Nebraska Symposium on Motivation

Michael D. Dodd
John H. Flowers *Editors*

The Influence of Attention, Learning, and Motivation on Visual Search

Springer

Nebraska Symposium on Motivation

Series Editor
Debra A. Hope
Lincoln, Nebraska, USA

For further volumes:
http://www.springer.com/series/7596

Michael D. Dodd • John H. Flowers
Editors

The Influence of Attention, Learning, and Motivation on Visual Search

Springer

Editors
Michael D. Dodd
University of Nebraska
Lincoln, Nebraska
USA

John H. Flowers
University of Nebraska
Lincoln, Nebraska
USA

ISBN 978-1-4614-4793-1 ISBN 978-1-4614-4794-8 (eBook)
DOI 10.1007/978-1-4614-4794-8
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2012950600

© Springer Science+Business Media New York 2012
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. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.
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.
While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

The volume editors for this 59th volume of the Nebraska Symposium on Motivation are Michael Dodd and John Flowers. The volume editors coordinated the symposium that led to this volume including selecting and inviting the contributors. My thanks go to Professors Dodd and Flowers and to our contributors for their outstanding presentations and chapters. As we learned during the Symposium, visual search is part of our daily experience. A better understanding of the underlying processes and limitations helps us solve everything from minor everyday problems we all face to security problems that could pose a threat to our way of life.

At the Symposium we honored Professor Flowers as he retired in 2011 following a long and distinguished career at the University of Nebraska-Lincoln. This volume helps round out a fruitful scientific career as well as marks the passing of the torch to the next generation of cognitive psychologists in our department, including Professor Dodd.

This Symposium series is supported by funds provided by the Chancellor of the University of Nebraska-Lincoln, Harvey Perlman, and by funds given in memory of Professor Harry K. Wolfe to the University of Nebraska Foundation by the late Professor Cora L. Friedline. We are extremely grateful for the Chancellor's generous support of the Symposium series and for the University of Nebraska Foundation's support via the Friedline bequest. This symposium volume, like those in the recent past, is dedicated to the memory of Professor Wolfe, who brought psychology to the University of Nebraska. After studying with Professor Wilhelm Wundt in Germany, Professor Wolfe returned to Nebraska, his native state, to establish the first undergraduate laboratory in psychology in the nation. As a student at Nebraska, Professor Friedline studied psychology under Professor Wolfe.

Lincoln, Nebraska
USA

Debra A. Hope
Series Editor

Acknowledgements

As is the case with any large scale event, the 59th annual Nebraska Symposium on Motivation and accompanying volume would not have been possible were it not for the tireless efforts of numerous individuals. We are truly grateful for everyone's investment and would like to acknowledge a number of people for the critical role they played. First, the Department of Psychology staff, particularly Claudia Price-Decker, Roxanne Earnest, Jamie Longwell, and Jodi Wiser. In addition to their normal role in ensuring that the day-to-day operations of the Psychology Department run smoothly, each of these individuals took time out of their busy schedules to assist in a number of tasks that were critical to the success of the event. These included, but are not limited to, arranging speaker travel, hotel and conference room accommodations, conference registration and administration, program and gift-bag preparation, and the scheduling of meals and refreshments during breaks in the day. Mark Mills, a cognitive area graduate student, volunteered his time to schedule and assist in ground travel for all speakers both to-and-from the airport and to-and-from the conference. Peter Levitov was also very helpful in terms of informing the administrative work and protocol we needed to follow to bring in our international speakers. Chancellor Harvey Perlman is to be thanked for the introductory remarks he offered at the beginning of the Symposium which acknowledged not just the great history of the event itself, but also celebrated the career of Symposium co-organizer, John Flowers who retired in 2011 after 39 years as a major contributor to our faculty.

It goes without saying that we are incredibly indebted to our speakers, all of whom were eager and willing to take part in the Symposium and clearly put a great deal of time into their talks—all of which were dynamic, informative, and interesting—and their subsequent chapters appearing in this volume. These chapters fully capture the spirit of their Symposium presentation while also allowing them to go into more depth on specific issues which are currently at the forefront of the study of visual search. Moreover, many of the speakers brought along their graduate students to present posters at the end of the first day of the conference. The poster session also featured UNL students and a number of other graduate student participants from out-of-state, making it one of the larger poster sessions in recent memory with representatives from a number of different regions of the country doing research related to visual search at a variety of levels. These individuals are to be thanked not only for the

time and effort they put into their presentations, but also for taking the Symposium speakers out to lunch on the second day of the conference, allowing for additional intellectual exchange.

The 59th Nebraska Symposium on Motivation was also the first to be webcast live meaning that anyone in the world could watch the presentations in real time as they were occurring in addition to being able to send questions in via e-mail for the speakers to address. On the first day of the conference alone, more than 300 individuals logged in from around the world to view the talks and we received very positive feedback from people who were unable to travel to the Symposium but thrilled to have the opportunity to view the talks nonetheless. We anticipate that this will now be the standard for all future symposia. This also allowed us to make the talks available to the speakers when it came time for them to write their chapters which is a huge benefit given how quickly the event unfolds. We are very appreciative of Mark Hendricks and his team in UNL Educational Media for their outstanding efforts with the web streaming.

Finally, we would like to highlight and give thanks to the contributions of Deb Hope, Series Editor for the Nebraska Symposium on Motivation. Were it not for Deb, we would not have known where to begin in organizing an event and volume of this magnitude. Her willingness to share her expertise and her experiences with past events made it very easy for us to address every issue that arose during the course of planning and execution of the event and the volume that now follows.

Contributors

Stephen H. Adamo completed his B.S. in Psychology in 2010 at the University of Arizona. Under the guidance of Dr. Mary Peterson, he investigated visual awareness in figure-ground perception through the use of binocular rivalry. After graduation, he joined the laboratory of Dr. Stephen R. Mitroff at Duke University, where he has since focused on exploring the spatial and temporal aspects of visual attention. Stephen uses eye tracking and a battery of individual-differences measures to study the mechanisms of visual search and the attentional blink.

Brian A. Anderson is a Ph.D. student in the Department of Psychological and Brain Sciences at Johns Hopkins University, where he studies the interplay between attention and reward learning under the direction of Steven Yantis. He received an M.S. in Psychology from Villanova University in 2009, where he studied top-down control of attention with Charles Folk. Brian is originally from Maine and received his undergraduate degree from the University of Maine at Augusta.

Matthew S. Cain received his S.B. in Brain and Cognitive Science from the Massachusetts Institute of Technology in 2002 and his Ph.D. in Psychology from the University of California, Berkeley in 2009. His graduate work in the laboratory of Dr. Arthur P. Shimamura investigated cognitive control, attention, and task switching and the effects of factors such as sleep and video-game experience on these processes. As a postdoctoral researcher in Dr. Stephen R. Mitroff's laboratory, Matthew has investigated how visual search and visual attention are influenced by local factors such as anxiety, environmental statistics, and stroboscopic training as well as by more global factors such as media exposure and other individual differences.

Kait Clark completed her B.S. in Psychology in 2008 at Saint Joseph's University, where she worked with Drs. Patrick Garrigan and John Jewell investigating visual memory and change detection. She then pursued graduate studies at Duke University under the mentorships of Drs. Stephen R. Mitroff and Marty G. Woldorff. Over the course of her graduate career, Kait has focused on the influences on visual search performance from a variety of directions. She has explored how certain forms of training and experience (lab-based practice, video-game experience, and radiological expertise) may augment search performance and how individual differences between subjects and motivational factors impact search speed and accuracy.

Michael Dodd received his Ph.D. in Cognitive Psychology from the University of Toronto in 2005 and was a Killam postdoctoral fellow at the University of British Columbia before joining the University of Nebraska faculty in 2007. His research encompasses many different aspects of human cognition, with a particular focus on visual attention (e.g., visual search; inhibition of return; object-based attention; apparent motion; sensory processing; scene perception; oculomotor programming; task-induced changes in eye movements), memory (false memory, retrieval-induced forgetting, directed forgetting), individual differences (influences of political temperament on cognition) and goal-directed activity, as well as the interactions between these cognitive.

Dr. John H. Flowers jointed the UNL faculty in 1972. He received his Ph.D. from Yale University in 1972 in experimental psychology. His primary research interests are in the general area of human information processing, particularly attention, implicit learning, and the perception of structure. His interest in the perception of structure has recently led to a research program on the use of sound as a means for representing data.

Andrew Hollingworth received his Ph.D. from Michigan State University in 2000 and is currently a professor at the University of Iowa. He studies how humans select, store, and use visual information to support efficient behavior. His research program is composed of projects involving multiple visual/cognitive systems, with a focus on how these systems are coordinated during visual tasks, including: eye movements and saccade target selection, attention capture, visual search, transaccadic memory and integration, visual working memory, object correspondence and updating, long-term memory for object and scenes, and spatial cognition.

Yoko Ishigami is a postdoctoral fellow at Brain Repair Centre (Dalhousie University). Yoko completed her Ph.D. and M.Sc. at Dalhousie University and her undergraduate studies at University of Victoria. Her research interests include: visual attention, its development/aging, its disorders, and its individual differences. She is specifically interested in exploring their roles, interactions, and limitations in different contexts.

Raymond M. Klein is a University Research Professor and chair of the Department of Psychology & Neuroscience at Dalhousie University. He is an experimental psychologist and cognitive neuroscientist whose research strategies are most strongly influenced by those of D. O. Hebb, Donald Broadbent and Michael Posner. His research interests include: Visual attention and its disorders, reading and dyslexia; applied cognitive psychology.

Patryk Laurent is a Postdoctoral Researcher in the Yantis Laboratory at Johns Hopkins University. He earned his Ph.D. in Neuroscience at the University of Pittsburgh in 2009, while using computational modeling and fMRI to study how reinforcement leads to the coordination of perception, cognition, and action. His current research focuses on understanding the neural circuits involved in reward learning.

Stephen R. Mitroff is an associate professor in the Department of Psychology & Neuroscience and the Center for Cognitive Neuroscience at Duke University. He is originally from the San Francisco Bay Area and received his undergraduate degree in Cognitive Science from the University of California, Berkeley. He received a Masters degree and his Ph.D. in Cognitive Psychology from Harvard University. He completed three years of postdoctoral research at Yale University prior to joining the faculty at Duke University in 2005. Dr. Mitroff's research interests focus on the study of visual cognition, examining the nature of visual memory, perception, and attention. He is currently working with the Department of Homeland Security and the Army Research Office to explore how experiences influence vision and attention. A primary focus is on individual differences to reveal how it is that some individuals can outperform others, how can we quickly and easily identify the best performers, and how can we then train those individuals to make them even better.

Jan Theeuwes is professor and head of the Department of Cognitive Psychology at the Vrije Universiteit Amsterdam. He is a member of the Royal Dutch Academy of Science (KNAW) and received the Bertelson Award in recognition for outstanding psychological research from the European Society for Cognitive Psychology. He has published more than 185 peer-reviewed journal articles on attention, working memory, face processing, cross-modal integration and eye movements. He is widely known for his work on attentional and oculomotor capture. In addition to his basic research, he is an expert in traffic behavior and one of the principal advisors to the Dutch Department of Transportation.

Nicholas Turk-Browne is Assistant Professor and Lawrence S. Brodie Preceptor of Psychology at Princeton University. His NIH-funded research uses brain imaging and behavioral studies to explore interactions between multiple aspects of cognition in an effort to develop an integrated understanding of the mind and brain. In particular, he examines the interaction between perception (how we represent the current environment) and memory (how we store this experience for future use). As an example, he has published extensively on visual statistical learning, our ability to unconsciously extract regularities from experience. These and other studies have been published in journals such as Psychological Science, Neuron, and Proceedings of the National Academy of Sciences. Turk-Browne is the recipient of the James B. Grossman Dissertation Prize from Yale University and the Early Researcher Award from the American Psychological Association.

Emma Wampler is a graduate student in the Department of Psychology at the University of Washington. She received her B.A. in Neuroscience at the Kenyon College in 2009. Her research interests focus on language learning and learning to read.

Jeremy Wolfe graduated summa cum laude from Princeton in 1977 with a degree in Psychology and went on to obtain his Ph.D. in 1981 from MIT, studying with Richard Held. His PhD thesis was entitled "On Binocular Single Vision." Wolfe remained at MIT until 1991. During that period, he published papers on binocular rivalry, visual aftereffects, and accommodation. In the late 1980s, the focus of the lab shifted to visual attention. Since that time, his research has focused on visual search and visual attention with a particular interest in socially important search tasks in areas such as medical image perception (e.g. cancer screening) and security (e.g. baggage screening). In 1991, Wolfe moved to Brigham and Women's Hospital where he is Director of the Visual Attention Lab and the Center for Advanced Medical Imaging. At Harvard Medical School, he is Professor of Ophthalmology and Professor of Radiology. His work is currently funded by the US National Institutes of Health, the Office of Naval Research, Toshiba Corboration, and Google. He has published 127 peer-reviewed papers, 1 textbook, and 31 book chapters. Dr. Wolfe has taught Psychology courses at MIT & Harvard and is Past-President of the Eastern Psychological Association, President of Division 3 of the American Psychological Association, and editor of the journal Attention, Perception and Psychophysics. He won the Baker Memorial Prize for teaching at MIT in 1989. He is a fellow of the AAAS, the American Psychological Association (Div. 3 & 6), the American Psychological Society, and a member of the Society for Experimental Psychologists. He lives in Newton, Mass.

