customers were less likely to engage with social proof messaging and support information. This lead to modifications of the designs for each country to promote the salience of the desired product attribute information, and to alter the way persuasive messaging through social proof was presented.

In follow-up testing we used the high fidelity visual design of the main product listing page for the same client. Eyetracking was used to help determine the form and placement of the link to detailed product information. The design intent here was an interesting one aiming to:

- encourage browsing across all listed products (to communicate sufficient flexibility and choice)
- encourage engagement with key persuasive messaging integrated with the listing (authority/trust messages, and answers to key concerns delivered via social proof).

Thus the placement of the link to more information for each individual product was non-standard as the designers did not want it to be the most salient and natural thing to focus on and respond to. Instead, they wanted it to be discoverable if desired. Several forms and placements of the button were made and tested, Fig. 21.7 shows the one that was found to achieve the intended result. Eyetracking (along with other observations) confirmed that the people did engage with the content as required.

21.7.2 Eyetracking Comparison of Potential Homepages

In a series of comparative evaluations of alternative potential homepages for a price comparison site 24 participants were tested, using a within subjects design.

One of the aims of the testing was to determine which page functions best in communicating the range of items on offer for unfamiliar customers of the site. One input to this was analysis of eyetracking for the first 20 s of engagement with the alternative designs (this was correlated with other activities, e.g., a post showing recall test) (Fig. 21.8).

21.7.3 Eyetracking Investigation of How Consumers Interact with Search Listings

In 2013 some research conducted for a car insurance client into the effectiveness of their Google Adwords. Observation of behavior and eyetracking analysis showed that many of the participants actively engaged with and indeed clicked the items (the paid advertisements) at the top of the listings and follow-up questioning found the surprising result that about 40% of people (from 100 tested) did not realize that Google Adwords are actually advertisements.

This research was repeated in 2014 under more controlled conditions, and tested the effect of Google's updated design for denoting whether an ad was an ad (the newer design uses the word 'Ad' in a bright yellow box). The results were similar with 36/108 people tested not realizing that the GoogleAdvert links were in fact adverts rather than search findings.

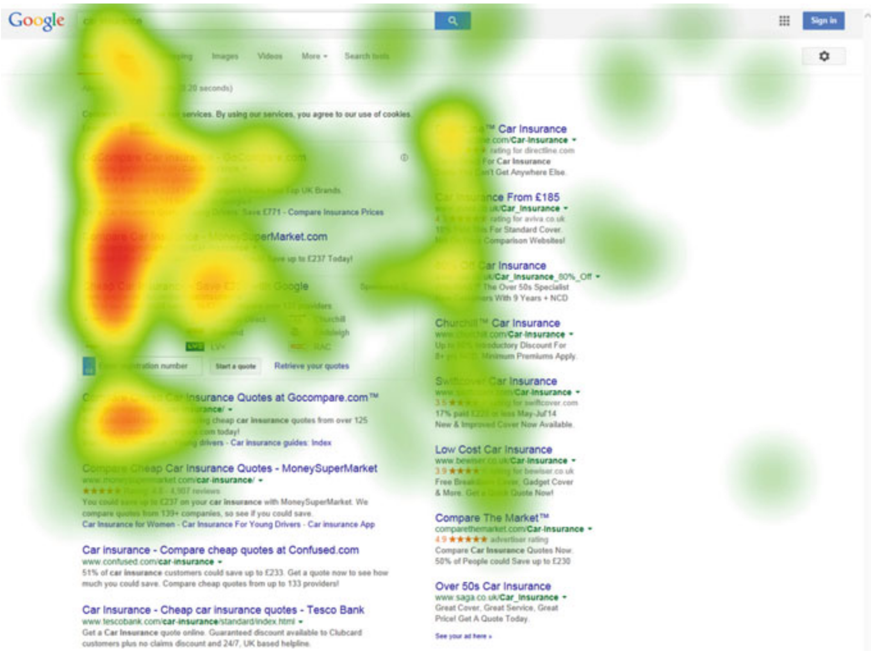
The study showed a clear difference for the way people engage with search engine result listings depending on their prior knowledge and expectations. The clearest result of which is engagement with the paid for advertisements. This is illustrated in Figs. 21.9 and 21.10.

21.8 Outlook and Conclusions

In this chapter we have introduced eyetracking as an applied research tool and how it can be used in evaluating web usability. It should be stressed that—like other methods in UX research—eyetracking should be treated as only one method among many, and is best seen as a way to tri-angulate the validity of findings in combination with other research data (observations, behavioral user metrics, questionnaire scores). However, eyetracking can contribute information that is difficult to obtain otherwise, such as which areas or design elements draw attention at which point in time.

There are necessary considerations when eyetracking is to be used. It demands skill and expertise in planning, execution, and analyzing the data, and can add considerably to the costs (and time) of UX projects. Eyetracking research has to be adapted to the relevant research questions and objectives and the testing environment. Therefore, the UX team employing eyetracking not only needs to be trained in the relevant technology and statistical methods, but also apt in advising their clients

Fig. 21.8 Hotspots illustrating average gaze behavior for the first 20 s exposure to alternative homepage designs. The first design was dominated by a ‘hero’ category (at the top of page) which took attention away from the rest of what was on offer. In the second design participants viewed the main menu categories and more in-page categories (products and services) and this viewing behavior correlated with their recall of what was on offer on the site



Car insurance - Compare cheap quotes at Confused.com
www.confused.com/car-insurance ▾
51% of car insurance customers could save up to £233. Get a quote now to see how much you could save. Compare cheap quotes from up to 133 providers!

Car Insurance - Cheap car insurance quotes - Tesco Bank
www.tescobank.com/car-insurance/standard/index.html ▾
Get a Car Insurance quote online. Guaranteed discount available to Clubcard customers plus no claims discount and 24/7, UK based helpline.

Compare The Market™
comparethemarket.com/Car-Insurance ▾
4.9 ★★★★★ advertiser rating
Compare Car Insurance Quotes Now.
50% of People could Save up to £230

Over 50s Car Insurance
www.saga.co.uk/Car_Insurance ▾
Great Cover, Great Service, Great Price! Get A Quote Today.

See your ad here »

Fig. 21.9 Gazeplot from the 36 people who did not realize that the top items in the google listings were advertisements

as to whether eyetracking makes sense for given research goals. However, with the costs of eyetracking equipment predicted to come down significantly in the near future, the range of its application will broaden.

We have not covered all topics related to eyetracking research in web usability, for example, eyetracking with mobile gadgets such as smart phones and tablets. However, most of the recommendations in this chapter will generalize to these devices as well (see, e.g., Cheng, 2011).

The application of eyetracking in UX research is evolving with technology, and new measures can now be added to provide new insight. For example, biometric measures such as skin conductance and face (emotion) recognition (via cameras) can supply information on users' state of arousal and their feelings. Combined with eyetracking data this will allow researchers to better determine, for example, whether certain visual elements are the cause of attention for aroused participants or users experiencing positive or negative emotions, or whether these states are the consequence of what users just saw.

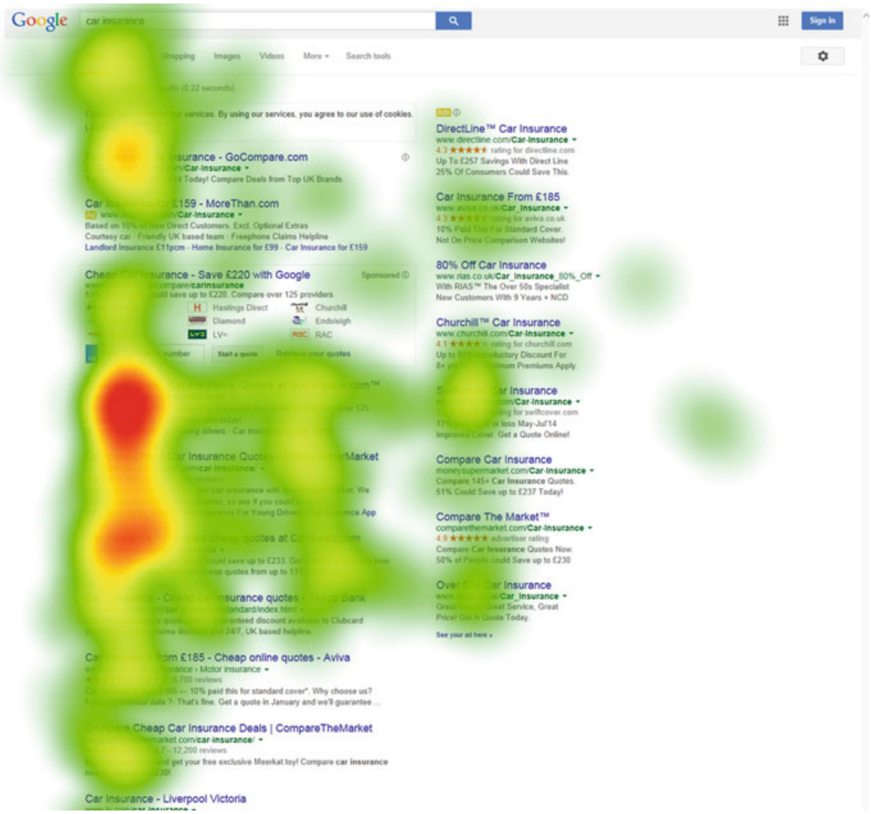


Fig. 21.10 Gazeplot from the 77 people who did know that the top items were advertisements

21.9 Suggested Readings

Bojko, A. (2013). *Eye Tracking the User Experience: A Practical Guide to Research* (1st edition). Brooklyn, New York: Rosenfeld Media.

– *A good hands-on guide, that is also very comprehensive.*

Nielsen, J., & Pernice, K. (2009). *Eyetracking Web Usability* (1 edition). Berkeley, CA: New Riders.

– *One of the first books on eye-tracking in UX—still worthwhile as a good introduction, and giving lots of examples.*

Norman, D. A. (1988). *The Psychology of everyday things*. New York: Basic Books.

– *A classic that paved the way for most of the industry today. Its findings and principles are still valid today. A must read.*

Sauro, J., & Lewis, J. R. (2012). *Quantifying the User Experience: Practical Statistics for User Research.* Amsterdam ; Waltham, MA: Morgan Kaufmann Publishers Inc.

– *Covers very comprehensively the statistical and sampling aspects that need to be considered.*

Cooper, A., Reimann, R., Cronin, D., Noessel, C. (2014). About face—the essentials of Interaction Design. Wiley, New Jersey.

– *A great book for in-depth understanding of how to design software and systems that work at a human level—reasonably advanced.*

21.10 Questions for Students

1. What is usability, and how does it differ from user experience?
2. A client is developing a commercial website and shows you an almost finished prototype/beta version. They ask you how eyetracking can improve it—what do you answer?
3. What are the advantages and disadvantages of "thinking aloud" protocol in an eye-tracking study? When would you recommend it?
4. Why should the user test team be preferably not be directly involved with the design and development of the product (or service) that needs to be tested?
5. You need to test whether people are confused by certain page-sections or navigation items. Which eye-tracking metrics could help you gather insight on this issue?
6. What are the pitfalls of relying too much on "heatmaps" when interpreting results from eyetracking research?
7. What are the main advantages of involving eyetracking in a usability test?
8. What are the main considerations in planning a usability test with eyetracking?

References

Alshammari, T., Alhadreti, O., & Mayhew, P. (2015). When to ask participants to think aloud: A comparative study of concurrent and retrospective think-aloud methods. *International Journal of Human Computer Interaction, 6*(3), 48–64.

Bias, R. G., & Mayhew, D. J. (2005). *Cost-justifying usability: An update for an internet age.* Burlington: Morgan Kaufmann.

Bojko, A. (2013). *Eye tracking the user experience: A practical guide to research* (1st ed.). Brooklyn, NY: Rosenfeld Media.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). Psychophysiological responses to changes in workload during simulated air traffic control. *Biological Psychology, 42*(3), 361–377.

Bruneau, D., Sasse, M. A., & McCarthy, J. D. (2002). The eyes never lie: The use of eye tracking data in HCI research. In *Proceedings of the CHI* (Vol. 2, p. 25). Retrieved from http://hornbeam.cs.ucl.ac.uk/hcs/people../documents/Angela%20Publications/2002/CHI_Physio.computing_Final%20(1)_revised.pdf.

Card, S. K., Newell, A., & Moran, T. P. (1983). *The psychology of human-computer interaction*. Mahwah, NJ: Erlbaum.

Chen, M. C., Anderson, J. R., & Sohn, M. H. (2001). What can a mouse cursor tell us more? Correlation of eye/mouse movements on web browsing. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (pp. 281–282). New York, NY: ACM. https://doi.org/10.1145/634067.634234.

Cheng, S. (2011). The research framework of eye-tracking based mobile device usability evaluation. In *Proceedings of the 1st International Workshop on Pervasive Eye Tracking & Mobile Eye-Based Interaction* (pp. 21–26). New York, NY: ACM. https://doi.org/10.1145/2029956.2029964.

Clients—What our clients say about us and our work. (n.d.). Retrieved May 10, 2015, from http://www.bunnyfoot.com/about/clients/.

Cowen, L., Ball, L. J., & Delin, J. (2002). An eye movement analysis of web page usability. In *People and Computers XVI-Memorable yet Invisible* (pp. 317–335). Berlin: Springer.

Definitions of User Experience and Usability. (n.d.). Retrieved May 10, 2015, from https://uxpa.org/resources/definitions-user-experience-and-usability.

Dumas, J. S., & Redish, J. C. (1999). *A practical guide to usability testing* (Rev Sub edition). Exeter; Portland, or: Intellect Ltd.

Eger, N., Ball, L. J., Stevens, R., & Dodd, J. (2007). Cueing retrospective verbal reports in usability testing through eye-movement replay (pp. 129–137). Presented at the Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI… but not as we know it (Vol. 1). British Computer Society.

Ehmke, C., & Wilson, S. (2007). Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI…But Not As We Know It* (Vol. 1, pp. 119–128). Swinton: British Computer Society. Retrieved from http://dl.acm.org/citation.cfm?id=1531294.1531311.

Elling, S., Lentz, L., & de Jong, M. (2011). Retrospective think-aloud method: Using eye movements as an extra cue for participants' verbalizations. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1161–1170). New York, NY: ACM.

Ericsson, K. A., & Simon, H. A. (1998). How to study thinking in everyday life: Contrasting think-aloud protocols with descriptions and explanations of thinking. *Mind, Culture, and Activity, 5*(3), 178–186.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*(2), 24–29.

Gilb, T. (1998). Optimizing software inspections. *Crosstalk*, *3*.

Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: Methods and constructs. *International Journal of Industrial Ergonomics, 24*(6), 631–645.

Goldberg, J. H., Stimson, M. J., Lewenstein, M., Scott, N., & Wichansky, A. M. (2002). Eye tracking in web search tasks: Design implications. In *Proceedings of the 2002 Symposium on Eye Tracking Research & Applications* (pp. 51–58). New York, NY: ACM.

Goldberg, J. H., & Wichansky, A. M. (2003). Eye tracking in usability evaluation: A practitioner's guide. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 493–516). Amsterdam: Elsevier.

Guan, Z., Lee, S., Cuddihy, E., & Ramey, J. (2006). The validity of the stimulated retrospective think-aloud method as measured by eye tracking. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1253–1262). New York, NY: ACM.

Hoffman, J. E. (1998). Visual attention and eye movements. *Attention, 31,* 119–153.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye Tracking: A comprehensive guide to methods and measures* (1st ed.). Oxford; New York: OUP Oxford.

Huang, J., White, R., & Buscher, G. (2012). User see, user point: Gaze and cursor alignment in web search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1341–1350). New York, NY: ACM. Retrieved from http://dl.acm.org/citation.cfm?id=2208591.

ISO 9241-11: Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability. (1998).

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, 2*(3), 4.

Jones, A., & Thoma, V. (2019). Determinants for successful Agile collaboration between UX designers and software developers in a complex organisation. *International Journal of Human—Computer Interaction*, 1–22.

Józsa, E. (2011). A potential application of pupillometry in web-usability research. *Social and Management Sciences, 18*(2), 109–115.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*(4), 441–480.

Klein, C., Betz, J., Hirschbuehl, M., Fuchs, C., Schmiedtová, B., Engelbrecht, M., … Rosenberg, R. (2014). Describing art—An interdisciplinary approach to the effects of speaking on gaze movements during the beholding of paintings. *PLoS ONE*, *9*(12), e102439. https://doi.org/10.1371/journal.pone.0102439.

Kuniavsky, M. (2010). *Smart things: Ubiquitous computing user experience design*. Amsterdam: Elsevier.

Kuniavsky, M., Goodman, E., & Moed, A. (2012). *Observing the user experience: A practitioner's guide to user research* (2nd ed.). Amsterdam; Boston, MA: Morgan Kaufmann.

Lavie, N., Lin, Z., Zokaei, N., & Thoma, V. (2009). The role of perceptual load in object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1346–1358. https://doi.org/10.1037/a0016454.

Marshall, S. P. (2002). The index of cognitive activity: Measuring cognitive workload. In *Proceedings of the 2002 IEEE 7th Conference on Human Factors and Power Plants, 2002* (pp. 7-5–7-9). https://doi.org/10.1109/HFPP.2002.1042860.

Mitzner, T. L., Touron, D. R., Rogers, W. A., & Hertzog, C. (2010). Checking it twice: Age-related differences in double checking during visual search. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 54, pp. 1326–1330). Thousand Oaks, CA: Sage.

Nielsen, J., & Gilutz, S. (2003). *Usability return on investment*. Fremont, CA: Nielsen Norman Group.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 249–256). New York, NY: ACM. https://doi.org/10.1145/97243.97281.

Nielsen, J., & Pernice, K. (2009). *Eyetracking web usability* (1st ed.). Berkeley, CA: New Riders.

Norman, D. A. (1988). *The psychology of everyday things*. New York, NY: Basic Books.

Petrie, H., & Bevan, N. (2009). The evaluation of accessibility, usability and user experience. *The Universal Access Handbook*, 10–20.

Pomplun, M., & Sunkara, S. (2003). Pupil dilation as an indicator of cognitive workload in human-computer interaction. In *Proceedings of the International Conference on HCI*.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of Human Computer Interaction, 1,* 211–219.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading (Reprint)*. New York, NY: Routledge.

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review, 105*(1), 125–157. https://doi.org/10.1037/0033-295X.105.1.125.

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: Beyond human—Computer interaction*. Hoboken, NJ: Wiley.

Rubin, J., & Chisnell, D. (2010). *Handbook of usability testing: How to plan, design, and conduct effective tests*. Hoboken, NJ: Wiley.

Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: Practical statistics for user research*. Amsterdam; Waltham, MA: Morgan Kaufmann Publishers Inc.

Shneiderman, B. (1986). *Designing the user interface: Strategies for effective human-computer interaction*. Addison-Wesley Longman Publishing Co., Inc. Retrieved from http://portal.acm.org/citation.cfm?id=6682.

Shneiderman, S. B., & Plaisant, C. (2005). *Designing the user interface* (4th ed.). Hoboken, NJ: Pearson Addison Wesley.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28*(9), 1059–1074.

Stone, D., Jarrett, C., Woodroffe, M., & Minocha, S. (2005). *User interface design and evaluation*. Amsterdam; Boston, MA: Morgan Kaufmann.

Task Scenarios for Usability Testing. (n.d.). Retrieved May 11, 2015, from http://www.nngroup.com/articles/task-scenarios-usability-testing/.

Thoma, V., Davidoff, J., & Hummel, J. (2007). Priming of plane-rotated objects depends on attention and view familiarity. *Visual Cognition, 15*(2), 179–210. https://doi.org/10.1080/13506280500155627.

Thoma, V., & Henson, R. N. (2011). Object representations in ventral and dorsal visual streams: fMRI repetition effects depend on attention and part-whole configuration. *NeuroImage, 57*(2), 513–525. https://doi.org/10.1016/j.neuroimage.2011.04.035.

Thoma, V., Hummel, J. E., & Davidoff, J. (2004). Evidence for holistic representations of ignored images and analytic representations of attended images. *Journal of Experimental Psychology: Human Perception and Performance, 30*(2), 257–267. https://doi.org/10.1037/0096-1523.30.2.257.

Thoma, V., & Lavie, N. (2013). Perceptual load effects on processing distractor faces indicate face-specific capacity limits. *Visual Cognition*, 1–24. https://doi.org/10.1080/13506285.2013.853717.

Thoma, V., & Williams, B. (2009). Developing and validating personas in e-Commerce: A heuristic approach. In *Lecture notes in computer science* (pp. 527–531).

Tullis, T., & Albert, B. (2013). *Measuring the user experience: Collecting, analyzing, and presenting usability metrics* (2nd ed.). Amsterdam; Boston, MA: Elsevier/Morgan Kaufmann.

UsabilityNet: Methods table. (n.d.). Retrieved May 10, 2015, from http://www.usabilitynet.org/tools/methods.htm.

What Is User Experience Design? Overview, Tools And Resources. (n.d.). Retrieved May 10, 2015, from http://www.smashingmagazine.com/2010/10/05/what-is-user-experience-design-overview-tools-and-resources/.

Willis, L. M., & McDonald, S. (2016). Retrospective protocols in usability testing: A comparison of Post-session RTA versus Post-task RTA reports. *Behaviour & Information Technology, 35*(8), 628–643. https://doi.org/10.1080/0144929X.2016.1175506.

Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Chapter 22
Eye Movements in Vehicle Control

Agnes Rosner, Thomas Franke, Frederik Platten and Christiane Attig

Contents

A. Rosner (✉)
Department of Psychology, Cognitive Decision Psychology, University of Zurich, Zurich, Switzerland
e-mail: agnesp.rosner@gmail.com

T. Franke
Institute for Multimedia and Interactive Systems, Engineering Psychology and Cognitive Ergonomics, University of Lübeck, Lübeck, Germany

F. Platten
BMW Group, Munich, Germany

C. Attig
Department of Psychology, Cognitive and Engineering Psychology, Technische Universität Chemnitz, Chemnitz, Germany

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_22

Abstract Measuring gaze behaviour is useful to understand the cognitive processes involved in vehicle control and to test new assistance technology. Most of the research on eye movements in vehicle control was conducted in the context of road traffic and will therefore be focused on in this chapter. In the following, we will first introduce the driving task and outline how the eye-tracking methodology can be used to get insights into the cognitive processes that guide driving behaviour. Furthermore, we will highlight important classical findings and recent developments in the field of eye movements in driving. These include eye movements during basic vehicle control tasks like steering, driving manoeuvres and detecting hazards in the road environment. Additionally, factors influencing task performance (e.g., effects of visual distraction, workload, fatigue, driving experience, and aging) that can be observed by applying the eye-tracking method will be introduced. Conducting an eye-tracking study in the driving context often takes place in complex and highly dynamic environments. Therefore, in the last part of this chapter, we will give a practical guideline of what is important in order to study eye movements in the context of vehicle control including an overview of the most commonly used parameters to describe gaze behaviour in the context of driving. We will sum up this introductory chapter with an outline for future research on the topic of eye movements in vehicle control.

22.1 Introduction and Learning Objectives

Vehicle control takes place in all kinds of transportation from road and air to rail and ship (Vollrath & Krems, 2011). Vehicle control includes operating a carriage and is one of the major research interests of traffic psychologists who aim to describe and explain behaviour in traffic and transport situations. Additionally, research on vehicle control is the prime example for ergonomically designing human-machine interactions (Bubb & Wohlfarter, 2013), as, especially in the case of road traffic, it is an activity with a high prevalence while at the same time it holds the risk of accidents.

Studying the interaction of a human with a machine (i.e., a car) is challenging, as a multitude of factors have to be taken into account. First, there are cognitive processes like perception, attention, comprehension, decision making, and response execution that can influence driving behaviour. Additionally, there are individual and situational factors. On the individual level this can be factors like prior driving experience, workload or fatigue. On the situational level, factors such as other road users, road and weather conditions can influence driving behaviour.

Already the interplay of cognitive, individual and task factors make driving a rather complex task. Moreover, the driving task requires fast processing of information in very short time intervals while at the same time operating in a highly dynamic environment (e.g., Castro, 2009). That is, driving is a task where the temporal dimension of the human perception is crucially involved and the interaction of perception and action (e.g., Neisser, 1976) can well be researched. It seems astonishing that this task can be performed rather easily by experienced drivers.

Besides the question of how people solve the complex task of controlling a vehicle, studying driving behaviour allows studying phenomena that can hardly be observed in more static environments with less complex information (e.g., reading, scene perception). One example is the concept of the *optical flow*, introduced by Gibson (1950). That is, the world seems to expand around a stable point in front of the driver (see Fig. 22.3). How the optical flow is used to control the vehicle can only be investigated in a dynamic environment that we can observe while driving a car.

Research in vehicle control allows studying human cognition in complex, highly dynamic environments, *and* the knowledge gained can directly be applied to real life. Every second person in the European Union owns a car. At the same time, about one million people get injured, and about 26,000 die in traffic accidents per year (BMVI, 2014). Studying human-machine interaction in vehicle control can help to identify failures and prevent accidents. Perceptual problems have been identified as one leading cause of traffic crashes (Lestina & Miller, 1994). Furthermore, in-car observation of driver behaviour preceding an actual crash supports the causal role of distraction and inattention (Klauer, Dingus, Neale, Sudweeks, & Ramsey, 2006).

Lastly, the knowledge gained about human-machine interaction during vehicle control can be delivered to related sciences, especially engineers and computer scientists and together used to develop efficient support systems (e.g., in ergonomic design, traffic infrastructure, and assistant systems).

In sum, driving is a complex, highly dynamic cognitive task. At the same time, understanding driver behaviour is of high practical relevance and has applications for a multitude of related sciences. For these reasons, it is important to have proper methods to understand how people interact with the vehicle, i.e., how vehicle control takes place. Here, we argue for using eye movements as a process-tracing technique to serve exactly this purpose. Figure 22.1 shows examples of eye movement recordings during driving.

After working through this chapter, students should be able to: (1) describe cognitive processes that take place in driving, (2) describe where drivers look when they steer, change lanes and scan their environment for hazards (i.e., what visual information they extract from their environment), (3) explain how factors influencing driving performance can be detected through examining eye movement behaviour, (4) name the most important eye movement metrics in the driving context as well as (5) be able to plan and conduct an empirical study regarding eye movements in vehicle control.

In the following, we will first give a historical overview regarding the research of eye movements in driving behaviour. Then, we will describe the driving task and its underlying cognitive processes and show how tracking eye movements can be useful to gain insights into the process of controlling a vehicle.

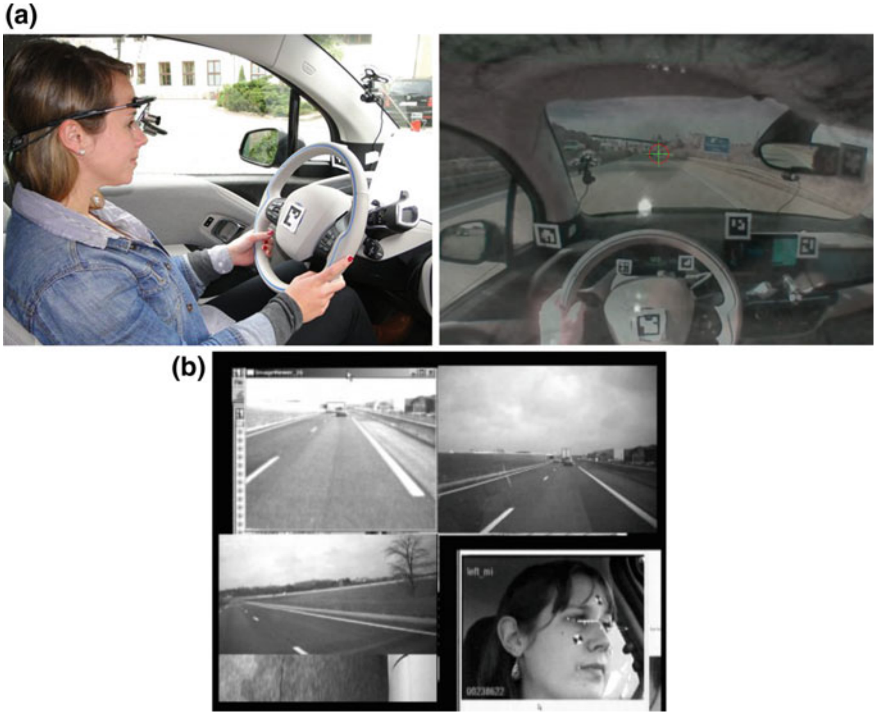


Fig. 22.1 **a** Eye-tracker setup (left side) and scene view with gaze recording (right side) with a head-mounted eye-tracking system (Ergoneers GmbH, Manching, Germany). Figures with kind permission from Maria Kreußlein. **b** Setup of a remote eye-tracking system (Seeing Machines, Tucson, USA) showing front view (top right and top left) and rear view (bottom left, upper part), and the participant's head with the glance direction of the eyes (bottom right). Figure from Henning (2010)

22.2 Historical Annotations

Tracking eye movements in the context of vehicle control was not possible before the development of mobile devices that allowed tracking the eyes in highly dynamic environments such as driving a car. Consequently, employing the eye-tracking method to study driving behaviour did not emerge until the 1950s, when Mackworth was the first to introduce a head-mounted system in order to record eye movements more flexible than by utilizing stationary systems (Land, 2007; Wade & Tatler, 2011). At that time, researchers were mainly interested in the question: Where do drivers look when driving? Even though the first empirical studies regarding drivers' gaze allocation can be traced back to the 1960s (Castro, 2009; Shinar, 2008), the usage of eye-trackers in driving research has particularly increased since the 1990s, when Land (1992) introduced a more lightweight mobile eye-tracking system (see Wade & Tatler, 2011). Since then, the eye-tracking method became more and more popular as can be observed in an enormous increase in research on drivers' eye movements during steering (e.g., Land & Lee, 1994) and lane changing (e.g., Salvucci & Liu, 2002) that led to important insights regarding driver distraction (e.g., Recarte & Nunes, 2009) and accident prevention (e.g., Borowsky, Shinar, & Oron-Gilad, 2010). Besides eye movements, also pupil dilation arose as a measure in the driving context. Since the 1960s (Hess & Polt, 1964), systematic research concerning pupil dilation revealed that it could be utilized as an indicator for mental workload. However, in the driving context, the utilization of pupillometry is still in its infancy with first domain specific studies emerging in the 2000s (Marquart, Cabrall, & de Winter, 2015; Platten, Schwalm, Hülsmann, & Krems, 2014).