Steven Yantis is Professor and Chair of the Department of Psychological and Brain Sciences at Johns Hopkins University. He received a Ph.D. in Experimental Psychology at the University of Michigan in 1985, and following a year as postdoctoral fellow at Stanford University, he joined the JHU faculty in 1986. His research interests include selective attention and cognitive control.

Contents

Introduction

Michael D. Dodd and John H. Flowers

Abstract The 59th Annual Nebraska Symposium on Motivation (The Influence of Attention, Learning, and Motivation on Visual Search) took place April 7–8, 2011, on the University of Nebraska–Lincoln campus. The symposium brought together leading scholars who conduct research related to visual search at a variety levels for a series of talks, poster presentations, panel discussions, and numerous additional opportunities for intellectual exchange. The Symposium was also streamed online for the first time in the history of the event, allowing individuals from around the world to view the presentations and submit questions. The present volume is intended to both commemorate the event itself and to allow our speakers additional opportunity to address issues and current research that have since arisen. Each of the speakers (and, in some cases, their graduate students and post docs) has provided a chapter which both summarizes and expands on their original presentations. In this chapter, we sought to a) provide additional context as to how the Symposium came to be, b) discuss why we thought that this was an ideal time to organize a visual search symposium, and c) to briefly address recent trends and potential future directions in the field. We hope you find the volume both enjoyable and informative, and we thank the authors who have contributed a series of engaging chapters.

Given that the Nebraska Symposium on Motivation is considered one of the most influential and well revered symposium series in psychology and beyond, it will not come as a surprise to the reader to discover that the topics covered often reflect not just significant areas of research within the sciences, but also issues that are widely relevant and of interest to the public at large. A cursory examination of the titles of the symposia over the past decade reveals a number of gripping and instantly relatable topics: false memory, ethnicity and youth health disparities, emotion and the law,

---

M. D. Dodd (✉)
University of Nebraska, 222 Burnett Hall,
68588 Lincoln, NE, USA
e-mail: mdodd2@unlnotes.unl.edu

J. H. Flowers
University of Nebraska, Burnett Hall 219,
68588–0308 Lincoln, NE, USA
e-mail: jflowers@unl.edu

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_1, © Springer Science+Business Media New York 2012

nicotine use, lesbian/gay/bisexual identity, prejudice and racism. As such, the goal of any symposium organizer and volume editor is to propose a topic that is substantial, interesting and easily relatable. To this end, there was never any question in our mind as to what the focus should be for the 59th annual Nebraska Symposium. Visual search is one of the most well researched topics of the past few decades and is one of the most, if not the most, ubiquitous behaviors that we engage in on a daily basis. Despite the importance of search to everyday functioning, however, it is noteworthy that the term "visual search" has not permeated the vocabulary of the public at large in any substantive way. Whereas numerous researchers across disciplines can immediately identify what visual search is, how important it is to our understanding of attention, and how it can be used to examine and inform numerous aspects of behavior, the term "visual search" does not have that same cache outside of the sciences. Despite this, everyone can instantly relate to the concept of visual search if you provide examples such as searching for your keys which you have misplaced; searching for a ripe piece of fruit in a supermarket; attempting to locate your car when you have forgotten where you have parked, and the nostalgia that accompanies the memory of searching for Waldo as a child (one of our volume contributors and Symposium speakers, Raymond Klein, was the first to use Where's Waldo displays when studying visual search in the laboratory). This creates an interesting conundrum: how can you convince an audience of the tremendous importance of visual search when very few people will actually be familiar with that term or realize that they engage in this behavior repeatedly throughout each day for the entire duration of their life?

At least part of the reason that the term "visual search" may not have caught on beyond the laboratory could be attributed to the fact that the earliest experimental paradigms for studying search could seem quite artificial. Participants would search for a sideways T amidst upright Ts and sideways Ls; or participants could be asked to detect whether a target letter appears in a stream of letters presented one at a time in a very rapid manner; alternatively it could be your goal to detect a red square in a display consisting of blue squares and red circles. These are not tasks that we ever find ourselves engaged with in the real world yet these early tasks provided important demonstrations of how attention and search success are influenced by factors such as target salience, the number and type of distractors in a display, and the relationship between targets and distractors. These concepts easily generalize to the real world even if they do not appear to at first blush: finding the ripest banana is quite easy if the majority of other bananas you are searching through are green but this task becomes considerably more difficult if most of the bananas you are looking through are yellow; finding a car in a parking lot is daunting if you have no idea where you parked, but will be considerably easier if you can remember the general area in which you parked and narrow your efforts to that specific region. It will be easier still if your car is bright red and the majority of other cars in the parking lots are black or blue. The reason Where's Waldo tasks are both challenging and rewarding is that they are designed to be particularly difficult based on what we have previously learned about search from these laboratory paradigms: the greater the number of distractors, the more difficult it will be to find Waldo; as the similarity between the target (Waldo) and the distractors (everything else) increases, so too does task difficulty (which explains why so many other characters and background items in a Where's Waldo book appear in the colors red and white. . . if everything else was purple and green it would be easy to detect our bespectacled friend). As such, the vast majority of individuals are intuitively aware of the factors that can make search easy or difficult, even if they have not heard the term "visual search" before.

Though the earliest visual search paradigms could seem artificial, they were exceptionally well controlled and provided a great deal of invaluable information into the factors that can influence search. Over time, various advancements have afforded researchers the opportunity to study search in a variety of different contexts. These include *technological* advances that allow much greater control of the timing and complexity of visual displays, and the simultaneous collection of multiple streams of behavioral and brain activity data. They also include advancement in our understanding of the relationship between *neurological processes and behavior* that point to potentially important relationships between such factors as motivation, memory, and reinforcement history and the manner in which visual search is carried out. As a consequence of these developments, search displays will now routinely consist of more realistic scenes, both static and dynamic, that are more akin to what we experience in the real world. Eyetracking technology has made it possible to study the manner in which individuals search in more naturalistic settings whereas advancements in neuroimaging have made it possible to determine the biological and cognitive processes engaged during search. With each new advancement and each new study, however, it has become increasingly clear that a staggering number of factors influence attention and search beyond those that were identified in earlier research. Moreover, it has also become clear that the importance of attention and search extends beyond the target-directed examples we have provided above. At a basic level, visual search is the process through which attention is allocated throughout the environment in order to process information that is relevant to an individual's goals or interests while simultaneously ignoring information that is irrelevant at best and potentially distracting at worst. In this way, the vast majority of visual tasks we perform are some form of search task. Watching television requires the viewer to focus attention on task relevant stimuli (the television) while ignoring irrelevant stimuli. The act of driving requires continuous shifts of attention to relevant stimuli (such as traffic lights, pedestrians, and street signs) and an attempt to block out potentially distracting stimuli (e.g. billboards, one's cellphone). Even reading this introduction requires the continuous movement of attention to specific stimuli (the words on the page) so as to extract relevant information while simultaneously ignoring other items in your environment that could attract attention. In this way one could argue that the vast majority of all visual behaviors are a form of search and an understanding of these behaviors—and how to make them efficient—can only be gleaned by considering the many different factors that influence search at multiple levels.

Given that visual search encompasses a variety of behaviors and is influenced by a staggering number of variables, our goal in putting together the symposium was to not merely invite scholars who are best known for doing research on visual search. Rather, we sought to bring together individuals who engage in research at a number of different levels relevant to search. As such, we scholars who study attention and search in both laboratory and applied settings; individuals who do not study search per se but study factors that are highly related to search performance (e.g. motivation/reward, statistical learning); individuals who study cognitive mechanisms which greatly influence the manner in which search is conducted (e.g. eye movements, object representation, memory and knowledge). In this volume alone you will read about a number of factors relating to search which are currently being studied such as a) selection in space and time (Klein), b) automatic processes (Theeuwes), c) the role of scene perception, scene memory, and schematic knowledge (Hollingworth), d) motivation/reward (Yantis), e) statistical learning (Turk-Browne), and f) satisfaction of search (Mitroff), and g) search termination (Wolfe).

In summary, the goal of the symposium was twofold. We sought to take stock of where we currently are in the study of search and how our understanding can be increased across a variety of areas but we also hoped to provide some insight into how our understanding of search will be shaped by future research and what questions seem likely to develop in the future. It seems clear that emerging technology involving visual displays of information for a wide variety of applications will make the study of search processes an important topic for applied research. While the precise role that visual search tasks will play in future research in cognitive psychology and neuroscience is somewhat more difficult to predict, the presentations of this symposium suggest that studying search behavior, in conjunction with applying other tools for assessing cognitive and behavioral activity, will continue to be a useful endeavor.

Searching in Space and in Time

Raymond M. Klein and Yoko Ishigami

Abstract Our conception of attention is intricately linked to limited processing capacity and the consequent requirement to select, in both space and time, what objects and actions will have access to these limited resources. Seminal studies by Treisman (*Cognitive Psychology*, *12*, 97–136, 1980) and Broadbent (*Perception and Psychophysics*, *42*, 105–113, 1987; Raymond et al. *Journal of Experimental Psychology: Human Perception and Performance*, *18*, 849–860, 1992) offered the field tasks for exploring the properties of attention when searching in space and time. After describing the natural history of a search episode we briefly review some of these properties. We end with the question: Is there one attentional "beam" that operates in both space and time to integrate features into objects? We sought an answer by exploring the distribution of errors when the same participant searched for targets presented at the same location with items distributed over time (McLean et al. *Quarterly Journal of Experimental Psychology, 35A,* 171–186, 1982) and presented all at once with items distributed over space (Snyder *Journal of Experimental Psychology, 92,* 428–431, 1972). Preliminary results revealed a null correlation between spatial and temporal slippage suggesting separate selection mechanisms in these two domains.

Keywords Attention · Limited processing capacity · Selection · Space · Time · Attentional blink

The concept of attention is at the nexus of search behavior. Two related aspects of attention are critical: limited processing capacity and the consequent requirement to select in space and time where and when to allocate our limited resources. The concept of selective attention both entails and represents these two related ideas. Attention may be captured by the target or by salient distractors. Lacking capture by the target, attention may be allocated individually to potential targets to allow information processes to determine whether these items are what we are looking for. It is the sine qua non of "limited capacity" that when mental processes are allocated to some things (e.g., items during search) they become less available for other activities.

R. M. Klein (✉) · Y. Ishigami
Dalhousie University, 6299 South Street, B3H3J5, Halifax, NS, Canada
e-mail: Ray.Klein@Dal.Ca

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_2,

The scientific study of search behavior is motivated by both basic and applied research questions. Basic research has been preoccupied by two inter-related efforts: Using studies of search to reveal the properties of attention and to reveal the nature of object recognition. Several "basic" questions are suggested by the combination of these efforts. For example, What is the nature of the limited capacity mechanisms that are captured or deliberately allocated when we search? What is the role of these mechanisms in the construction of object representations? And, What is the role of objects and object properties in the control of attention.

Basic research on search behavior has led to fundamental principles (Treisman and Gelade 1980; Wolfe et al. 1989) and empirical generalizations (e.g., Duncan and Humphreys 1989; Wolfe 1998) that can be used to characterize and improve real-world search behavior (e.g., Berbaum et al. 1990; Wolfe et al. 2007). This kind of applied psychological science was a pre-occupation of Donald Broadbent, who is inextricably linked to the archetypical theory of attention that often bears his name ("Broadbent's Filter Theory", Broadbent 1958). And while we know that his theory, which provides a departure point for so many subsequent theories of attention, is wrong in detail, it is a fitting tribute to Broadbent's applied inclinations that Neville Moray (1993) would later say of Broadbent's filter theory that: "Whatever the deep structure of attention may be, its surface performance is, in the vast majority of cases, well described by a single, limited capacity channel" (page 113). Consequently, from an applied perspective we can not only aim to improve search performance per se but also to decrease the demands it places on limited capacity resources.

A Taxonomy of Attention

With roots in a program of research begun by Michael **Posner** over 40 years ago (Posner and Boies 1971) three isolable functions of attention—alertness, orienting, and executive control—have been identified and linked to specific neural networks (Posner and Peterson 1990; Fan et al. 2005). In the domain of space, where selection has been referred to as orienting and most of the research has been on visual orienting, two important distinctions were first made by Posner (1980) and have since been highlighted in work from Klein's laboratory (for a review, see Klein 2009). One concerns whether selection is accomplished by an overt reorientation of the receptor surface (an eye movement) or by a covert reorientation of internal information processing mechanisms. The other concerns whether the eye movement system or attention is controlled primarily by exogenous (often characterized as bottom-up or reflexive) means or by endogenous (often characterized as top-down or voluntary) means.

Helmholtz provided the first demonstration that attention could be shifted covertly and consequently independently of the direction of gaze. When control is purely endogenous, (Klein 1980; Klein and Pontefract 1994) and others (e.g., Hunt and Kingstone 2003; Schall and Thompson 2011) have demonstrated that such shifts of

Fig. 1 A taxonomy of attention proposed by Klein and Lawrence (2011)

		Mode of allocation	
		Exogenous	Endogenous
Domain of allocation	Space		
	Time		
	Sense		
	Task		

attention are not accomplished via sub-threshold programming of the oculomotor system. On the other hand, when orienting is controlled exogenously, by bottom-up stimulation, it is difficult to disentangle activation of covert orienting from activation of the oculomotor programs.