22.3 Structuring the Driving Task

What is driving about? As an introduction to the driving task, imagine the following situation:

> You want to drive back home from work in your car. After getting ready for driving and looking outside and in the side and back mirrors for anything that could prevent you from driving, you coordinate movements of the gear stick, clutch, and accelerator at a precise moment called by the engine speed to start driving. Meanwhile, you look continuously outside the front window to monitor the lane and other cars. You want to listen to radio news. Therefore, you look at the radio button to switch on the radio. Later, you see an intersection and detect a stop sign telling you to stop at the intersection. After stopping, you again look outside to see if you can cross safely and start driving. There is a very slow driving car in front of you. You start looking if there is oncoming traffic and, if not, start overtaking.

It already becomes clear from this example scenario that the visual modality is crucially involved in vehicle control (see also Sivak, 1996) and therefore, eye movements are a useful method to explore the driving task. But what is the driving task about?

There have been several attempts to categorize the elements of the driving task (e.g., Alexander & Lunefeld, 1986; Michon, 1985; for an overview see Engström & Hollnagel, 2007). One recent model was proposed by Hollnagel and Woods (2005). *The Extended Control Model (ECOM)* is a conceptual framework that aims to account for driving behaviour on different behavioural levels, from low-level vehicle handling to high-level navigation and route planning, and the relation between the levels. It does so by introducing a hierarchical structure of the driving task consisting of *goals* a driver wants to achieve, for instance, trip planning, obstacle avoidance or vehicle control. The underlying assumption is that achieving a certain goal state takes place by processing *control* (Engström & Hollnagel, 2007). Control is defined as the ability to direct and manage the development of events (Hollnagel & Woods, 2005). The ECOM consists of four control layers: tracking, regulation, monitoring and targeting. Each control layer describes different sub goals of the driving task that can take place simultaneously but with different time frames. The goals for control processes on a given layer are determined by the control process one layer up. An overview is presented in Table 22.1.

Table 22.1 The four control layers, their goals and examples for each layer according to the extended control model (Hollnagel & Woods, 2005)

Control layer	Control task	Example
Tracking	Compensatory	Keeping the distance to the lead car
Regulating	Anticipatory	Deliberatively taking a larger margin to the lead car
Monitoring	Condition monitoring	Monitoring the traffic signals in order to use the correct lane
Targeting	Goal setting and navigation	Choosing a route which is not congested

A specific driving task can be accompanied by characteristic eye movement behaviours. For example, in order to keep the distance to a leading car and, if needed, to adjust the distance, it will be necessary to look at the car in front. Monitoring includes watching vehicle state functions, such as fuel and oil level indication, route guidance, and other traffic information. Thus, eye movements to in-vehicle devices like the fuel indicator give insight into when drivers pay attention towards tasks on the monitoring level. And maybe even more interesting, when they do not look at those devices, for instance, because the driving situation does not allow for control activities on this layer. Also, information processing on the targeting level might become visible by observing a driver's eye movements. For instance, when a driver is looking at the navigation system or looking for an alternative route.

22.4 Cognitive Processes in Driving

Models like the one by Hollnagel and Woods (2005) can be useful in order to structure the driving task and deliver a functional classification of subtasks and goals. This is important for designing assistant and support systems (so-called *In-Vehicle Information Systems*, *IVIS*). Furthermore, those models are widely used in traffic research, which is another reason why we introduced them here. However, it has to be considered that these models are not built to explain the cognitive processes involved in driving, e.g., how a human solves the task of driving a car (see Vollrath & Krems, 2011). Therefore, the following section will describe cognitive processes that are assumed to take place during driving, from information selection, information integration to action selection and execution. This description provides us with a framework in which to describe the relevance of eye movements to study vehicle control. Beforehand, a large range of driving-related tasks (e.g., distance keeping, hazard perception) depend on the optimal allocation of visual attention, which can be assessed via eye-tracking.

22.4.1 Step 1: Information Selection

In comparison to viewing static stimuli like a text or a picture, the visual scenery in driving rapidly changes with very short intervals of covering and uncovering objects in the scenery (e.g., Castro, 2009). Thus, driving is a highly dynamic task that requires extensive processing of visuospatial information (Sivak, 1996) at a high temporal resolution.[1] However, not all information can be processed at the same time. Therefore, the driver has to select relevant information. Models have been developed to describe the deployment of attention and thus which information is selected and gazed at in dynamic environments (Senders, 1964; Wickens, 2002; Wickens, Goh, Helleberg, Horrey, & Talleur, 2003; Wickens, & McCarley, 2007) and in the context of vehicle control (Salvucci, 2006; Wierwille, 1993). One of the most influential models of applied selective attention was proposed by Wickens et al. (2003). According to this model, four factors are important in order to describe the deployment of attention (Fig. 22.2):

(1) *Salience*: Object features that attract a driver's attention. For instance, object colour: red stop signals will attract a driver's attention and lead to an eye movement to the stop signal more than grey ones; motion: fast approaching traffic will attract attention more than slowly approaching traffic.

(2) *Effort*: Cognitive resources necessary to attend to an object. For example, the further apart an object is from a currently attended region, the lesser the chance that this object will be looked at.

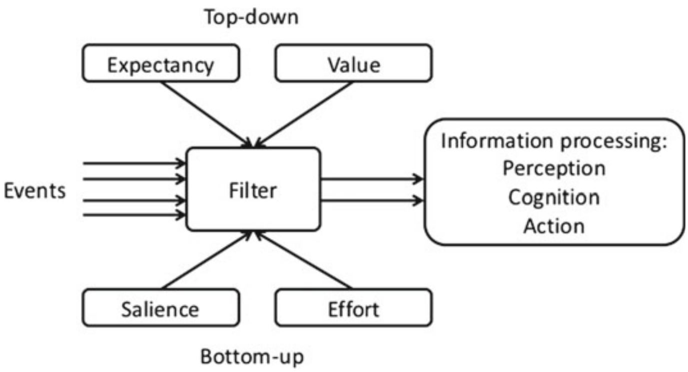


Fig. 22.2 The SEEV (S: Salience, E: Effort, E: Expectancy, V: Value) model of applied selective attention (modified after Wickens & McCarley, 2007) describes the deployment of attention during driving

---

[1] Although the information a driver uses is predominantly visual (Sivak, 1996), the driver also processes information from other perceptual systems. For example, it has been shown that the driver uses auditory information in order to judge the driving speed (Evans, 1970) or vestibular information during lane changing (Macuga, Beall, Kelly, Smith, & Loomis, 2007). However, in this chapter, we focus on visual information.

(3) *Expectancy*: A driver's visual sampling of information is associated with the frequency with which relevant information occurs at that location. Thus it reflects information bandwidths (see also Senders, 1964). For instance, a driver will have the expectation that a traffic sign will be located at the side of the lane and thus sample this location more often with his eyes.

(4) *Value*: Some information is more important than others, e.g., a shop next to the road is less important than other road users. An optimal visual scanning strategy should include the information bandwidth and the expected value of perceiving this information. For instance, the value of perceiving a vehicle that has cut across a driver's path is much greater than correctly perceiving the items on a display in a storefront window (Carbonell, 1966).

Salience and effort are classified as bottom-up processes (data-driven), whereas expectancy and value are seen as top-down (concept-driven) influences. Thus, the attentional selection is guided by both bottom-up and top-down processes.

Eye-tracking gives valuable insights into the attentional deployment during driving, as predicted by the SEEV model. For instance, although a route sign might be irrelevant for the driving task, it still grasps a driver's attention, which may then lead to failures. Thus, tracking a driver's eye movements can give valuable insights into when and under what circumstances driver distraction may take place.

22.4.2 Step 2: Information Integration

Studying which information is perceived and how attention is allocated in order to select relevant information is of vital interest for understanding the cognitive processes involved in driving and to increase driver safety (Lee, 2008; Underwood, 2007). However, driving safely is not only achieved by selecting relevant information. It must be interpreted correctly and considered in planning and controlling driving behaviour. In other words, the selected information (e.g., the road or other car drivers) need to be combined with the driver's goals and prior knowledge in order to guide information selection and processing, to anticipate future events and select the correct actions. A concept that has become rather popular and aims at describing and integrating the underlying cognitive processes in driving is called *situation awareness* (Baumann & Krems, 2007; Durso, Rawson, & Girotto, 2007; Endsley, 1995; Endsley & Garland, 2000; Rousseau, Tremblay, & Breton, 2005). Situation awareness describes a knowledge state that means knowing what is going on in a given situation. Situation awareness draws on the existence of a *situation model*, a mental representation that includes three levels of information: (1) on the *perceptual* level the situational model consists of information about status, attributes, and dynamics of the current situation, (2) on the *comprehension* level perceptual information is interpreted and a meaning is derived, and (3) the level of *projection* generates predictions about future events based on the comprehension of the situation. Understanding the cognitive processes involved in controlling a vehicle requires the assessment of the situation model (see also Baumann & Krems, 2009; Krems & Baumann, 2009). As an example, remember the driving situation described at the beginning of this chapter. There was a slow car driving in front of the driver. The driver *perceives* this slow driving car. The driver must interpret or *comprehend* this information according to its relevance to the driver's goals. For instance, in order to drive safely, the driver needs to keep an intended speed and distance to other road users. In the following, the driver has to *predict* that some behavioural adaption is required to stick to the goal of driving safely. Tracking eye movements can be a useful tool to study information processing in memory (see also Scholz, Mehlhorn & Krems, 2016). An overview of studies using eye movements to track which information a driver integrates will be provided in Sect. 22.6.

22.4.3 Step 3: Action Selection and Execution

The chain of events does not stop at selecting the correct information, comprehending the situation, and deciding for future behaviour. Action plans that are the result of integrating information in a situation model need to be executed (e.g., braking or accelerating). Baumann and Krems (2009) use the model by Norman and Shallice (1986) to describe the voluntary, attentional control of actions by schemata-driven processes that guide the selection of (competing) actions in the context of driving. The actual driving performance also depends on situational factors and individual abilities. For instance, workload (e.g., driving an empty rural road vs. driving on a six-lane highway) may impede task performance as well as experience (novice drivers vs. experienced drivers) and the physiological state of the driver (e.g., fatigue). Eye movements can reveal influences of individual as well as task factors. We will come back to this point in Sect. 22.6.

22.5 What the Driver's Eye Movements Can Tell About Cognitive Processes

Eye movements have been shown to be an ideal dependent variable to measure cognitive processes involved in reading, scene perception, and visual search (e.g., Irwin, 2004; Rayner, 2009; see also Chap. 20 of this book). Eye movement behaviour is also applicable and useful in order to investigate the cognitive processes involved in the complex task of controlling a vehicle. Why is this the case? We have seen in the previous section that a wide range of driving-related tasks, from information selection and integration in a situation model to the anticipation of future events and hazards, are to a large extent dependent on the *optimal allocation of visual attention*. It is this allocation of visual attention that can be observed through overt eye movements (Deubel & Schneider, 1996; Hoffman & Subramaniam, 1995; Shepard, Findlay, & Hockey, 1986; for an overview see Theeuwes, Belopolsky, & Olivers, 2009). Thereafter, a salient event in the visual environment (e.g., a traffic sign, a hazard) captures the driver's attention. Attentional capture precedes eye movements; that is, attention can be shifted covertly to new information in the environment. However, being attracted by an object in the visual world will most likely lead to an eye movement to the salient stimulus. In order to describe the relation between eye movements and the allocation of attention, Holmqvist et al. (2011) used the analogy of a rubber band. Thereafter, stretching the rubber band to one point (the point of cognitive processing) means that the other end of the rubber band (the eye fixations) will naturally follow. In sum, the fixation location and the duration a location is fixated can be used as an approximation of what information is actually processed. Eye movements can both reveal information selection as well as information processing, i.e., the subject of situation assessment. Furthermore, they can reveal which future actions will be executed, for instance, when a lane change will take place (Beggiato & Krems, 2013; Henning, 2010; Salvucci & Liu, 2002).

The assumption that gaze position and gaze duration are coexistent with processing information from that location has been stated early by Just and Carpenter (1976) and became famous under the title *eye-mind hypothesis*. Also for the context of driving it has been shown that the eye-mind assumption is robust, meaning that what the driver is looking at reflects which information the driver is processing (Crundall & Underwood, 2011).

However, there have been exceptions to the eye-mind assumption that should not be left unnoticed. Before we hand over to a review of literature that used eye-tracking to study different aspects of vehicle control we want to discuss challenges to the eye-mind assumption that are relevant for the context of driving (Irwin, 2004).

In cognitive psychology, a phenomenon called *inattentional blindness* has been explored that refers to the failure to see something that is present in a display (for comprehensive reviews see Mack, 2003; Mack & Rock, 1998; Simons, 2000). For instance, Simons and Chabris (1999) showed that people whose task was to count basketball passes by players wearing white shirts, while at the same time ignoring players wearing black shirts, failed to notice when a person wearing a gorilla suit entered the display. A similar phenomenon has been observed by people driving cars, the so-called *looked-but-failed-to-see error* (Hills, 1980; Treat, 1980). That is, car drivers have actually been looking in the direction where another road user was, but have not seen the driver, pedestrian or bicyclist, which then leads to an accident. In an analysis of over 3000 driver errors, Brown (2005) found that 10% of all errors were due to look-but-failed-to-see errors. This error is not committed because of overlooking the other road user. Rather it reflects the driver's mental processing of information, like misinterpreting the visual information or a lack of detecting the critical event (Herslund & Jörgensen, 2003). Thus, processing was separated from the fixation location (i.e., the visual input from the fixated area was not processed and consequently, was not integrated into the situation model). This finding is challenging for using eye movements as an indicator of cognitive processing.

Another phenomenon that is especially salient in the driving context is peripheral information processing (e.g., Crundall, Underwood, & Chapman, 1999; Horrey, 2009; Mourant & Rockwell, 1972; Summala, Nieminen, & Punto, 1996). Most of the time, a driver fixates a rather small visual area straight ahead (Rockwell, 1972). This area is called the *focus of expansion* with a size of about ±4° of visual angle. Within this area, the scenery appears stationary. The area around this point seems to expand (Fig. 22.3). Objects that appear in the periphery as the driver moves past them can be processed by peripheral vision. For example, peripheral vision allows experienced drivers to adjust lane-keeping while performing a distracting foveal task which keeps their eye turned away from the road (Summala et al., 1996).

In sum, information during driving is not only processed by foveal vision, but also by peripheral vision (Crundall et al., 1999; Vollrath & Krems, 2011). In order to describe the area of information intake in driving, Mackworth (1965) coined the term *functional field of view*. The functional field of view describes the visual area around the fixation point from which information can be acquired within one eye fixation. A similar concept has been proposed by Ball, Beard, Roenker, Miller, and Griggs (1988), who termed it *useful field of view*. To complicate things even further, the size of the functional field of view is not fixed (e.g., the size of the fovea), but is influenced by factors such as processing demands or age (Crundall et al., 1999). In the latter case, it has been shown that the size decreases with increasing age (Ball et al., 1988; Maltz & Shinar, 1999). As a consequence, when interpreting eye movement data, it should be taken into account that the size of the functional field of view can be larger than the visual area under foveal vision, which corresponds with the eye position that is tracked by current eye-tracking systems (see also Horrey, 2009). However, recording eye movements still provide valuable information. Concerning foveal vision, it is of interest to better understand why drivers look at objects and fail to recognize that an action is required. This question can only be explored when eye movements are recorded. The same holds true for peripheral vision. If drivers react

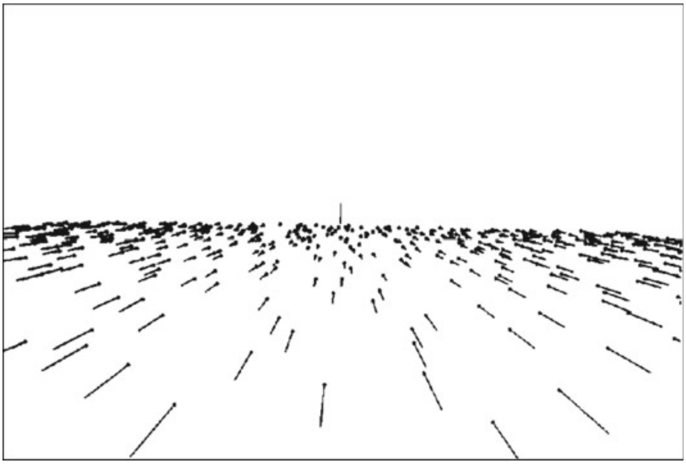


Fig. 22.3 Optic flow from the perspective of a driver. Arrows indicate associated velocity vectors. Figure taken from Warren, Morris, and Kalish (1988)

to cues which they did not fixate, they must have perceived them with peripheral attention. Again, this can only be shown by tracking eye movements. So especially for these two cases, gaze behaviour measurement provides valuable information to all interested in understanding vehicle control processes.

22.6 Where Do Drivers Look When Driving?

In the following, key insights in research on gaze allocation in vehicle control are described. Thereby, we mostly focus on the more recent research in the field. For book chapters that give a more detailed account of earlier studies, see for example Chaps. 17–20 in Underwood (1998). Here, we describe research findings on four core areas in the field: (1) eye movements in steering behaviour and lane keeping, (2) eye movements in controlling the driving manoeuvre of changing lanes, (3) effects of visual distraction, workload and fatigue on eye movements and (4) effects of driving experience and aging on eye movements. Hence, in part (1) and (2) we describe general eye movement patterns in driving and in part (3) and (4) we explore factors that lead to differences in eye movement patterns caused by situational and individual factors.

22.6.1 How Eye Movements Guide Lane Keeping

Think back to our driver in the beginning. How does s/he make sure to stay in the lane? Which information does the driver use to control steering behaviour? In a seminal paper Donges (1978) proposed a two-level model of steering behaviour to explain how drivers handle this task using two kinds of control information: *stabilizing information* and *guidance information*. Based on earlier related research (Shinar, McDowell, & Rockwell, 1977) he proposed that there was an anticipatory component in drivers' steering behaviour (i.e., curve negotiation starts well before a curved road section). Hence, drivers would not only extract stabilization information (i.e., visual cues that allow the estimation of instantaneous deviations between the actual vs. the desired path of the vehicle). They would additionally use guidance information from the visual perception of the road scene (i.e., visual cues that allow preparing for upcoming changes in road curvature) to keep their vehicle in the lane. Consequently, Donges defined two subcomponents in his model: a compensatory closed-loop control component, which compensates errors in path curvature and heading angle as well as lateral deviation. And an anticipatory open-loop control component, which controls steering wheel angle based on the desired path curvature at a certain anticipation time. Both control components act together to produce the steering wheel angle at each specific point in time. To test his model, Donges instructed participants to drive exactly on the centre line of a simulated winding road and assessed their steering behaviour. Indeed, he found drivers started curve negotiation around 1 s before the curve. Moreover, the steering wheel angle predicted by the model fitted quite well with the real driving behaviour of the participants.

In the year 1994, Land and Lee characterized a specific visual cue that drivers use as guidance information (i.e., for the anticipatory control): the tangent point (see also Land & Horwood, 1995). For example, in a curve to the right the tangent point is the continuously moving position on the lane edge in the current visual impression (i.e., when looking at the lane edge) where it stops to go to the left and begins to go to the right (i.e., the line of sight is tangential to the lane edge, see Fig. 22.4).

Land and Lee recorded drivers' fixations while driving on a curvy road in Scotland. They found that most of the drivers' fixations fell within a relatively small area around exactly this tangent point (see Fig. 22.5). Typically, drivers directed their gaze to the tangent point 1–2 s before they entered the curve (i.e., similar value to the anticipation time identified by Donges, 1978).

In recent years, however, several studies presented results that question whether the tangent point is indeed the only point that drivers use to extract guidance information (e.g., Lehtonen, Lappi, & Summala, 2012; Robertshaw & Wilkie, 2008). Indeed drivers seem to be more flexible regarding which far points they use (e.g.,

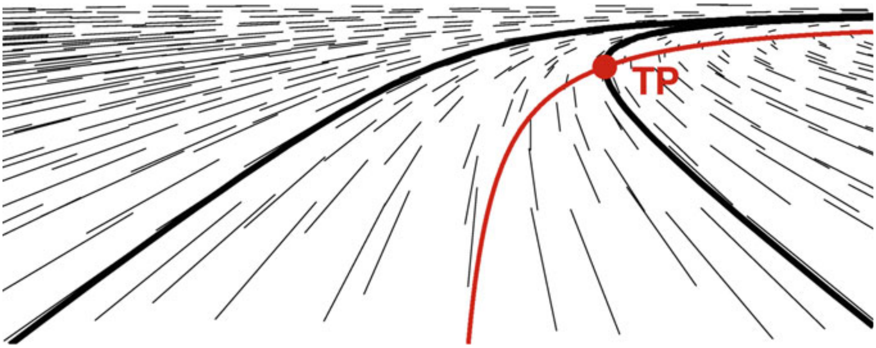


Fig. 22.4 An illustration of the tangent point from the driver's point of view. Figure taken from Authié and Mestre (2012)

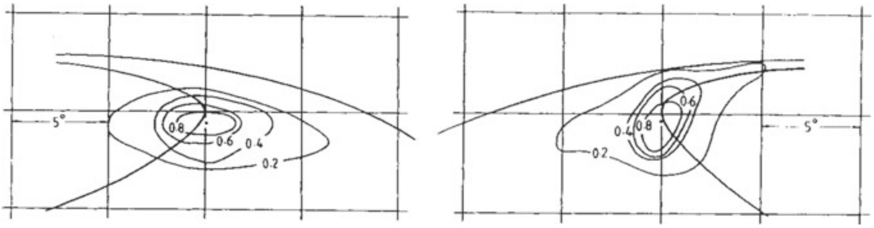


Fig. 22.5 Distribution of drivers' fixations during right bends in Land and Lee (1994). Note that this study was performed on a single-lane street. On a more common two-lane street the fixations in left curves would rather fall around the centre line tangent point (e.g., Lappi, Lehtonen, Pekkanen, & Itkonen, 2013)

also the centre of the road; Salvucci & Gray, 2004; Wilkie & Wann, 2003) and it is not unlikely that they use several far points in parallel (Lappi, 2014). Moreover, not only the far/tangent point seems to guide steering. In their two-point visual control model of steering, Salvucci and Gray (2004) also include a near point, which is located at the center of the road at a nearby distance, and serves to monitor lateral position and stability.

This line of research demonstrates how our knowledge about steering behaviour has continuously advanced and which role eye movement research has played in this context. The idea that our driver from the example at the beginning of this chapter uses two kinds of information (stabilization information and guidance information, i.e., information from a near-zone and a far-zone of the road ahead) has become a commonly accepted model (Land, 2006; Salvucci & Gray, 2004).

However, as already pointed out in Sect. 22.5, research suggests that especially in lane keeping drivers do not only use their foveal (e.g., by fixating on the tangent point) but also their peripheral vision (Summala et al., 1996) as the latter supports spatial orientation and postural control (Horrey, 2009). In fact, drivers with impaired foveal vision did not perform worse than healthy controls when driving in city traffic (Lamble, Summala, & Hyvärinen, 2002).

Despite our knowledge regarding the exact visual strategies that drivers use to stay in lane on a winding road, there are still many open questions. As Lappi (2014) points out in his comprehensive review: "What visual cues are used, how they are represented, and how eye movements are used to sample them are among the outstanding questions in understanding the visual basis of vehicle-assisted locomotion. Despite more than 100 years of eye movement research […] we are still only beginning to answer these questions" (p. 17).

22.6.2 How Drivers Change Lanes

Think back to our introductory example. Our driver will not only have to control the vehicle in order to stay in lane. The driver will most likely avoid other obstacles on the road, for example by overtaking a slower or even parking vehicle. At some point during the trip, the driver might change lanes in order to be in the right lane at the next junction or even to initiate a turning manoeuvre.

Lane changes constitute a type of manoeuvre that has received considerable attention in research. Lane changes are a major source of road crashes (Sen, Smith, & Najm, 2003), because they are typically accompanied by relatively demanding interactions with other vehicles. Moreover, unintended derivations from the driving lane, for example caused by drowsiness, can lead to severe crashes (Kircher, Uddman, & Sandin, 2002). Advanced driver assistant systems have been developed to support lane change manoeuvres. They warn the drivers of conflicting vehicles in the target lane, for example in the blind spot (Chun et al., 2013) or they support lane keeping or prevent run-off-the road accidents (Navarro, Mars, & Young, 2011). However, such systems would be considerably enhanced if they could predict the driver's intention to change lanes. Hence, studies have been conducted to better understand the course of a lane change, also focusing on possible eye movement indicators that could predict lane changes.

In a driving simulator study Salvucci and Liu (2002; see also Salvucci, Liu, & Boer, 2001) analysed drivers' eye movements while performing a lane change and focused on the time course of lane changes. Participants drove on a highway which consisted of two lanes in each direction with moderate traffic. To segment the data, participants were asked to verbally indicate when they intended to do a lane change (defined as starting point of the lane change manoeuvre) and when they had finished the lane change. This was defined as the lane change segment (i.e., the time period or data segment between the starting and end point of the lane change). Data was collected about vehicle control behaviour and eye movement behaviour. To enable an aggregated analysis of multiple lane changes the data was re-sampled as follows. The data stream of one lane change segment was divided into 10 sub-segments with equal time length and the data for each sub-segment was averaged for each of the dependent variables (e.g., average lane position in segment 3 of 10). Equivalently, 8 sub-segments were created for the period before and after the lane change (e.g., if a lane change took 5 s, one sub-segment *during* the lane change equalled 0.5 s. Additionally, 8 sub-segments, with the same duration of 0.5 s were added to represent the time *before and after* the lane change). In the study by Salvucci and Liu, each sub-segment represented on average 0.5 s of data.

Regarding the eye movement data, three areas of interest (AOIs) were distinguished: glances to the lane where the driver was before the lane change (i.e., start lane of the lane change), glances to the end lane of the lane change, and glances to the rear mirror. A ratio of the time spent looking at these three areas was finally computed (*gaze dwell ratio*, see Table 22.2) for each of the sub-segments. The results can be seen in Fig. 22.6. Already around 4 s before the lane change, the gaze dwell ratio to the start lane began to decrease in favour of glances to the rear mirror. Around the onset of the lane change, these glances to the rear mirror reached their peak and the glances to the end lane started to increase. The authors particularly point out how fast drivers shifted their primary visual attention (i.e., gaze) from the start lane to the end lane, already well before the vehicle had reached the end lane.

What do these results tell us about how drivers guide their steering behaviour during a lane change? The authors interpret the early increase in glances to the end lane as an indication that drivers start to use visual cues from the end lane to guide their steering manoeuvre towards this lane. They incorporated this notion in a computational model of steering behaviour (Salvucci & Gray, 2004; Salvucci, 2006): in a nutshell, this model assumes that whenever drivers decide to change lanes, they simply switch the near and far points that guide steering (see also Sect. 22.6.1) from the current lane to the destination lane. Interestingly, this basic model already shows a considerable fit with behavioural data (Salvucci & Gray, 2004; Salvucci, 2006).

Of course, the study of Salvucci and Liu (2002) has not been the only one which has examined glance behaviour during lane changes. Others have, for example, focused on aspects such as glance sequences or transition probabilities from one location to another (Lethaus & Rataj, 2007; Olsen, Lee, & Wierwille, 2005; Tjierina, Garrott,

Table 22.2 Overview and description of eye movement parameters in driving related studies

Name	Description	Example studies
Position measures		
Total glance/dwell time/duration	Cumulated time of all glances on an AOI; this parameter gives a good overview of the glance distribution	Beggiato and Krems (2013), Crundall et al. (2006), Victor et al. (2005)
Mean glance time/duration	Average of glances (total glance time divided by number of glances), if the difficulty of information processing/interpretation should be analysed this is a valid parameter	Crundall et al. (2012), Lehtonen et al. (2014), Mourant and Rockwell (1972)
Glance density	Standard deviation of fixation locations. This parameter indicates gaze concentration (i.e., the extent to which the fixations are spread)	Victor et al. (2005), Reimer et al. (2012), Wang et al. (2014)
Time to first glance	Time until participant fixates the first time on an AOI; this parameter is useful to compare driver reactions to events (for example the reaction to a warning signal)	Crundall et al. (2006), Petzold et al. (2013), Huestegge, Skottke, Anders, Müsseler, and Debus (2010)
Longest/shortest glance duration	Longest/shortest glance; this can give insights to extraordinary reactions in particular situations	Wikman, Nieminen, and Summala (1998), Martens and Fox (2007), Simons-Morton, Guo, Klauer, Ehsani, and Pradhan (2014)
Gaze dwell ratio	Ratio of total glance time to a specific AOI; this ratio shows over all participants the distribution between the AOIs and can show the change of attentional focus over time	Salvucci et al. (2001), Salvucci and Liu (2002)
Pupillometry/pupil diameter	The measurement of pupil dilation in response to an event; changes in the pupil diameter represent changes in cognitive processing (e.g., increased mental workload)	Recarte and Nunes (2000), Schwalm et al. (2008), Konstantopoulos, Chapman, and Crundall (2010)
Numerosity measures		
Number of glances/glance frequency	Number of single glances on an AOI; the relevance of an AOI can be estimated with this parameter	Beggiato and Krems (2013), Underwood et al. (2003), Victor et al. (2005)

(continued)

Table 22.2 (continued)

Name	Description	Example studies
Glance percentage	Percentage of glances on an AOI (number of glances on an AOI divided by the total number of glances multiplied by 100); with this parameter it is possible to get a good picture of the importance of the different AOIs for the driver	Lethaus and Rataj (2007), Victor et al. (2005), Falkmer and Gregersen (2005)
Saccadic rate	The number of saccades per time unit (e.g., second, minute); a decrease of this parameter indicates an increase of mental workload, fatigue and task complexity	Dishart and Land (1998), Crabb et al. (2010), Cardona and Quevedo (2013)
Movement measures		
Scan path analysis	Using a transition matrix, it is possible to detect systematic regularities in sequences of glance locations	Underwood et al. (2003, 2005)
Transition probability	Probability of a fixation change from one area to another; this parameter shows if the participants changed often from one particular AOI to another AOI	Lee, Olsen, and Wierwille (2004), Olsen et al. (2005), Lethaus and Rataj (2007)
Latency and distance measures		
Saccadic latency	Latency between the onset of a stimulus presentation and onset of the saccadic eye movement; this parameter indicates where a stimulus must be located to be quickly noticed or it is used as a measure to infer driver status (e.g., sleepiness)	Wittmann et al. (2006), Ahlstrom et al. (2013)

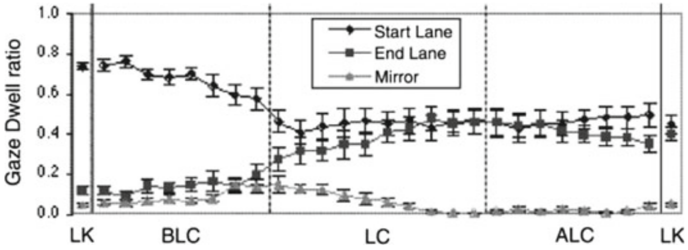


Fig. 22.6 Data on gaze dwell ratios before a lane change (BLC), during lane change (LC), and after the lane change (ALC); LK is lane keeping. The data points are not fixed (not e.g., one data point per second) but adaptive to the varying length of the lane change (see text for detailed description). One average one data point represents 0.5 s. Figure taken from Salvucci and Liu (2002)

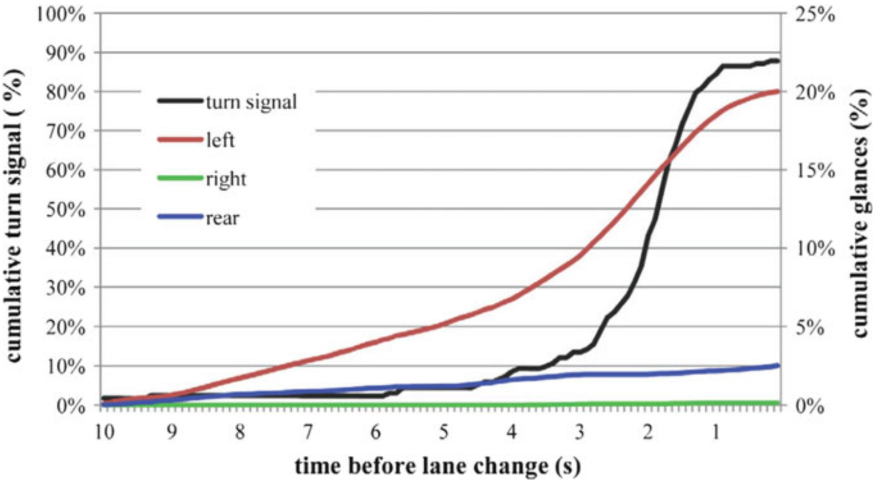


Fig. 22.7 Share of cumulative glance duration to different regions (and turn signal usage) preceding a lane change to the left lane triggered by a slow vehicle ahead. Glances to the left mirror increase prior to the lane change. Figure taken from Beggiato and Krems (2013)

Stoltzfus, & Parmer, 2005). For instance, an on-road study by Beggiato and Krems (2013) also showed an increased glance activity around 2–4 s before the lane change, similar to what Salvucci and Liu (2002) found (Fig. 22.7).

Given these results, can lane changes be predicted based on eye movement behaviour? There is an increasing body of evidence that suggests that this can be done to some degree. For example, Henning (2010) found that within several behavioural indicators for lane changes, glances to the left mirror constituted the best single predictor. Furthermore, Lethaus, Harris, Baumann, Köster, and Lemmer (2013) demonstrated that lane changes can be predicted by gaze data as early as around 3 s before the onset of the manoeuvre. Moreover, some researchers have shown that already the measurement of head motion has been found to provide reliable information to predict lane changes (Doshi & Trivedi, 2009). However, given the many challenges that still have to be solved, it will probably take some more time until eye-trackers (or head-trackers) will monitor drivers and predict their manoeuvre intentions in an everyday environment.

22.6.3 *The Role of Visual Distraction, Workload and Fatigue*

Drivers do not only allocate their attention to task-relevant information. Visual distraction is an important cause of road crashes (Liang, Lee, & Yekhshatyan, 2012; Recarte & Nunes, 2009; Underwood, 2007). Research has shown that factors like salient environmental stimuli (e.g., a ringing mobile phone, traffic signs) can lead to drivers allocating their eyes away from the street (Recarte & Nunes, 2009). Glances away from the street that are above 1.8 s have been regarded as particularly critical (Wierwille, 1993) and therefore found their way into safety regulations (Alliance of Automobile Manufacturers Driver Focus-Telematics Working Group, 2003).

Think of our driver from the example in the beginning. This driver might be in a situation where multiple objects in the environment attract visual attention, for example roadside advertisements. Under which condition will such distracting visual information lead the driver to shift attention away from the driving-related elements in the road scene? One influencing variable could be the position of the advertisement. In a video clip study Crundall, van Loon, and Underwood (2006) examined the difference between roadside advertisements that were placed on the street-level (e.g., at a bus stop) versus those that were elevated (i.e., raised-level). The general pattern was that street-level advertisements had longer mean fixation durations than raised-level advertisements, particularly when participants were asked to scan for hazards in the road scene while watching the video clips. The authors concluded that street-level advertisements were situated in the *safety inspection window*. Hence, they have to be actively filtered out when scanning the road scene for hazards (i.e., drivers have to actively avoid looking at them) while raised-level advertisements have to be “filtered in” (for an illustration see Fig. 22.8). Hence, if salient information that is unrelated to the driving task is placed within the safety inspection window, such information will likely be much more distracting. This finding goes hand in hand with the assumptions of the SEEV model (Wickens et al., 2003; see Sect. 22.4.1): the driver will shift his/her attention more easily to objects that are salient (which a large roadside advertisement usually is) and that are located closer to the currently attended region and, thus, can be detected with less effort (compare the street-level advertisement to the raised-level advertisement). In addition to these two bottom-up processes, the two top-down processes expectancy and value also play a role: an object will be detected with a higher probability if the driver expects important information at the object’s location (which is more probable in the location of the street-level advertisement) and if the object carries valuable information (even though the street-level advertisement is not informative for the driving process per se, the bus stop could become relevant for the driver, e.g., because of waiting pedestrians who might want to cross the road).

Box 1

Secondary tasks: In vehicle control studies secondary tasks (in addition to the primary driving task) are often used to either (1) heighten the overall workload or (2) simulate a distracting driving situation. The amount of available cognitive resources for the driving task is reduced by the amount invested into the secondary task (Sweller, Ayres, & Kalyuga, 2011). By measuring eye movements, the attentional deployment between the primary and secondary tasks can be investigated. By measuring the performance in the secondary task, the cognitive resources that are not available for the primary task can be estimated. Measuring performance and eye movements leads to a more holistic understanding of driving behaviour (see also Box 2 in Sect. 22.7).

A multitude of artificial secondary tasks have been developed for research on distraction while car driving in order to differentiate between different kinds of distraction (Petzoldt, 2011; Petzoldt, Brüggemann, & Krems, 2014). Three of the most frequently used are the surrogate reference task (SuRT; Mattes, 2003), the critical tracking task (CTT, Platten, 2013) or the peripheral detection task (PDT, Vollrath & Krems, 2011). These tasks can be used to simulate In-Vehicle Information Systems (IVIS), they are rather simple and their intended effect on attentional allocation is well understood. In addition to tasks that simulate IVIS, there is a broad array of cognitive tasks that heighten the task load for participants (e.g., one-back task, Kirchner, 1958).

One particularly distracting activity that many drivers engage in while driving is mobile phone conversation and indeed there is a plethora of studies examining

Fig. 22.8 The road-level (above) and raised-level (below) advertisements from the study of Crundall et al. (2006). A density plot from the study of Chapman and Underwood (1998) is included in the figure to illustrate that road-level advertisements are in the safety inspection window while raised-level advertisements are not

the distracting effects of this secondary task (Caird, Willness, Steel, & Scialfa, 2008; Drews & Strayer, 2009). For a short introduction to the concept of secondary task see Box 1. While policy regulations address the physical distraction caused by conversion with handheld mobile phones in several countries, research has shown that there is still distraction associated with hands-free mobile phone conversation. Using the eye-tracking methodology, Strayer, Drews, and Johnston (2003) examined memory for road elements (in fact again advertisement billboards) when conversing with a hands-free mobile phone in a driving simulator. Indeed, the authors found that hands-free mobile phone conversation had no marked effect on fixation patterns (i.e., fixation probability and fixation duration of billboards). In other words, drivers did not look at the advertisements differently depending on whether they conversed at the mobile phone or not. However, the recognition memory of fixated advertisements was much worse in the condition with mobile phone conversation. This can be seen as an instance of the looked-but-failed-to-see-error mentioned in the beginning of this chapter. Hence, under the condition of higher cognitive distraction, even when drivers look at objects in the environment, they often fail to see them (i.e., fail to process the seen information).

Apart from that, there have been studies which have shown that increased cognitive load caused, for example, by a secondary task can lead to changes in eye movement patterns. Namely, gaze tends to concentrate more towards the road centre under conditions of high cognitive load (e.g., Harbluk, Noy, Trbovich, & Eizenman, 2007; Recarte & Nunes, 2003). This gaze concentration has been described as *tunnel vision* or *visual tunnelling* (Recarte & Nunes, 2003). For example, Victor, Harbluk, and Engström (2005) examined on-road and driving simulator data looking for two dependent measures of gaze concentration: percent of glances to the road centre and standard deviation of gaze (see also Table 22.2). Both measures indicated that secondary tasks lead to a higher concentration of glances to the centre of the road. In reaction to this relatively well-replicated finding, measures of gaze concentration have become an accepted method to assess cognitive demand (Reimer, Mehler, Wang, & Coughlin, 2012; Victor et al., 2005; Wang, Reimer, Dobres, & Mehler, 2014). However, several indicators have been found to be applicable: a recent study concludes that horizontal gaze dispersion (i.e., standard deviation of horizontal gaze position) is a particularly sensitive indicator for cognitive demand (Wang et al., 2014).

There also exist eye-tracking measures beyond eye movements that are indicators of workload. A promising approach to measure mental workload is provided by pupillometry (see also chapter by Laeng & Alnaes, in this volume). Pupillometry measures changes in the size of the pupil over time. Pupil size is highly correlated with mental workload (Hess & Polt, 1964). Especially the *index of cognitive activity* (Marshall, 2007) has been successfully applied to the driving context (Platten et al., 2014; Schwalm, Keinath, & Zimmer, 2008).

Last but not least, not only high workload but also low-arousal driver states like fatigue and mind wandering have impacts on drivers' eye movements. For instance, He, Becic, Lee, and McCarley (2011) asked their participants in a driving simulator study to indicate periods of mind wandering. When comparing those episodes with the rest of the driving sequence, the horizontal dispersion of the participants' gaze was found to be smaller during mind wandering than during attentive driving. Similar effects have been found to be related to driver fatigue. Again, eye-tracking indicators have been found to be useful to detect the drivers' sleepiness. For example, blinks and certain saccadic patterns have been related to drivers' sleepiness (Schleicher, Galley, Briest, & Galley, 2008). To sum up, both high and low arousal driver states can be associated with a narrower concentration of gaze patterns.

22.6.4 The Role of Driving Experience and Aging

Besides variations due to driver states, eye movement patterns also change over longer time scales. Research on steering behaviour for example suggests that while usage of the far point to guide steering (i.e., anticipatory information from a point further down the road) increases with driving experience, near-point related behaviour is not linked to experience (Dishart & Land, 1998; Lehtonen, Lappi, Koirikivi, & Summala, 2014; Mourant & Rockwell, 1972).

Many studies have been conducted on the influence of experience on hazard perception and visual search strategies (e.g., Chapman, Underwood, & Roberts, 2002; Falkmer & Gregersen, 2005; Mourant & Rockwell, 1972; Underwood, 2007). These suggest that experienced drivers, compared to novice drivers, are better in identifying hazards, especially those that are not visible directly in the road ahead. Experienced drivers show more anticipatory glances to areas that are important to assess potential risks (Underwood, 2007). For example, Underwood, Chapman, Brocklehurst, Underwood, and Crundall (2003) examined two- and three-fixation scan paths (i.e., probabilities of eye movements from one of eleven parts of the road scene to another part) of experienced versus novice drivers. The results showed that experienced drivers exhibited more flexible/variable scan patterns while novice drivers showed more stereotyped scan paths (for an in depth-treatment of scan paths see chapter by Foulsham, in this volume).

Furthermore, in a more recent study Crundall et al. (2012) examined in drivers with different levels of driving experience (three groups) how long it took them to look at hazards (i.e., identify hazards/show appropriate visual scanning strategies). It was found that more experienced drivers in general fixated a higher share of the critical stimuli in the driving scenarios (i.e., hazards and hazard precursors) and also did this within a shorter time span.

The effects of driving experience leading to more flexible scanning strategies have also been demonstrated in motorcycle riders (Hosking, Liu, & Bayly, 2010): drivers with higher driving and motorcycle riding experience adapted their glance patterns more to the presence of hazards (i.e., showed more flexible visual search patterns) compared to drivers with less experience. As a side effect, such results also have practical implications for conducting driving simulator studies: for example, Alberti, Shahar, and Crundall (2014) showed that only experienced drivers made use of wide field of view displays (i.e., directed their gaze towards areas with high eccentricity) in a driving simulator.

Given these positive effects of driving experience, the question arises whether effective glance strategies can also be trained already during driver education (i.e., before drivers get their license). In one recent study that tested for this possibility (Petzoldt, Weiß, Franke, Krems, & Bannert, 2013) a computer based training was used to teach learner drivers hazard perception skills. Specifically, the skill to anticipate hazard situations and look at hazard indicators as well as further relevant information at the right time (e.g., to identify relevant areas which indicate that a lane change manoeuvre could become necessary to avoid a hazard) was trained. Then, in a driving simulator study three groups of participants were compared: a control group without any learning intervention, a group that received a paper-based training, and a third group that received the computer-based training. Different test scenarios were included in the simulator tracks (e.g., a congestion at a gas station that finally led the lead vehicle which aimed to enter the gas station to brake and partly block the driving lane). It was analysed how long participants needed from the onset of the situation until their first glance to the hazard indicator (e.g., the congested gas station) as well as until completing the specified relevant glance sequence (e.g., looking if lane change will be possible given that the lead vehicle wants to enter the gas station). For both indicators learner drivers who had undergone the computer based training took less time (i.e., showed the relevant glance patterns more quickly) than the other two groups of participants. Hence, it is also possible, at least to some degree, to already teach novice drivers expert-like effective glance strategies.

Last but not least, drivers' eye movements can also change with age, yet not in every situation. Movie clip studies using relatively small video displays (less than 60° horizontally) have found hazard perception to be similar for young and old drivers for several situations. These studies lead to the conclusion that at least general hazard perception skills (i.e., scanning strategies) do not necessarily deteriorate with age (Borowsky et al., 2010). In particular, if visual impairments are eliminated as a potential factor for age-related differences (i.e., by matching the driver groups for visual impairments), visual scanning behaviour of older and younger experienced drivers has been found to be quite similar. For instance, Underwood, Phelps, Wright, van Loon, and Galpin (2005) found no major age-related differences for inspection of objects in the scene and the variance of horizontal search (i.e., more or less extensive search). Yet, the time needed to react to hazards can increase with age (e.g., Borowsky et al., 2010).

However, studies in real traffic and driving simulators with wide field of view (e.g., >130° of visual angle horizontally) have identified that older drivers have more difficulties at intersections compared to younger drivers. For example, Bao and Boyle (2009) found a reduced scanning behaviour to the left and right (i.e., to regions with higher eccentricity) for older drivers when negotiating an intersection. Furthermore, Romoser, Pollatsek, Fisher, and Williams (2013) found, for different intersection scenarios, that older drivers less likely scanned their visual environment for potential threats. For example, at an intersection where participants had to give way to crossing traffic (they had to drive straight through the intersection) older drivers focused less on the far left and/or far right regions and instead spend more time looking at the central region. This shows that results can change considerably with different experimental setups and stimuli.

Finally, there is also evidence showing that eye movement related time-sharing performance decreases with older age (e.g., Wikman & Summala, 2005). Specifically, older drivers have been found to direct more long, and therefore riskier (see Sect. 22.6.3), glances to secondary tasks (e.g., in-vehicles displays), which makes overlooking road hazards more likely and impairs steering performance.

22.7 Experimental Design and Common Eye-Tracking Measures

In the last section, we provided an overview of studies using the eye-tracking methodology to study eye movements in vehicle control. Tracking eye movements allows deep insights in how a driver solves the driving task. The eye-tracking method also reflects influences on the situational and individual level. However, this also challenges the usage of the eye-tracking method when studying driving behaviour. For instance, Fisher, Pradhan, Pollatsek, and Knodler (2007) showed that both trained and untrained novice drivers had longer glance durations regarding a potential hazard in a driving simulator than in a comparable field study. Although the overall correspondence between results obtained in the driving simulator and in the field was remarkable in this study, the experimental setting for an eye-tracking study in the context of vehicle control should be carefully chosen. Therefore, in the following sections we focus on special requirements that come at hand with studying eye movements in vehicle control. Additionally, we will give an overview of the most commonly used eye-tracking metrics. As we pointed out in the chapter introduction, vehicle control is a highly interdisciplinary research area. Therefore, Box 2 provides examples for research questions and outlines how to set up an experimental study in the context of vehicle control.

Box 2

Research questions and hypotheses: Central in planning studies is the clarification of the research question, i.e., defining the target of your study. The main research question should be formulated as a falsifiable hypothesis (Bortz, 2002).

Remember the example from the beginning of this chapter when the driver was reaching the intersection. Given the assumption that an intersection is a cognitively more demanding situation than driving on a rural road, one could assume higher effort values (remember the SEEV model by Wickens et al., 2003) when driving through the intersection as compared to the rural road. Higher effort could be measured by observing pupil dilation (see Table 22.2).

Another example: Imagine you are an automotive manufacturer and your task is to produce an In-Vehicle Information System (IVIS) that guides the attention of a driver back to the road when the driver is distracted by a secondary task (see Box 1). There exist different options of how to attract a driver's attention (e.g., a salient light stimulus, an auditory signal). The automotive supplier wants to know which warning signal has the higher potential to reallocate the driver's attention back to the road. The automotive supplier may have the hypothesis (based on some theoretical consideration or practical experience) that the auditory signal might work better than the visual signal.

Design: When a falsifiable hypothesis is deducted from the initial research question, the researcher can start to define (1) the experimental conditions (the variations of the independent variables), (2) the measurement variables (the dependent variables), (3) the experimental setting, and (4) if the study has a within- or between-subjects design.

(1) **Independent variables**: The experimental conditions are manipulated and should therefore be defined in an early stage. Derived from the hypothesis the interesting factors in the experiment should be specified. For instance, in the second example, the hypothesis states that an auditory signal is better suited to attract a driver's attention back to the road than a visual signal. Therefore, at least three different levels in the factor "warning signal" should be varied: "auditory", "visual" and "no signal". The third condition "no signal" is important in order to test the effects of both conditions in comparison to a **baseline**. The auditory signal condition should lead to a higher guidance rate than the visual signal condition.

(2) **Dependent variables**: The dependent variables are the variables that are measured. It has to be clarified which behavioural parameter is expected to differ between the treatment conditions. In the second example, the goal is to assess a driver's attention. As should have become clear at this point, one way of measuring attentional allocation is by measuring eye movements. For instance, we would assume that drivers change their focus of attention faster from inside the car onto the road (time to first glance; see Sect. 22.7.1) when hearing an auditory signal than when viewing a visual signal and when receiving no signal at all. Importantly, **method triangulation** (i.e., measurement of several variables all indicating about the same behaviour) is warranted in order to strengthen the results obtained with the eye-tracker and to exclude alternative explanations. In our driving experiment it would be reasonable to measure parameters of the driving behaviour itself to get an idea whether the intended manipulation has led to optimized driving behaviour and enhanced driver safety (e.g., lane departure events, time to collision to the vehicle in front of the driver, secondary task performance).

(3) **Experimental setting**: A proper definition of the experimental setting is crucial in order to achieve the study goals. However, defining a proper

driving scenario is rather challenging due to the high complexity of driving situations (e.g., Vollrath & Schießl, 2004). Driving scenarios can roughly be classified by: **road type** (e.g., rural road, inner city road, highway), **traffic density** (e.g., other vehicles), **traffic behaviour** (e.g., high risk behaviour of other drivers), and **driving simulator or field test** (Krems & Petzoldt, 2011).

Concerning the last aspect, driving simulators offer the great opportunity to simulate and test behaviour in hazardous situations without a real risk for the participants. Thus, a driving simulator should be chosen if dangerous situations are tested in the study. Another advantage of driving simulators lies in the high degree of experimental control over the environmental conditions (e.g., traffic density, behaviour of other road users). By using a driving simulator, it is possible to create highly controlled experimental situations to your research question. This helps especially if these situations are hard to find or control in the real world. If the research focus lies, for instance, on gaze behaviour in curves, a driving course could be created with a large number of curves with varying radius, length or speed limits.

In comparison, the important advantage of field studies is their high ecological validity. Furthermore, in field studies influencing factors can be found that were not known to the experimenter in the beginning of the experiment—and due to this lack of knowledge were not programmed into the driving simulator.

In sum, if the research area is rather unknown and the study goal is exploratory in nature, conducting a field study is advised. However, when a high degree of experimental control is needed, a driving simulator might be advantageous. For an in-depth comparison between driving simulators and real world studies see Vollrath and Krems (2011).

(4) **Within- or between-subjects design**: In a within-subjects (or repeated measures) design, each participant runs through each condition of the independent variable(s), while in a between-subjects (or independent groups) design, the participants are divided in groups with each participant running through a different condition of the independent variable (McQueen & Knussen, 2006). In the example of the automotive manufacturer, a within-subjects design would mean that each participant experiences all levels of the independent variable (i.e., “auditory”, “visual” and “no signal”), and a between-subjects design would require dividing the participants in three groups according to the three levels, while a single participant only runs through one of the three levels (e.g., only the “auditory” condition).

Additionally, several other points have to be taken into account before the study can be executed (e.g., defining number of participants, assigning participants to conditions and defining the appropriate test statistics, see e.g., Schinka & Velicer, 2003).

To summarize all design issues with the named example: We would prefer the setup of a driving simulator, as this would allow us testing risky driving scenarios. Additionally, we would need a distracting secondary task that has the potential of simulating a real world distraction (e.g., choosing a radio channel; see Box 1). From our experience a within-subject design is most reasonable for driving related studies, because of the high variance in subjects' behaviour (see also Greenwald, 1976 for a detailed discussion). The broad frame of the study is set after defining the experimental conditions, the dependent variables, and the experimental setting.

22.7.1 Eye-Tracking Measures

Several parameters can be measured in a driving related study. Which ones are most suitable depends on your research question (see Box 2). Before specifying the most important parameters, it is necessary to note that a standardized taxonomy of the eye movement concepts and measures is missing yet and different designations for similar concepts are used in different disciplines (Holmqvist et al., 2011). When observing eye movements, three general behaviours can be distinguished: fixations, saccades and smooth pursuit movements (Fisher, Pollatsek, & Horrey, 2011). When fixating on an object, the eye remains basically still. However, besides *fixation*, also *gaze* or *glance* are used to describe this type of eye movement behaviour,[2] while the term *dwell* refers specifically to the amount of *time* that is spent fixating during one visit of an area of interest (AOI; Holmqvist et al., 2011). In contrast, *saccades* are rapid eye movements that take between 10 and 100 ms (Duchowsky, 2007) and typically occur between fixations, with the exception of *microsaccades*: these very rapid and small eye movements occur during fixations (Holmqvist et al., 2011). *Smooth pursuit movements* take place when moving objects are fixated or when the observer is moving relative to his/her environment (Duchowsky, 2007; Fisher et al., 2011). However, smooth pursuit movements are difficult to measure. Thus, even though the eyes are moving when fixating, for instance, a roadside sign while driving by, it is common to measure smooth pursuit movements as fixations in the driving context (Fisher et al., 2011). In order to interpret the measured eye movements correctly, it is important to know that information is mainly encoded during fixations, not during saccades (with the exception of microsaccades), a phenomenon known as saccadic suppression (Fisher et al., 2011; Rayner, 2009).

In general, four main eye movement categories can be differentiated: movement measures, position measures, numerosity measures and latency and distance measures (Holmqvist et al., 2011). The different measures can only be briefly discussed

---

[2] However, there are slight differences between these terms, e.g., a gaze consists of multiple fixations on the same object (Crundall & Underwood, 2011).

in this chapter; for an in-depth treatment of eye movement measures see the excellent book by Holmqvist et al. (2011).

Movement measures give insight into eye movement characteristics (i.e., direction, amplitude, duration, velocity, and acceleration) during a predefined period of time (Holmqvist et al., 2011). Thus, these measures can answer questions concerning distances completed with a single saccade, the sequential order of glances towards different AOIs or the length of a scan path.

Position measures reveal information about glance position characteristics; thus, these measures can answer the questions where a person looks and how long they fixate a specific area (Holmqvist et al., 2011). Fixation duration measures provide insight into the difficulty of information extraction or interpretation (Fitts, Jones, & Milton, 1950; Jacob & Karn, 2003). The pupil dilation (measured by pupillometry) is another position measure and used as an indicator of mental workload, with an increase in pupil size indicating higher workload (see Sect. 22.6.3).

Numerosity measures refer to numbers, proportions or rates of eye movement events (e.g., fixations, saccades, microsaccades; Holmqvist et al., 2011). In the driving context, gaze frequency measures are commonly used to show the relevance of a spatial area (e.g., the tangent point during steering).

Finally, latency and distance measures provide information about eye movement events in relation to other events (Holmqvist et al., 2011). Latency refers to a delay in time (e.g., time between first fixation of a red traffic light and the onset of a braking event) while distance refers to spatial distances (e.g., distance between an actual saccadic landing point and the intended landing position; the intended saccadic landing position can be a signal in the PDT).

As mentioned, selecting the appropriate measure depends on your research question and hypotheses. For instance, if the research question is whether relevant information can be extracted from a visual array, the total glance time (e.g., Beggiato & Krems, 2013), the mean number of glances (e.g., Underwood et al., 2003) and the mean glance time while performing a secondary task (e.g., Metz, Schömig, & Krüger, 2011) are commonly used metrics (see also Bubb & Wohlfarter, 2013; Poole & Ball, 2006). Table 22.2 provides an overview of common eye movement parameters in driving related studies. As position and numerosity measures are used more frequently than movement and latency/distance measures in the driving context, these measures are emphasized.

22.7.2 Areas of Interest

When analysing gaze data, the field of view can be divided into different AOIs (see Fig. 22.9). As an example, think back to the study by Salvucci and Liu (2002; Sect. 22.6.2): to identify glances during a lane change, three relevant AOIs were determined: (1) the start lane of the lane change, (2), the end lane of the lane change, and (3), the rear mirror. After defining the AOIs, glances into these areas can be analysed and compared (e.g., how much time did the driver spent looking at a secondary

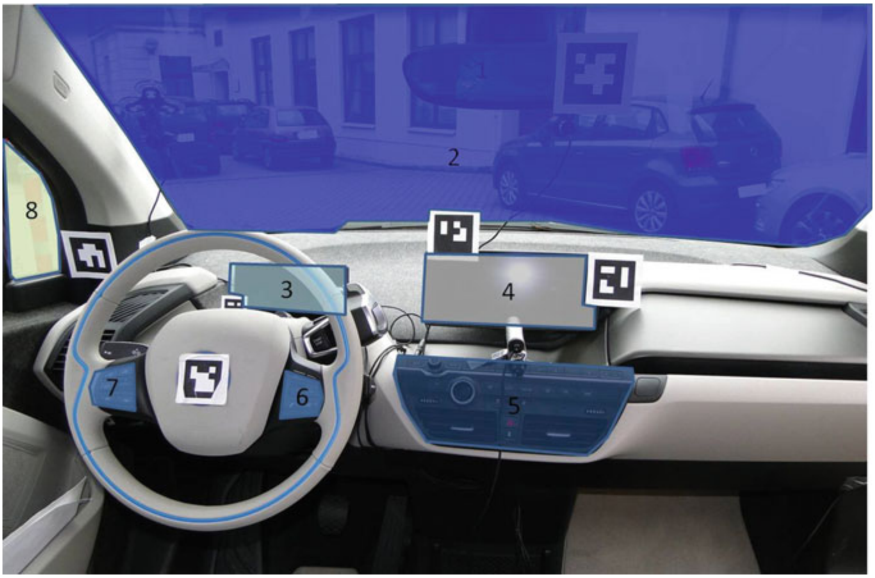


Fig. 22.9 Eight different AOIs are marked in this cockpit. The position of the AOIs is defined by the position of markers (the five black and white patterns). Figure with kind permission from Maria Kreußlein. AOIs added by the authors

task versus on the road). For defining AOIs one important question is whether it is enough to distinguish if a participant has looked at the dashboard or the driving scenery or whether it is needed to distinguish between small areas on the dashboard to answer the research questions. The AOIs should already be defined when planning the study, as they affect study design and depend on the eye-tracking system (see Sect. 22.8). In Fig. 22.9 there are eight AOIs separated from each other. These areas were defined this way because these are the areas in which either the driving scenery has to be monitored or an interaction with the car takes place (Area 4 for infotainment, Area 3 for checking speed, Area 6 and 7 for operating a secondary task, etc.). Additionally, also the driving scenery can be separated into road, sidewalk and environment.

As mentioned earlier, one of the biggest challenges for eye-tracking systems are highly dynamic changes in the driving scenery that are typical in the context of vehicle control (e.g., each participant takes another course through the driving scenario, objects in the field of view are continuously changing their position relative to the vehicle). Static AOIs can hardly be used to the data analysis of moving objects. Thus, manual coding of gaze events is still the common method to address the research questions in the driving context. Manual video coding is of course highly time consuming and risky in terms of human reliability. One way of controlling for lapses that could occur due to manual coding is to code the data by at least two independent raters and to calculate inter-rater reliabilities (see Klauer, Perez & McClafferty, 2011; Rauh, Franke & Krems, 2015).

22.8 Special Requirements of Vehicle Control Studies

In the following section affordances for tracking eye movements in a vehicle control study and advantages of the eye-tracking methodology for conducting research within this specific environment will be discussed.

22.8.1 Hardware

The hardware of eye-tracking systems can be roughly divided into two categories: head mounted systems and remote systems (see chapter by Hutton, in this volume). Head mounted systems are placed onto the head of the participant, either with a headband or a spectacle frame. Remote systems are located in front of the participant, filming the participant from a distance. In applied studies a major goal is to get data with a high degree of external validity. Consequently, the eye-tracking system should be as unobtrusive as possible. At the same time, the needed precision and the measured data per second by the system should be taken into account. This also depends on the study design. If the focus is on very short gaze events (e.g., microsaccades), the system should collect data at a higher frame rate, i.e., more measurement points per second than when analysing longer gaze events (e.g., fixations). Head mounted systems offer often a higher frame rate and a better precision.

Another technology that both provides high precision and high external validity is the usage of eye-tracking glasses which are worn like a common pair of glasses and are equipped with high resolution cameras (Mele & Federici, 2012). For applications of eye-tracking glasses in the driving context, see research by Costa et al. (2014) or Muttard et al. (2011). Furthermore, if it is important to differentiate between very small AOIs, the eye-tracker needs to be precise (to produce highly reliable results) and calibrated accurately (i.e., minimal distance between the true gaze position and the recorded gaze position; Holmqvist et al., 2011). Imagine an eye-tracking study in which it is important to distinguish if the participant looks at the speedometer or the fuel gauge while both are positioned behind the steering wheel. Fixations on either one of these two AOIs can only be correctly classified if the eye-tracker tracks with a high precision and is calibrated accurately.