In the domain of covert orienting, Klein has emphasized the importance of distinguishing between whether control is (primarily) endogenous or exogenous because different resources or mechanisms seem to be recruited to the selected location or object when the two different control systems are employed. This assertion was first supported by the following double dissociation: (1) When exogenously controlled, attention interacts with opportunities for illusory conjunctions and is additive with non-spatial expectancies, and (2) when endogenously controlled, attention is additive with opportunities for illusory conjunctions and interacts with non-spatial expectancies (Briand and Klein 1987; Briand 1998; Handy et al. 2001; Klein and Hansen 1990; Klein 1994). Several other dissociations discovered by others reinforce Klein's conclusion that different resources are recruited when orienting is controlled endogenously versus exogenously (for reviews, see Klein 2004, 2009).

Thinking about the importance of this distinction in the world of orienting led Klein and Lawrence to propose an alternative taxonomy (Klein and Lawrence 2011), illustrated in Fig. 1, in which two modes of control (endogenous and exogenous) operate in different domains time, space, modality, task, etc.). Searching entails the endogenous and exogenous control of attention in space and time. In contrast to the literature using Posner's cuing paradigm, however, in typical search tasks the endogenous/exogenous distinction is often not made explicit. In spatial search, for example, perhaps this is because even when search is hard (the target does not exogenously capture attention) we typically do not experience volitional control of the search process—of the sequence of decisions about where to look next for the target. It has been suggested that these "decisions" are typically made by low-level subroutines (Klein and Dukewich 2006). It seems likely that the endogenous control of search is instantiated before the search episode begins based on the observer's knowledge about properties of the target (setting up a template matching process) and distractors (e.g. establishing attentional control settings to implement guided search).

Natural History of a Search Episode

A typical search episode begins with some specification of what the target is; usually some information about the nature of the material to be searched through for the target; perhaps some useful information on how to find it; and, critically, what to do when it is found. The human searcher is thought to incorporate these tasks- or goal-oriented elements into a mental set, program or strategy so that their performance will optimize their payoffs. In Broadbent's theory (1958) an important component of this process was "setting the filter" so that task-relevant items (targets) would have access to limited capacity processing mechanisms while task irrelevant items would be excluded. Duncan (1981) would later provide a useful recasting of Broadbent's ideas. Instead of "filtering" he referred to a "selection schedule" and, recognizing the many empirical demonstrations that an unselected stimulus could nevertheless activate complex internal representations, he suggested that the limitation has more to do with availability for reporting an item than the quality or nature of an item's internal representation. We see subsequently proposed endogenous control mechanisms such as attentional control settings (ACS) (Folk et al. 1992) and "task-set reconfiguration" (Monsel 1996) as firmly rooted in these earlier ideas.

During the search episode the efficient performer must represent the target and the feature(s) that will distinguish the target from the distractors. Representations activated by the spatial search array or temporal search stream are compared against these representations to determine if the target is present and if so to report its properties according to observer's goals. This comparison process might take place one at a time or in parallel across the items in the search array or stream.

Two paradigms for exploring the information processing dynamics of searching will be emphasized in this chapter. These paradigms were developed to study, in relatively pure form, searching in space and in time. Searching in space entails the allocation of attention to items distributed in space and presented at the same time. Searching in time entails the allocation of attention to items distributed in time and presented at the same location. With a few exceptions (e.g., Arend et al. 2009; Keele et al. 1988; McLean et al. 1982; Vul and Rich 2010) searching in space and time has been studied separately, usually in studies with a similar objective: understanding the role of attention in detecting, identifying, or localizing targets. We believe that it will be empirically fruitful and theoretically timely for these somewhat separate efforts to be combined. And, it will be useful, because in the real world searching often combines these two pure forms.

Searching in Space

There are many studies from before 1980 that used a wide variety of spatial search tasks, The spatial search paradigm emphasized here (see Fig. 2) was imbued with excitement by Anne **Treisman's** (Treisman and Gelade 1980; Treisman and Schmidt 1982) use of it to provide support for her feature integration theory in which spatial

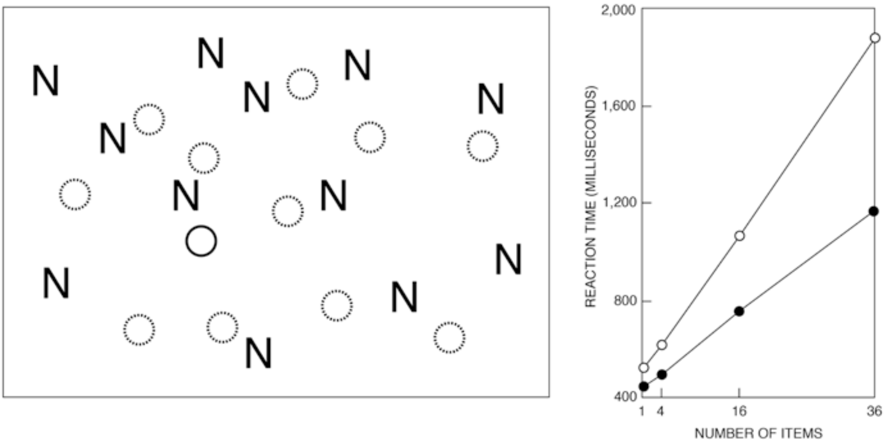


Fig. 2 A prototypical "present/absent" search task (is there a solid "O" in the display?) is illustrated on the *left*. Typical results illustrated on the *right* showing reaction time to make the decision (*open symbols* = target absent trials; *filled symbols* = target present trials) as a function of the number of items in the display. (Adapted from Treisman 1986)

attention is the binding agent for otherwise free-floating features. When observers are asked to indicate whether a target is present in an array of distractors, two dramatically different patterns are frequently reported. In one case (i.e., difficult search—target is not defined by a single unique feature), illustrated in Fig. 2, reaction time for both target absent and target present trials is a roughly linear function of the number of distractors and the slope for the target absent trials is approximately twice that of the present trials. This pattern is intuitively compatible with (indeed predicted by) a serial self-terminating search (SSTS) process in which each item (or small groups of items) is compared against a representation of the target and this process is repeated until a match is found or until the array has been exhausted. In the other case (not illustrated) (i.e., easy search—target is defined by a single unique feature), reaction time is unaffected by the number of distractor items. Phenomenologically, instead of having to search for the target, it "pops out" of the array.

This model task and the theory Treisman inferred from its use have been remarkably fruitful in generating: modifications of the model task (e.g., the preview-search paradigm of Watson and Humphreys 1997; the dynamic search paradigm of Horowitz and Wolfe 1998), theoretical debates (such as: are so-called "serial" search patterns like that illustrated in Fig. 2 caused by truly sequential or by parallel processes; and, when search is a sequential process of inspections, how much memory is there about rejected distractors, see Klein and Dukewich 2006, for a review), empirical generalizations (e.g., Wolfe's 1998, review; the search surface of Duncan and Humphreys 1989), and conceptual contributions (e.g., the guided search proposal of Wolfe et al. 1989; the foraging facilitator proposal of Klein 1988).

The model task and the theory of Treisman encouraged Klein and Dukewich (2006) to address the question whether search is primarily driven by serial or parallel mechanisms. While rooted in basic research on spatial search, we believe that their advice applies equally to searching in time and to real-world search behavior:

> When there is more than one good strategy to solve a problem it seems reasonable to assume that nature may have figured out a way to take advantage of both. . . .We recommend that future research seek to determine, rather than which strategy characterizes search, "when" and "how" the two strategies combine. (Klein and Dukewich 2006, p. 651)

Searching in Time

In the mid-1960's Molly **Potter** discovered that people could read when the text was presented using rapid serial visual presentation (RSVP), that is with words presented one after the other at the same location in a rapid sequence. A few decades later this mode of stimulus presentation began to be used as a tool for exploring the consequences of limited processing capacity, particularly for dealing with multiple "targets" in streams of unrelated items (Broadbent and Broadbent 1987; Weichselgartner and Sperling 1987). Broadbent and Broadbent (1987), for example, showed how difficult it is to identify two targets when they are in close succession.

The difficulty identifying subsequent items after successfully identifying an earlier one was subsequently named an "attentional blink" by Raymond et al. (1992). The blink and the task for exploring it that was developed by **Broadbent, Raymond** and **Shapiro** propelled this paradigm to the center stage of attention research. In the seminal paradigm of Raymond et al. (1992) (see Fig. 3, left/bottom), multiple letters are presented rapidly and sequentially at the same location (in RSVP). In the sequence of letters, all but one of which are black, there are two targets (separated by varying numbers of distractors) and the observer has two tasks: Report the identity of the white letter and report if there was an X in the stream of letters after the white letter.

One possible weakness of this particular paradigm (often called "detect X") is that the "blink" it generates and measures may have quite different sources: double speeded identification and switching the mental set (the selection schedule or filter setting) from color to form ("white" to "X"). A more general paradigm (that is more like Broadbent's) is often used to avoid such switching. Chun and Potter (1995) used one version of this paradigm (Fig. 3, right/top) in which the observers task is to report the identity of two letter targets that are embedded in a stream of digits.

As with the spatial search paradigm, these methods for exploring "searching in time" using one or more targets embedded in a stream of rapidly presented items, have been remarkably fruitful in generating: modifications of the model task, theoretical debates, empirical generalizations, and conceptual contributions (e.g., Dux and Marios 2009 for a review).

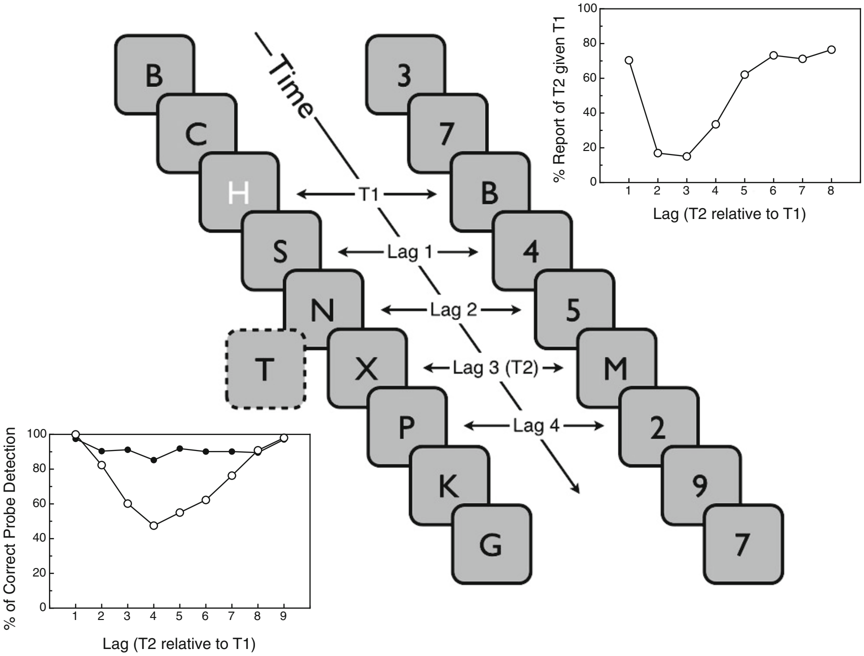


Fig. 3 Two different methods that have been used to explore the attentional blink. Both entail presenting a sequence of individual alphanumeric items using RSVP (with about 100 ms separating item onsets). The stream on the *left* illustrates the "detect-X" task pioneered by Raymond et al. (1992). After a random number of black letters the first target, a white letter (T1), is displayed. Then, at varying lags after the presentation of T1 an X (T2 or probe) might or might not (this alternative is shown in the *box* with the *dashed line*) be presented. At the end of the stream the observer reports the identity of the white letter and whether or not an X had appeared in the stream. Typical results from this task are shown in the inset at the bottom. *Open symbols* show the probability of correctly reporting that an X was present as a function of its position following a white letter when that letter had been correctly identified. *Filled symbols* show the same results when there was no requirement to report the white letter. The stream on the *right* illustrates the paradigm developed by Chun and Potter (1995) and used by many others. Here there is a stream of items in one category (digits) in which two targets from another category (letters) are embedded. At the end of the stream the observer's task is to report the identities of the two targets. Typical results (accuracy of T2 reports when T1 was identified correctly) are shown in the inset at the top

Searching in Space and Time: Some Comparisons

The Nature of the Stimuli

It seems likely that if a certain kind of stimulus pops out in a spatial search it might also do so in temporal search and vice versa. Duncan and Humphreys (1989) identified two principles that interact in determining the difficulty of searching in space for a target among distractors. One factor is: How similar is the target to the distractors?