22.8.2 Data Synchronization

To get a complete picture of the participant's behaviour it is crucial to match the recorded data from the different input devices (e.g., eye-tracking system, driving data from the vehicle, behavioural data from the secondary task) and to bring them in a temporal coherence. Thereby a common time stamp in all different systems or using a system that integrates all data streams can be helpful. Think back to our

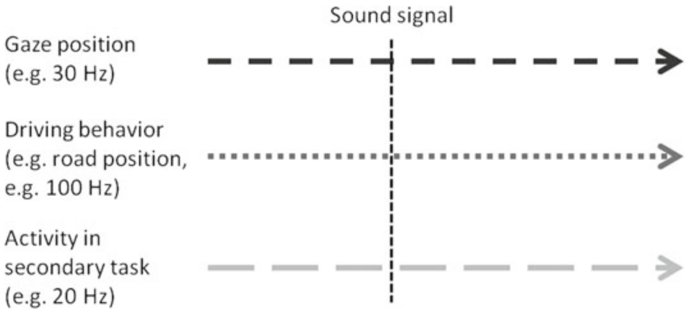


Fig. 22.10 Three sets of data with a common time stamp to get a complete picture of driver reaction to a signal. The length of the dots in the lines exemplifies the different recording frequencies

example with the automotive manufacturer, who plans to implement a warning system that directs the driver's attention back to the road. In this case, we would have one time stamp in (1) the recorded driving behaviour, (2) the recorded gaze data and (3) the recorded secondary task performance in the exact moment the acoustic warning signal occurs (see Fig. 22.10). By having these time stamps in all data sources, it can be analysed how long it took until the driver fixates on the road (time to first glance), if they changed their driving behaviour in that moment, and if they stopped operating the secondary task.

22.8.3 Binocularity

Many eye-trackers take the data of both eyes (binocular gaze tracking). By using the lateral disparity of the eyes, spatial depth information can be gathered. Binocular gaze tracking is especially valuable in the context of studying eye movements in vehicle control. For instance, head-up displays (HUD) use the front window to show driving-relevant information. Binocular gaze tracking allows to differentiate between glances to the HUD or through the HUD to the road (e.g., Milicic, 2010).

22.8.4 Lighting

The lighting conditions can change rapidly while driving. In a driving simulator, however, the lighting conditions are rather stable. When assessing eye movements in a real road environment, the eye-tracking system must be able to track the position of the eyes in situations where brightness (e.g., driving in direct sunlight) rapidly changes to darkness (e.g., driving into a tunnel) or shadows (e.g., driving along an avenue). Several systems include measuring the eye position with an infrared light, which can be disturbed by changing lighting conditions. A possible solution to this problem can be technology based on a pulsed infrared light; thereby the recorded eye position is detected more stable and independent to light changes of the environment.

22.8.5 Piloting

Before conducting the main study with as many participants as planned, it is highly recommended to completely carry out the study and the data analysis of, at least, one participant. Driving studies (simulator studies as well as field studies) are rather extensive and costly, both in terms of timing and budget. Consequently, pilot testing can help to avoid a waste of time and money. Furthermore, it does not only help to get a first impression of the data. More importantly, it allows testing if the toolchain for data processing works and delivers the expected data.

22.9 Conclusions

We have seen that the assessment of eye movements can give valuable insights into the driving process as well as infer a driver's current status. This is possible because eye movements, gaze behaviour and pupil dilation reflect cognitive processes which influence driving performance (e.g., attention allocation, mental workload) and can therefore be used as objective measures to assess driving process-related information. Using measures of gaze allocation, research has shown that drivers fixate specific areas that guide steering and that lane change manoeuvres can be indicated by certain fixations. Moreover, specific eye movement patterns and pupil dilation can detect drivers' mental workload, fatigue or mind wandering which is important in terms of accident prevention. Lastly, also the experience and age of a driver affect eye movement behaviour. This knowledge enables fruitful approaches for designing tutor or assistance systems that support effective attention allocation and hazard detection and, thus, help preventing accidents.

When researching driving behaviour using the eye movement methodology, it is important to specify a concrete research question in order to derive an adequate study design and choose the right eye movement measures. As we have seen, a plethora of different measures exists and it is crucial to know their interpretations (i.e., which part of the information process they reflect). Moreover, assessing eye movements when studying vehicle control involves special considerations concerning hardware, data acquisition and synchronization.

In conclusion, even though the eye-mind hypothesis has its constraints (which becomes apparent by, e.g., the looked-but-failed-to-see-error), assessing eye movements remains a worthwhile possibility to gain knowledge about drivers' cognitive processes and states which makes it hardly surprising that it has been used so frequently when studying driver behaviour.

22.10 Suggested Readings

Bubb, H., & Wohlfarter, M. (2013). Eye-tracking data analysis and neuroergonomics. In M. Fafrowicz, T. Marek, W. Karwowski, & D. Schmorrow (Eds.), *Neuroadaptive systems: Theory and applications* (pp. 255–310). Boca Raton, FL: CRC Press.

– *The chapter provides a detailed introduction into cognitive processes involved in the driving task*

Cacciabue, P. C. (2007). *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems.* London, United Kingdom: Springer.

– *This book gives detailed descriptions of the driving task and current computational models used to explain driver behaviour*

Castro, C. (2009). *Human factors of visual and cognitive performance in driving.* Boca Raton, FL: CRC Press.

– *An in-depth introductory book on what goes on when driving.*

22.11 Questions Students Should Be Able to Answer

- How can the driving task be characterized?
- What are advantages and challenges for using eye-tracking in the context of vehicle control?
- Following the work by Wickens et al. (2003) which information is likely to being selected from the environment?
- Where does a driver look when steering into a curve? How has this been investigated?
- Do older drivers drive differently than younger drivers?
- What has to be taken into account when studying eye movements in the context of vehicle control?
- When to do research in the driving simulator and when is a field test warranted?
- What are the most common eye-tracking metrics used in the context of vehicle control?
- Why is it useful to run a pilot test prior to data collection when planning to conduct a driving study?

Acknowledgements The authors would like to thank Esko Lehtonen and Mark Vollrath for valuable comments on a previous version of this chapter.

References

Ahlstrom, C., Nyström, M., Holmqvist, K., Fors, C., Sandberg, D., Anund, A., et al. (2013). Fit-for-duty test for estimation of drivers' sleepiness level: Eye movements improve the sleep/wake predictor. *Transportation Research Part C: Emerging Technologies, 26,* 20–32.

Alberti, C. F., Shahar, A., & Crundall, D. (2014). Are experienced drivers more likely than novice drivers to benefit from driving simulations with a wide field of view? *Transportation Research Part F: Traffic Psychology and Behaviour, 27,* 124–132.

Alexander, G. J., & Lunefeld, H. (1986). *Driver expectancy in highway design and traffic operations* (Report No. FHWA-TO-86–1). Washington, DC: U.S. Department of Transportation.

Alliance of Automobile Manufacturers Driver Focus-Telematics Working Group. (2003). *Statement of principles, criteria and verification procedures on driver interactions with advanced in-vehicle information and communication systems*. Retrieved from http://www.autoalliance.org.

Authié, C. N., & Mestre, D. R. (2012). Path curvature discrimination: Dependence on gaze direction and optical flow speed. *PLoS ONE, 7*(2), 1–12.

Ball, K. K., Beard, B. L., Roenker, D. L., Miller, R. L., & Griggs, D. S. (1988). Age and visual search: Expanding the useful field of view. *Journal of the Optical Society of America A, 5,* 2210–2219.

Bao, S., & Boyle, L. N. (2009). Age-related differences in visual scanning at median-divided highway intersections in rural areas. *Accident Analysis and Prevention, 41,* 146–152.

Baumann, M. R. K., & Krems, J. F. (2007). Situation awareness and driving: A cognitive model. In P. C. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems* (pp. 253–265). London, United Kingdom: Springer.

Baumann, M. R. K., & Krems, J. F. (2009). A comprehension based cognitive model of situation awareness. In *Lecture Notes in Computer Science* (Vol. 5620, pp. 192–201). Berlin, Germany: Springer.

Beggiato, M., & Krems, J. F. (2013, November). *Sequence analysis of glance patterns to predict lane changes on urban arterial roads*. Paper presented at 6. Tagung Fahrerassistenz—Der Weg zum automatischen Fahren, Munich, Germany.

Borowsky, A., Shinar, D., & Oron-Gilad, T. (2010). Age, skill, and hazard perception in driving. *Accident Analysis and Prevention, 42,* 1240–1249.

Bortz, J., & Döring, N. (2002). *Forschungsmethoden und Evaluation* (3rd ed.). Berlin, Germany: Springer.

Brown, I. (2005). *Review of the 'looked but failed to see' accident causation factor* (Road Safety Research Report No. 60). London, United Kingdom: Department for Transport.

Bubb, H., & Wohlfarter, M. (2013). Eye-tracking data analysis and neuroergonomics. In M. Fafrowicz, T. Marek, W. Karwowski, & D. Schmorrow (Eds.), *Neuroadaptive systems: Theory and applications* (pp. 255–310). Boca Raton, FL: CRC Press.

Bundesministerium für Verkehr und digitale Infrastruktur (BMVI). (2014). *Verkehr in Zahlen 2014/2015*. Hamburg, Germany: DVV Media Group.

Caird, J. K., Willness, C. R., Steel, P., & Scialfa, C. (2008). A meta-analysis of the effects of cell phones on driver performance. *Accident Analysis and Prevention, 40,* 1282–1293.

Carbonell, J. R. (1966). A queuing model of many-instrument visual sampling. *IEEE Transactions on Human Factors in Electronics, HFE-7*(4), 157–164.

Cardona, G., & Quevedo, N. (2013). Blinking and driving: The influence of saccades and cognitive workload. *Current Eye Research, 39,* 239–244.

Castro, C. (2009). *Human factors of visual and cognitive performance in driving*. Boca Raton, FL: CRC Press.

Chapman, P. R., & Underwood, G. (1998). Visual search of driving situations: Danger and experience. *Perception, 27,* 951–964.

Chapman, P. R., Underwood, G., & Roberts, K. (2002). Visual search patterns in trained and untrained novice drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 5,* 157–167.

Chun, J., Lee, I., Park, G., Seo, J., Choi, S., & Han, S. H. (2013). Efficacy of haptic blind spot warnings applied through a steering wheel or a seatbelt. *Transportation Research Part F: Traffic Psychology and Behaviour, 21,* 231–241.

Costa, M., Simone, A., Vignali, V., Lantieri, C., Bucchi, A., & Dondi, G. (2014). Looking behavior for vertical road signs. *Transportation Research Part F: Traffic Psychology and Behaviour, 23,* 147–155.

Crabb, D. P., Smith, N. D., Rauscher, F. G., Chisholm, C. M., Barbur, J. L., Edgar, D. F., & Garwey-Heath, D. F. (2010). Exploring eye movements in patients with glaucoma when viewing a driving scene. *PLoS One, 5*(3), e9710.

Crundall, D., Chapman, P. R., Trawley, S., Collins, L., van Loon, E., Andrews, B., & Underwood, G. (2012). Some hazards are more attractive than others: Drivers of varying experience respond differently to different types of hazard. *Accident Analysis & Prevention, 45*, 600–609.

Crundall, D., van Loon, E., & Underwood, G. (2006). Attraction and distraction of attention with roadside advertisements. *Accident Analysis and Prevention, 38,* 671–677.

Crundall, D., & Underwood, G. (2011). Visual attention while driving. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 137–148). San Diego, CA: Academic Press.

Crundall, D., Underwood, G., & Chapman, P. R. (1999). Driving experience and the functional field of view. *Perception, 28,* 1075–1087.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research, 36,* 1827–1837.

Dishart, D. C., & Land, M. F. (1998). The development of the eye movement strategies of learner drivers. In G. Underwood (Ed.), *Eye guidance in reading and scene perception*. Kidlington, Umited Kingdom: Elsevier Science Ltd.

Donges, E. (1978). A two-level model of driver steering behaviour. *Human Factors, 20,* 691–707.

Doshi, A., & Trivedi, M. M. (2009). On the roles of eye gaze and head dynamics in predicting driver's intent to change lanes. *IEEE Transactions on Intelligent Transportation Systems, 10,* 452–463.

Drews, F. A., & Strayer, D. L. (2009). Cellular phones and driver distraction. In M. A. Regan, J. D. Lee, & K. L. Young (Eds.), *Driver distraction* (pp. 169–190). Boca Raton, FL: CRC Press.

Duchowski, A. (2007). *Eye tracking methodology: Theory and practice*. London, United Kingdom: Springer.

Durso, F. T., Rawson, K. A., & Girotto, S. (2007). Comprehension and situation awareness. In F. T. Durso (Ed.), *Handbook of applied cognition* (2nd ed., pp. 163–194). Chichester, United Kingdom: Wiley.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors, 37,* 65–84.

Endsley, M. R., & Garland, D. J. (2000). *Situation awareness: Analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Engström, J., & Hollnagel, E. (2007). A general conceptual framework for modelling behavioural effects of driver support functions. In P. Cacciabue (Ed.), *Modelling driver behaviour in automotive environments: Critical issues in driver interactions with intelligent transport systems* (pp. 61–84). London, United Kingdom: Springer.

Evans, L. (1970). Speed estimation from a moving automobile. *Ergonomics, 13*(2), 219–230.

Falkmer, T., & Gregersen, N. P. (2005). A comparison of eye movement behaviour of inexperienced and experienced drivers in real traffic environments. *Optometry and Vision Science, 82,* 732–739.

Fisher, D. L., Pollatsek, A., & Horrey, W. J. (2011). Eye behaviors: How driving simulators can expand their role in science and engineering. In D. L. Fisher, M. Rizzo, J. K. Caird, & J. D. Lee (Eds.), *Handbook of driving simulation for engineering, medicine, and psychology* (pp. 18.1–18.22). Boca Raton, FL: CRC Press.

Fisher, D. L., Pradhan, A. K., Pollatsek, A., & Knodler, M. A. (2007). Empirical evaluation of hazard anticipation behaviors in the field and on driving simulator using eye tracker. *Transportation Research Record: Journal of the Transportation Research Board, 2018,* 80–86.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review, 9*(2), 24–29.

Gibson, J. J. (1950). *Perception and the visual world*. Boston, MA: Houghton Mifflin.

Greenwald, A. G. (1976). Within-subjects design: To use or not to use? *Psychological Bulletin, 83,* 314–320.

Harbluk, J. L., Noy, Y. I., Trbovich, P. L., & Eizenman, M. (2007). An on-road assessment of cognitive distraction: Impacts on drivers' visual behaviour and braking performance. *Accident Analysis and Prevention, 39,* 372–379.

He, J., Becic, E., Lee, Y., & McCarley, J. S. (2011). Mind wandering behind the wheel: Performance and oculomotor correlates. *Human Factors, 53,* 13–21.

Henning, M. (2010). *Preparation for lane change manoeuvres: Behavioural indicators and underlying cognitive processes* (Doctoral dissertation). Technische Universität Chemnitz, Chemnitz, Germany.

Herslund, M., & Jörgensen, N. O. (2003). Looked-but-failed-to-see-errors in traffic. *Accident Analysis and Prevention, 35,* 885–891.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143,* 1190–1192.

Hills, B. L. (1980). Vision, visibility and perception in driving. *Perception, 3,* 434–467.

Hoffman, J. E., & Subramaniam, B. (1995). The role of visual attention in saccadic eye movements. *Perception and Psychophysics, 57,* 787–795.

Hollnagel, E., & Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: CRC Press.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford, United Kingdom: Oxford University Press.

Horrey, W. J. (2009). On allocating the eyes: Visual attention and in-vehicle technologies. In C. Castro (Ed.), *Human factors of visual and cognitive performance in driving* (pp. 151–166). Boca Raton, FL: CRC Press.

Hosking, S. G., Liu, C. C., & Bayly, M. (2010). The visual search patterns and hazard responses of experienced and inexperienced motorcycle riders. *Accident Analysis and Prevention, 42,* 196–202.

Huestegge, L., Skottke, E., Anders, S., Müsseler, J., & Debus, G. (2010). The development of hazard perception: Dissociation of visual orientation and hazard processing. *Transportation Research Part F: Traffic Psychology and Behaviour, 13,* 1–8.

Irwin, D. E. (2004). Fixation location and fixation duration as indices of cognitive processing. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 105–133). New York, NY: Psychology Press.

Jacob, R. J. K., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radach, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (Vol. 2, pp. 573–605). Amsterdam, Netherlands: Elsevier.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8,* 441–480.

Kircher, A., Uddman, M., & Sandin, J. (2002). *Vehicle control and drowsiness* (Report No. VTI meddelande 922A). Linköping, Sweden: Swedish National Road and Transport Research Institute.

Kirchner, W. K. (1958). Age differences in short-term retention of rapidly changing information. *Journal of Experimental Psychology, 55,* 352–358.

Klauer, S. G., Dingus, T. A., Neale, V. L., Sudweeks, J. D., & Ramsey, D. J. (2006). *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data* (Report No. DOT HS 810 594). Washington, DC: National Highway Traffic Safety Administration.

Klauer, S. G., Perez, M., & McClafferty, J. (2011). Naturalistic driving studies and data coding and analysis techniques. In B. E. Porter (Ed.), *Handbook of traffic psychology* (pp. 73–85). San Diego, CA: Academic Press.

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility: Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident Analysis and Prevention, 42,* 827–834.

Krems, J. F., & Baumann, M. R. K. (2009). Driving and situation awareness: A cognitive model of memory-update processes. In *Lecture Notes in Computer Science* (Vol. 5619, pp. 986–994). Berlin, Germany: Springer.

Krems, J. F., & Petzoldt, T. (2011). Tools and procedures for measuring safety-relevant criteria. In Y. Barnard, R. Risser, & J. F. Krems (Eds.), *The safety of intelligent driver support systems* (pp. 93–109). Farnham, United Kingdom: Ashgate.

Lamble, D., Summala, H., & Hyvärinen, L. (2002). Driving performance of drivers with impaired central visual field acuity. *Accident Analysis and Prevention, 34,* 711–716.

Land, M. F. (1992). Predictable eye-head coordination during driving. *Nature, 359,* 318–320.

Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25,* 296–324.

Land, M. F. (2007). Fixation strategies during active behaviour: A brief history. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 75–95). Oxford, United Kingdom: Elsevier.

Land, M. F., & Horwood, J. (1995). Which parts of the road guide steering? *Nature, 377,* 339–340.

Land, M. F., & Lee, D. N. (1994). Where we look when we steer. *Nature, 369,* 742–744.

Lappi, O. (2014). Future path and tangent point models in the visual control of locomotion in curve driving. *Journal of Vision, 14,* 1–22.

Lappi, O., Lehtonen, E., Pekkanen, J., & Itkonen, T. (2013). Beyond the tangent point: Gaze targets in naturalistic driving. *Journal of Vision, 13*(13), 1–18.

Lee, J. D. (2008). Fifty years of driving safety research. *Human Factors, 50,* 521–528.

Lee, S. E., Olsen, E. C. B., & Wierwille, W. W. (2004). *A comprehensive examination of naturalistic lane-changes* (Report No. DOT HS 809 702). Washington, DC: National Highway Traffic Safety Administration.

Lestina, D. C., & Miller, T. R. (1994). Characteristics of crash-involved younger drivers. In *38th Annual proceedings of the association for the advancement of automotive medicine* (pp. 425–437). Des Plaines, IL: Association for the Advancement of Automotive Medicine.

Lethaus, F., Harris, R. M., Baumann, M. R. K., Köster, F., & Lemmer, K. (2013). Windows of driver gaze data: How early and how much for robust predictions of driver intent? In M. Tomassini, A. Antonioni, F. Daolio, & P. Buesser (Eds.), *Adaptive and natural computing algorithms* (pp. 446–455). Berlin, Germany: Springer.

Lethaus, F., & Rataj, J. (2007). Do eye movements reflect driving manoeuvres? *IET Intelligent Transport Systems, 1,* 199–204.

Lehtonen, E., Lappi, O., Koirikivi, I., & Summala, H. (2014). Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accident Analysis and Prevention, 70,* 195–208.

Lehtonen, E., Lappi, O., & Summala, H. (2012). Anticipatory eye movements when approaching a curve on a rural road depend on working memory load. *Transportation Research Part F: Traffic Psychology and Behaviour, 15,* 369–377.

Liang, Y., Lee, J. D., & Yekhshatyan, L. (2012). How dangerous is looking away from the road? Algorithms predict crash risk from glance patterns in naturalistic driving. *Human Factors, 54,* 1104–1116.

Mack, A. (2003). Inattentional blindness: Looking without seeing. *Current Directions in Psychological Science, 12,* 180–184.

Mack, A., & Rock, I. (1998). *Inattentional blindness*. Cambridge, MA: MIT Press.

Mackworth, N. H. (1965). Visual noise causes tunnel vision. *Psychonomic Science, 3,* 67–68.

Macuga, K. L., Beall, A. C., Kelly, J. W., Smith, R. S., & Loomis, J. M. (2007). Changing lanes: Inertial cues and explicit path information facilitate steering performance when visual feedback is removed. *Experimental Brain Research, 178*(2), 141–150.

Maltz, M., & Shinar, D. (1999). Eye movements of younger and older drivers. *Human Factors, 41,* 15–25.

Marquart, G., Cabrall, C., & de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. *Procedia Manufacturing, 3,* 2854–2861.

Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space and Environmental Medicine, 78,* 165–175.

Martens, M. H., & Fox, M. (2007). Does road familiarity change eye fixations? A comparison between watching a video and real driving. *Transportation Research Part F: Traffic Psychology and Behaviour, 10,* 33–47.

Mattes, S. (2003). The lane-change-task as a tool for driver distraction evaluation. In H. Strasser, K. Kluth, H. Rausch, & H. Bubb (Eds.), *Quality of work and products in enterprises of the future* (pp. 57–60). Stuttgart, Germany: Ergonomia.

McQueen, R. A., & Knussen, C. (2006). *Introduction to research methods and statistics in psychology*. Harlow, United Kingdom: Pearson.

Mele, M. L., & Federici, S. (2012). Gaze and eye-tracking solutions for psychological research. *Cognitive Processing, 13,* 261–265.

Metz, B., Schömig, N., & Krüger, H. (2011). Attention during visual secondary tasks in driving: Adaptation to the demands of the driving task. *Transportation Research Part F: Traffic Psychology and Behaviour, 14,* 369–380.

Michon, J. A. (1985). A critical review of driver behaviour models: What do we know? What should we do? In L. A. Evans & R. C. Schwieg (Eds.), *Human behaviour and traffic safety* (pp. 487–525). New York, NY: Plenum Press.

Milicic, N. (2010). *Sichere und ergonomische Nutzung von Head-Up Displays im Fahrzeug* (Doctoral dissertation). Technische Universität München, Munich, Germany.

Mourant, R. R., & Rockwell, T. H. (1972). Strategies of visual search by novice and experienced drivers. *Human Factors, 14,* 325–335.

Muttard, J. W., Peck, L. R., Guderian, S., Bartlett, W., Ton, L. P., Kauderer, C., et al. (2011). Glancing and stopping behavior of motorcyclists and car drivers at intersections. *Transportation Research Record: Journal of the Transportation Research Board, 2265,* 81–88.

Navarro, J., Mars, F., & Young, M. S. (2011). Lateral control assistance in car driving: Classification, review and future prospects. *IET Intelligent Transport Systems, 5,* 207–220.

Neisser, U. (1976). *Cognition and reality*. San Francisco, CA: W. H. Freeman.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1–18). New York, NY: Plenum Press.

Olsen, E. C. B., Lee, S. E., & Wierwille, W. W. (2005). Eye glance behaviour during lane changes and straight-ahead driving. *Transportation Research Record: Journal of the Transportation Research Board, 1937*, 44–50.

Petzold, T. (2011). *Theoretical and methodological issues in driver distraction* (Doctoral dissertation). Technische Universität Chemnitz, Chemnitz, Germany.

Petzold, T., Brüggemann, S., & Krems, J. F. (2014). Learning effects in the lane change task (LCT)—Realistic secondary tasks and transfer of learning. *Applied Ergonomics, 45,* 639–646.

Petzoldt, T., Weiß, T., Franke, T., Krems, J. F., & Bannert, M. (2013). Can driver education be improved by computer based training of cognitive skills? *Accident Analysis and Prevention, 50,* 1185–1192.

Platten, F. (2013). *Analysis of mental workload and operating behaviour in secondary tasks while driving* (Doctoral dissertation). Technische Universität Chemnitz, Chemnitz, Germany.

Platten, F., Schwalm, M., Hülsmann, J., & Krems, J. F. (2014). Analysis of compensative behaviour in demanding driving situations. *Transportation Research Part F: Traffic Psychology and Behaviour, 26,* 38–48.

Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. In C. Ghaoui (Ed.), *Encyclopedia of human computer interaction* (pp. 211–219). Hershey, PA: Idea Group Reference.

Rauh, N., Franke, T., & Krems, J. F. (2015). Understanding the impact of electric vehicle driving experience on range anxiety. *Human Factors, 57,* 177–187.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62,* 1457–1506.

Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial-imagery tasks on eye fixations while driving. *Journal of Experimental Psychology: Applied, 6,* 31–43.

Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination, and decision making. *Journal of Experimental Psychology: Applied, 9,* 119–137.

Recarte, M. A., & Nunes, L. M. (2009). Driver distractions. In C. Castro (Ed.), *Human factors of visual and cognitive performance in driving* (pp. 75–88). Boca Raton, FL: CRC Press.

Reimer, B., Mehler, B., Wang, Y., & Coughlin, J. F. (2012). A field study on the impact of variations in short-term memory demands on drivers' visual attention and driving performance across three age groups. *Human Factors, 54,* 454–468.

Robertshaw, K. D., & Wilkie, R. M. (2008). Does gaze influence steering around a bend? *Journal of Vision, 8*(4), 1–13.

Rockwell, T. H. (1972). Skills, judgment and information acquisition in driving. In T. W. Forbes (Ed.), *Human factors in highway traffic safety research* (pp. 133–164). New York, NY: Wiley.

Romoser, M. R. E., Pollatsek, A., Fisher, D. L., & Williams, C. C. (2013). Comparing the glance patterns of older versus younger experienced drivers: Scanning for hazards while approaching and entering the intersection. *Transportation Research Part F: Traffic Psychology and Behaviour, 16,* 104–116.

Rousseau, R., Tremblay, S., & Breton, R. (2005). Defining and modeling situation awareness: A critical review. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 3–21). Burlington, United Kingdom: Ashgate.

Salvucci, D. D. (2006). Modeling driver behaviour in a cognitive architecture. *Human Factors, 48,* 362–380.

Salvucci, D. D., & Gray, R. (2004). A two-point visual control model of steering. *Perception, 33,* 1233–1248.

Salvucci, D. D., & Liu, A. (2002). The time course of a lane change: Driver control and eye-movement behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour, 5,* 123–132.

Salvucci, D. D., Liu, A., & Boer, E. R. (2001, August). *Control and monitoring during lane changes*. Paper presented at the Vision in Vehicles IX, Brisbane, Australia.

Schinka, J. A., & Velicer, W. F. (2003). *Handbook of psychology, Vol. 2: Research methods in psychology*. Hoboken, NJ: Wiley.

Schleicher, R., Galley, N., Briest, S., & Galley, L. (2008). Blinks and saccades as indicators of fatigue in sleepiness warnings: Looking tired? *Ergonomics, 51,* 982–1010.

Scholz, A., Mehlhorn, M., & Krems, J. F. (2016). Listen up, eye movements play a role in verbal memory retrieval. *Psychological Research, 80,* 149–158.

Schwalm, M., Keinath, A., & Zimmer, H. (2008). Pupillometry as a method for measuring mental workload within a simulated driving task. In D. de Waard, F. O. Flemisch, B. Lorenz, H. Oberheid, & K. A. Brookhuis (Eds.), *Human factors for assistance and automation* (pp. 75–88). Maastricht, Netherlands: Shaker Publishing.

Sen, B., Smith, J. D., & Najm, W. G. (2003). *Analysis of lane change crashes* (Report No. DOT-VNTSC-NHTSA-02-03). Cambridge, MA: U.S. Department of Transportation, Research and Special Programs Administration.

Senders, J. W. (1964). The human operator as a monitor and controller of multidegree of freedom systems. *IEEE Transaction on Human Factors in Electronics, HFE-5,* 2–5.

Shepard, M., Findlay, J. M., & Hockey, R. J. (1986). The relationship between eye movements and spatial attention. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology, 38,* 475–491.

Shinar, D. (2008). Looks are (almost) everything: Where drivers look to get information. *Human Factors, 50,* 380–384.

Shinar, D., McDowell, E. D., & Rockwell, T. H. (1977). Eye movements in curve negotiation. *Human Factors, 19,* 63–71.

Simons, D. J. (2000). Attentional capture and inattentional blindness. *Trends in Cognitive Sciences, 4,* 147–155.

Simons, D. J., & Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception, 28,* 1059–1074.

Simons-Morton, B. G., Guo, F., Klauer, S. G., Ehsani, J. P., & Pradhan, A. K. (2014). Keep your eyes on the road: Young driver crash risk increases according to duration of distraction. *Journal of Adolescent Health, 54,* S61–S67.

Sivak, M. (1996). The information that drivers use: Is it indeed 90% visual? *Perception, 25,* 1081–1089.

Strayer, D. L., Drews, F. A., & Johnston, W. A. (2003). Cell phone-induced failures of visual attention during simulated driving. *Journal of Experimental Psychology: Applied, 9,* 23–32.

Summala, H., Nieminen, T., & Punto, M. (1996). Maintaining lane position with peripheral vision during in-vehicle tasks. *Human Factors, 38,* 442–451.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York, NY: Springer.

Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica, 132,* 106–114.

Tijerina, L., Garrott, W. R., Stoltzfus, D., & Parmer, E. (2005). Eye glance behaviour of van and passenger car drivers during lane change decision phase. *Transportation Research Record: Journal of the Transportation Research Board, 1937,* 37–43.

Treat, J. R. (1980). A study of precrash factors involved in traffic accidents. *HSRI Research Review, 10,* 1–35.

Underwood, G. (Ed.). (1998). *Eye guidance in reading and scene perception*. Oxford, United Kingdom: Elsevier.

Underwood, G. (2007). Visual attention and the transition from novice to advanced driver. *Ergonomics, 50,* 1235–1249.

Underwood, G., Chapman, P., Brocklehurst, N., Underwood, J., & Crundall, D. (2003). Visual attention while driving: Sequences of eye fixations made by experienced and novice drivers. *Ergonomics, 46,* 629–646.

Underwood, G., Phelps, N., Wright, C., van Loon, E., & Galpin, A. (2005). Eye fixation scanpaths of younger and older drivers in a hazard perception task. *Ophthalmic and Physiological Optics, 25,* 346–356.

Victor, T. W., Harbluk, J. L., & Engström, J. A. (2005). Sensitivity of eye-movement measures to in-vehicle task difficulty. *Transportation Research Part F: Traffic Psychology and Behaviour, 8,* 167–190.

Vollrath, M., & Krems, J. F. (2011). *Verkehrspsychologie. Ein Lehrbuch für Psychologen, Ingenieure und Informatiker.* Stuttgart, Germany: Kohlhammer.

Vollrath, M., & Schießl, C. (2004). Belastung und Beanspruchung im Fahrzeug—Anforderungen an Fahrerassistenzsysteme. In VDI (Ed.), *Integrierte Sicherheit und Fahrerassistenzsysteme* (pp. 343–360). Düsseldorf, Germany: VDI.

Wade, N. J., & Tatler, B. W. (2011). Origins and applications of eye movement research. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 17–44). Oxford, United Kingdom: Oxford University Press.

Wang, Y., Reimer, B., Dobres, J., & Mehler, B. (2014). The sensitivity of different methodologies for characterizing drivers' gaze concentration under increased cognitive demand. *Transportation Research Part F: Traffic Psychology and Behaviour, 26,* 227–237.

Warren, W. H., Morris, M. W., & Kalish, M. (1988). Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 646–660.

Wickens, C. D. (2002). Multiple resources and performance predictions. *Theoretical Issues in Ergonomic Science, 3,* 159–177.

Wickens, C. D., & McCarley, J. S. (2007). *Applied attention theory*. CRC Press. https://www.crcpress.com/Applied-Attention-Theory/Wickens-McCarley/p/book/9781420063363.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W., & Talleur, D. A. (2003). Attentional models of multi-task pilot performance using advanced display technology. *Human Factors, 45,* 360–380.

Wierwille, W. W. (1993). Visual and manual demands of in-car controls and displays. In J. B. Peacock & W. Karwowski (Eds.), *Automotive ergonomics: Human factors in the design and use of the automobile* (pp. 299–320). London, United Kingdom: Taylor & Francis.

Wikman, A., Nieminen, T., & Summala, H. (1998). Driving experience and time-sharing during in-car tasks on roads of different width. *Ergonomics, 41,* 358–372.

Wikman, A., & Summala, H. (2005). Aging and time-sharing in highway driving. *Optometry and Vision Science, 82,* 716–723.

Wilkie, R. M., & Wann, J. P. (2003). Eye-movements aid the control of locomotion. *Journal of Vision, 3,* 677–684.

Wittmann, M., Kiss, M., Gugg, P., Steffen, A., Fink, M., Pöppel, E., & Kamiya, H. (2006). Effects of display position of a visual in-vehicle task on simulated driving. *Applied Ergonomics, 37*, 198–199.

Chapter 23
Eye Movements and Human-Computer Interaction

Päivi Majaranta, Kari-Jouko Räihä, Aulikki Hyrskykari and Oleg Špakov

Contents

P. Majaranta (✉) · K.-J. Räihä · A. Hyrskykari · O. Špakov
Tampere University, Tampere, Finland
e-mail: paivi.majaranta@tuni.fi

C. Klein and U. Ettinger (eds.), *Eye Movement Research*,
Studies in Neuroscience, Psychology and Behavioral Economics,
https://doi.org/10.1007/978-3-030-20085-5_23

Abstract Gaze provides an attractive input channel for human-computer interaction because of its capability to convey the focus of interest. Gaze input allows people with severe disabilities to communicate with eyes alone. The advances in eye tracking technology and its reduced cost make it an increasingly interesting option to be added to the conventional modalities in every day applications. For example, gaze-aware games can enhance the gaming experience by providing timely effects at the right location, knowing exactly where the player is focusing at each moment. However, using gaze both as a viewing organ as well as a control method poses some challenges. In this chapter, we will give an introduction to using gaze as an input method. We will show how to use gaze as an explicit control method and how to exploit it subtly in the background as an additional information channel. We will summarize research on the application of different types of eye movements in interaction and present research-based design guidelines for coping with typical challenges. We will also discuss the role of gaze in multimodal, pervasive and mobile interfaces and contemplate with ideas for future developments.

Keywords Eye input · Gaze interaction · Gaze-controlled applications · Attentive interfaces

23.1 Introduction and Learning Objectives

This chapter presents the basics of using gaze in human-computer interaction in real time, either as an explicit input method (intentionally controlled gaze-based applications) or as an additional information channel in the background (gaze-aware, attentive applications). We will discuss how different types of eye movements such as fixations, saccades, smooth pursuit or even vergence can be used in human-computer interaction (HCI). We also give a brief introduction to human-computer interaction and user-centered design. The general design principles for HCI are relevant for any interface, including gaze-based interfaces.

After reading this chapter, the reader should:

- Understand the pros and cons in using the eyes for interaction in real time.
- Recognize common pitfalls involved in using the eyes for interaction and know how to cope with the typical challenges such as inaccuracy and Midas touch.
- Have an idea how human-centered design principles and user interface guidelines could be applied in the context of gaze-based applications.
- Have an overview of the technical design and implementation of gaze-aware applications.

In addition, we give some pointers to further information at the end of the chapter, including links to online examples and advanced information about implementation issues for programming-oriented readers.

23.1.1 Human-Computer Interaction

Human-computer interaction (HCI) is a field of research that has studied human factors in engineering since the bloom of personal computers in the early 1980s. Nowadays, HCI covers many subfields that study how people interact with computers and technology in general, incorporating insights from cognitive psychology, linguistics, perception, ergonomics, etc. “Human-computer interaction is a discipline concerned with the design, evaluation and implementation of interactive computing systems for human use and with the study of major phenomena surrounding them” (Hewett et al., 1992). Traditionally, interaction between humans and computers has been done with keyboard and mouse as the input methods and a computer screen as the output device. Today computers are found embedded into everyday appliances and we are constantly surrounded by this technology. The choice of available input and output methods has increased significantly over the years. Vision is one of the modalities that can be used either for perception (perceiving the output) or giving commands to the system (input using gaze). More information about human-computer interaction is available in the free online Encyclopedia of Human-Computer Interaction (Soegaard & Dam, 2012).

General design principles include for example *feedback*, referring both to the immediate response on user’s actions as well as the feedback given on successfully accomplished tasks (Shneiderman, 1998). Other important principles include *visibility* and *consistency* (Norman, 1998). Visibility means the user should be aware of the system’s current status and have some idea what to do next. The term “visibility” should be interpreted broadly here, also covering non-visual ways to make the user aware of the possible functionalities (e.g., spoken instructions). Maintaining consistency means, for instance, that given actions produce the same results, making it possible to learn and predict system behavior within and between systems with similar elements. For more information about general interaction design principles, see e.g. Preece, Rogers, and Sharp (2002).

There are also other key principles that are relevant for this chapter, such as *accessibility*. There should be alternative ways to find information and use a system, to allow people with different needs and abilities to use it. Ensuring that the intended target users can use the system requires a *user-centered approach*. The key principles of *human-centered design* apply also in the design and implementation of gaze-based interfaces. Users can be involved from the very beginning, for example, by applying *participatory design* (Muller & Kuhn, 1993). An iterative design and evaluation process is the key to success. Principles of good usability also apply in the design of gaze-based interfaces; for more information, see Heikkilä and Ovaska (2012).

Box 1: Human-Centered Design

Human-centered design (HCD), or *user-centered design* (UCD), considers the needs, wishes and limitations of the user of the product or service during the design process. Six key principles to ensure user-centered design for interactive systems (from ISO standard 9241-210:2010):

- The design is based upon an explicit understanding of users, tasks and environments.
- Users are involved throughout design and development.
- The design is driven and refined by user-centered evaluation.
- The process is iterative.
- The design addresses the whole user experience.
- The design team includes multidisciplinary skills and perspectives.

23.1.2 Gaze-Based Human-Computer Interaction

Gaze is a powerful interaction method. Gaze direction shows to whom a speaker's utterance is targeted at, or what is the object of our interest. Gaze also typically precedes action and gives valuable hints of our intent (Land & Hayhoe, 2001). This makes gaze an attractive alternative input method for human-computer interaction.

Gaze-based HCI is still fairly rare in everyday applications but it has been used in assistive technology for several decades (Istance, 2016; Majaranta & Bulling, 2014; Majaranta et al., 2012). People who are otherwise completely paralyzed may still maintain the control of their eye movements. The control of the eyes is also often preserved in progressive conditions such as amyotrophic lateral sclerosis (ALS) until the late phases of the disease. Access to an eye tracking communication device can significantly improve the quality of life of such users (Caligari, Godi, Guglielmetti, Franchignoni, & Nardone, 2013). Eye tracking also holds the potential for being the easiest and most comfortable alternative means of communication for people with a wide variety of disabilities (Donegan et al., 2009). For example, if manual control or speaking requires considerable effort and induces compulsory movements, a simple "look-to-select" eye tracking application may provide a far more relaxed method for communication and control (Donegan, 2012).

Gaze input holds potential also for mainstream computing. For the general public, gaze could provide an alternative control method, e.g. hands-free pointing and highlighting in presentations (Špakov, Siirtola, Istance, & Räihä, 2016). Gaze can be used to increase immersion in games (Sundstedt, 2011), to control aiming in shooting games (Isokoski, Joos, Špakov, & Martin, 2009), to control the avatar in virtual worlds (Istance, Vickers, & Hyrskykari, 2009), or to provide remote guidance (Akkil, James, Isokoski, & Kangas, 2016).

Information from eye movements can also be used subtly in the background, without disturbing the viewer's natural gaze behavior. The simplest form of *eye-awareness* could be that an application notices the presence of eyes, without the knowledge of the actual gaze direction or target of the gaze. For example, if a cell phone's front camera sees the user's eyes, it can deduce that the user is probably looking at it, even if the cell phone does not know the exact location of the gaze on the screen (Akkil et al., 2015). If information on the gaze direction and scan path (Bischof et al., this volume; Foulsham, this volume) is available, the system "knows" much more about the user's interests and cognitive processes. In gaming, such *gaze-awareness* could, for example, make horror games scarier by letting the player see scary elements in the corner of their eye but make them disappear just before the user can see them sharply (Dechant, Heckner, & Wolff, 2013). The reading experience can be enhanced by offering automatic translation or explanations of difficult terms (Hyrskykari, 2006) or by interactively providing visual and auditory effects to enliven the story as the reading progresses (Biedert, Schwarz, Hees, & Dengel, 2010).

The boundary between explicit gaze input and implicit gaze-aware interfaces is not always clear. People may learn to take advantage of the proactive features of a system and start to use them intentionally to enhance the eye tracking features of an application. For example, if translations to foreign words are provided automatically based on natural gaze behavior during reading, readers may learn to intentionally prolong their gazes at hard words to get them translated (Hyrskykari, Majaranta, & Räihä, 2003).

Majaranta and Bulling (2014) introduced a "continuum" of eye tracking applications for human-computer interaction (Fig. 23.1). It represents a simplified categorization of different ways to use eye tracking in HCI; as already discussed, the boundaries of the categories are not clear cut. At one end of the continuum, off-line recordings of unintentional, natural eye movements are used passively for diagnostic applications. For example, eye tracking is a valuable tool for evaluating user interfaces (Thoma and Dodd, this volume). The information gathered from natural gaze behavior can be used for understanding and modeling user behavior. This information can then be used, for example, for activity recognition (Bulling, Ward, Gellersen, & Tröster, 2011), which takes us towards more active application of gaze data. The next step is to exploit the information in real time to infer user intent and predict user behavior. Such gaze-aware, *attentive applications* change their behavior based

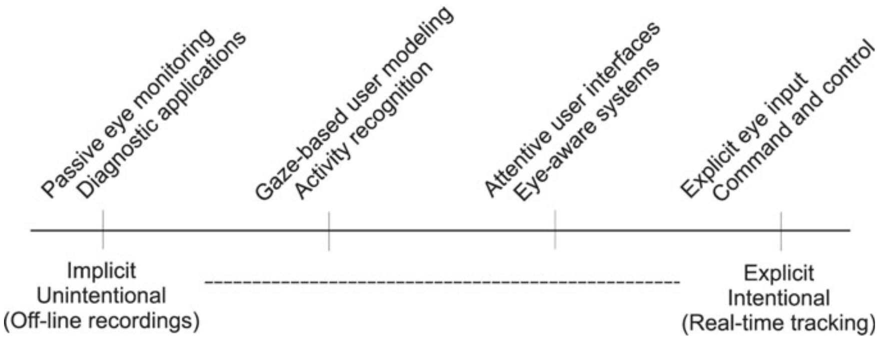


Fig. 23.1 Continuum of eye tracking applications [slightly modified from Majaranta and Bulling (2014)]

on the user's natural gaze behavior and can react more naturally, at the right time (Vertegaal & Shell, 2008). Finally, the eyes can be used for explicit command and control, which requires real-time eye tracking and intentional control of the eyes. The degree of explicit eye control varies, especially when gaze is combined with other modalities (Feit et al., 2017).

23.2 Historical Annotations

Eye movements have been used for communication by people with disabilities long before computer-based assistive technology existed (Goossens' & Crain, 1987). The development of computing technology enabled real-time tracking and analysis of eye behavior. This also made it possible to build assistive technology based on eye tracking. In fact, early work on eye movements in HCI initially focused primarily on disabled users (Jacob & Karn, 2003).

23.2.1 Early Experimental Eye-Controlled Communication Aids

Several eye-controlled electronic communication aids were introduced in the 1970s (Foulds & Fincke, 1979; Rinard & Rugg, 1976). One of the earliest systems, "Eye-Letter-Selector" (Frietman & Tausch, 1978), had phototransistors attached to spectacle frames and a remote electronic control unit with a screen. The system could detect eye movements to left and right, allowing eyes to be used as simple switches (Ten Kate, Frietman, Stoel, & Willems, 1979). Letters were shown in an electronic box that could give feedback on the current selection. Since only simple left and right commands were available, the system provided a column-row scanning procedure to ease the typing. The user selected the currently highlighted letter by an eye gesture to the right. A glance to the left could be used for cancelling the last action. Using eyes as switches (Grauman, Betke, Lombardi, Gips, & Bradski, 2003) and combining eye movements with scanning are relevant methods also for today's systems (Biswas & Langdon, 2013).

Another early system, EYECOM (Rosen & Durfee, 1978), detected several (rough) gaze directions by the optics mounted in spectacles. The system included eight recognizable directions and an "Enter" in the middle. Eight 3 × 3 character matrices, including a total of 64 characters, were shown around the central "enter area". A letter was coded by first looking at the direction of the matrix where the correct letter was located and then looking at the direction corresponding with the location of the letter in the matrix (see Fig. 23.2). Feedback for successful selection was given by a brief flash of two LEDs mounted on the spectacles frames. Similar 2-step coding with directional "gaze gestures" have since then been applied in several

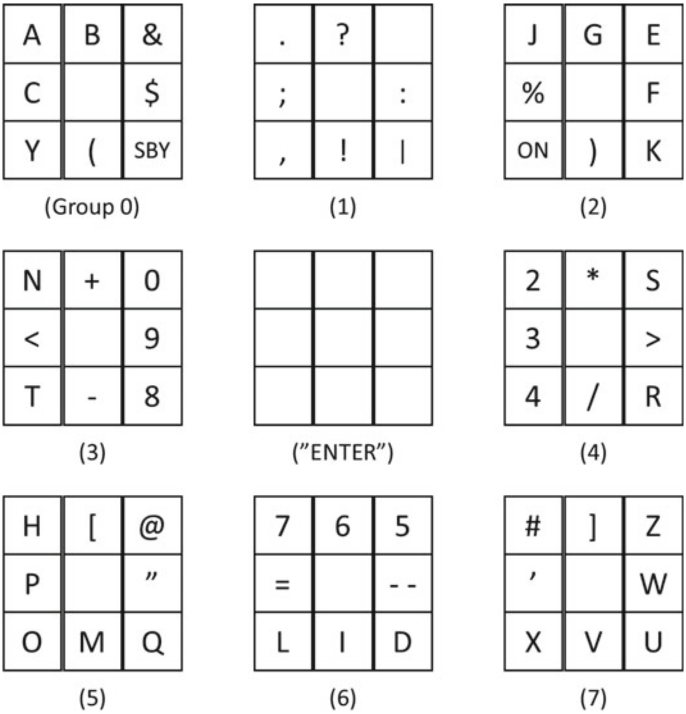


Fig. 23.2 Illustration of the eight 3 × 3 character matrix sets [adapted from Rosen and Durfee (1978)]. For example, looking first at left and then up would type a character "+" when gaze returns to the center

systems and the technique is still useful for example in low-cost systems that use webcams (Bee & André, 2008; Khan, Heynen, & Snuggs, 1999).

23.2.2 *First Commercial VOG Systems*

In the 1980s, video-based eye tracking or video-oculography (VOG, see Hutton, this volume) sparked off wider interest in eye input, and many systems were developed (Hutchinson, White, Martin, Reichert, & Frey, 1989; Levine, 1981; Yamada & Fukuda, 1987). All of them could be used to enter text and some even included a speech synthesizer to speak out the typed text (Friedman, Kiliany, Dzmura, & Anderson, 1982). To cope with inaccurate tracking, the systems adopted methods where typing was possible even if only a few objects were shown at a time. For example, the first versions of the ERICA system (originally distributed by Eye Response Technologies Inc., today part of DynaVox Inc.) only included six keys at a time (Hutchinson et al., 1989). The user had to make two selections to enter one letter: first select a group of letters and then the desired letter in the group. Letter and word prediction techniques were developed (Frey, White, & Hutchinson, 1990) to speed up the communication. Another way to speed up communication was to include common sentences in the selection set (Friedman et al., 1982). These techniques are still useful if an accurate or precise enough calibration cannot be obtained (Donegan, Gill, & Ellis, 2012; Donegan et al., 2005, 2009).

The LC Technologies Eyegaze system (Chapman, 1991) is one of the early commercial systems (Cleveland, 1997) that is still on the market (www.eyegaze.com). It uses direct pointing by gaze, meaning that the user can select the desired target simply by looking directly at it. In addition to facilitating text entry using gaze, the system also implements a set of other useful functions, such as enabling the control of appliances and lights, gameplay, reading electronic documents, and even dialing a telephone by eyes alone.

The potential benefits as well as challenges of using gaze for general human-computer interaction were discussed already in the 1980s. Ware and Mikaelian (1987) showed that pointing by gaze can be very fast—though not as accurate as by using the mouse. Jacob (1990, 1991) discussed the importance of distinguishing normal visual perception from intentional gaze input. Using eye position as a direct substitute for a mouse creates the *Midas touch*[1] *problem*: "Everywhere you look, another command is activated". Jacob noted that people's natural tendency to view the objects of interest longer could be used for implicit input. He compared using "dwell time" (a prolonged fixation) to using a key press for selection, noting that a long dwell time mitigates the speed advantage of eye movements, but dwell can be more convenient for simple information tasks. By careful, iterative interface design, he was able to implement a system where selections were made by gaze using a very short dwell time of 150–250 ms, which actually corresponds with the typical fixation time for visual perception (see Alexander & Martinez-Conde, this volume). Wrong selections could be cancelled easily by looking at the correct object. The problem of Midas touch is a major challenge of gaze input even today. Over the years, a number of solutions have been examined (Velichkovsky, Sprenger, & Unema, 1997); many of them will be discussed later in this chapter.

Another pioneering experiment was conducted by Starker and Bolt (1990), who explored how information from normal viewing could be used to make inferences about the user's interest and consequently change how the system functioned based on the user's gaze behavior. Their experimental software implemented a 3D version of the world of The Little Prince story (de Saint-Exupéry, 1943). The software analyzed the user's gaze patterns and changed the order of the narration if the user started to pay more attention to specific objects. All this was done without explicit commands from the user. Gaze is a valuable information channel for non-command (Nielsen, 1993) and proactive interfaces (Hyrskykari et al., 2003).

Zhai, Morimoto, and Ihde (1999) demonstrated with their MAGIC (MAnual and Gaze Input Cascaded) pointing how proactive information from normal viewing could be combined with manual control. Gaze was used to warp the mouse cursor

---

[1] The phrase "Midas touch" originates from Greek mythology, where King Midas was said to be able to turn everything he touched into gold.

near the point the user was looking at (but not exactly, due to tracking inaccuracy; see Fig. 23.3). Manual adjustment with conventional mouse was used to fine-tune the position. This way they got the best of both: the high speed of gaze pointing and the fine accuracy of manual pointing. More information and examples of using gaze as an input method for both command and non-command interfaces can be found in (Hyrskykari, Majaranta, & Räihä, 2005).

Natural eye movements have also been used implicitly to adapt the level of information shown to the user based on their visual focus. Such foveated rendering of gaze-contingent displays is useful, for example, in virtual reality applications where high resolution details are shown on the user's current area of interest (Duchowski, 1998) or ease visual discomfort of stereo displays (Duchowski et al., 2014). For an overview of related research and application areas, see Reingold et al. (2003) and Duchowski, Cournia, and Murphy (2004). Gaze-contingent displays are located in the implicit end of the continuum of eye tracking applications illustrated in Fig. 23.1. This chapter focuses more on the explicit end, discussing how gaze can be used interactively for explicit control or as an additional channel in eye-aware and attentive interfaces.

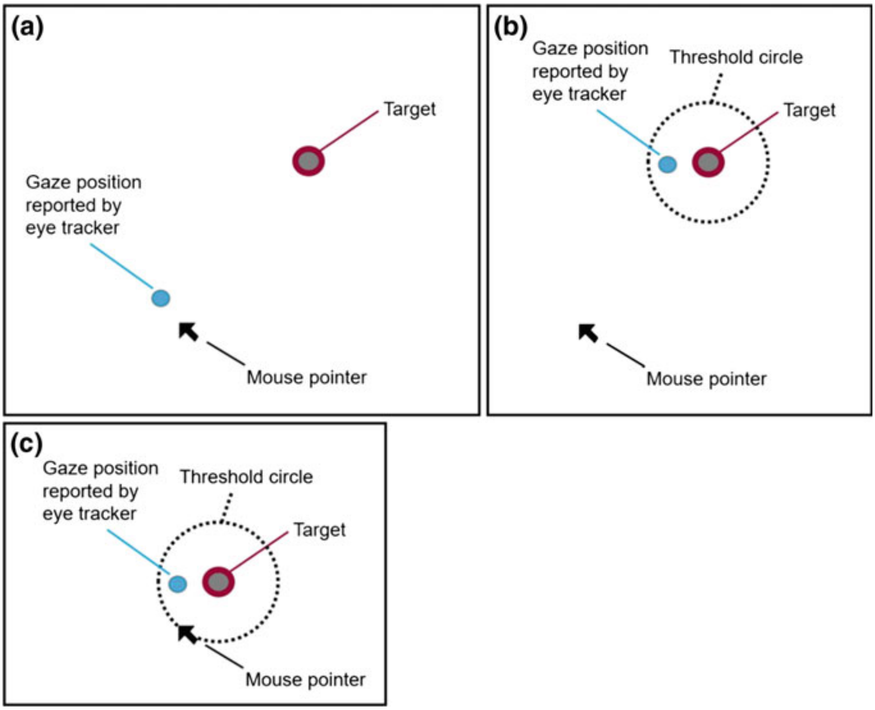


Fig. 23.3 MAGIC pointing. **a** The user wants to move the mouse cursor to the target. Calibration is slightly off. **b** The user looks at the target. Gaze does not land on the target, but close enough, within a predefined threshold. **c** The mouse pointer is automatically warped close to the target, and moving it to the target is then done by using the mouse

23.2.3 EOG-Systems for Eye-Based Interaction

In addition to video-based systems, also systems based on electro-oculography (EOG, see Hutton, this volume) were used for interactive purposes. For example, Gips, Olivieri, and Tecce (1993) developed an EOG-based system, EagleEyes (www.bc.edu/bc-web/schools/carroll-school/sites/eagle-eyes.html), which has been and is still being used by people with disabilities for controlling a computer, typing, drawing, playing games etc.

EOG trackers are especially well suited for mobile and outdoor interactions since they require less power (meaning longer battery life) and can cope with ambient light (Bulling, Roggen, & Tröster, 2009; Ishimaru et al., 2015). Since EOG gives information from relative eye rotations, they are especially well suited for steering tasks. For example, they could be used to control a wheelchair (Barea, Boquete, Mazo, & López, 2002) or a robot (Chen & Newman, 2004). Nevertheless, video-based systems have been considered to be better for interactive applications because of their unobtrusiveness and non-invasiveness (Morimoto & Mimica, 2005). However, this might change because the EOG technology has developed so that the electrodes can now be hidden in the nose piece of eye glass frames (Ishimaru et al., 2015) or in ear plugs (Manabe, Fukumoto, & Yagi, 2015).

23.3 Using the Eyes as an Input Method

Eye tracking technology is introduced in more detail by Hutton (this volume). The same eye tracking hardware that is used for research can also be used for interaction purposes. However, there are some additional requirements.

23.3.1 Eye Tracker as an Input Device

Using the eye tracker as an input device requires that the data can be transferred to the computer in real time. The gaze data must then be also analyzed in real time, so that the application can react to the user's eye movements. However, the requirements on the quality and details of the data depend on the application. Sometimes exact gaze coordinates are not needed; even information on the presence of eyes may be sufficient, and then also more affordable and less accurate trackers can be adequate.

Both desktop-mounted and head-mounted trackers are used for gaze-based HCI. The exact type and features of the tracker depend on the type of the application and the needs of the user. For example, in laboratory experiments the researcher may require the participant to use a head support or sit very still to preserve best quality in eye movement recordings. The tracker may also be calibrated as often as needed between sessions (see Hutton (this volume) for more information about calibration methods). In real-life scenarios, the user wants to concentrate on the task at hand, without needing to pay attention to the input device (Jacob, 1995; Nielsen, 1993). Non-invasiveness, tolerance to head movements, overall robustness, long-lasting calibration and ease of setup and use are appreciated (Hansen, Hansen, Johansen, & Elvesjö, 2005; Morimoto & Mimica, 2005).

People with disabilities have their own set of requirements for the device as it may be used extensively in their daily activities (Borgestig et al., 2016a). Issues to be considered include portability, maintenance and technical support, good ergonomics, a range of mounting equipment for attaching and positioning the system to be useable from bed or wheelchair, possibility to select which eye to track (e.g., because of divergent squint), customized calibration and ability to adjust the procedure (e.g., because only partial eye movement is retained), tolerance to compulsory body movements or eye nystagmus, ability to connect the system to other assistive systems such as environmental control, and ability to adjust feedback (Donegan, Gill, & Ellis, 2012; Donegan et al., 2005; Holmqvist & Buchholz, 2012).

Our focus is on the design of the software for gaze-based software. Hardware and ergonomics are out of the scope of this chapter.

23.3.2 Benefits and Challenges of Gaze Input

Using gaze in HCI is attractive, because gaze conveys the *focus of our visual interest*. We look at things we wish to interact with (Land & Furneaux, 1997). Thus, pointing at the objects of interest with gaze is quite natural, easy and effortless. On the other hand, gaze is easily distracted and most of the normal eye movements occur unconsciously. Consequently, persistent voluntary control may feel unnatural and may require some practice, especially for people with severe impairments (Borgestig, Sandqvist, Parsons, Falkmer, & Hemmingsson, 2016b).

By following normal gaze behavior, a computer program "knows" more about what is going on in the user's mind. This has been called the "eye-mind assumption" (Just & Carpenter, 1980). Gaze is *proactive* in nature, giving hints on what we may do next. Such information can be applied in the background, without requiring the user to explicitly change their gaze behavior. However, *inferring relevance* from the gaze is not easy. The problem comes from ambiguous interpretation. As shown already by Yarbus (1967, p. 172), gaze behavior on the same target is very different depending on the task. Methods that apply machine learning to predict interaction intention from gaze behavior have been developed (Bednarik, Vrzakova, & Hradis, 2012; Kandemir & Kaski, 2012). For example, changes in fixation and saccade positions and durations, as well as changes in the pupil dilation, are related to changes in the user's state. Nevertheless, knowledge of the context, environment, objects, task, as well as additional information from the user's actions (such as speech or body movement) is highly beneficial in the interpretation (Ajanki et al., 2011; Hyrskykari et al., 2003; Istance & Hyrskykari, 2012).

In both cases, be it voluntary pointing or inferred intention, the big challenge is to avoid false activations, i.e., the Midas touch problem. In voluntary gaze input, unintentional selections can be avoided by adding a separate switch that is used to select the object pointed by gaze (Majaranta, Bates, & Donegan, 2009b). However, if the user is only able to move their gaze, other methods like dwell time or gaze gestures are needed (discussed in the next section).

When the eyes point at the focus of the attention, it is called “overt attention”: others can see where the person’s visual focus is targeted at. Since the eyes are always “on” and constantly moving, one should take into account the possibility that the attention may not be focused on the target the eyes happen to point at. Paying attention without moving ones’ eyes is called “covert attention”. For example, a person may pay attention to what is happening in the periphery or outside of their visual field, but does not want to look at it for one reason or the other. In addition, visual attention is actually always slightly ahead of the eye, planning where to move next (Findlay, 2005). Since there is no way to know with absolute certainty which visual information the user is processing at the moment, gaze-aware systems rely on the eye-mind assumption for simplicity. The problem of covert attention is attenuated by informing the user of the system status, by providing feedback on the system’s reactions to the gaze. For a more detailed discussion on covert attention in HCI, see Cantoni and Porta (2014).

Another advantage of gaze is that it is *very fast* (Ware & Mikaelian, 1987). Gaze reaches the object before we move the mouse cursor on the object (Zhai et al., 1999) or before we touch it on the tablet (Weill-Tessier, Turner, & Gellersen, 2016). However, this “look-first-then click” pattern may be more intricate; different people may show different strategies (Smith, Ho, Ark, & Zhai, 2000) and gaze and cursor are not always aligned (Huang, White, & Buscher, 2012). For example, gaze may already leave for the next target before we click an object, or the user may start to move the mouse towards the known target location already before fixating it (Bieg, Chuang, Fleming, Reiterer, & Bülthoff, 2010). The speed can also be challenging, if the system is supposed to react to the rapid gaze movements fluently, without too much system delay (Kangas et al., 2014a).