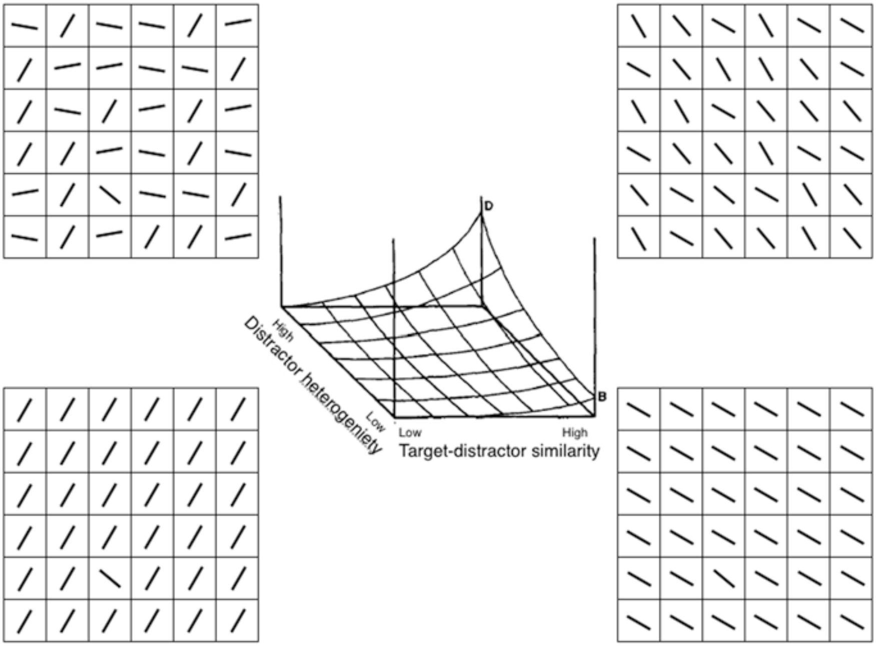


Fig. 4 Center: The "search surface" (adapted from Duncan and Humphreys 1989) represents the difficulty of finding a target (height of the surface is the predicted slope of the reaction time/set size function) as a function of two properties of the search array: target distractor similarity and distractor heterogeniety. Corners: Sample search arrays illustrating the four corners of the search surface. The line with the obviously unique slope in the *lower left* panel is the target in all four panels. The target is easily found when it is accompanied by a homogenous array of distractors of a very different orientation (*lower left*)

The other is: How heterogeneous are the distractors? How these factors interact to determine search difficulty (see Fig. 4) was described by Duncan and Humphreys (1989); neither factor alone makes searching particularly hard, but when combined they interact and conspire to make search extremely difficult. Would searching in time (in RSVP) show the same relationship? While there are hints that this might be true, there are no dedicated studies that we are aware of.

There are a variety of other stimulus features for which we could pose a similar question: If your own name pops out of an RSVP stream and even escapes the attentional blink will it also pop out in spatial search? Will socially important stimuli such as faces (emotional or otherwise) capture attention in both space and time? Given the history of this symposium, we can ask "What does motivation have to do with it?" For example, would pictures of food be easier to find when you are hungry than after you have just eaten? Will attention be captured by stimuli that have been previously rewarded?

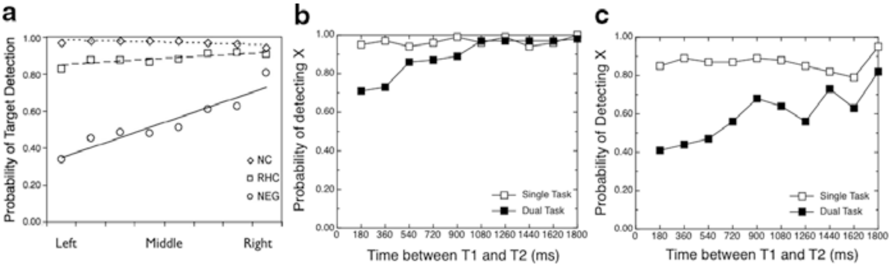


Fig. 5 Spatial and temporal processing in patients suffering from neglect and control participants. **a** Probability of report [by normal controls (NC), control patients with right hemisphere lesions (RHC), and patients suffering from neglect following damage to the right hemisphere (NEG)] of target letters and numbers among non-alphanumeric distractors presented in a 20 by 30 degree spatial array in peripersonal and extrapersonal space (from Butler et al. 2009). **b** and **c** Probability of detecting an X in the "detect X" paradigm illustrated in Fig. 3. *Unfilled squares* represent performance when participants were not required to report the white letter in the stream (single task). *Filled squares* represent performance on the "detect X" task (second target) when participants were required to report the white letter (first target). **b** data from normal controls. **c** data from patients with neglect (Data in **b** and **c** are from Husain et al. (1997); figures **b** and **c** are adapted from Husain and Rorden (2003)

The Participants

There are many participant factors that could be explored. We would expect searching in space and time to show similar benefits from training and expertise, for example. The same expectation would apply to developmental changes. Exploring the efficiency of spatial search across the lifespan, Hommel et al. (2004) found a U-shaped function with less efficient performance at the extremes. Based on their findings, if you have recently turned 25 or so, you are at your peak. A similar pattern, though perhaps with a slightly older "optimum" age, was reported for the magnitude of the attentional blink by Georgiou-Karistianis et al. (2007) Looking at patients with focal brain damage or known neurological problems would provide an arena for comparison that could have relevance to the neural systems involved in search. Examples described here are from studies of patients with unilateral neglect, a disorder commonly associated with parietal lesions. In spatial search tasks patients with neglect are slower and less likely to find targets, particularly when these are present in the neglected hemifield (e.g., Butler et al. 2009; Eglin et al. 1989). The right-to-left gradient of increasing omissions (see Fig. 5a) might be related to a difficulty disengaging attention from attended items toward items in the neglected field (for a review, see Losier and Klein 2001). Poor performance, particularly repeated reports of targets (cf Butler et al. 2009), might be attributed, in part, to defective spatiotopic coding of inhibition of return (IOR) which depends on an intact right parietal lobe (Sapir et al. 2004). This would converge with the proposal that the function of IOR is to encourage orienting to novelty (Posner and Cohen 1984) and, consequently, to discourage reinpsections (Klein 1988). Using an RSVP task, Husain et al. (1997)

showed that the attentional blink was longer and deeper in patients suffering from visuo-spatial neglect due to damage to the right hemipshere. In this study, all the items were presented at fixation. Consequently, this temporal deficit might be a more general version of the aforementioned disengage deficit: difficulty disengaging attention from any item on which it is engaged.

The Role(s) of Endogenous Attention in Time and Space

As noted earlier the concept of limited capacity seems to play an important role in both kinds of search. When searching in space, one reflection of this limit is seen in the relatively steep slopes that characterize difficult searches (searches for which the target does not pop-out). As noted earlier, one way to explain steep slopes is in terms of the amount of time required for an attentional operator to sequentially inspect individual items in the array or to sequentially inspect regions (when it is possible for small sets of nearby items to be checked simultaneously) until the target is located. When searching in time this is seen as an attentional blink—in the period immediately following the successful identification of a target, some important target-identifying resources appear to be relatively unavailable.

An interesting difference that characterizes at least the standard versions of these tasks is that stimuli in RSVP are data limited: every item is both brief and masked while in a typical spatial search episode the stimulus array is neither brief nor masked. With multiple items displayed all at the same time, spatial search is characteristically resource limited. That noted, several researchers (e.g., Dukewich and Klein 2005; Eckstein 1998) have explored spatial search using limited exposure durations. And, while in this chapter we are concentrating on relatively pure examples of searching in space and time, there have also been some highly productive hybrids (such as the dynamic search condition of Horowitz and Wolfe 1998 2003,).

The ideas of attentional control settings and contingent capture seem to operate similarly in both space and time. In spatial search it has been demonstrated that attentional capture is contingent on the features one is searching for (Folk et al. 1992) as well as the locations where targets will be found (Ishigami et al. 2009; Yantis and Jonides 1990). Capture by distracting non-targets that share features with the target has also been demonstrated in temporal search (Folk et al. 2008).

Another aspect of attentional control concerns its intensity (Kahneman 1973). For example, in his review of IOR, Klein (2000) proposed that the strength of attentional capture by task-irrelevant peripheral cues would depend directly on the degree to which completing the target task requires attention to peripheral onsets. As a consequence of increased capture, attentional disengagement from the cue and therefore the appearance of IOR would be delayed.

A similar mechanism was uncovered in our studies of the attentional blink. The initial question we (McLaughlin et al. 2001) posed was whether difficulty to identify the first target (T1), when varied randomly from trial-to-trial, would affect blink magnitude. We used the target-mask, target-mask paradigm (which, it must be noted,

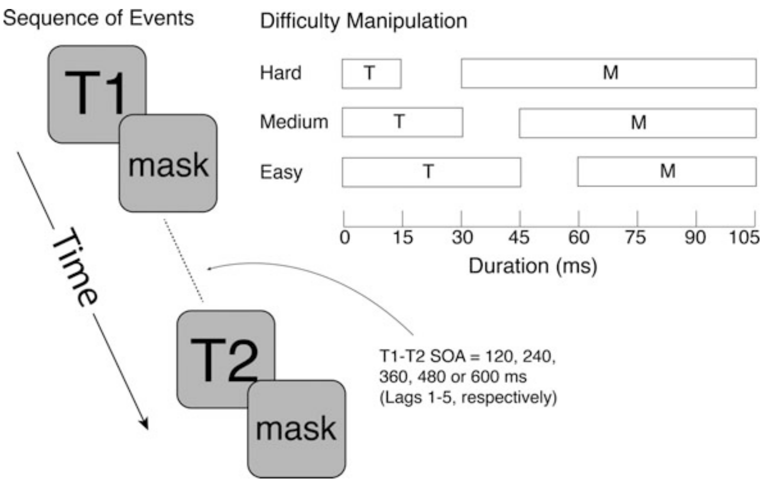


Fig. 6 Methods used by McLaughlin et al. (2001) to explore the effect of the difficulty of target (T) processing upon the magnitude of the blink using a target-mask, target-mask paradigm to induce and measure the blink. The difficulty of either T1 (first target) or T2 (second target) was manipulated by varying the relative durations of the target and mask (M)

demonstrates that it is not necessary to use RSVP streams to explore searching in time) pioneered by Duncan et al. (1994). As shown in the bottom panel of Fig. 6 we varied how much data was available about either T1 or T2 (second target) in order to implement an objective, quantifiable and data-driven difference in target identification difficulty. We designed the experiment so as to avoid any location or task switching (the task was simply to report the two letters). Despite the success of our data-driven manipulation of T1 difficulty, the answer to this question was a resounding “NO” (see top panel of Fig. 7)[1]. When we manipulated the difficulty of T2, this had dramatic effects on T2 performance and no effect on T1 (bottom panel of Fig. 7).

Why would such a dramatic difference in difficulty of T1 have no effect on the blink? We suggested that this was because the blink is about the effort the participant expects to have to exert in advance of the trial—an ACS that is about how much processing resources might be needed to perform the task. Because we randomly intermixed the 3 difficulty levels, and because (apparently) resources are not (or cannot be) re-allocated in real time when T1 is presented, all trials would have been subjected to the same ACS. We tested this proposal, in a subsequent paper (Shore et al. 2001), by comparing the results when the same data-driven manipulation of T1 difficulty was mixed or blocked. As predicted by an ACS view, when we blocked difficulty there was a significant effect of T1 difficulty on the magnitude of the AB (particularly between the hard and medium/easy conditions, See Fig. 8).

---

[1] Also note the absence of lag-1 sparing. This occurred, despite the very short amount of time between T1 and T2 at lag 1, because T1 had nevertheless been masked—see Fig. 6).

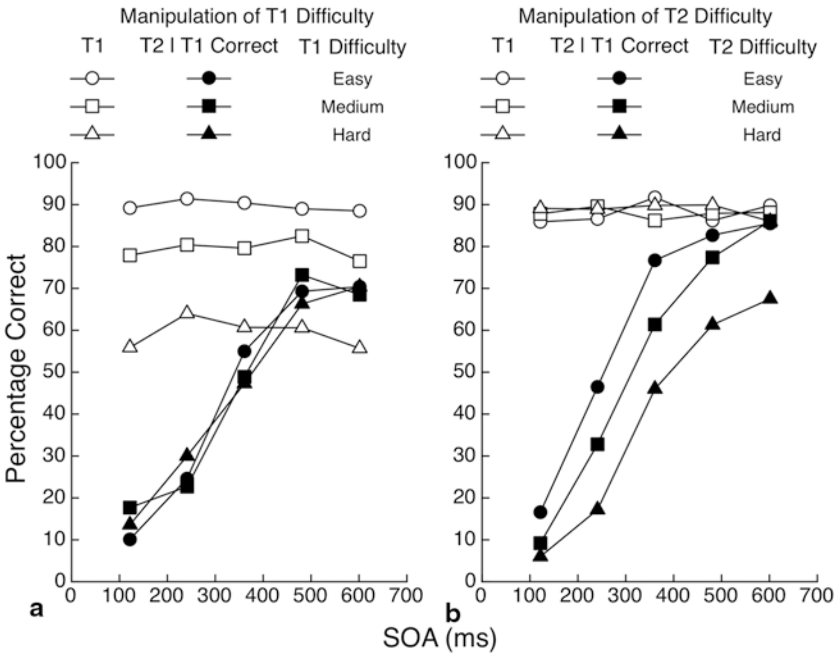


Fig. 7 Results from McLaughlin et al. (2001). (See Fig. 6 for explanation of the difficulty manipulation)

There may be a related "strategic" effect in both the spatial and temporal search literatures. Smilek et al. (2006), in a paper entitled: "Relax! Cognitive strategy influences visual search" seemed to show that simply telling their participants not to try so hard reduced their slopes (i.e., increased their search efficiency). Similarly,

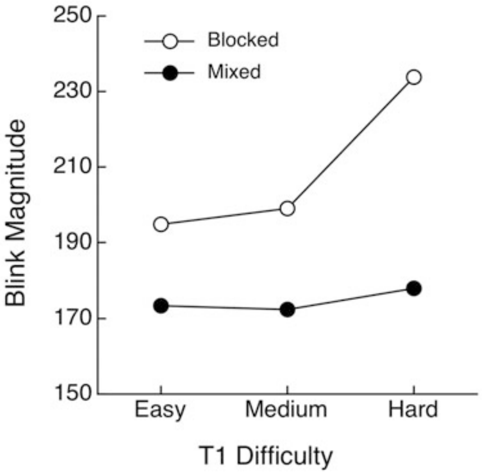


Fig. 8 Results from Shore et al. (2001). Magnitude of the attentional blink as a function of T1 difficulty and whether T1 difficulty could be predicted (*blocked*) or not (randomly intermixed, as in McLaughlin et al. 2001)

Olivers and Nieuwenhuis (2005) reported that relaxing by listening to music could reduce the attentional blink.