A major challenge for gaze input is the *limited accuracy and precision* of the measured point of gaze. Accuracy means how well the measured gaze point matches the real gaze position. Precision, on the other hand, refers to the amount of variation of successive samples within a fixation. There is always some noise in measuring our gaze, even during a fixation (Hansen & Ji, 2010). If the mouse cursor is bound to the eye movement, it will not stay stable on the target. Therefore, many systems apply slight smoothing by averaging consecutive gaze data points. Furthermore, the problem of maintaining good quality gaze data is much more serious outdoors and with mobile systems. Varying light conditions may substantially detract the robustness of the system (Holmqvist et al., 2011).

The active research into eye-based interaction techniques during the last few decades has resulted in a multitude of eye-aware applications and prototypes. Different types of eye movements and their properties can be utilized in these techniques. In the following paragraph we will give examples, categorized by the type of eye movement on which they are based.

23.4 Eye Movement Types and Their Application in HCI

All of the basic types of eye movements such as fixations, saccades or smooth pursuit can be exploited voluntarily in gaze interaction (Huckauf & Urbina, 2008; Møllenbach, Hansen, & Lillholm, 2013; Skovsgaard, Räihä, & Tall, 2012). We will introduce the basic eye movements and their use for interaction below. However, it should be noted that real interactions are a combination of several types of eye movements.

23.4.1 Fixations and Dwell Time

Prolonged fixation on a target, referred to as "*dwell time*", is the most used method for making selections by gaze alone. Differentiation from visual inspection is made based on the duration of the dwell time, i.e., selection takes place when the time spent on the target exceeds the predefined threshold. The threshold depends on the task and the user. In any case, the duration should exceed the normal viewing time for the current object in the current context. For example, an experienced gaze typist may adjust the dwell time threshold to be very short to speed up the repetitive task of entering text letter by letter. Dwell times as short as 200–400 ms are possible (Räihä & Ovaska, 2012). Such short dwell times require a fairly accurate tracker (Räihä, 2015). Tasks that require consideration and have higher penalty for false selections, such as problem solving tasks, require longer times, typically 1 s or more (Bednarik, Gowases, & Tukiainen, 2009).

In longitudinal gaze typing experiments (Majaranta, Ahola, & Špakov, 2009a; Räihä & Ovaska, 2012), it was found that participants shortened the dwell duration already in the first or second session. The dwell time threshold can also be adjusted automatically, e.g., based on the speed at which the typists leave keys (Špakov & Miniotas, 2004). Sudden or large automatic adjustments can be problematic, because the users learn the rhythm inherent in the dwell time and may exploit it in their typing (Majaranta, MacKenzie, Aula, & Räihä, 2006). A recent technique (Mott, Williams, Wobbrock, & Morris, 2017) aims at moderate automatic adjustment in order to maintain the rhythm.

Dwell-free methods have gained interest lately (Kurauchi, Feng, Joshi, Morimoto, & Betke, 2016; Pedrosa, Pimentel, Wright, & Truong, 2015). Such fixation-based dwell-free methods assume that the system has enough information about the context and the task to be able to differentiate fixations for perception from intended selections. Dwell-free typing is possible by comparing fixation patterns to known words (Salvucci, 1999). The user fixates on the letters in the correct order but does not need to dwell on them, which can potentially increase the typing speed (Kristensson & Vertanen, 2012). With inaccurate gaze tracking, the method is prone to errors, such as a fixation landing on the neighboring key. Advanced algorithms can be used to handle missing or wrong letters (Liu, Lee, & McKeown, 2016). Dwell-free methods are also used in attentive applications that exploit people's natural tendency to fixate somewhat longer or more often on objects of interest (Räihä, Hyrskykari, & Majaranta, 2011).

23.4.2 Saccades and Gaze Gestures

Gaze gestures are eye movements that follow a defined pattern, interpreted as a command (Drewes & Schmidt, 2007). They can be simple, even based on a single stroke (Møllenbach, Hansen, Lillholm, & Gale, 2009) or complex, including several strokes or "legs" (Istance, Hyrskykari, Immonen, Mansikkamaa, & Vickers, 2010). Simple gestures are fast and easy but they may be confused with natural eye movements. Complex gestures are rare in nature but they take time to learn (Huckauf & Urbina, 2008) and can also be harder to execute, which makes them error prone (Köpsel, Majaranta, Isokoski, & Huckauf, 2016).

The gestures can be bound to certain locations, such as gestures made by crossing predefined screen areas in a certain order. With such gestures, it is possible to implement quite an extensive set of commands, such as those used for text entry (Wobbrock, Rubinstein, Sawyer, & Duchowski, 2008) or for controlling a computer (Porta & Turina, 2008). Or, the gestures can be off-screen (Isokoski, 2000), which frees the screen for other purposes. Simple gestures that start from the screen and then go off-screen and back by crossing one of the display borders have been used for mode change during gaze-based gaming (Istance, Bates, Hyrskykari, & Vickers, 2008), and controlling a mobile phone (Kangas et al., 2014b) or a smart wrist watch (Akkil et al., 2015). Figure 23.4 illustrates an on-screen gesture implementation in a game.

Location free gestures that are recognized based on changes in the angle of the eye movement are insensitive to spatial accuracy problems and can be done even with uncalibrated trackers (Drewes & Schmidt, 2007). Such relative gestures can easily be tracked with EOG-based trackers, useful in mobile settings (Bulling, Roggen, & Tröster, 2009; Ishimaru et al., 2015).

Rozado, San Agustin, Rodriguez and Varona (2012) defined *gliding gaze gestures* a separate category, meaning gestures where gaze is glided along a predefined path. In their experiment, they found that saccadic gestures outperformed the gliding gestures in user performance and satisfaction. Saccadic gestures were also easier to recognize, with less false positives. Gliding gestures (though not named as such) have been earlier applied by Heikkilä and Räihä (2009) in drawing tasks, where they asked people to draw different shapes by following their outlines by gaze. Curved shapes were especially hard. Drawing a reasonable circle by "gliding" the gaze required many small saccades and was slow.

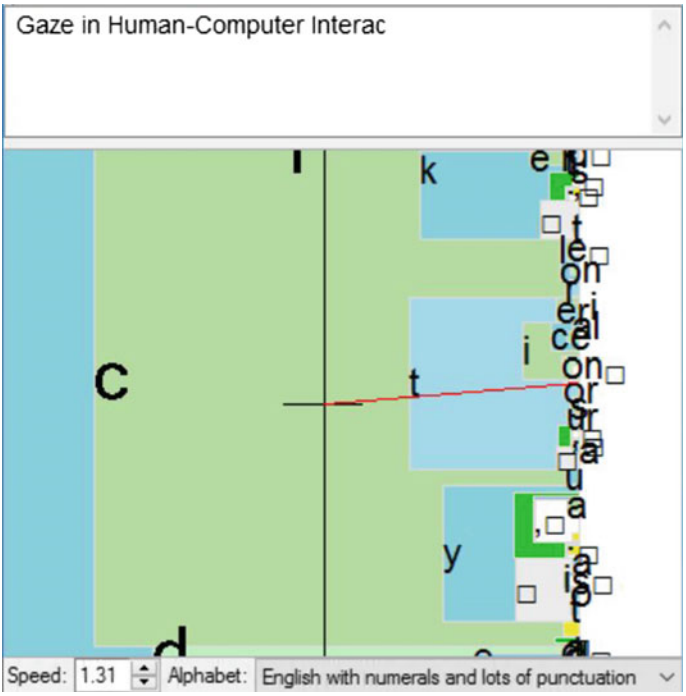
Continuous pointing gestures is a special sub-category of gaze gestures. What makes them special is the continuous (smooth) eye movement, instead of discrete saccades. Gaze-controlled Dasher (Ward & MacKay, 2002) is a well-known example. With Dasher, one can write with continuous eye movements by pointing at the desired

Fig. 23.4 On-screen gesture for gaming. The semi-transparent sectors around the avatar help the user to focus on the gaze-reactive areas. The white numbers and gesture path are for illustration only (Vickers, Istance, & Hyrskykari, 2009)

characters. The interface zooms in, by increasing the area of the desired character thus making it easier to select. The character is selected when it crosses the central line, followed by a new set of characters, with the most probable next letters occupying more space thus making them again easier to select (see Fig. 23.5). Continuous gestures are useful in various zooming interfaces and steering tasks, for example, inspecting large images (Adams, Witkowski, & Spence, 2008) or steering an avatar (Vickers, Istance, & Hyrskykari, 2013).

Anti-saccades have been experimented with as another potential way to make selections by gaze. The user starts by first looking at the button, which causes feedback to appear on one side of the button. Using anti-saccades, the user selects the button by looking at the opposite direction of the feedback that just appeared. The idea is to avoid unintentional selections as the user needs to resist the natural reaction to look at the feedback (Huckauf & Urbina, 2011). Perhaps because of the unnaturalness of anti-saccades, we have not seen them applied elsewhere.

Being able to benefit from the information in *microsaccades* (Alexander & Martinez-Conde, this volume) in interactive applications is beyond the state-of-the-art of current gaze estimation techniques. However, they seem to contain some information that might be useful in HCI, such as potentially useful clues of our emotional state (Laretzaki et al., 2011) or covert attention shifts (Hafed & Clark, 2002). As the quality of the trackers continues to improve, this is an area for future research.



Gaze in Human-Computer Interac

Speed: 1.31 Alphabet: English with numerals and lots of punctuation

Fig. 23.5 Dasher in action. Dasher has predicted that the most likely letter to follow “c” in “Interaction” is “t” and given it’s the largest screen space. The most likely follow-up letters are also already visible

23.4.3 Smooth Pursuit Eye Movements for Interaction

Smooth pursuit occurs when we follow (pursue) a smoothly moving object. It is near to impossible to voluntarily move the eyes smoothly (Lorenceau, 2012). As we learned above, “gliding” produces small saccades instead of smooth movement. However, by following a moving target, we can easily draw a smoothly curved circle by gaze (Tchalenko, 2001).

Smooth pursuit can also be used for interaction: finding a correlation between the object movement and the user’s eye movements can be used to select the desired object (Vidal, Bulling, & Gellersen, 2013). Location information (or calibration) is not needed, if the mapping is based on the object’s movement pattern and speed. With this method, it is possible to select one specific object out of many, by testing which object’s movement pattern best matches with the user’s eye movements (Fig. 23.6).

Smooth pursuit interaction is especially suitable for spontaneous interaction in public spaces where random passers-by cannot be expected to go through a calibration phase (Khamis, Alt, & Bulling, 2015; Velloso, Wirth, Weichel, Esteves, & Gellersen, 2016). In addition, it is useful in any interaction situations where calibration is hard to do or hard to maintain, such as for entering pin codes (Cymek et al.,

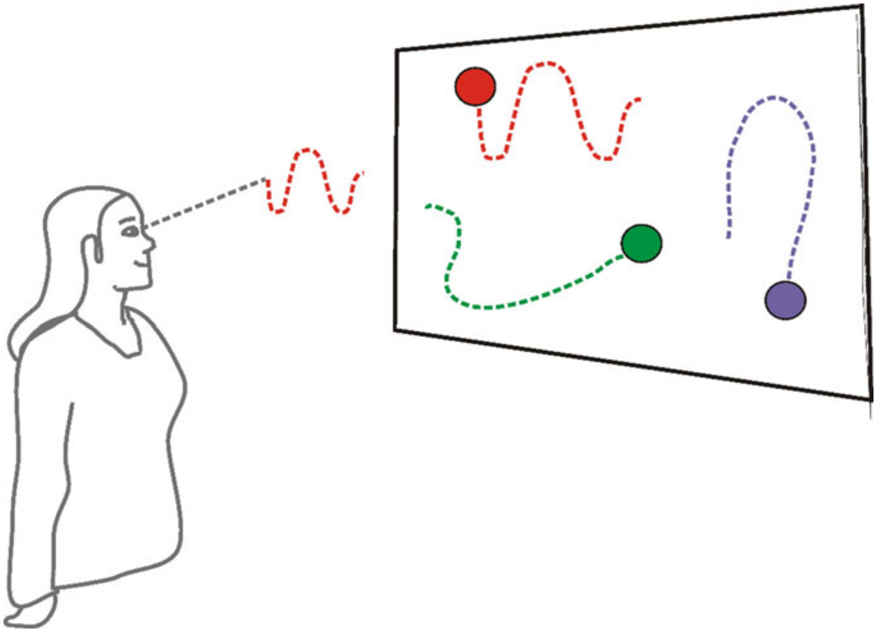


Fig. 23.6 In smooth pursuit interaction, an object is selected by following it with gaze [re-drawn based on the original by Vidal et al. (2013)]

2014) or controlling a smart wrist watch (Esteves, Velloso, Bulling, & Gellersen, 2015). However, smooth pursuit can also be useful for interacting with the conventional user interface widgets (Špakov, Isokoski, Kangas, Akkil, & Majaranta, 2016). Smooth pursuit movement is easy to suppress, thus making it easy to avoid unintentional activations (Schenk, Tiefenbacher, Rigoll, & Dorr, 2016).

23.4.4 Voluntary Vergence

Convergence (moving eyes inwards to see objects closer) or *divergence* (looking at objects far away) can be used voluntarily for gaze interaction. For example, while focusing on an object of interest, the user moves their gaze behind the object at will, thus using voluntary divergence to select the object of interest (Kudo et al., 2013). Forcing oneself to move the viewpoint forward can be learned but the opposite, moving the viewpoint nearer by simply crossing the eyes, may be easier. Kirst and Bulling (2016) carried out an experiment with voluntary convergence and found it a feasible technique to make selections. All participants were able to use the technique after practice but their experiences varied from "completely awkward" to "surprisingly easy". Convergence is a fairly robust selection method in 2D interfaces, where it would not happen by accident. One probably would not make exaggerated convergences in 3D interfaces either, but smaller-scaled divergences might occur.

23.4.5 Voluntary Pupil Dilation

Pupil dilation is one (though unnatural) way to implement eyes only interaction in highly restricted settings. Even if not natural, it is possible to learn to control pupil dilation by physical and psychological self-regulation (Ekman, Poikola, Mäkäräinen, Takala, & Hämäläinen, 2008b). However, people vary in their ability to successfully control pupil dilation. Even for those who manage to do it, it is possible only over brief periods (Ehlers, Strauch, Georgi, & Huckauf, 2016). Such "magical" control method could be used, for example, in gaze-based games (Ekman, Poikola, & Mäkäräinen, 2008a). However, the technique should be used cautiously as it depends on many external factors, most notably the lighting conditions. For more information on pupil behavior, see Laeng and Alnæs (this volume).

For attentive applications, information from natural pupil size changes may provide useful information of the user's emotional and cognitive state (Partala & Surakka, 2003).

23.4.6 Voluntary Blinks and Winks

Similar to using gaze for interaction versus visual inspection, using the blinks for control induces the Midas touch problem. *Intentional blinks* can be separated from natural blinks by using prolonged blinks (Grauman, Betke, Lombardi, Gips, & Bradski, 2003). A blink can be used together with gaze pointing to select the focused item, or as an additional switch that implements a specific function, e.g. a "stop" command (Heikkilä & Räihä, 2012). A blink can also be used as a switch, combined with automatic scanning that moves the focus from item to item (MacKenzie & Ashtiani, 2011). Most commercially available eye-control systems include blinks as one of the optional selection methods (Majaranta et al., 2009b).

In addition to being an active control method, blink detection may also be used in attentive interfaces, for example, to detect alertness (González-Ortega, Antón-Rodríguez, Martínez-Zarzuela, & Díez-Higuera, 2013) or to avoid computer vision syndrome where eyes get exhausted by too much screen time without blinking (Han, Yang, Kim, & Gerla, 2012).

Winks (closing only one eye) should be fairly robust as they require intentional action from the user. On the other hand, producing winks may be hard for some people. If a person can wink, wink with one eye can implement left click and the other be used for right click (Missimer & Betke, 2010).

23.4.7 *Eye-Based Head Gestures*

Eye-based head gestures detected from data provided by eye trackers may bring a new dimension to the gaze-based interaction. Head gesture detectors utilize eye position in a remote (Špakov & Majaranta, 2012) or head-mounted (Mardanbegi, Hansen, & Pederson, 2012) camera view that some eye tracking systems provide along the gaze point. The user views the target (using the eyes as a pointing device) and executes the head gesture while keeping the eyes on the target. A nodding gesture may be an especially convenient method to select the target under visual focus.

23.4.8 *Summary of Eye Movement Types and Their Application*

Table 23.1 summarizes the ways different eye movements have been used in human-computer interaction. Their advantages and challenges are listed together with indications of their usage in gaze-based interaction.

23.5 Design Considerations

Above, we saw how different eye movement types can be exploited in human-computer interaction. Here we will discuss some design principles based on the lessons learnt in the research articles reviewed in this chapter.

General principles for good usability can also be applied for gaze-based interfaces. For example, when designing and evaluating gaze-based systems, one might consult the widely accepted usability heuristics (see Nielsen, 1994). These usability heuristics substantially affect the system's ease of use and learnability. Some of these principles are discussed below, with notions specific to the gaze input point of view.

Box 2: Usability heuristics (Nielsen, 1994)

- visibility of the system status,
- match between the system and real world (e.g. speaking the user's language),
- user control and freedom, consistency and standards,
- error prevention,
- recognition rather than recall (e.g. minimize memory load by making actions visible),
- flexibility and efficient use,
- aesthetic and minimalist design,
- help users recognize, diagnose, and recover from errors,
- provide help and documentation.

Table 23.1 Eye movement types in gaze-based interaction

	Advantages	Challenges of restrictions
Fixations and dwell select		
Category		
Fixed dwell time threshold	Simple to use	Finding the best threshold
Adjustable dwell time threshold	Can be adjusted to user and task	Too frequent adjustments can affect the interaction rhythm
Dwell-free selections	Potential for fast interaction	Restricted to text entry
Saccades and gaze gestures		
Number of strokes		
Simple gestures (single or few strokes)	Fast	Can be confused with natural eye movements
Complex gestures	Don't get confused with natural eye movements	Need learning, can be error prone
Continuous pointing	Useful for zooming and steering tasks	Not useful for other tasks
Location of strokes		
In a fixed part of the screen	Don't get confused with viewing the screen content	Require shift of attention from target to control
Anywhere on the screen	Can be used with a variety of trackers, including EOG-based	Need to be distinguished from natural eye movements
Off-screen	Can be used with a variety of display devices	Limited set of commands
Less used or newer methods		
Category		
Smooth pursuit	Does not require calibration; can be used for calibration; suitable for ubiquitous use of eye tracking	Slower than traditional techniques
Voluntary vergence	Robust, does not happen by accident	Can be awkward for users
Voluntary pupil dilation		Not possible for every user; only works for brief moments; affected by (changing) lighting conditions
Voluntary blinks and winks	Can be easily separated from natural blinks; can be used together with the traditional methods	Can only work as switches
Eye-based head gestures	Simple and intuitive selection method; allows keeping the gaze on the target while interacting	Requires ability to move the head (not possible for paralyzed users)

23.5.1 Controllability and Non-interfering Design

Finding a suitable eye movement type for an application affects how easy it is to control, and how well it enables the user to concentrate on the task instead of the interface. Which types of eye movement work best depends on the task, context, and the user. Dwell select (selection when dwell time exceeds a threshold) is well suited for desktop applications that rely on point and click interactions. However, dwell select is perhaps not the best option for public displays that are used by visitors who are not willing to go through a calibration procedure to get the pointing accuracy required. In such a context, a more feasible option could be the one applied in GazeHorizon (Zhang, Chong, Müller, Bulling, & Gellersen, 2015), which maps uncalibrated horizontal eye movements to sideways scrolling. Keeping the gaze at the center does nothing, allowing the user to look at the object in the center without a risk of Midas touch. Figure 23.7 illustrates the idea.

Some diseases may restrict the available eye movement types. Involuntary nystagmus may prevent focusing on small targets or inability to move eyes horizontally

Fig. 23.7 GazeHorizon (Zhang et al., 2015) makes any content looked at by the user drift to the center. A hint ("Look here!") is given on both sides of the screen to attract the user to glance on the side. Photo courtesy of Yanxia Zhang, 2015 (c) Yanxia Zhang

may restrict the interaction to vertical eye movements alone. According to Donegan (2012), even the most complex cases can be solved and a functional eye control method can be found with persistent user-centered design and iteration.

Perhaps the most important part of being in control of a system based on eye movements is the ability to cope with the Midas touch problem. In gaze-controlled systems, using a combination of gaze pointing and a separate switch for selection is a good solution (Kumar, Paepcke, & Winograd, 2007; Stellmach & Dachselt, 2012). The switch can be anything, whatever is available or preferred by the user: speech, foot button, puff control, frown, smile, blink, nod, etc.

Alternatively, the system can exploit selection methods with multiple steps. A gaze-reactive button could have two separate areas, one for the command and a separate selection area next to it. The selection area could also be rolled out only after a fixation on the object area is detected, as in NeoVisus (Tall, 2008). The idea of separate areas for viewing and selection is also exploited in the context switching method (Morimoto & Amir, 2010). The interface is replicated, for example, so that the keyboard is duplicated. After the user finds the desired key, it can be selected by looking at the same key on the duplicated screen, i.e. by switching context. This enables viewing the whole screen without fear of the Midas touch problem but requires a duplicate amount of screen space.

Gaze gestures with specific patterns can be used to make selections. Gestures can be used independently or combined with dwelling. Istance et al. (2008) experimented with gaze-alone control of virtual worlds and games, by combining gaze pointing with quick glances outside of the borders of the screen for changing the mode. When the locomotion mode is activated, an avatar on the screen walks to the direction pointed by gaze. When the mouse action mode is activated, the avatar stays in place and the gaze can be used for selecting objects. These and other methods for selection are discussed in more detail by Skovsgaard et al. (2012) in their chapter on computer control by gaze.

One important issue to consider with eye control is the possibility to turn it temporarily off. The user may wish to visually investigate the screen without the fear of false activations. For people with disabilities, it is important that the eye control can then also be turned on again, without a need to call for an assistant—which may be impossible if the user relies on gaze alone. In addition, ability to recalibrate easily is desired. Ideas of how these can be implemented can be found in the report by Donegan et al. (2006).

Since eye movements are partly unconscious and easily distracted, and also because the technology is not always totally robust, an easy way to cancel any gaze-initiated actions is crucial for good usability. Some systems have been implemented so that the system's reaction to gaze is immediate and there is no big penalty for wrong eye movements. For example, in Dasher (Ward & MacKay, 2002) cancelling is part of the continuous control method: selection is done by pointing at the letters on the right side of the screen and cancelling by pointing at the left side of the screen.

In attentive interfaces, poorly implemented proactivity and unwanted assistance may substantially distract and annoy the user. In addition to spatial accuracy (reacting on the correct target), also temporal accuracy is important. The system should recognize when the user needs help. In developing the iDict reading assistant, Hyrskykari, Majaranta, Aaltonen and Räihä (2000) combined knowledge from eye movements with lexical and syntactic analysis. The frequency of the words and word length were used together with time spent on reading to decide if translation was needed. Thus, rare words got translated faster than common words. With grammatical knowledge of the text the system was also able to provide better quality assistance: it could, for instance, translate verbs and nouns correctly.

Wrong interpretations are bound to happen sometimes. The negative effect of the false alarm can be diminished by careful design of the interface. The gaze-evoked action can perhaps be implemented so that the user can ignore it, instead of being required to stop their task and forced to react. For example, Hyrskykari et al. (2003) provided automatic translations to difficult words in a subtle manner that allowed users to continue reading, sometimes even without noticing that a translation had been given.

23.5.2 Feedback and Visibility of the System Status

Since the eyes are primarily used for perception, the system should clearly inform the user of the system status and the system's response to the eye movements. Proper feedback not only helps in preventing the Midas touch problem but it can also significantly improve user performance and satisfaction (Majaranta, MacKenzie, Aula, & Räihä, 2006).

First, the user should know if the system is following the user's gaze correctly. This cannot be achieved if the system does not see the user's gaze in the first place. To help the user in finding an optimal position for tracking of the eyes, a desktop application could visualize good positioning during calibration (Donegan et al., 2006), or a gaze-aware public screen could give hints on where the passer-by should stand for the remote camera to see their eyes (Alt, Bulling, Gravanis, & Buschek, 2015).

Without any feedback on the focused object, the user would not know if the gaze tracker has correctly recognized the desired target. One option could be showing a cursor that is following the user's gaze. That can, however, be distracting, especially if the calibration is a bit off; it may be hard to prevent oneself from automatically glancing at the moving cursor, which would then move the cursor further as the calibration is off (Jacob, 1991). Thus, in some cases it is better to, for example, highlight the target itself. On the other hand, recent research (Graupner & Pannasch, 2014) has shown that the cursor movement may not actually be as distracting as previously thought and showing the cursor may even improve interaction at least in some cases, such as gaze typing, where seeing a semi-transparent cursor may help the user in understanding why the soft keyboard is not reacting as they expect.

If dwell time is used for commands, the user's gaze on the target only initiates the action. The actual selection happens after the required dwell time threshold has passed. Furthermore, dwell select requires that the user's gaze stays on the target during the whole dwell duration. Thus, the user cannot move their eyes away from the key, for example, to check if the desired letter has already appeared in the text input field. Feedback on the progression of the dwell duration may help the user maintain the focus on the key long enough for it to be selected (Majaranta et al., 2006). It should be noted, however, that it is also possible to implement the dwell time accumulation so that the duration does not reset if gaze data samples land outside of the key (Hansen & Hansen, 2006; Špakov, 2008). This can happen either because of inaccuracies with the eye tracker or if the user unintentionally glances outside of the focused target. Even in this case, it is useful to give feedback on the remaining dwell duration (see an example in Fig. 23.8).

In addition to feedback on the focused object, clear feedback on the action is important as well. The feedback on selection can be shown on the target itself, especially if the dwelling or smooth pursuit requires the user to look at the target anyways. However, with short dwell times where the user moves the gaze away from the target very fast or when rapid gaze gestures are used, other feedback modalities may be useful. For example, if a person controls a mobile phone by off-screen gaze gestures, haptic feedback on the hand-held phone can inform the user of the successful action (Kangas et al., 2014b). Haptic feedback may also be preferred in situations

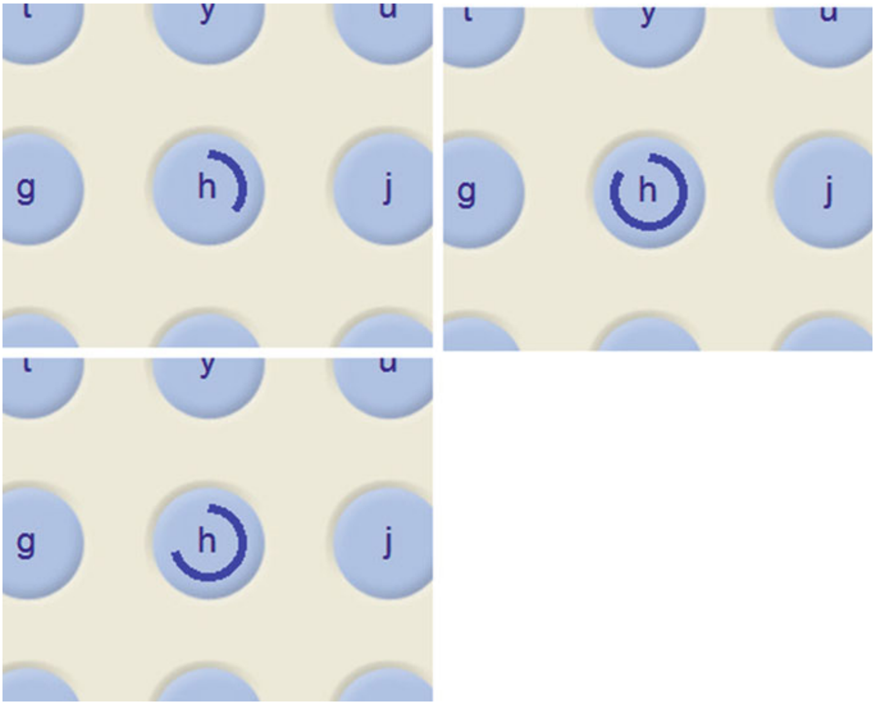


Fig. 23.8 An animated closing circle indicates the progression of dwell duration in AltTyping (available at https://www.sis.uta.fi/~csolsp/downloads.php). If cumulative dwell is used, the circle goes backwards while gaze is out and continues when the gaze returns into the key

where privacy is needed; haptic feedback can only be felt by the person wearing or holding the device. In head-mounted mobile systems, information on the gazed object could be whispered in the user's ear (Eaddy, Blasko, Babcock, & Feiner, 2004). Even with dwell-based systems, users may appreciate additional auditory or haptic feedback since it can be heard or felt even during a blink (Majaranta et al., 2016).

Since eye movements can be very fast, a gaze interface designer should pay attention to the system delay (latency). Eye trackers need to deal with noise in the signal. Even high-end devices cannot avoid the natural jitter of the eyes (Alexander & Martinez-Conde, this volume). The devices may therefore smooth the raw gaze signal, which affects the delay, and may affect the interaction. If the delay between the gaze event and the corresponding feedback is too long, it may confuse the user and hinder performance (Kangas et al., 2014a). When gaze is used for mouse emulation, smoothing the gaze data may help in stabilizing the cursor movement, making it easier to place the mouse cursor over a target and keeping it there. However, if the cursor is used for controlling action games or for fast gaze writing with Dasher, too much smoothing can make the system react too slowly (Donegan et al., 2006). Launching mouse events with gaze gestures may also be problematic because of the fast speed of the saccades. Istance et al. (2010) pointed out that there is a variable delay after the gaze enters a region and before a mouse-over event is generated, which makes the use of Windows events instead of gaze events unreliable.

In attentive applications that are based on natural eye movements one should find a balance between feedback that is not disturbing but that still informs the user of the system status. For example, an application that provides automatic translations depends on good tracking accuracy of the current line of text and the word that is focused. The system could, for example, underline the current line so that the user immediately notices if a wrong line is being tracked, and could easily correct it by pressing the arrow key (Hyrskykari, Majaranta, & Räihä, 2003).

23.5.3 Layout, Color Design, and Typography

The principles of visual usability, aesthetics and minimalistic design are well applicable for gaze-based interfaces. Avoiding unnecessary information is an important part of successful gaze-based user interface design. Limited accuracy and precision mean that the objects should be fairly big to be easily selectable by gaze pointing. How much information can be comfortably fit on the screen depends also, of course, on the screen size. If very small screens are used with eye movements, one should consider using off-screen targets or gestures (see the section of different types of eye movements for examples).

For desktop size screens, it may be useful if the content is located separately, away from the gaze-reactive controls as it allows the user to view the content without needing to worry about accidental activations (Penkar, Lutteroth, & Weber, 2012). Since the accuracy is often better in the center, placing the frequently used objects in the center is useful. Alternatively, easy access to all objects can be facilitated by increasing the size of the objects near the sides of the screen where the accuracy is not as good as in the center (Tuisku, Surakka, Rantanen, Vanhala, & Lekkala, 2013).

One option is to use off-set menus or context sensitive gaze widgets that appear when needed. If a person dwells on a web page link or menu, a gaze-reactive popup menu with fairly large buttons could appear (Yeoh, Lutteroth, & Weber, 2015). Instead of trying to hit one of the small, closely located menu options, the user can comfortably select one of the big buttons that corresponds to the desired link or menu option (see Fig. 23.9). In addition to such "semantic zooming", also graphical zooming can be applied (Skovsgaard et al., 2012). In practice, many gaze control systems provide a menu bar for accessing gaze control options, including a magnification option in addition to mouse button functions.

Gaze-based text entry using virtual keyboards is a special case that has been studied a lot. In text entry, not only the size of the keys matters but also the keyboard layout. Direct pointing with dwell select on a full-sized keyboard is more efficient and preferred by people with disabilities compared to a layout with only a few large items combined with word prediction (Pannasch, Malischke, Storch, & Velichkovsky, 2008). The order of the keys (QWERTY vs. alphabetical vs. frequency-based) in the full keyboard layout does not have significant effect on efficiency, though the familiar

Main Page

Contents [hide]

1 COGAIN - Communication by Gaze Interaction
1.1 I don't know anything about Eye gaze
1.2 I know about Eye Gaze, but I want more information
1.3 I am a student/researcher in HCI
1.4 I want to cor

COGAIN - Communication by Gaze Interaction

This is the Knowledge Base of the COGAIN Network. Here you can find information, photos, videos, guides, reports, links and more. You are also welcome to join our community and help us build knowledge about Gaze Interaction.

For information and news about the COGAIN Association, check our website at www.cogain.org.

I don't know anything about Eye gaze

You can start checking out the following pages:

- User Involvement, filled with information and useful tips for new users!
- FAQ, frequently asked questions (about COGAIN and the eye tracker as communication aids)
- Training, educational materials for (self)training of gaze interaction and eye tracking related issues
- Media, for a list of media appearances, press releases, brochures
- COGAIN forum

I know about Eye Gaze, but I want more information

You can start checking out the following pages:

Fig. 23.9 Screen capture of the "semantic zooming" of the MyTobii browser by Tobii. *Note* MyTobii is not available anymore but is replaced with other solutions, see www.tobii.com

QWERTY may be preferred (Raupp, 2013). Text entry and on-screen keyboards are also needed with other applications, e.g., filling in web forms. In such cases, the user may want to see at least part of the web page at the same time as entering text. Then either a method based on gaze gestures (Wobbrock, Rubinstein, Sawyer, & Duchowski, 2008) or a scrolling keyboard (Špakov & Majaranta, 2009) that only shows one or a few rows of the keyboard at a time can be useful. Design issues of gaze-based text entry are discussed in more detail by Majaranta (2009).

When designing the layout for a gaze-based interface, it is worth considering how the system can passively help the user focus their gaze. For example, it may be useful to include a visible fixation spot inside a gaze-reactive key (Špakov & Miniotas, 2005). Otherwise, the user may look at the corners of the button instead of focusing on the center. Eye drawing is difficult on a blank surface. A grid of small dots (preferably with adjustable density) helps the user in focusing their gaze, for example for planning the start and end of a line (Hornof, Cavender, & Hoselton, 2003).

Chitty (2013) proposed design guidelines for eye-controlled technology based on a literature review, surveys, expert interviews and user testing with people with disabilities. Many of the guidelines have already been discussed above (feedback, layout, size). In addition, she proposes guidelines for the color scheme and typography. She found out that, in extended use, dark backgrounds are less straining to the eyes than white backgrounds. Color preferences, however, varied between people. For typography, Chitty recommends using fairly large font sizes and typefaces that are familiar to the user. Also here, the best option is to give the freedom of choice to the user.

23.5.4 Acceptability, Social Norms, and User Experience

Social acceptability and user experience significantly affect the user's willingness to use the technology. This applies to all users, both those with special needs as well as able bodied people.

Mele and Federici (2012) conducted a literature review of eye tracking technology and found that more emphasis is required on the user experience. In most studies, the focus is on the technology design and users are considered as a measurement object to prove that the technology works. User experience and social norms affect the acceptability and utilization of communication aids. These systems are partly underused because self-consciousness prevents people from using them in public (Hodge, 2007), even though they could be of great benefit. Poor user experience in the beginning may also significantly affect the motivation to use the system. Thus, one should devote enough time during the introductory period for solving problems, learning to use the system and making sure the content is motivating—especially in the case with children (Holmqvist, Thunberg, & Dahlstrand, 2017).

When gaze is used in human-human communication, a screen-based communication system can be obstructive for face-to face-communication (Chitty, 2013). It would be best if the user can maintain eye contact as much as possible also during technology-assisted eye communication.

The ability to maintain gaze contact in human-human interaction is also important for people wearing smart glasses. Akkil et al. (2016) studied user expectations towards everyday gaze interaction with smart glasses. Even though participants were generally positive for the idea of using gaze-aware smart glasses, they were concerned about the social aspects of their use. For example, they worried about looking weird, if they use abnormal eye movements in public. Dwelling on people to get more information about them via the smart glasses was considered invasive; "people are not products". Interaction with objects was seen more positively. If gaze gestures are used, people would want the system to help them remember the gestures and provide clear feedback on their use. Instead of automatically offering help, it was suggested that the glasses could politely ask if the user wants more information about the focused item.

23.5.5 Customizability and User-Centered Design

Chitty (2013) as well as Donegan et al. (2006, 2009, 2012) emphasize the importance of being able to customize the interface to suit the needs and wishes of the user. Especially considering people with disabilities, there is no one-size-fits-all solution. There is a great diversity in the needs and user characteristics. One should also keep in mind that many of the diseases are progressive and the condition may vary from day to day. Thus having a one-time setup is not enough.

Successful design requires understanding the users and their needs (Donegan et al., 2009) and expectations (Akkil et al., 2016). It is also important that the actual intended users are involved, otherwise the results may not be valid (Istance, Vickers, & Hyrskykari, 2012). Perhaps the best way to make sure the system really meets the needs of the users is to apply the participatory design method, where the users are part of the design team from the very beginning (Donegan, 2012; Hornof, 2009).

User experience and usability are essential if one wishes that the eye control technology is adapted in everyday use, and not left only as a laboratory experiment or a niche technology only used by those who do not have options.

23.6 Development of Gaze-Responsive Applications and Web Pages

This section provides a quick introduction and discusses some of the challenges related to the development of gaze-controlled applications to those who wish to implement gaze-based applications of their own. More detailed, up-to-date information and examples can be downloaded from Oleg Špakov's web pages (http://

www.sis.uta.fi/~csolsp/downloads.php). The examples include source code that can be modified, thus providing a convenient access point to implementation of gaze-responsive applications. The implementation is fairly straightforward but basic programming skills are required.

23.6.1 Towards Tracker-Independent Applications

As we saw above, during the last decades, tens of prototypes were constructed in the academic institutes, and dozens of commercial systems were released to the market. Most of these systems are supplemented by dedicated software, usually developed by the system manufacturer. Some use a closed protocol for communicating with the hardware, while other systems allow developers to access eye tracking data using a specific API (application programming interface). The variety of APIs and absence of a common (standard) protocol for accessing eye tracking data makes the development and support of gaze-responsive tools and applications more expensive and time-consuming. Several protocols were proposed to standardize the communication between tracking systems and applications (for example, “Open Gaze API” from GazePoint), but none was accepted. This issue remains unsolved even today, when the eye tracking technology steps into its mature phase with cheap and reliable off-the-shelf systems introduced in the mass market.

Until the market players develop a standard protocol, the common method to construct software that supports multiple eye tracking systems is based on using modular architecture. A separate module needs to be developed for each supported system (unless some of them share the same API). While each module utilizes a system-specific API to access eye tracking data, all modules expose the same programming interface to the application core, which then communicates with all eye tracking systems using the same set of methods.

An early example that demonstrates a shift toward a modular architecture is ILAB, a tool for Matlab to analyze recorded gaze data (Gitelman, 2002). This tool does not collect data directly from eye tracking systems; rather, it reads data collected previously and saved in a file. For that ILAB has separate modules to parse gaze data recorded by several eye tracking systems. The parsed data is stored in a single format that is used in all analysis and computations, independently of the system used for collecting it.

OGAMA (Open Gaze And Mouse Analyzer[2]) is another application with modular architecture for gaze data recording and analysis (Voßkühler, Nordmeier, Kuchinke, & Jacobs, 2008). It was implemented originally to support only one eye tracker but now it supports eye trackers from nine commercial manufacturers and two open-source eye trackers. The application records gaze data directly from an eye tracker, and if the tracker’s API allows configuring and calibrating no other tools are needed to work with OGAMA. The application is published as open-source, therefore the

---

[2] Available from http://www.ogama.net/ (July 2017).

modules for data access and communication with eye tracking systems can be reused in software (written in C# programming language).

ETU-Driver (the Eye-Tracking Universal Driver) makes a step forward in facilitating the development of system-independent gaze-responsive applications. It was first developed as a part of iComponent (Špakov, 2008), software similar to OGAMA, and it contained an abstraction layer and modules to work with multiple eye tracking systems transparently from the application core. Later it was extracted into a separate tool to serve as a middleware layer that can be integrated into any software that can embed COM/ActiveX objects (Bates & Špakov, 2006). The tool is distributed with usage examples in several programming languages (C++, C#) that may serve as templates to start developing gaze-responsive applications.[3]

ETU-Driver supports plugins that extend its capabilities. For example, the "Fixation Detector" plugin consumes gaze points, detects fixations, and streams back fixation start and end events. It may be useful in case an application bases gaze interaction on fixations rather than on gaze points, but the eye tracking system in use does not detect these events internally. Another plugin streams gaze data into the network. There are plugins to transmit gaze data over Bluetooth connection, control mouse cursor by gaze, show notifications when data is not available or of poor quality, etc.

The existing solutions that emulate common gaze tracking API have some shortcomings: a fixed data structure restricts access to system-specific data. The issue may be illustrated in the following example. Normally, an eye tracker outputs a gaze point that consists of 2 or 3 numerical values. The wearable Dikablis eye tracker from Ergoneers (www.ergoneers.com) estimates up to 17 gaze points: one ordinary gaze point that is relative to the scene camera, and up to 16 gaze points relative to 16 detectable markers (a.k.a. 2D barcodes). At the moment, these 16 gaze points cannot be transmitted to the application core using tools like ETU-Driver, thus a solution with more flexible data structure is still missing.

Most of the eye tracking software and tools have been developed so far for traditional PCs with ×86 CPU architecture, primarily running the Microsoft Windows operating system. Exceptions are rare, but evolution of high-level (interpreted) programming languages and rapid expansion of mobile devices with ARM processors already changes the trend. A good example is PyGaze,[4] an open-source eye tracking software that can be used on MS Windows, Linux and Mac OS X operating systems. It is written in Python and currently supports three commercial systems. The tool allows real-time data collection and usage in Python scripts written by users.

23.6.2 Gaze on the Web

While traditional applications have access to gaze data either using a system API or tools like ETU-Driver, web-based applications running in browsers do not have such

---

[3] Available from https://www.sis.uta.fi/~csolsp/downloads.php.

[4] Available from http://www.pygaze.org/ (July 2017).

access. The share of these applications tends to increase, therefore it is important to have a way to stream gaze points to a browser. Such connectivity is available using WebSockets technology, as first shown by Wassermann et al. (2012), or using eye trackers based on web-camera image analysis directly in the web page code (Papoutsaki et al., 2016). Recently, Kumar, Menges, Müller, and Staab (2017) developed a browser based on Chromium that is integrated with an eye tracker and provides access to gaze data from a page code out-of-the-box. With gaze points accessed from JavaScript code, gaze-responsive web pages and applications introduce new opportunities in interaction with online media using gaze.

One perspective direction in bringing gaze to the web is social media. Today's web applications allow people to share their activities, such as collaborative text editing, drawing, and gaming. Sharing the focus of attention in a virtual world may have a strong impact: it is simply faster and more convenient to share the point of attention using gaze rather than the mouse (Qvarfordt, Beymer, & Zhai, 2005). The attention can also be collected through some period of time and shared later (say, while observing static content such as images). Further, gaze data could be accumulated from every observer and, for example, visualized online as heat maps on shared media. This might be especially important in online learning/teaching, which is nowadays rapidly developing.

Dedicated JavaScript libraries available today, such as GazeHook,[5] support the development of gaze-responsive web pages by implicitly executing all necessary basic actions, such as connecting to the web-socket server and converting data received from a WebSocket to JavaScript objects, as well as providing many useful methods to deal with gaze data online. For example, the *gazeTarget.js* library[6] implements several algorithms to detect user focus on web page elements. It may show gaze pointer and visual feedback when using dwell-time selection algorithms, play "click" sound on object selection, correct gaze point by analyzing small head movements (Špakov, Isokoski, & Majaranta, 2014), etc. The library has several built-in utilities, such as a gaze-controlled keyboard that can be shown when the gaze focuses on a text input field, calibration verification tool, automatic page scrolling (similar to the one shown in Fig. 23.10), nodding gesture detector, and free-form head gestures recording and recognition.

Similar functionality is provided by the GazeTK framework (Heil, Renner, & Pfeiffer, 2015): it consists of the component that can handle data from multiple eye trackers and stream it to a browser extension BlickBrowser via WebSocket that enables gaze data access from the page code. The framework implements page auto-scrolling, gaze-enabled main and page menus, smart zoom and link selection, etc. Additionally, it exposes a set of JavaScript functions for developers to add gaze interactivity on custom elements.

---

[5] Available from http://ux.fiit.stuba.sk/GazeHook (July 2017).

[6] Available from https://github.com/lexasss/etudriver-web (July 2017).

Transparent scrolling panels with 4 speed zones shown as icons

Preview of the page shown in the active (focused) zone

Welcome to ECEM 2013

Submitted by nils on Tue, 05/29/2012 - 14:37

ECEM 2013 is the 17th European Conference on Eye Movements, with the original aims of the very first ECEM, 'to exchange information on current research, equipment and software', at the forefront. In 2013, it is held in Lund, Sweden.

Conference: 11th to 16th August 2013

Methods workshops: 7th to 10th August 2013 (see here)

Fig. 23.10 Scrolling panels with semi-transparent areas to indicate varying scrolling speeds and page thumbnail at the focused area for in-place inspection of the current scroll location (© O. Špakov)

23.7 Conclusion and Future Directions

In this chapter, we have given an introduction to the application of eye movements in human computer interaction. We have seen how information from natural eye movements can be used subtly in the background as an additional information channel and how to use voluntary gaze input for communication and computer control. We have focused on gaze alone, but one should remember that there are many factors that determine which modality works best for a particular user. Also, many times there is no need to use only gaze. Reviews of using gaze together with other modalities are discussed by Jacob and Stellmach (2016) and extensively by Qvarfordt (2017).

Knowledge of the user, context and task are essential for interpreting the gaze behavior. Noninterfering design should be a guiding principle for attentive interfaces, to avoid distraction and frustration caused by unwanted proactivity. Key issues to consider when designing gaze-based interfaces are the Midas touch problem and feedback. The user should maintain the feeling of being in control. For example, the system should provide a possibility to cancel erroneous actions. In addition to technical implementation, user experience and usability affect the potential utility and acceptance of the novel technology.

Developments in the technology have significantly increased its availability and usability. Eye trackers have already for a long time been used for daily communication by people with severe disabilities and the number of potential beneficiaries is increasing rapidly as the technology continues to improve (Donegan et al., 2009). The new web technologies along with the ever more accurate web cameras open new possibilities for using eye input with browsers as the platform.

Simple eye-awareness can already been found in commercial mobile phones (Samsung, 2013, p. 30), and some off-the-shelf games already support gaze interaction (Tobii, 2016). At the same time, there is active research on pervasive and mobile eye tracking (Mardanbegi, Jalaliniya, Khamis, & Majaranta, 2016), extending their use to public environments and wearable devices. We hope this chapter serves as an entry point for those who wish to take advantage of this exciting technology in modern interfaces.

23.8 Key Events and Suggested Readings

The first conference series dedicated to research on eye movements, the European Conference on Eye Movements (ECEM), was launched in 1981 with Rudolf Groner as the driving force. ECEM is organized biannually, and it has grown into the biggest gathering of eye movement researchers. Although the original focus was on psychological studies, applied eye tracking has in the last decade had a growing role in the programme. The abstracts of the papers are published in the *Journal of Eye Movement Research* (https://www.bop.unibe.ch/jemr).

The first international conference to explicitly concentrate on eye tracking applications, the ACM Eye Tracking Research and Applications Symposium (ETRA), was established in 2000, with Andrew Duchowski as the General Chair. Since 2000, ETRA has been organized biannually. A large part of the papers related to applied use of eye tracking are also published in general HCI conferences such as the ACM SIGCHI Conference of Human Factors in Computing Systems (CHI). There are also some relevant workshops that have been organized for several years, such as The Scandinavian Workshop on Applied Eye Tracking (SWAET, first organized in 2006) and the International Workshop on Pervasive Eye Tracking and Mobile Eye-Based Interaction (PETMEI, www.petmei.org, first organized in 2011; Bulling, Duchowski, & Majaranta, 2011). In addition, there has been an increasing number of related workshops with varying foci (e.g. on gaze & games and gaze in intelligent user interfaces).

The first book to thoroughly discuss the applied use of gaze, "Eye tracking methodology: Theory and practice", was published by Duchowski (2007). The book (now in its third edition) gives an overview of gaze tracking technology and also reviews gaze-based applications. A different point of view is taking in the book by Land and Tatler (2009), who leave the lab and study how eyes behave in real world activities. This is useful knowledge to take into account when designing gaze-based applications. A third key book in this area is "Eye tracking: A comprehensive guide to methods and measures" by Holmqvist et al. (2011). This method book is useful to anyone conducting eye tracking experiments, also including interactive applications.

Finally, we wish to mention one more milestone from the history of gaze-based HCI, or *gaze interaction*. The European Commission's funded Network of Excellence on Communication by Gaze Interaction (COGAIN, www.cogain.org), started in 2004. It gathered Europe's leading expertise in eye tracking, involving researchers, eye tracking manufacturers and assistive technology industry, and assistive technology professionals to work together on gaze-based assistive technologies. Users and their representatives were also involved from the very beginning (Bates, Donegan, Istance, Hansen, & Räihä, 2007). COGAIN eventually involved also members from the USA and Japan. COGAIN worked to better integrate of the eye tracking technology and assistive software, to develop eye tracking standards and to bring down the cost of the systems. Most results achieved during the EU-funded period (2004–2009) are reported in the "COGAIN book" (Majaranta et al., 2012) and others published in the COGAIN website (http://wiki.cogain.org/index.php/COGAIN_Reports). The work continues in the form of an association. It provides a wealth of information on gaze-based assistive technologies, downloadable software, links to open-source eye trackers, and bibliographies on its public web pages, see http://wiki.cogain.org/.

For those that need more information on designing interactive software in general, the Interaction Design Foundation provides free electronic books and other online materials related to interaction design, such as The Encyclopedia of Human-Computer Interaction by Mads Soegaard and Rikke Dam (Eds.), see https://www.interaction-design.org/literature.

23.9 Questions Students Should Be Able to Answer

What is the "Midas touch problem" and what can be done to avoid it?
What kind of challenges are involved in using gaze for mouse emulation?
Why is feedback especially crucial in gaze-based human-computer interaction? (Think, for example, differences in pressing a button with a finger versus using dwell time to select it.)
What are the key design principles to keep in mind when designing visually attentive interfaces?

Bibliography

Adams, N., Witkowski, M., & Spence, R. (2008). The inspection of very large images by eye-gaze control. In *Proceedings of the working conference on advanced visual interfaces (AVI '08)* (pp. 111–118). New York, NY: ACM.

Ajanki, A., Billinghurst, M., Gamper, H., Järvenpää, T., Kandemir, M., Kaski, S. … Tossavainen, T. (2011). An augmented reality interface to contextual information. *Virtual Reality, 15*(2), 161–173.

Akkil, D., James, J. M., Isokoski, P., & Kangas, J. (2016). GazeTorch: Enabling gaze awareness in collaborative physical tasks. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '16)* (pp. 1151–1158). New York, NY: ACM.

Akkil, D., Kangas, J., Rantala, J., Isokoski, P., Špakov, O., & Raisamo, R. (2015). Glance awareness and gaze interaction in smartwatches. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '15)* (pp. 1271–1276). New York, NY: ACM.

Akkil, D., Lucero, A., Kangas, J., Jokela, T., Salmimaa, M., & Raisamo, R. (2016). User expectations of everyday gaze interaction on smartglasses. In *Proceedings of the 9th Nordic conference on human-computer interaction (NordiCHI '16).* New York, NY: ACM.

Alt, F., Bulling, A., Gravanis, G., & Buschek, D. (2015). GravitySpot: Guiding users in front of public displays using on-screen visual cues. In *Proceedings of the 28th annual ACM symposium on user interface software & technology (UIST '15)* (pp. 47–56). New York, NY: ACM.

Barea, R., Boquete, L., Mazo, M., & López, E. (2002). System for assisted mobility using eye movements based on electrooculography. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 10*(4), 209–218.

Bates, R., Donegan, M., Istance, H. O., Hansen, J. P., & Räihä, K.-J. (2007). Introducing COGAIN: Communication by gaze interaction. *Universal Access in the Information Society, 6*(2), 159–166.

Bates, R., & Špakov, O. (2006). *Implementation of COGAIN gaze tracking standards.* Communication by Gaze Interaction (COGAIN, IST-2003-511598): Deliverable 2.3. Retrieved from http://wiki.cogain.org/index.php/COGAIN_Reports.

Bednarik, R., Gowases, T., & Tukiainen, M. (2009). Gaze interaction enhances problem solving: Effects of dwell-time based, gaze-augmented, and mouse interaction on problem-solving strategies and user experience. *Journal of Eye Movement Research, 3*(1), 1–10.

Bednarik, R., Vrzakova, H., & Hradis, M. (2012). What do you want to do next: A novel approach for intent prediction in gaze-based interaction. In *Proceedings of the symposium on eye tracking research and applications (ETRA '12)* (pp. 83–90). New York, NY: ACM.

Bee, N., & André, E. (2008). Writing with your eye: A dwell time free writing system adapted to the nature of human eye gaze. In R. Goebel, J. Siekmann, & W. Wahlster (Eds.), *Perception in multimodal dialogue systems, LNCS 5078* (pp. 111–122). Berlin Heidelberg: Springer.

Biedert, R. B., Schwarz, S., Hees, J., & Dengel, A. (2010). Text 2.0. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '10)* (pp. 4003–4008). New York, NY: ACM.

Bieg, H.-J., Chuang, L. L., Fleming, R. W., Reiterer, H., & Bülthoff, H. H. (2010). Eye and pointer coordination in search and selection tasks. In *Proceedings of the symposium on eye tracking research & applications (ETRA '10)* (pp. 89–92). New York, NY: ACM.

Biswas, P., & Langdon, P. (2013). A new interaction technique involving eye gaze tracker and scanning system. In *Proceedings of the 2013 conference on eye tracking South Africa* (pp. 67–70). New York, NY: ACM.

Borgestig, M., Sandqvist, J., Ahlsten, G., Falkmer, T., & Hemmingsson, H. (2016a). Gaze-based assistive technology in daily activities in children with severe physical impairments–An intervention study. *Developmental Neurorehabilitation, 20*(30), 129–141. https://doi.org/10.3109/17518423.2015.1132281.

Borgestig, M., Sandqvist, J., Parsons, R., Falkmer, T., & Hemmingsson, H. (2016b). Eye gaze performance for children with severe physical impairments using gaze-based assistive technology—A longitudinal study. *Assistive Technology, 28*(2), 93–102.

Bulling, A., Duchowski, A. T., & Majaranta, P. (2011). PETMEI 2011: The 1st international workshop on pervasive eye tracking and mobile eye-based interaction. In *Proceedings of the 13th international conference on ubiquitous computing (UbiComp '13)* (pp. 627–628). New York, NY: ACM.

Bulling, A., Roggen, D., & Tröster, G. (2009). Wearable EOG goggles: Eye-based interaction in everyday environments. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '09)* (pp. 3259–3264). New York, NY: ACM.

Bulling, A., Ward, J. A., Gellersen, H., & Tröster, G. (2011b). Eye movement analysis for activity recognition using electrooculography. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 33*(4), 741–753.

Caligari, M., Godi, M., Guglielmetti, S., Franchignoni, F., & Nardone, A. (2013). Eye tracking communication devices in amyotrophic lateral sclerosis: Impact on disability and quality of life. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, 14*(7–8), 546–552.

Cantoni, V., & Porta, M. (2014). Eye tracking as a computer input and interaction method. In *Proceedings of the 15th international conference on computer systems and technologies (CompSysTech '14)* (pp. 1–12). New York, NY: ACM.

Chapman, J. E. (1991). Use of an eye-operated computer system in locked-in syndrome. In *Proceedings of the sixth annual international conference on technology and persons with disabilities (CSUN '91).* Los Angeles, CA.

Chen, Y., & Newman, W. S. (2004). A human-robot interface based on electrooculography. In *Proceedings of 2004 IEEE international conference on robotics and automation (ICRA '04)* (pp. 243–248). IEEE.

Chitty, N. (2013). *User fatigue and eye controlled technology.* OCAD University. Toronto, Ontario, Canada: Licensed under the Creative Commons. Retrieved from http://openresearch.ocadu.ca/143/1/Chitty_MRP.pdf.

Cleveland, N. R. (1997). The use of a commercially available eyetracking system by people with severe physical disabilities. In *The 9th European conference on eye movements (ECEM '04).* Ulm, Germany.

Cymek, D. H., Venjakob, A. C., Ruff, S., Lutz, O. H.-M., Hofmann, S., & Roetting, M. (2014). Entering PIN codes by smooth pursuit eye movements. *Journal of Eye Movement Research, 7*(4), 1.

De Saint-Exupéry, A. (1943). *The little prince*. Reynal & Hitchcock.

Dechant, M., Heckner, M., & Wolff, C. (2013). Den Schrecken im Blick: Eye Tracking und Survival Horror-Spiele. In *Mensch & computer 2013* (pp. 539–542). Oldenbourg Verlag.

Donegan, M. (2012). Participatory design: The story of Jayne and other complex cases. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 55–61). Hershey, PA: IGI Global.

Donegan, M. M., Corno, F., Signorile, I., Chió, A., Pasian, V., Vignola, A. … Holmqvist, E. (2009). Understanding users and their needs. *Universal Access in the Information Society, 8*(4), 259–275.

Donegan, M., Gill, L., & Ellis, L. (2012). A client-focused methodology for gaze control assessment, implementation and valuation. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 279–286). Hershey, PA: IGI Global.

Donegan, M., Oosthuizen, L., Bates, R., Daunys, G., Hansen, J., Joos, M. … Signorile, I. (2005). *D3.1 User requirements report with observations of difficulties users are experiencing.* Communication by Gaze Interaction (COGAIN), IST-2003-511598: Deliverable 3.1. Retrieved from http://wiki.cogain.org/index.php/COGAIN_Reports.

Donegan, M., Oosthuizen, L., Daunys, G., Istance, H., Bates, R., Signorile, I. … Majaranta, P. (2006). *D3.2 Report on features of the different systems and development needs.* Communication by Gaze Interaction (COGAIN), IST-2003-511598: Deliverable 3.2. http://wiki.cogain.org/index.php/COGAIN_Reports.

Drewes, H., & Schmidt, A. (2007). Interacting with the computer using gaze gestures. In *Human-Computer Interaction—INTERACT 2007* (pp. 475–488). Berlin Heidelberg: Springer.

Duchowski, A. T. (1998). Incorporating the viewer's point-of-regard (POR) in gaze-contingent virtual environments. In *The engineering reality of virtual reality, SPIE proceedings* (Vol. 3295, pp. 332–343).

Duchowski, A. T. (2007). *Eye tracking methodology: Theory and practice*. London, UK: Springer.

Duchowski, A. T., Cournia, N., & Murphy, H. (2004). Gaze-contingent displays: A review. *CyberPsychology & Behavior, 7*(6), 621–634.

Duchowski, A. T., House, D. H., Gestring, J., Wang, R. I., Krejtz, K., Krejtz, I. … Bazyluk, B. (2014). Reducing visual discomfort of 3D stereoscopic displays with gaze-contingent depth-of-field. In *Proceedings of the ACM symposium on applied perception (SAP '14)* (pp. 39–46). New York, NY: ACM.

Eaddy, M., Blasko, G., Babcock, J., & Feiner, S. (2004). My own private kiosk: Privacy-preserving public displays. In *Eighth international symposium on wearable computers (ISWC 2004)* (pp. 132–135). IEEE.

Ehlers, J., Strauch, C., Georgi, J., & Huckauf, A. (2016). Pupil size changes as an active information channel for biofeedback applications. *Applied Psychophysiology and Biofeedback, 41*(3), 331–339.

Ekman, I., Poikola, A., & Mäkäräinen, M. (2008a). Invisible Eni: Using gaze and pupil size to control a game. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '08)* (pp. 3135–3140). New York, NY: ACM.

Ekman, I., Poikola, A., Mäkäräinen, M., Takala, T., & Hämäläinen, P. (2008b). Voluntary pupil size change as control in eyes only interaction. In *Proceedings of the symposium on eye tracking research and applications (ETRA '08)* (pp. 115–118). New York, NY: ACM.

Esteves, A., Velloso, E., Bulling, A., & Gellersen, H. (2015). Orbits: Gaze interaction for smart watches using smooth pursuit eye movements. In *Proceedings of the 28th annual ACM symposium on user interface software & technology (UIST '15)* (pp. 457–466). New York, NY: ACM.

Feit, A. M., Williams, S., Toledo, A., Paradiso, A., Kulkarni, H., Kane, S., & Morris, M. R. (2017). Toward everyday gaze input: Accuracy and precision of eye tracking and implications for design. In *Proceedings of the 2017 CHI conference on human factors in computing systems (CHI '17)* (pp. 1118–1130). New York, NY: ACM.

Findlay, J. (2005). Covert attention and saccadic eye movements. In L. Itti, G. Rees, & J. K. Tsotsos (Eds.), *Neurobiology of attention* (pp. 114–116). Amsterdam: Elsevier.

Foulds, R. A., & Fincke, R. W. (1979). A computerized line of gaze system for rapid non-vocal individuals. In *National computer conference 1979 personal computing proceedings* (p. 81).