Binding of Targets in Space and Time

We will end this section by describing one empirical strategy for comparing searching in space and in time. The background comes from two papers that reported interesting "slippage" of targets in space and time. The first, by Snyder (1972), was about searching in space; the second by McLean et al. (1982), was about searching in time. In Snyder's study multiple items were presented briefly at the same time in different locations whereas in McLean et al. (1982) multiple items were presented rapidly in time at the same location. For present purposes we will emphasize the conditions in which the participant's task was to report the identity of a target letter that was defined by color. As we will see, both studies reported a certain amount of sloppiness of the attentional beam (or window); whether the errors were true illusory conjunctions is not so important as their distribution in space and time.

In Snyder's spatial search task, 12 letters were placed in a circular arrangement on cards for presentation using a tachistoscope[2]. On each trial the participant had to verbally report the name of a uniquely colored letter and then report its position (using an imaginary clockface: 1–12). Stimulus duration was adjusted on an individual basis so that accuracy of the letter identification was about 50 % (regardless of accuracy of the letter localization). The key finding for present purposes was that when reporting identities[3], errors were more likely to be spatially adjacent to the target letter than further away. Snyder found a similar pattern of spatial slippage when the feature used to identify the target was form-based (a broken or inverted letter).

In McLean et al.'s temporal search task, the target color varied from trial to trial and the participant's task was to report the identity of the single item presented in the target color. (In another condition the participant reported the color of a target defined by its identity.) Each stream, created photographically using movie film, consisted of 17 letters rendered using 5 different identities and 5 different colors. Films were projected on the screen with SOAs of 67 ms (15 flames/s). The key finding for present purposes was an excess of temporally adjacent intrusion errors relative to reports of items in the stream temporally more distant from the target (interestingly, immediate post-target intrusions were more likely than immediate pre-target intrusions).

If there were one attentional beam that operates in both space and time to integrate features into objects[4], and if there are individual differences in the efficacy of this

---

[2] Some readers may find this surprising, but even though Posner's laboratory (which is where these experiments were conducted) was in the forefront of using computers for psychological research, in 1971 there was almost no possibility of computerized presentation with color displays.

[3] Snyder (1972) used 'legitimate' trials for the analyses reported in his paper. By his definition legitimate trials are trials for which the reported location falls within ±1 of the location of reported identity.

[4] This is the beam controlled exogenously by bottom-up stimulation (see also, Briand and Klein 1987). To be sure, and as described earlier, the ACS or selection schedule was put into operation by endogenous control mechanisms.

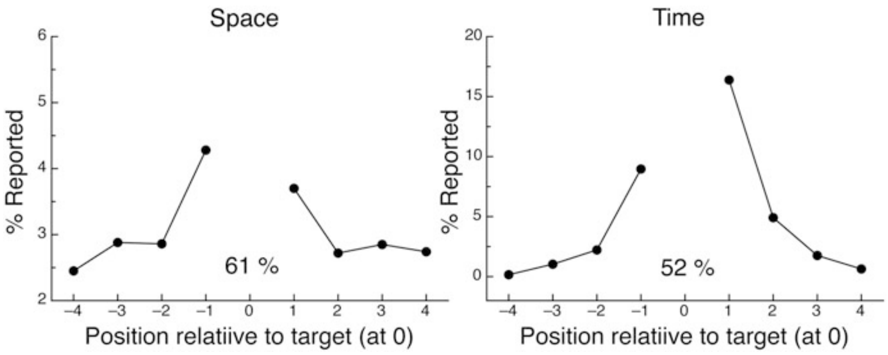


Fig. 9 Results from Ishigami and Klein (2011). Observers were searching in space (*left panel*) and time (*right panel*) for a target of a pre-specified color. Accurate reports of the target's identity are indicated in the percentages indicated above relative position = 0. The remaining data are the percentage of erroneous reports of items from the array (that were not the target) as a function of the distance (in space and time) of these items relative to the target. Positive positions are, relative to the target, clockwise in space and after in time

beam, then we would expect the spatial and temporal sloppiness that was reported by Snyder (1972) and McLean et al. (1982) to be correlated across individuals. To test this idea, data on spatial and temporal search must be obtained from the same participants. We have begun to explore this possibility and will briefly report some of our preliminary findings.

In our first project we tested 46 participants on spatial and temporal search tasks that were closely matched to those of Snyder (1972) and McLean et al. (1982). The order of tasks was counterbalanced. In order to ensure that there would be a sufficient number of errors while performance would be substantially above chance, for each task we titrated the exposure duration so that overall accuracy in reporting the target's identity was in the 50–60 % correct range. The key results are illustrated in Fig. 9.

We were quite successful in achieving the overall level of accuracy we were aiming for (50–60 % correct). While the scales are different (there were fewer errors in the spatial task) the patterns are similar in space and time, and the key findings from Snyder (1972) and McLean et al. (1982) were replicated: errors are more likely to come from positions adjacent to the target. Moreover, in space there were more counterclockwise than clockwise errors; and in time there were more post- than pre-target errors. For each participant and task we computed a measure of "slippage" that was the average rate of near errors (± 1) minus the average rate of far errors (all other erroneous reports from the presented array). The correlation between spatial and temporal slippage was very close to zero ($r_{44} = 0.03$) suggesting that the attentional beam that attaches identities to locations may not be the same beam that attaches identities to time[5].

---

[5] When we applied Snyder's exclusion criteria (i.e., legitimate trials, see footnote 3) to both our spatial and temporal tasks, the correlation was marginally significant, $r = 0.34$, $p = 0.051$, but becomes non-significant when a single outlier is removed ($r = 0.20$). For a confident conclusion, further research is required.

Conclusion

We have discussed the concept of attention—selection made necessary by limited processing capacity—and some of its manifestations in spatial and in temporal search behavior. As described in the chapter, searching in space and time has been typically studied separately predominantly with an objective to understand the role of attention in detecting, identifying, or localizing targets. However, in the real-world, we are often searching for targets that are surrounded by distractors in space and all of this happens in scenes that unfold over time (e.g., looking for a particular exit on a highway when driving; or your child in a busy playground). We described above our first attempt to compare searching in time and space in the same individuals. Preliminary results revealed a null correlation between spatial and temporal slippage suggesting different selection mechanisms in these two domains. We plan next to experimentally balance two tasks (space and time) so that we can have firmer conclusion about this relationship and merge our two tasks so that we can explore searching in space and time simultaneously.

In the course of this chapter we have raised several questions: Will the principles (Duncan and Humphreys 1989) that determine the difficulty of searching in space generalize to searching in time? Are the same brain regions responsible for spatial and temporal search (e.g., Arend et al. 2009)? Do attentional control settings work in the same way in spatial and temporal search? Is the binding of features to space and to time implemented by one "beam" or by independent "beams," each operating in its own domain? Answers to these questions which, in some cases, the literature is beginning to provide, will have important theoretical and practical implications.

References

Arend, I., Rafal, R., & Ward, R. (2009). Spatial and temporal deficits are regionally dissociable in patients with pulvinar lesions. *Brain: A journal of neurology*, *131*(Pt 8), 2140–2152.

Berbaum, K. S., Franken Jr., E. A., Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., Barloon, T. J., Behlke, F. M., et al. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology*, *25*(2), 133.

Briand, K. A. (1998). Feature integration and spatial attention: More evidence of a dissociation between endogenous and exogenous orienting. *Journal of Experimental Psychology: Human Perception and Performance, 24,* 1243–1256.

Briand, K., & Klein, R. M. (1987). Is Posner's beam the same as Treisman's glue?: On the relationship between visual orienting and feature integration theory. *Journal of Experimental Psychology: Human Perception & Performance*, *13*(2), 228–247.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Broadbent, D. E., & Broadbent, M. H. (1987). From detection to identification: Response to multiple targets in rapid serial visual presentation. *Perception & Psychophysics, 42,* 105–113.

Butler, B. C., Lawrence, M., Eskes, G. A., & Klein, R. M. (2009). Visual search patterns in neglect: Comparison of peripersonal and extrapersonal space. *Neuropsychologia*, *47,* 869–878.

Chun, & Potter (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 109–127

Dukewich, K., & Klein, R. M. (2005). Implications of search accuracy for serial self-terminating models of search. *Visual Cognition, 12,* 1386–1403.

Duncan, J. (1981). Directing attention in the visual field. *Perception & Psychophysics*, *30*(1), 90–93.

Duncan, J., & Humphreys, G. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature*, *369*(6478), 313–5.

Dux, P. E., & Marois, R. (2009). The attentional blink: A review of data and theory. *Attention, Perception & Psychophysics*, *71*(8), 1685–1700.

Eckstein, M. P. (1998). The lower visual search efficiency for conjunctions is due to noise and not serial attentional processing. *Psychological Science, 9*(2), 111–118.

Eglin, M., Robertson, L. C., & Knight, R. T. (1989). Visual search performance in the neglect syndrome. *Journal of Cognitive Neuroscience, 12,* 542–5.

Fan, J., McCandliss, B. D., Fossella, J., Flombaum, J. I., & Posner, M. I. (2005). The activation of attentional networks. *NeuroImage, 26,* 471–479.

Folk, C. L., Leber, A. B., & Egeth, H. E. (2008). Top-down control settings and the attentional blink: Evidence for nonspatial contingent capture. *Visual Cognition*, *16*(5), 616–642.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030–1044.

Georgiou-Karistianis, N., Tang, J., Vardy, Y., Sheppard, D., Evans, N., Wilson, M., Gardner, B., Farrow, M., & Bradshaw, J. (2007). Progressive age-related changes in the attentional blink paradigm. *Aging, Neuropsychology, and Cognition, 14*, 213–226.

Handy, T. C., Green, V., Klein, R. M., & Mangun, G. R. (2001). Combined expectancies: event-related potentials reveal the early benefits of spatial attention that are obscured by reaction time measures. *Journal of Experimental Psychololgy: Human Perception Performance, 27*(2), 303–317.

Hommel, B., Li, K. X. H., & Li, S.-C. (2004). Visual search across the lifespan. *Developmental Psychology, 40*(4), 545–558.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(6693), 575–576.

Horowitz, T., & Wolfe, J. (2003). Memory for rejected distractors in visual search? *Visual Cognition, 10*(3), 257–298.

Hunt, A., & Kingstone, A. (2003). Covert and overt voluntary attention: Linked or independent? *Cognitive Brain Research*, *18,* 102–105.

Husain, M., & Rorden, C. (2003). Nonspatially lateralized mechanisms in hemispatial neglect. *Nature Reviews Neuroscience, 4,* 26–36.

Husain, M., Shapiro, K., Martin, J., & Kennard, C. (1997). Abnormal temporal dynamics of visual attention in spatial neglect patients. *Nature, 385,* 154–156.

Ishigami, Y., & Klein, R. M. (2011). Attending in space and time: Is there just one beam? *Canadian Journal of Experimental Psychology, 65*(4), 30.

Ishigami, Y., Klein, R. M., & Christie, J. (2009). Using illusory line motion to explore attentional capture. *Visual Cognition*, *17*(3), 431–456.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs: Prentice Hall.

Keele, S. W., Cohen, A., Ivry, R., Liotti, M., & Yee, P. (1988). Tests of a temporal theory of attentional binding. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(3), 444–452.

Klein, R. M. (1980). Does oculomotor readiness mediate cognitive control of visual attention. In R. Nickerson (Ed.), *Attention and Performance VIII* (pp. 259–276). Hillsdale: Erlbaum.

Klein, R. M. (1988). Inhibitory tagging system facilitates visual search. *Nature*, *334*(6181), 430–431.

Klein, R. M. (1994). Perceptual-motor expectancies interact with covert visual orienting under endogenous but not exogenous control. *Canadian Journal of Experimental Psychology, 48,* 151–166.

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

Klein, R. M. (2009). On the control of attention. *Canadian Journal of Experimental Psychology, 63,* 240–252.

Klein, R. M., & Dukewich, K. (2006). Does the inspector have a memory? *Visual Cognition, 14,* 648–667.

Klein, R. M., & Hansen, E. (1990). Chronometric analysis of spotlight failure in endogenous visual orienting. *Journal of Experimental Psychology: Human Perception & Performance, 16*(4), 790–801.