Frey, L., White, K. J., & Hutchinson, T. (1990). Eye-gaze word processing. *IEEE Transactions on Systems, Man, and Cybernetics, 20,* 944–950.

Friedman, M., Kiliany, G., Dzmura, M., & Anderson, D. (1982). The eyetracker communication system. *Johns Hopkins APL Technical Digest, 3,* 250–252.

Frietman, E., & Tausch, P. (1978). *An eye-movement controlled communication appliance: From Eye-letter-selector to handi-writer.* A Progress and Evaluation Report at request of the Foundation of Multiple Sclerosis.

Gips, J., Olivieri, P., & Tecce, J. (1993). Direct control of the computer through electrodes placed around the eyes. In *Human–computer interaction: Applications and case studies (Proceedings of HCI International '93)* (pp. 630–635). Amsterdam: Elsevier.

Gitelman, D. (2002). ILAB: A program for post-experimental eye movement analysis. *Behavior Research Methods, Instruments, & Computers, 34*(4), 605–612.

González-Ortega, D. D.-P., Antón-Rodríguez, M., Martínez-Zarzuela, M., & Díez-Higuera, J. (2013). Real-time vision-based eye state detection for driver alertness monitoring. *Pattern Analysis and Applications, 16*(3), 285–306.

Goossens', C., & Crain, S. (1987). Overview of nonelectronic eye-gaze communication techniques. *Augmentative and Alternative Communication, 3*(2), 77–89.

Grauman, K., Betke, M., Lombardi, J., Gips, J., & Bradski, G. R. (2003). Communication via eye blinks and eyebrow raises: Video-based human-computer interfaces. *Universal Access in the Information Society, 2*(4), 359–373.

Graupner, S.-T., & Pannasch, S. (2014). Continuous gaze cursor feedback in various tasks: Influence on eye movement behavior, task performance and subjective distraction. In *HCI international 2014—Posters' extended abstracts* (pp. 323–329). Cham: Springer.

Hafed, Z. M., & Clark, J. J. (2002). Microsaccades as an overt measure of covert attention shifts. *Vision Research, 42*(22), 2533–2545.

Han, S., Yang, S., Kim, J., & Gerla, M. (2012). EyeGuardian: A framework of eye tracking and blink detection for mobile device users. In *Proceedings of the twelfth workshop on mobile computing systems & applications (HotMobile '12)* (Article 6, 6 pages). New York, NY: ACM.

Hansen, D. W., & Hansen, J. P. (2006). Robustifying eye interaction. In *Proceedings of the computer vision and pattern recognition workshop (CVPRW '06)*. IEEE.

Hansen, D. W., & Ji, Q. (2010). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(3), 478–500.

Hansen, J. P., Hansen, D. W., Johansen, A. S., & Elvesjö, J. (2005). Mainstreaming gaze interaction towards a mass market for the benefit of all. In *Proceedings of the 11th international conference on human-computer interaction (HCII '05).* Mahwah, NJ: Lawrence Erlbaum.

Heikkilä, H., & Ovaska, S. (2012). Usability evaluation of gaze interaction. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye tracking: Advances in assistive technologies: Advances in assistive technologies* (pp. 255–278). Hershey, PA: IGI Global.

Heikkilä, H., & Räihä, K.-J. (2009). Speed and accuracy of gaze gestures. *Journal of Eye Movement Research, 3*(2), 1–14.

Heikkilä, H., & Räihä, K.-J. (2012). Simple gaze gestures and the closure of the eyes as an interaction technique. In *Proceedings of the symposium on eye tracking research and applications (ETRA '12)* (pp. 147–154). New York, NY: ACM.

Heil, D., Renner, P., & Pfeiffer, T. (2015). *GazeTK/Blinker: OpenSource framework for gaze-based interaction for the web*. Poster presented at the 18th European Conference on Eye Movements (ECEM '15). Vienna, Austria. More information at http://gazetk.eyemovementresearch.com/.

Hewett, T. T., Baecker, R., Card, S., Carey, T., Gasen, J., Mantei, M. … Verplank, W. (1992). *ACM SIGCHI curricula for human-computer interaction.* New York, NY: ACM.

Hodge, S. (2007). Why is the potential of augmentative and alternative communication not being realized? Exploring the experiences of people who use communication aids. *Disability & Society, 22*(5), 457–471.

Holmqvist, E., & Buchholz, M. (2012). A model for gaze control assessments and evaluation. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 36–47). Hershey, PA: IGI Global.

Holmqvist, E., Thunberg, G., & Dahlstrand, M. P. (2017). Gaze-controlled communication technology for children with severe multiple disabilities: Parents and professionals' perception of gains, obstacles, and prerequisites. *Assistive Technology, 30*(4), 201–208. https://doi.org/10.1080/10400435.2017.1307882.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. Oxford: Oxford University Press.

Hornof, A. J. (2009). Designing with children with severe motor impairments. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '09)* (pp. 2177–2180). New York, NY: ACM.

Hornof, A., Cavender, A., & Hoselton, R. (2003). EyeDraw: A system for drawing pictures with eye movements. In *Proceedings of the 6th international ACM SIGACCESS conference on computers and accessibility (Assets '04)* (pp. 86–93). New York, NY: ACM.

Huang, J., White, R., & Buscher, G. (2012). User see, user point: Gaze and cursor alignment in web search. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '12)* (pp. 1341–1350). New York, NY: ACM.

Huckauf, A., & Urbina, M. H. (2008). On object selection in gaze controlled environments. *Journal of Eye Movement Research, 2*(4), 4,1–7.

Huckauf, A., & Urbina, M. H. (2011). Object selection in gaze controlled systems: What you don't look at is what you get. *ACM Transactions on Applied Perception, 8*(2), Article 13, 14 pages.

Hutchinson, T., White, K., Martin, W., Reichert, K., & Frey, L. (1989). Human–computer interaction using eye-gaze input. *IEEE Transactions on Systems, Man, and Cybernetics, 19,* 1527–1534.

Hyrskykari, A. (2006). *Eyes in attentive interfaces: Experiences from creating iDict, a gaze-aware reading aid.* Dissertation, University of Tampere. Retrieved from http://urn.fi/urn:isbn:951-44-6643-8.

Hyrskykari, A., Majaranta, P., Aaltonen, A., & Räihä, K.-J. (2000). Design issues of iDICT: A gaze-assisted translation aid. In *Proceedings of the symposium on eye tracking research and applications (ETRA '00)* (pp. 9–14). New York, NY: ACM.

Hyrskykari, A., Majaranta, P., & Räihä, K.-J. (2003). Proactive response to eye movements. *Human-Computer Interaction—INTERACT 2003* (pp. 129–136). Amsterdam: IOS press.

Hyrskykari, A., Majaranta, P., & Räihä, K.-J. (2005). From gaze control to attentive interfaces. In *Proceedings of human-computer interaction international (HCII 2005)* (Vol. 2*)*. Las Vegas, NY.

Ishimaru, S., Kunze, K., Tanaka, K., Uema, Y., Kise, K., & Inami, M. (2015). Smart eyewear for interaction and activity recognition. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '15)* (pp. 307–310). New York, NY: ACM.

Isokoski, P. (2000). Text input methods for eye trackers using off-screen targets. In *Proceedings of the symposium on eye tracking research and applications (ETRA '00)* (pp. 15–21). New York, NY: ACM.

Isokoski, P., Joos, M., Špakov, O., & Martin, B. (2009). Gaze controlled games. *Universal Access in the Information Society, 8*(4), 323–337.

Istance, H. (2016). *An investigation into gaze-based interaction techniques for people with motor impairments.* Dissertation, Loughborough University. Retrieved from https://dspace.lboro.ac.uk/2134/23911.

Istance, H., Bates, R., Hyrskykari, A., & Vickers, S. (2008). Snap clutch, a moded approach to solving the Midas touch problem. In *Proceedings of the symposium on eye tracking research and applications (ETRA '08)* (pp. 221–228). New York, NY: ACM.

Istance, H., & Hyrskykari, A. (2012). Gaze-aware systems and attentive applications. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *Gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 175–195). Hershey, PA: IGI Global.

Istance, H., Hyrskykari, A., Immonen, L., Mansikkamaa, S., & Vickers, S. (2010). Designing gaze gestures for gaming: An investigation of performance. In *Proceedings of the symposium on eye tracking research and applications (ETRA '10)* (pp. 323–330). New York, NY: ACM.

Istance, H., Vickers, S., & Hyrskykari, A. (2009). Gaze-based interaction with massively multiplayer on-line games. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '09)* (pp. 4381–4386). New York, NY: ACM.

Istance, H., Vickers, S., & Hyrskykari, A. (2012). The validity of using non-representative users in gaze communication research. In *Proceedings of the symposium on eye tracking research and applications (ETRA '12)* (pp. 233–236). New York, NY: ACM.

Jacob, R. J. (1990). What you look at is what you get: Eye movement-based interaction techniques. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '90)* (pp. 11–18). New York, NY: ACM.

Jacob, R. J. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get. *ACM Transactions on Information Systems, 9*(2), 152–169.

Jacob, R. J. (1995). Eye tracking in advanced interface design. In W. Barfield & T. A. Furness (Eds.), *Virtual environments and advanced interface design* (pp. 258–288). New York, NY: Oxford University Press Inc.

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In J. Hyönä, R. Radah, & H. Deubel (Eds.), *The mind's eye: Cognitive and applied aspects of eye movement research* (pp. 573–605). Amsterdam: Elsevier.

Jacob, R., & Stellmach, S. (2016). What you look at is what you get: Gaze-based user interfaces. *Interactions, 23*(5), 62–65.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kandemir, M., & Kaski, S. (2012). Learning relevance from natural eye movements in pervasive interfaces. In *Proceedings of the 14th ACM international conference on multimodal interaction (ICMI '12)* (pp. 85–92). New York, NY: ACM.

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2014a). Delayed haptic feedback to gaze gestures. In *Proceedings of EuroHaptics 2014* (pp. 25–31). Berlin Heidelberg: Springer.

Kangas, J., Akkil, D., Rantala, J., Isokoski, P., Majaranta, P., & Raisamo, R. (2014b). Gaze gestures and haptic feedback in mobile devices. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '14)* (pp. 435–438). New York, NY: ACM.

Khamis, M., Alt, F., & Bulling, A. (2015). A field study on spontaneous gaze-based interaction with a public display using pursuits. In *Proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2015 ACM international symposium on wearable computers* (pp. 863–872). New York, NY: ACM.

Khan, D., Heynen, J., & Snuggs, G. (1999). Eye-controlled computing: The VisionKey experience. In *Proceedings of the fourteenth international conference on technology and persons with disabilities (CSUN'99).* Los Angeles, CA. Retrieved from http://www.eyecan.ca/.

Kirst, D., & Bulling, A. (2016). On the verge: Voluntary convergences for accurate and precise timing of gaze input. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '16)* (pp. 1519–1525). New York, NY: ACM.

Köpsel, A., Majaranta, P., Isokoski, P., & Huckauf, A. (2016). Effects of auditory, haptic and visual feedback on performing gestures by gaze or by hand. *Behavior & Information Technology, 35*(12), 1044–1062.

Kristensson, P. O., & Vertanen, K. (2012). The potential of dwell-free eye-typing for fast assistive gaze communication. In *Proceedings of the symposium on eye tracking research and applications (ETRA '12)* (pp. 241–244). New York, NY: ACM.

Kudo, S., Okabe, H., Hachisu, T., Sato, M., Fukushima, S., & Kajimoto, H. (2013). Input method using divergence eye movement. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '13)* (pp. 1335–1340). New York, NY: ACM.

Kumar, C., Menges, R., Müller, D., & Staab, S. (2017). Chromium based framework to include gaze interaction in web browser. In *Proceedings of the 26th international conference on world wide web companion (WWW '17 Companion)* (pp. 219–223). International World Wide Web Conferences Steering Committee, Geneva, Switzerland.

Kumar, M., Paepcke, A., & Winograd, T. (2007). EyePoint: Practical pointing and selection using gaze and keyboard. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '08)* (pp. 421–430). New York, NY: ACM.

Kurauchi, A., Feng, W., Joshi, A., Morimoto, C., & Betke, M. (2016). EyeSwipe: Dwell-free text entry using gaze paths. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '16)* (pp. 1952–1956). New York, NY: ACM.

Land, M. F., & Furneaux, S. (1997). The knowledge base of the oculomotor system. *Philosophical Transactions of the Royal Society of London B: Biological Sciences, 352*(1358), 1231–1239.

Land, M., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25), 3559–3565.

Land, M., & Tatler, B. (2009). *Looking and acting: vision and eye movements in natural behaviour*. Oxford University Press.

Laretzaki, G., Plainis, S., Vrettos, I., Chrisoulakis, A., Pallikaris, I., & Bitsios, P. (2011). Threat and trait anxiety affect stability of gaze fixation. *Biological Psychology, 86*(3), 330–336.

Levine, J. (1981). *An eye-controlled computer*. Yorktown Heights, NY: IBM Thomas J. Watson Research Center.

Liu, Y., Lee, B.-S., & McKeown, M. J. (2016). Robust eye-based dwell-free typing. *International Journal of Human-Computer Interaction, 32*(9), 682–694.

Lorenceau, J. (2012). Cursive writing with smooth pursuit eye movements. *Current Biology, 22*(16), 1506–1509.

MacKenzie, I. S., & Ashtiani, B. (2011). BlinkWrite: Efficient text entry using eye blinks. *Universal Access in the Information Society, 10*(1), 69–80.

Majaranta, P. (2009). *Text entry by eye gaze.* Dissertation, University of Tampere. Retrieved from http://urn.fi/urn:isbn:978-951-44-7787-4.

Majaranta, P., Ahola, U.-K., & Špakov, O. (2009a). Fast gaze typing with an adjustable dwell time. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '09)* (pp. 357–360). New York, NY: ACM.

Majaranta, P., Bates, R., & Donegan, M. (2009b). Eye tracking. In C. Stephanidis (Ed.), *The universal access handbook* (pp. 36:1–20). Boca Raton, FL: CRC Press.

Majaranta, P., Aoki, H., Donegan, M., Hansen, D. W., Hansen, J. P., Hyrskykari, A., & Räihä, K.-J. (Eds.). (2012). *Gaze interaction and applications of eye tracking: Advances in assistive technologies*. Hershey, PA: IGI Global.

Majaranta, P., & Bulling, A. (2014). Eye tracking and eye-based human-computer interaction. In S. Fairclough, & K. Gilleade (Eds.), *Advances in physiological computing* (pp. 39–65). Springer.

Majaranta, P., Isokoski, P., Rantala, J., Špakov, O., Akkil, D., Kangas, J., & Raisamo, R. (2016). Haptic feedback in eye typing. *Journal of Eye Movement Research, 9*(1), 1–14.

Majaranta, P., MacKenzie, I. S., Aula, A., & Räihä, K. J. (2006). Effects of feedback and dwell time on eye typing speed and accuracy. *Universal Access in the Information Society, 5*(2), 199–208.

Majaranta, P., & Räihä, K.-J. (2002). Twenty years of eye typing: Systems and design issues. In *Proceedings of the symposium on eye tracking research and applications (ETRA '02)* (pp. 15–22). New York, NY: ACM.

Manabe, H., Fukumoto, M., & Yagi, T. (2015). Conductive rubber electrodes for earphone-based eye gesture input interface. *Personal and Ubiquitous Computing, 19*(1), 143–154.

Mardanbegi, D., Jalaliniya, S., Khamis, M., & Majaranta, P. (2016). The 6th international workshop on pervasive eye tracking and mobile eye-based interaction. In *Proceedings of the 2016 ACM international joint conference on pervasive and ubiquitous computing: Adjunct publication (UbiComp '16 Adjunct).* New York, NY: ACM.

Mardanbegi, D., Hansen, D. W., & Pederson, T. (2012). Eye-based head gestures. In *Proceedings of the symposium on eye tracking research and applications (ETRA '12)* (pp. 139–146). New York, NY: ACM.

Mele, M. L., & Federici, S. (2012). A psychotechnological review on eye-tracking systems: Towards user experience. *Disability and Rehabilitation: Assistive Technology, 7*(4), 261–281.

Missimer, E., & Betke, M. (2010). Blink and wink detection for mouse pointer control. In *Proceedings of the 3rd international conference on PErvasive technologies related to assistive environments (PETRA '10)* (Article 23, 8 pages). New York, NY: ACM.

Møllenbach, E., Hansen, J. P., & Lillholm, M. (2013). Eye movements in gaze interaction. *Journal of Eye Movement Research, 6*(2), 1–15.

Møllenbach, E., Hansen, J. P., Lillholm, M., & Gale, A. G. (2009). Single stroke gaze gestures. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '09)* (pp. 4555–4560). New York, NY: ACM.

Morimoto, C. H., & Amir, A. (2010). Context switching for fast key selection in text entry applications. In *Proceedings of the 2010 symposium on eye tracking research & applications (ETRA '10)* (pp. 271–274). New York, NY: ACM.

Morimoto, C. H., & Mimica, M. R. (2005). Eye gaze tracking techniques for interactive applications. *Computer Vision and Image Understanding, 98*(1), 4–24. https://doi.org/10.1016/j.cviu.2004.07.010.

Mott, M. E., Williams, S., Wobbrock, J. O., & Morris, M. R. (2017) Improving dwell-based gaze typing with dynamic, cascading dwell times. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '17)* (pp. 2558–2570). New York, NY: ACM.

Muller, M. J., & Kuhn, S. (1993). Participatory design. *Communications of the ACM, 36*(6), 24–28.

Nielsen, J. (1993). Noncommand user interfaces. *Communications of the ACM, 36*(4), 83–99.

Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. Mack (Eds.), *Usability inspection methods* (pp. 25–62). New York, NY: Wiley.

Norman, D. (1998). *The design of everyday things*. New York, NY: Basic Books.

Pannasch, S. H., Malischke, S., Storch, A., & Velichkovsky, B. M. (2008). Eye typing in application: A comparison of two systems with ALS patients. *Journal of Eye Movement Research, 2*(8), 6, 1–8.

Partala, T., & Surakka, V. (2003). Pupil size variation as an indication of affective processing. *International Journal of Human-Computer Studies, 59*(1), 185–198.

Papoutsaki, A., Sangkloy, P., Laskey, J., Daskalova, N., Huang, J., & Hays, J. (2016). Webgazer: Scalable webcam eye tracking using user interactions. In *Proceedings of the twenty-fifth international joint conference on artificial intelligence (IJCAI'16)* (pp. 3839–3845). AAAI Press.

Pedrosa, D., Pimentel, M. D., Wright, A., & Truong, K. N. (2015). Filteryedping: Design challenges and user performance of dwell-free eye typing. *ACM Transactions on Accessible Computing, 6*(1), 3 (37 pages).

Penkar, A. M., Lutteroth, C., & Weber, G. (2012). Designing for the eye: Design parameters for dwell in gaze interaction. In *Proceedings of the 24th Australian computer-human interaction conference (OzCHI '12)* (pp. 479–488). New York, NY: ACM.

Porta, M., & Turina, M. (2008). Eye-S: A full-screen input modality for pure eye-based communication. In *Proceedings of the symposium on eye tracking research and applications (ETRA '08)* (pp. 27–34). New York, NY: ACM.

Preece, J., Rogers, Y., & Sharp, H. (2002). *Interaction design: Beyond human-computer interaction*. New York, NY: Wiley.

Qvarfordt, P. (2017). Gaze-informed multimodal interaction. In S. Oviatt, B. Schuller, P. R. Cohen, D. Sonntag, G. Potamianos, & A. Krüger (Eds.), *The handbook of multimodal-multisensor interfaces* (pp. 365–402). New York, NY: ACM.

Qvarfordt, P., Beymer, D., & Zhai, S. (2005). RealTourist—a study of augmenting human-human and human-computer dialogue with eye-gaze overlay. *Human-computer interaction—INTERACT 2005* (pp. 767–780). Berlin Heidelberg: Springer.

Reingold, E. M., Loschky, L. C., McConkie, G. W., & Stampe, D. M. (2003). Gaze-contingent multiresolutional displays: An integrative review. *Human Factors, 45*(2), 307–328.

Räihä, K.-J. (2015). Life in the fast lane: Effect of language and calibration accuracy on the speed of text entry by gaze. *Human-computer interaction—INTERACT 2015* (pp. 402–417). Berlin Heidelberg: Springer.

Räihä, K.-J., Hyrskykari, A., & Majaranta, P. (2011). Tracking of visual attention and adaptive applications. In C. Roda (Ed.), *Human attention in digital environments* (pp. 166–185). Cambridge: Cambridge University Press.

Räihä, K.-J., & Ovaska, S. (2012). An exploratory study of eye typing fundamentals: Dwell time, text entry rate, errors, and workload. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '12)* (pp. 3001–3010). New York, NY: ACM.

Raupp, S. (2013). *Keyboard layout in eye gaze communication access: Typical vs. ALS.* Communication sciences and disorders. East Carolina University. Retrieved from http://hdl.handle.net/10342/4254.

Rinard, G., & Rugg, D. (1976). An ocular control device for use by the severely handicapped. In *1986 Conference on systems and devices for the disabled* (pp. 76–79). Boston, MA.

Rosen, M., & Durfee, W. (1978). Preliminary report on EYECOM, an eye-movement system for non-vocal communication. In *Proceedings of the 5th conference on systems and devices for the disabled.* Houston, TX.

Rozado, D., San Agustin, J., Rodriguez, F. B., & Varona, B. (2012). Gliding and saccadic gaze gesture recognition in real time. *ACM Transactions on Interactive Intelligent Systems, 1*(2), 10:1–27.

Salvucci, D. D. (1999). Inferring intent in eye-based interfaces: Tracing eye movements with process models. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '99)* (pp. 254–261). New York, NY: ACM.

Samsung. (2013). *User manual GT-19295.* Samsung Electronics (www.samsung.com).

Schenk, S., Tiefenbacher, P., Rigoll, G., & Dorr, M. (2016). SPOCK: A smooth pursuit oculomotor control kit. In *Extended abstracts of the conference on human factors in computing systems (CHI EA '16)* (pp. 2681–2687). New York, NY: ACM.

Shneiderman, B. (1998). *Designing the user interface*. Reading, MA: Addison-Wesley.

Skovsgaard, H., Räihä, K.-J., & Tall, M. (2012). Computer control by gaze. In P. Majaranta, H. Aoki, M. Donegan, D. W. Hansen, J. P. Hansen, A. Hyrskykari, & K.-J. Räihä (Eds.), *gaze interaction and applications of eye tracking: Advances in assistive technologies* (pp. 78–102). Hershey, PA: IGI Global.

Smith, B. A., Ho, J., Ark, W., & Zhai, S. (2000). Hand eye coordination patterns in target selection. In *Proceedings of the symposium on eye tracking research and applications (ETRA '00)* (pp. 117–122). New York, NY: ACM.

Soegaard, M., & Dam, R. F. (2012). *The encyclopedia of human-computer interaction.* The Interaction Design Foundation. Retrieved from https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed.

Špakov, O. (2008). *iComponent—Device-independent platform for analyzing eye movement data and developing eye-based applications.* Dissertation, University of Tampere. Retrieved from http://urn.fi/urn:isbn:978-951-44-7321-0.

Špakov, O. (2011). Comparison of gaze-to-objects mapping algorithms. In *Proceeding of the 1st conference on novel gaze-controlled applications (NGCA'11)* (Article 6, 8 pages). New York, NY: ACM.

Špakov, O. (2012). Comparison of eye movement filters used in HCI. In *Proceedings of the symposium on eye tracking research and applications (ETRA '12)* (pp. 281–284). New York, NY: ACM.

Špakov, O., Isokoski, P., Kangas, J., Akkil, D., & Majaranta, P. (2016). PursuitAdjuster: An exploration into the design space of smooth pursuit-based widgets. In *Proceedings of the symposium on eye tracking research and applications (ETRA '16)* (pp. 287–290). New York, NY: ACM.

Špakov, O., Isokoski, P., & Majaranta, P. (2014). Look and lean: Accurate head-assisted eye pointing. In *Proceedings of the symposium on eye tracking research and applications (ETRA '14)* (pp. 35–42). New York, NY: ACM.

Špakov, O., & Majaranta, P. (2009). Scrollable keyboards for casual eye typing. *PsychNology Journal, 7*(2), 159–173.

Špakov, O., & Majaranta, P. (2012). Enhanced gaze interaction using simple head gestures. In *Proceedings of the 2012 ACM conference on ubiquitous computing (UbiComp '12)* (pp. 705–710). New York, NY: ACM.

Špakov, O., & Miniotas, D. (2004). On-line adjustment of dwell time for target selection by gaze. In *Proceedings of nordic conference on human-computer interaction (NordiCHI '04)* (pp. 203–206). New York, NY: ACM.

Špakov, O., & Miniotas, D. (2005). EyeChess: A tutorial for endgames with gaze-controlled pieces. In *Proceedings of COGAIN 2005* (pp. 16–18). Communication by Gaze Interaction (COGAIN).

Starker, I., & Bolt, R. A. (1990). A gaze-responsive self-disclosing display. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '90)* (pp. 3–10). New York, NY: ACM.

Stellmach, S., & Dachselt, R. (2012). Look & touch: Gaze-supported target acquisition. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '12)* (pp. 2981–2990). New York, NY: ACM.

Sundstedt, V. (2011). *Gazing at games: An introduction to eye tracking control.* Morgan & Claypool Publishers.

Tall, M. (2008). *NeoVisus: Gaze interaction interface components*. M.Sc. thesis, Lund University. Retrieved from https://www.researchgate.net/publication/235007493_Neovisus_Gaze_Interaction_Interface_Components.

Tchalenko, J. (2001). Free-eye drawing. *Point: Art and Design Research Journal, 11,* 36–41.

Ten Kate, J., Frietman, E., Stoel, F., & Willems, W. (1979). Eye-controlled communication aids. *Medical Progress Through Technology, 8*(1), 1–21.

Tobii. (2016, June 28). *Apps*. Retrieved from Tobii tech/experience: http://www.tobii.com/xperience/apps/.

Tuisku, O., Surakka, V., Rantanen, V., Vanhala, T., & Lekkala, J. (2013). Text entry by gazing and smiling. In *Advances in human-computer interaction*, Article ID 218084, 13 pages.

Velichkovsky, B., Sprenger, A., & Unema, P. (1997). Towards gaze-mediated interaction: Collecting solutions of the "Midas touch" problem. *Human-computer interaction—INTERACT'97* (pp. 509–516). New York, NY: Springer.

Velloso, E., Wirth, M., Weichel, C., Esteves, A., & Gellersen, H. (2016). AmbiGaze: Direct control of ambient devices by gaze. In *Proceedings of the 2016 ACM conference on designing interactive systems (DIS '16)* (pp. 812–817). New York, NY: ACM.

Vertegaal, R., & Shell, J. S. (2008). Attentive user interfaces: The surveillance and sousveillance of gaze-aware objects. *Social Science Information, 47*(3), 275–298.

Vickers, S., Istance, H., & Hyrskykari, A. (2009). Selecting commands in 3D game environments by gaze. In *Proceedings of 5th conference on communication by gaze interaction (COGAIN)* (pp. 55–59). The Technical University of Denmark, Lyngby, Denmark.

Vickers, S., Istance, H., & Hyrskykari, A. (2013). Performing locomotion tasks in immersive computer games with an adapted eye-tracking interface. *ACM Transactions on Accessible Computing (TACCESS), 5*(1), 2:1–33.

Vidal, M., Bulling, A., & Gellersen, H. (2013). Pursuits: Spontaneous interaction with displays based on smooth pursuit eye movement and moving targets. In *Proceedings of the 2013 ACM international joint conference on pervasive and ubiquitous computing* (pp. 439–448). New York, NY: ACM.

Voßkühler, A., Nordmeier, V., Kuchinke, L., & Jacobs, A. (2008). OGAMA—open gaze and mouse analyzer: Open source software designed to analyze eye and mouse movements in slideshow study designs. *Behavior Research Methods, 40*(4), 1150–1162.

Ward, D. J., & MacKay, D. J. (2002). Fast hands-free writing by gaze direction. *Nature, 414,* 838.

Ware, C., & Mikaelian, H. H. (1987). An evaluation of an eye tracker as a device for computer input2. In *Proceedings of the SIGCHI/GI conference on human factors in computing systems and graphics interface (CHI '87)* (pp. 183–188). New York, NY: ACM.

Wassermann, B., Hardt, A., & Zimmermann, G. (2012). Generic gaze interaction events for web browsers. WWW12 workshop: Emerging Web Technologies, Facing the Future of Education. In *Proceedings of the 21st international conference on World Wide Web* (Vol. 9). https://doi.org/10.13140/2.1.1871.9362.

Weill-Tessier, P., Turner, J., & Gellersen, H. (2016). How do you look at what you touch?: A study of touch interaction and gaze correlation on tablets. In *Proceedings of the symposium on eye tracking research and applications (ETRA '16)* (pp. 329–330). New York, NY: ACM.

Wobbrock, J. O., Rubinstein, J., Sawyer, M. W., & Duchowski, A. T. (2008). Longitudinal evaluation of discrete consecutive gaze gestures for text entry. In *Proceedings of the symposium on eye tracking research and applications (ETRA '08)* (pp. 11–18). New York, NY: ACM.

Yamada, M., & Fukuda, T. (1987). Eye word processor (EWP) and peripheral controller for the ALS patient. *IEEE Proceedings A—Physical Science, Measurement and Instrumentation, Management and Education—Reviews, 134*(4), 328–330.

Yarbus, A. L. (1967). *Eye movements and vision*. New York, NY: Plenum Press.

Yeoh, K. N., Lutteroth, C., & Weber, G. (2015). Eyes and keys: An evaluation of click alternatives combining gaze and keyboard. *Human-computer interaction—INTERACT 2015* (pp. 367–383). Berlin Heidelberg: Springer.

Zhai, S., Morimoto, C., & Ihde, S. (1999). Manual and gaze input cascaded (MAGIC) pointing. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '99)* (pp. 246–253). New York, NY: ACM.

Zhang, Y., Bulling, A., & Gellersen, H. (2013). Sideways: A gaze interface for spontaneous interaction with situated displays. In *Proceedings of the SIGCHI conference on human factors in computing systems (CHI '13)* (pp. 851–860). New York, NY: ACM.

Zhang, Y., Chong, M. K., Müller, J., Bulling, A., & Gellersen, H. (2015). Eye tracking for public displays in the wild. *Personal and Ubiquitous Computing, 19*(5), 967–981.

Zhang, Y., Müller, J., Chong, M. K., Bulling, A., & Gellersen, H. (2014). GazeHorizon: Enabling passers-by to interact with public displays by gaze. In *Proceedings of the 2014 ACM international joint conference on pervasive and ubiquitous computing (UbiComp '14)* (pp. 559–563). New York, NY: ACM.