Klein, R. M., & Lawrence, M. A. (2011). The modes and domains of attention. In M. I. Posner (Ed.) *Cognitive Neuroscience of Attention* (2nd edn.). New York: Guilford Press.

Klein, R. M., & Pontefract, A. (1994). Does oculomotor readiness mediate cognitive control of visual attention? Revisited! In C. Umiltà & M. Moscovitch (Eds.) *Attention & Performance XV: Conscious and Unconscious Processing* (p. 333–350). Cambridge: MIT Press.

Losier, B. J., & Klein, R. M. (2001). A review of the evidence for a disengage operation deficit following parietal lobe damage. *Neuroscience and Biobehavioral Reviews, 25,* 1–13.

McLean, J. P., Broadbent, D. E., & Broadent, M. H. P. (1982). Combining attributes in rapid serial visual presentation tasks. *Quarterly Journal of experimental Psychology, 35A,* 171–186.

McLaughlin, E. N., Shore, D. I., & Klein, R. M. (2001). The attentional blink is immune to masking induced data limits. *Quarterly Journal of Experimental Psychology, 54A,* 169–196.

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93–148). Howe: Erlbaum.

Moray, N. (1993). Designing for attention. In A. Baddeley & L. Weiskrantz (Eds.) *Attention: selection, awareness, and control* (pp. 111–134). Oxford: Oxford University Press.

Olivers, C. N. L., & Nieuwenhuis, S. (2005). The beneficial effect of concurrent task-irrelevant mental activity on temporal attention. *Psychological Science, 16*(4), 265–269.

Posner M. I. (1980). Orienting of attention. *Quarterly Journal of Experimental Psycholology A, 32,* 3–25

Posner, M. I., & Boies, S. (1971). Components of attention. *Psychological Review, 78,* 391–408.

Posner, M. I., & Cohen, Y. (1984). Components of attention. In H. Bouma & D. Bowhuis (Eds), *Attention and Performance X* (p. 531–556). Hillside: Erlbaum.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25–42.

Raymond, J. E., Shapiro, K. L., & Arnell, K. M. (1992). Temporary suppression of visual processing in an RSVP task: An attentional blink? *Journal of Experimental Psychology: Human Perception and Performance, 18,* 849–860.

Sapir, A., Hayes, A., Henik, A., Danziger, S., & Rafal, R. (2004). Parietal lobe lesions disrupt saccadic remapping of inhibitory location tagging. *Journal of Cognitive Neuroscience, 16*(4), 503–509.

Schall, J. D., & Thompson, K. G. (2011). Neural mechanisms of saccade target selection evidence for a stage theory of attention and action. In M. I. Posner (Ed.) *Cognitive Neuroscience of Attention* (2nd edn.). New York: Guilford Press.

Shore, D. I., McLaughlin, E. N., & Klein, R. M. (2001). Modulation of the attentional blink by differential resource allocation. *Canadian Journal of Experimental Psychology, 55,* 318–324.

Smilek, D., Enns, J. T., Eastwood, J. D., & Merikle, P. M. (2006). Relax! Cognitive strategy influences visual search. *Visual Cognition, 14,* 543–564.

Snyder, C. R. R. (1972). Selection, inspection and naming in visual search. *Journal of eperimental Psychology, 92,* 428–431.

Townsend, J. T. (1971). A note of the identifiability of parallel and serial processes. *Perception and Psychophysics, 10,* 161–163.

Treisman, A. (1986). Features and objects in visual processing. *Scientific American 255,* 114B–125B.

Treisman, A. M., & Gelade, G. (1980). A feature integrauon theoryof perception. *Cognitive Psychology, 12,* 97–136

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*(1), 107–141.

Vul, E., & Rich, A. N. (2010). Independent sampling of features enables conscious perception of bound objects. *Psychological Science, 21*(8), 1168–1175.

Watson, D. G., & Humphreys, G. W. (1997). Visual marking: Prioritizing selection for new objects by top-down attentional inhibition of old objects. *Psychological Review, 104,* 90–122

Weichselgartner, E., & Sperling, G. (1987). Dynamics of automatic and controlled visual attention. *Science, 238,* 778–780.

Wolfe, J. M. (1998). What can 1 million trials tell us about visual search? *Psychological Science, 9*(1), 33–39.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: an alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433.

Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of experimental psychology. General*, *136*(4), 623–38.

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

Automatic Control of Visual Selection

Jan Theeuwes

Abstract This paper seeks out to reduce the role of the homunculus, the 'little man in the head' that is still prominent in most psychological theories regarding the control our behaviour. We argue that once engaged in a task (which is a volitional act), visual selection run off more or less in an automatic fashion. We argue that the salience map that drives automatic selection is not only determined by raw physical salience of the objects in the environment but also by the way these objects appear to the person. We provide evidence that priming (feature priming, priming by working memory and reward priming) sharpens the cortical representation of these objects such that these objects appear to be more salient above and beyond their physical salience. We demonstrate that this type of priming is not under volitional control: it occurs even if observers try to volitionally prepare for something else. In other words, looking at red prepares our brain for things that are red even if we volitionally try to prepare for green.

Keywords Attention · Selection · Automatic processing · Priming · Reward

Introduction

"*The eye is the window of the soul*" is one of the famous declarations of Leonardo Da Vinci. Nowadays such a quote would mean that the eyes provide a window on the brain—particularly how the brain controls where we attend now and where we will attend next. Visual selection is critical for the interaction with the world as it present threats and opportunities and drives our behaviour. It has been argued that we are not passive receivers of the world around us; instead we actively interact and visual selection (including eye movements) reflect how our brain resolves the competition between external stimulation from the environment and internal motivations such as our goals and intentions. As stated by William James in his seminal book *The Principles of Psychology* "*each of us literally chooses, by his ways of attending*

J. Theeuwes (✉)
Department of Cognitive Psychology, Vrije Universiteit,
Van der Boechorststraat 1, 1081 BT, Amsterdam, Netherlands
e-mail: J.Theeuwes@psy.vu.nl

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_3,

to things, what sort of universe he shall appear himself to inhabit" (James 1890, p. 424). Because our cognition depends on what we perceive, understanding cognition requires an understanding of visual selection processes.

Even though intuitively it may seem that seeing is a matter only of opening our eyes, it has become clear that visual selection involves an active process of selecting some and ignoring other information. We are not passive recorders of the outside world but merely active explorers in search for relevant information. The amount of information impinging on retina is enormous and there is a constant need to prioritize and select information from the environment. One of the most fundamental questions in cognitive research is: who controls selection, the gateway that determines what we think, see, and remember and how we act? On the one hand, selection may be forced upon us. When a bicyclist rapidly moves toward us, we attend to it even though we were not planning to do so. Our system seems to be wired such that important, salient objects get attended and acted upon even when we had no intention to do so. On the other hand, we have strategies to attend to those objects in the environment that are important for our current goals. If our goal is to safely navigate our car through busy traffic we attend to road markings, traffic lights and stop signs.

Since the early 1990s, a hot debate has emerged regarding the extent to which selection is controlled by the person in an active, volitional, top-down way or controlled by the characteristics of the features in the environment in a passive, automatic, "bottom-up" way (see reviews Burnham 2007; Corbetta et al. 2002; Itti and Koch 2001). Even though in the last 20 years, we have described compelling conditions in which salient events capture attention (Theeuwes 1991, 1992; van Zoest et al. 2004b) or the eyes (Theeuwes et al. 1999); (Theeuwes et al. 1998) in a bottom-up way, the majority view is that visual selection is fully under volitional top-down control. In other words, at any point in time, we determine what we select from the environment (Bacon and Egeth 1994; Folk et al. 1994; Wolfe 1994). Indeed, at any time it feels like we are controlling what we are searching for and looking at, for example, when searching for your favourite coffee in the supermarket or when searching for your car at the parking lot. How top-down and bottom-up control interact to produce selection has been debated **vigorously** (Burnham and Neely 2007; Corbetta and Shulman 2002; Itti and Koch 2001). Little progress has been made towards a resolution, with the debate shifting to the peculiarities of experimental designs, and to questions of which source of control dominates in which situation. Perhaps, the bottom-up versus top-down dichotomy is not the most productive way to conceptualize the question of how the gateway to cognition is controlled.

The Alternative: Automatic Control In the current paper, we abandon the classic bottom-up/top-down controversy and suggest a new perspective. Even though it may feel that we are in control of what we select from the environment, as an alternative we pose that once we engage in a task (which obviously is a volitional act) selection processes run off in an automatic way without much, if any, conscious control. Once a task is set into motion, selection is beyond top-down control mainly because there is no conscious feedback from what we are selecting. We claim that our ability to perform volitional control over selection also known as 'executive functions', ‘supervisory attentional system’ and ‘the Will’ is very limited. In this sense, our goal is to shrink the size of the black box of executive control, reducing in this way the cage of the homunculus, the ‘little man in the head’ that in many psychological theories appears to control our behaviour. Our conjecture is that visual selection processes do not require conscious choice or top-down intervention. The engagement to perform a task is enough to trigger an automatic selection repertoire that is controlled by and large by the (acquired) characteristics of the environmental input. This would imply that volitionally controlled selection is essentially not very different from automatic selection. In fact, the distinction between automatic and volitional control may even be misleading and therefore detrimental to a proper understanding of visual selection processes. The present approach is reminiscent of classic behaviourists, who suggested that higher-order behaviour is directly controlled by stimuli. However, unlike behaviourists we focus on the processes in mind and brain that explain behaviour. Automatic selection processes are fast, do not require conscious deliberations and occur with little or no awareness. People do not seem to have introspective access to these processes or volitional control over them. It is likely that such automatic selection processes solve problems of evolutionary importance, such as the immediate detection of threat in the environment or the tendency to select objects that were rewarded in the past. One of the major discussions in twentieth century psychology, philosophy and more recently neuroscience is concerned with the question of free will and the extent to which complex higher-order mental process are under volitional control. Mainstream psychology still adheres to the position that complex behaviour including visual selection is mediated by the person’s purpose and his or her active interpretation of the environment. Our view is that visual selection proceeds without conscious deliberation and choice and is much less flexible as previously assumed.

Our claim is that we only have limited, if any, possibility to modulate our internal control processes once we have committed ourselves to perform a task. To change our behaviour we need to execute the task with a new setting several times, which ultimately will result in a new set of parameters for automatic behaviour. Changing visual selection on-line *at will* is not possible. For example, selection is automatically driven by priming (influence of past experience), by the content of our working memory (things on our mind), and by reward (influences of the consequences of our behaviour). Our conjecture that selection may not be a deliberate choice but rather an automatic modus operandi puts William James’ words quoted above (“each of us literally chooses, by his ways of attending to things, what sort of universe he shall appear himself to inhabit”; (James 1890, p. 424), in a completely different perspective. It is not a choice, we cannot “choose the universe we want to inhabit”, but instead automatic, nonconscious selection processes create the universe for us.

Indeed, if we want to change our selection priorities on-line, we should be aware what we are selecting at any given time. The question is how often we are in fact aware of where we attend or even where our eyes are. With respect to overt selection involving eye movement it has been argued that the eye movement machinery runs off more or less automatically inspecting the environment in a fast and efficient way. Eye movements are “cheap”, quickly made without much effort. In one of our eye movement studies (Theeuwes et al. 1998) our eye movement trackers showed that people fixated particular salient objects. However, none of the observers reported the presence of the salient object. When asked, all observers said that they never looked at it. It seems that our eye movement system does a lot of things we are not aware of. This is also illustrated by our experiments (Godijn and Theeuwes 2002b) in which observers had to make an eye movement to a target goal. During the saccade (in which people are temporarily blind due to saccadic suppression) the target goal was switched to another location. The eyes eventually still ended up at the target location even when it was quite different from the original location. Again, observers were not aware that the location of the target changed. These studies suggest that once the eye movement system is engaged, it runs to an end without any volitional top-down control.

Even though our view on automatic selection processes may appear radical, in this paper we discuss the viability of such a conjecture. The idea that selection may be fully automatic has been largely ignored so far, possibly because it feels like we are consciously and systematically processing incoming information. This may all be an illusion. In a classic study of Libet et al. (1983) participants had to lift their finger "whenever they felt the urge to do so". The time at which they felt the urge to make the action occurred some 100 ms after the first detectable change in brain activity. If, as this study illustrates, awareness comes after movement selection then presumably awareness has no role in making that selection.

Although our proposition goes against mainstream, it is consistent with the perspective that in general, behaviour is not controlled in a volitional way, i.e., by actively choosing and controlling actions. A recent influential review Bargh and Chartrand (1999) indicates that our ability to exercise intentional control is in fact quite limited despite the fact that "*much of contemporary psychological research is based on the assumption that people are consciously and systematically processing incoming information in order to construe and interpret their world and to plan and engage in courses of action"* (p. 462). It is also reminiscent of research in the area known as task switching in which participants have to switch at will their response to a particular stimulus. Task-switching studies consistently find that latencies are longer and error rates higher on 'switched' than on 'repeated' trials (Rogers and Monsell 1995). This is true even when participants are allowed long intervals between the cue and the stimulus to volitionally prepare for the upcoming task. These so-called 'residual switch cost' cannot be eliminated except by performing the task a few trials. In other words, true volitional control may not exist: one needs a few trials to establish the task control settings which then allows a more or less automatic execution of the task. Switching task settings from trial to trial *at will* may not be possible.

Below we first define covert and overt visual selection; then we discuss the viability of the notion of automatic selection on basis of four different mechanisms: bottom-up extraction, implicit memory, explicit memory and reward (see Fig. 1). As noted, the purpose of the present paper is to see how far we can go in reducing the role of endogenous control in visual selection.

Defining Visual Selection

In order to behave in a goal directed manner, it is important that we select only relevant information from the environment and ignore information that is irrelevant, particularly when this information disrupts our actions. This process of selecting part of simultaneous sources of information by either enhancing processing of some objects or/and by suppressing information of others can be accomplished either covertly or overtly. When selection occurs covertly only attention (and not the eyes) is directed at a location in space (Posner 1980). For example, without moving one's eyes by directing attention to the right side of the visual field one is able to detect an approaching car from a side street. When selection occurs overtly not only attention but also the eyes are moved to a particular location in space. Even though in every-day life attention and eye movements are usually correlated, attention precedes the overt movement of the eyes and therefore attention and eye movements may be dissociated. Theories of attention are concerned with how people select information to provide the basis for responding and with how information irrelevant to that response is dealt with.

Covert Selection

It is generally agreed that visual selection involves two functionally independent stages of processing (e.g., Broadbent 1958; Neisser 1967; Treisman and Gelade 1980). An early visual stage, sometimes referred to as pre-attentive, operates in parallel across the visual field and a later stage, often referred to as attentive, can deal with only one (or a few items) at the same time. Even though many modern theories of visual selection do not speak about a strict dichotomy between these two stages, in basically all past and present theories of visual attention this basic architecture is still present (e.g., Itti and Koch 2001; Koch and Ullman 1985; Li 2002; Wolfe 1994). It is assumed that visual selection depends principally on the outcome of the early stage of visual processing. Processing occurring during the initial wave of stimulation through the brain determines which element is selected and is passed on to the second stage of processing. In line with the two-stage approach, passing on an item to the second stage of processing implies that this item has been selected for further processing (e.g., Broadbent 1958; Neisser 1967; Treisman and Gelade 1980). This means that from all objects that are present in the visual field (and are available at the pre-attentive stage of processing), only the object that is passed onto the final stage of processing will affect decision-making and responding. This passing on from the initial stage of pre-attentive processing to attentive processing is what is considered to be selection.

This same operation can be considered from the well-known biased competition viewpoint (Desimone and Duncan 1995). According to this view, attention biases the competitive interactions occurring at the early stage of processing (possibly the pre-attentive stage) such that attended stimuli receive priority over unattended stimuli. Attentional effects on resolving this competition are the result of bottom-up and top-down factors. The bottom-up signal depends on the (acquired) properties (see Sect. Evidence for Automatic Selection) of the stimulus field. Objects that are highly salient and stand out from the background (i.e., a red poppy in a green field) will immediately receive attention priority. In this case, the visual system is biased towards salient stimuli that resolve the competition simply on the basis of the bottom-up input (see e.g., Hickey et al. 2010a; Mathot et al. 2010; Reynolds and Chelazzi 2004; van Zoest et al. 2004a).

Another way to bias the competition between objects is through top-down volitional feedback signals that depend on the goals, intentions and expectations of the observer. For example, directing spatial attention in a volitional top-down way to a location in space increases the sensory gain for features at that location (e.g., Theeuwes and Van der Burg 2007) and appears to alter the apparent stimulus contrast (e.g., Carrasco et al. 2004). In other words, directing attention to a location in space results in a greater neuronal sensitivity for objects appearing at that location (i.e., a decreased threshold). As a metaphor visual attention has been compared to a spotlight that "selects" parts of the visual world around us (e.g., Posner 1980). This type of selection is endogenous and is often referred to as goal-driven selection.

As noted, we adhere the classic notion that visual *selection* is the passing of information from the initial stage of preattentive processing to attentive processing. Note that in some conditions the preattentive, parallel stage of processing may play no role. In such a condition, there may be no salience calculation across the visual field and an object is selected purely on the basis of spatial information. For example, in case of endogenous cueing in which observers direct their attention to a location in space before the target is presented (c.f. Posner et al. 1980) the salience of the other elements in the visual field hardly plays a role. Indeed, previous studies have shown that when observers direct attention to a location in space (by means of an endogenous arrow cue) before the display is presented, irrelevant abrupt onsets cease to capture attention (Theeuwes 1991; Yantis and Jonides 1990). Also, when search is serial (or partly serial), preattentive processing plays no or only a minor role because due to the serial nature of the task, attention is focused on a restricted spatial area thereby circumventing preattentive processing outside that area (Belopolsky and Theeuwes 2010; Belopolsky et al. 2007).

We claim that "location" is not just like any feature that helps separating the target from noise but has a special status for selection (Theeuwes and Van der Burg 2007). In fact, there is evidence that even for the detection of the simplest feature, spatial attention is needed. In a recent study of Theeuwes et al. (2008) observers had to indicate whether a colour singleton (a red object between green objects) was present or not; there was no need to identify the target. The data indicated that even for such a simple detection response, focal attention needed to be shifted to the location of the singleton. These findings are inconsistent with claims from the main stream theories of visual attention such as feature integration theory (FIT Treisman and Gelade 1980) and the more modern version of FIT (Müller et al. 1995; Müller et al. 2003; Wolfe 1994), which assume that when observers need to detect a single feature singleton, they can check a pooled response from the relevant feature map for the presence of activity anywhere in that map. According to these theories detecting a pop-out target does not require the involvement of focal attention. Contrary to this assumption, the study of Theeuwes et al. (2006) suggests that spatial attention—the gateway of visual selection is always necessary in order to respond to target. A response without directing spatial attention—that is without selection is not possible.

Overt Selection

Selection is “overt” when people not only covertly direct attention to a location in space but also make a subsequent eye movement to that location. In normal circumstances (outside the lab) covert and overt orienting are highly correlated. Saccadic eye movements make it possible to quickly build an accurate representation of the visual environment, as they bring the fovea, the part of the retina with the highest acuity, to objects of interest. To appreciate the relationship between covert and overt orienting: by means of covert orienting we are able to discern in periphery for example that something big, squared and blue is present; however a saccade has to be made to that location to identify the text that is written on this blue traffic sign. By means of covert attention we are able to identify basic features such as colours, shapes, luminance and movement. We need to direct the fovea by means of eye (or head) movements to be able to resolve information that requires a high spatial resolution (such as reading). In everyday life situations, covert attention may be captured by a salient event in the periphery (or may be directed endogenously to a location in space). Typically such capture of attention is followed by a shift of the eyes to the location to which attention was initially captured. Note however that it is possible to direct attention to a location in space without the execution of a subsequent eye movement. It is assumed that attention and saccade programming are causally related, but a separate go-signal is required to trigger the saccade that has been programmed (e.g.,Deubel and Schneider 1996). Therefore, attention may move while the eyes remain fixated (e.g., Posner 1980).

Even though there appears to be a strong overlap between shifting attention and shifting the eyes (for a review, see Awh et al. 2006), it should be noted that there is one important difference. The eye movement system is basically an all-or-nothing system: the eyes can either go to one or to another location in space, but not to both at the same time. In other words: when there is competition between two objects, ultimately one object wins this competition and the eyes will go to that location. It should be realized that even though the eyes cannot go to two locations at the same time, the eyes can quickly switch between locations. Contrary to saccades, attention can be directed to two or more locations at the same time. For example, before an eye movement sequence of two saccades was made, Godijn and Theeuwes (2003) showed that spatial attention was directed to these two locations simultaneously just before the saccade was executed (see for a similar result Baldauf and Deubel 2008). Also, in conditions in which no eye movements have to be made, it is also possible to obtain a division of spatial attention between two non-consecutive locations (e.g., Kramer and Hahn 1995).

Over the past 20 years a great deal of research has been conducted to determine the relationship between (bottom-up and top-down) shifts of attention and saccades (Klein 1980; Posner and Petersen 1990). Since attention and saccades both have the goal of selecting the relevant portions of a visual scene, the idea that attention and saccades are related is intuitively appealing. In this view attention and saccades are related on the basis of their common function. That is, in order to further process and respond to an object, both orienting systems are typically directed to the same object, although in principle their focus may be dissociated. There are two viewpoints describing the relationship between attention and eye movements. According to one view, spatial attention constitutes a high level, supra-modal cognitive function that interacts with the low level, specialized sensory and motor processing systems only for the purpose of input and output and that can be functionally distinguished from them (Hunt and Kingstone 2003; Klein 1980; Posner and Petersen 1990). The alternative is the notion that spatial attention can be viewed as a direct product of these low-level processing systems and, more specifically, as a product of active interactions with the environment through eye movements. The influential premotor theory (Rizzolatti et al. 1987; Rizzolatti et al. 1994) posits that a shift of spatial attention involves all the steps necessary for making a saccade, except for the actual motor execution (see also Klein 1980; Klein and Pontefract 1994). A "grounding" of spatial attention in the oculomotor system predicts an interdependent relationship between covert and overt attentional orienting. In fact, there is quite some evidence indicating a close link between the covert-attention and oculomotor systems. Behavioural studies have shown that the allocation of attention affects saccade trajectories (Sheliga et al. 1994; Van der Stigchel et al. 2007; Van der Stigchel and Theeuwes 2007), have demonstrated a coupling between saccade preparation and spatial attention (Deubel and Schneider 1996; Hoffman 1986; Kowler et al. 1995) and have shown that the ability to make eye movements can affect covert attention (Craighero et al. 2004). In addition, neurophysiological studies have shown that sub threshold stimulation of several oculomotor structures, such as the frontal eye fields (FEF) and superior colliculus (SC), results in enhanced visual sensitivity at the corresponding retinotopic location (Cavanaugh and Wurtz 2004; Moore and Fallah 2001; Muller et al. 2005).

In a recent study we (Belopolsky and Theeuwes 2009) have made a distinction between the *shifting* of attention to a location and *maintaining* of attention at a location. The idea is that shifting of attention to a location results in an obligatory activation of a specific oculomotor program, which is consistent with the premotor theory (Rizzolatti et al. 1987). However, the voluntary maintenance of covert attention that occurs after attention is shifted is more flexibly (economically) related to the oculomotor system.

As noted eye movements may be directed in a top-down way to a location in space or alternatively may be captured by salient events in a bottom-up way. After the eyes have moved to a location, they may be disengaged quickly or slowly depending on the processing that will take place after the eyes arrived. According to the classic "immediacy assumption" which was developed on the basis of reading research (Just and Carpenter 1980), the fixation duration represents the time it takes to process the information. For reading, this implies that readers fixate high frequency words for a much shorter time duration than low frequency words, because in the later case more processing time is needed. For visual search, this implies that when an observer is searching for a particular target (say a red x), it will take longer to disengage the eyes from an object that resembles the target (say a red k) than from an object that does not resemble the target (e.g., a green O). Recently, Theeuwes and colleagues (Born et al. (2011); Mulckhuyse et al. 2009) addressed the issue of oculomotor disengagement in visual search studies. They showed that even though the initial capture of the eyes to a distractor was very much bottom-up (driven by the salience of the distractor) the subsequent oculomotor disengagement (i.e., how long does the eye remain fixated at the distractor) was very much under top-down control as fixation duration was determined by the amount of resemblance of the distractor to the target observers were looking for.

Evidence for Automatic Selection

Even though it is generally accepted that there is massive volitional top-down control of selection, our proposal is that (a) selection is not under volitional control but instead runs off in an automatic fashion without much, if any, conscious control and (b) during selection there is no conscious feedback regarding the selection processes. Consciousness regarding the volitional nature of these selection processes may be an emergent property that gives us the feeling that we are in control of what we select. Our claim is that this type of automatic selection runs off without much cortical processing and is not penetrable by volitional top-down set. An example of such processing is the well-known priming effect (Graf and Schacter 1985) which is defined as a nonconscious influence of past experience on current performance. Even though the classic priming experiments do not involve visual search (prime and target are typically presented at the same location), they provided compelling evidence that the occurrence of a prime may produce automatic facilitation as well as inhibition. Similarly Posner (Posner 1978) argued that a stimulus automatically may activate habitual pathways which may in turn result in enhanced processing of stimuli that share that pathway. Posner argued that the facilitation was automatic (Posner 1978). Subsequent inhibitory processes were associated with "conscious" or "intentional" attention. This latter claim is consistent with the notion that top-down processing can only have an effect relatively "late" in processing through recurrent feedback processing.

If one wants to argue that selection is automatic then it is important to define "automaticity". Automaticity is a concept with a long-standing history in psychology dating back to James (1890) and Wundt (1887). It is not the place here to provide a detailed review of research on automaticity (for a recent review see Moors and De Houwers 2006) but in general one speaks of automatic processing when processing takes place without much, if any, capacity. Typically, automaticity is associated with the resource theory suggesting that on the one hand, automatic processes draw minimal resources while on the other hand, non-automatic processing (cf. controlled processing) uses a lot of attentional resources. Another aspect that is prominent is the notion that automatic processes are hard to control intentionally. The most famous example is the Stroop task (Stroop 1935) in which participants have to attend to one aspect of the stimulus (the colour of the ink) while ignoring another aspect of the stimulus (the word). The classic finding is that people have trouble ignoring the written word as it interferes with the naming of the colour of the ink. On the basis of these and many other demonstrations it is argued that automatic processes cannot be controlled by the person.

Shiffrin and Schneider (1977) argue for a dual-mode model of information processing. Information processing is based on the activation of nodes from long-term memory which will be temporarily stored in short term memory. Processing is automatic when the activation in short term memory takes place on the basis of stimulus-input without much attentional demand. The strong version of this view claims that once started the stimulus-driven activation cannot be avoided. So called controlled (non-automatic) processes are under the control of the person and are established intentionally and volitionally by the person. Importantly, because this theory is mainly based on visual search experiments, Schneider and Shiffrin (1977; Shiffrin and Schneider 1977) claim that automatic processing is fast and occurs in parallel while controlled processing is slow, serial and effortful. Also, they claim that this automatic processing may be unconscious because the individual nodes are only active very briefly in short-term memory. On the other hand, control processes are conscious because the person has to volitionally activate the nodes in short term memory.

In recent years, it has become clear that the all-or-none distinction between automatic versus non-automatic processes may be incorrect (Bargh 1992; Logan 1985). For example, the Stroop effect that was generally considered to be the hallmark of automatic processing is reduced when spatial attention is directed away from the target (e.g., Treisman and Kahneman 1981). Even though Shiffrin and Schneider's theory (1977) is basically an all-or-none theory they allowed for some control by claiming that the initiation (the willingness to do a task) in under the person's control. However, once started the task runs to completion without any further top-down guidance. Interestingly for the current discussion, Shiffrin and Schneider distinguish two different mechanism of attentional allocation. One the one hand, a person can volitionally allocate spatial attention to a stimulus. One the other hand, attention may be automatically drawn to a stimulus as a result of "prior automatic relevance detection". This latter notion is obviously very similar to what was labelled in more recent years as attentional capture.

In the following sections we review the evidence for automatic selection. Fig. 1 provides an overview of our approach. Traditionally, salience (the box labelled bottom-up extraction) and the subsequent salience map are considered to be determined by the properties of the visual image. Salience is computed on the basis of the detection of locations whose local visual attributes significantly differ from the surrounding image attributes, along some dimension or combination of dimensions (Itti and Koch 2001). This approach dates back to the Feature Integration Theory (Treisman and Gelade 1980) which claimed that a number of simple visual feature

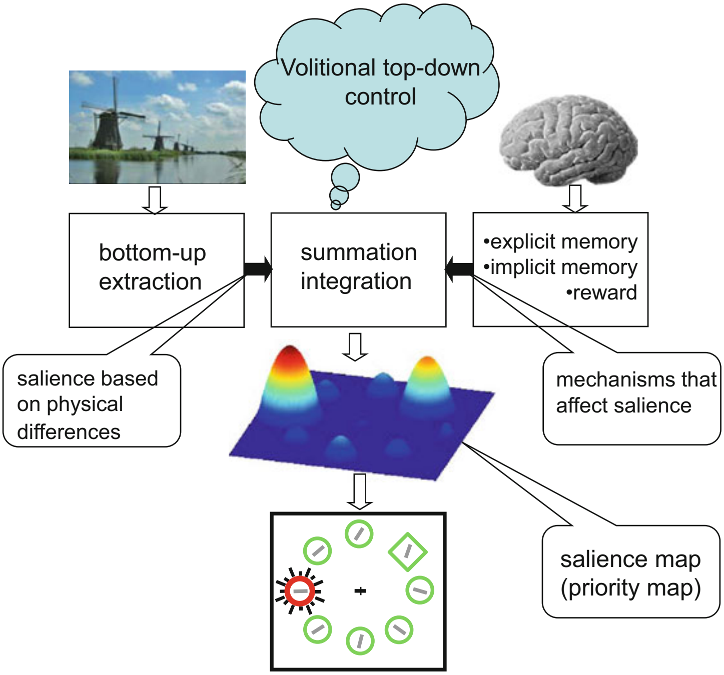


Fig. 1 Bottom-up extraction representing the raw salience of the objects in the environment interacts with internal representations at the observer. This interaction makes up the salience map (priority map) which drives visual selection in an automatic, winner take all fashion

dimensions are represented in the early stages of cortical visual processing such as colour, edge orientation, luminance, or motion direction. We review evidence that bottom-up salience plays a crucial role in visual selection, and argue that (most) studies that appear to provide evidence for volitional top-down selection are in fact demonstrations of automatic selection. In addition to the traditional bottom-up 'image attribute' salience which is linked to the physical properties of the objects in the environment, we argue that explicit and implicit memory and reward contingencies can alter the salience of the objects above and beyond their raw physical salience. In other words, selection may be driven in an automatic way by salience that goes beyond the classic basic feature properties (difference in colour, shape, luminance, etc.) of the stimuli in the environment.

Bottom-up Salience Based Selection

In the early 1990, we (Theeuwes 1991, 1992, 1994b) conducted a series of experiments investigating the extent to which observers are able to control visual selection processes. The basic question that Theeuwes addressed was whether one can select from the environment those stimuli needed for the task at hand even in conditions in which the stimulus that needs to be selected is less salient than stimuli that were irrelevant for the task. In those days this question was partly inspired by applied questions from the field of traffic where the question was posed whether a driver can select a relevant traffic sign when simultaneously a very salient distracting billboard screams for attention (for a discussion on the implication for driving see e.g., Theeuwes and Hagenzieker 1993).

When a salient object or event is selected even when the observer tries to select something else one speaks of attentional capture (Theeuwes 2010b). When not only attention is captured but the event or object also triggers a subsequent saccade to its location one typically speaks of oculomotor capture (Theeuwes et al. 1998; Theeuwes et al. 1999). It is important to note that it is not trivial to deduce true attentional capture. Even though it seems reasonable to assume that a unique feature singleton (such as a red element in a display of green elements) captures attention in a purely bottom-up way, such a claim is not necessarily correct. Indeed, when the feature singleton is also the element that observers are instructed to look for, one cannot determine whether this immediate selection of the feature singleton is the result of bottom-up or top-down control. As pointed out by Yantis and Egeth (1999), one can only speak of selection in a purely stimulus-driven fashion when the stimulus feature in question is completely task-irrelevant, so that there is no incentive for the observer to attend to it deliberately. As asserted by Yantis and Egeth (1999): "*If an object with such an attribute captures attention under these conditions, then and only then can that attribute be said to capture attention in a purely stimulus-driven fashion*" (p. 663).

Attentional Capture

Background Theeuwes (1991, 1992, 1994b) developed the so-called *additional singleton task* to investigate attentional control. The logic underlying this paradigm is simple: participants perform a visual search task in which two salient singletons are simultaneously present. One singleton is the target; the other singleton is a distractor. Participants are told that the irrelevant singleton is never the target, implying that there was no reason to attend to this item from a top-down point of view. This condition is then compared to a condition in which such an irrelevant singleton is not present. In this visual search task, observers search for one specific clearly defined salient singleton while another irrelevant singleton is simultaneously present. Figure 2 gives an example of a display. In one of the versions of this task, observers consistently searched throughout the whole experiment for a green diamond singleton. In the distractor condition, one of the green circles was made red, representing the colour distractor singleton. The crucial finding of the additional singleton search task is that reaction time (RT) in the condition in which a unique colour irrelevant distractor singleton was present (in this case the red circle) was higher than when such a distractor was not present (see Fig. 2).

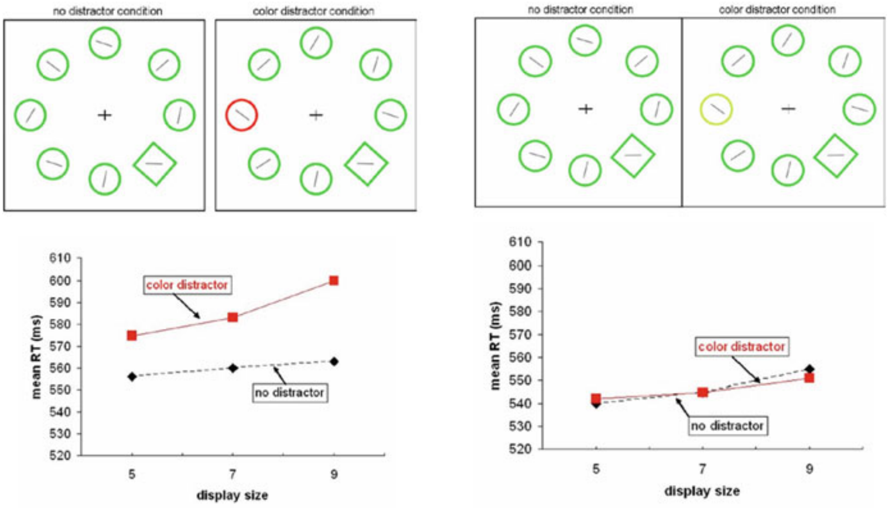


Fig. 2 Stimuli and data from Theeuwes (1992). Observers searched consistently for a shape singleton, a *green diamond*, presented among a variable number of *circles*. Observers responded to the orientation (horizontal or vertical) of the line segment presented within the green diamond. On the *left*: The *red irrelevant circle* captures attention and causes a Reaction Time (RT) increase because it was more salient than the green diamond (the target). On the *right*: Finding the shape singleton is not affected by the presence of the colour singleton because the colour singleton is now less salient than the target singleton (the *green diamond*). These results indicate that even though observers always search for a diamond singleton, top-down set cannot prevent the selection of the colour singleton. Selection appears to be completely controlled by the salience of the stimuli in the visual field. This result is taken as evidence for bottom-up attentional capture

It is also important to note that search functions are basically flat (see Fig. 2) indicating that the target popped out from the display. This is important because it implies that search is conducted in parallel across the visual display. Pop-out tasks have been implicated to subserve the early visual processing and single unit studies have shown the involvement of primary visual cortex in mediating bottom-up pop-out saliency computations (e.g., Nothdurft et al. 1999). Using pop-out search tasks makes it possible to determine the initial selection priorities (i.e., what grabs attention first?).

One important aspect of the experiments of Theeuwes (1994a, 1992, 1994b) that is often not taken into account is that the irrelevant singleton only causes an RT increase when the distractor is more salient than the target. When the colour distractor was made less salient (see Fig. 2; right panels), there was no measurable effect of its presence. In other words, if the target one is looking for is the most salient element, a less salient element does not affect performance. On the basis of these findings Theeuwes (1991, 1992, 1994b) postulated his notion of stimulus-driven capture, arguing that the bottom-up salience signal of the stimuli in the visual field determines the selection order. The increase in search time in conditions in which an irrelevant singleton was present was explained in terms of attentional capture.

Because the irrelevant colour singleton was selected exogenously (that is, captured spatial attention), it took more time before the target singleton could be selected and before a response could be emitted. Given the observation that selectivity completely depended on the relative salience of the target and the distractor singleton, it was argued early visual pre-attentive processing is only driven by bottom-up factors.

It is important to highlight particular aspects of the additional singleton task. First, due to the design of the experiments, observers never have an incentive to attend the colour distractor singleton. Because observers always search for a shape singleton, the colour singleton distractor is always task irrelevant. If one obtains capture by the colour singleton, it is obvious that this occurs independent of the intentions of the observer. Second, observers search for the green diamond target but respond to the orientation of the line segment inside of it. In such a compound search task (Duncan 1985) one is able to disentangle factors that affect the selection of the target from those affecting the response. Stated differently, observers search for one aspect (the diamond shape) but respond to another aspect (the line orientation inside the target). Because we use compound search, the response requirements remain the same over the various conditions ensuring that the RT effects caused by the presence of the colour distractor are due to perceptual interference and not to response interference. Third, the target and distractor singleton are always simultaneously present which guarantees that there is competition between the elements. According to the biased competition view of attention (Desimone and Duncan 1995), one can only see the effects of attention when there is competition between the elements. Mathot et al. (2010) showed that when a target and a distractor are presented sequentially, there is hardly any competition between a target and onset distractor singleton.

Note that in the original additional singleton paradigm (Fig. 2) and in other versions of this paradigm (the singleton cueing paradigm see Fig. 4) the line segment inside the target singleton that observers responded to was either horizontal or vertical while the line segments in the nontarget elements were slightly titled. Some have argued that this setup makes it possible for observers to directly search for the vertical or horizontal line segment while ignoring the circle and diamond shapes. However, control experiments showed that such a strategy is not available. If observers would use such a strategy, search becomes very slow and serial with search slopes of about 88 ms/item (see Experiment 1 and 2, Theeuwes 1991). Since the data (Fig. 2) indicate that in the additional singleton paradigm search is efficient and conducted in parallel (involving preattentive processing), it is clear that observers did not and possibly could not use the strategy.

Since its introduction in 1991, the basic findings of additional singleton task has been often replicated. For example, Bacon and Egeth (1994), replicated Theeuwes 1992 and demonstrated that it did not matter whether the distractors conditions were varied within or between blocks. Kumada (1999) examined between dimension (e.g., colour and orientation) and within dimension (e.g., orientation) interference in a simple and compound version and showed interference in the within dimension condition both in compound and simple search. In the between dimension condition, there was only interference in the compound search condition. Leber and Egeth (2006) used an extensive training scheme and replicated Theeuwes' basic findings.

Mounts (2000) used a display similar to those of Theeuwes (1992) and showed that the identification of a letter was slowed by the presence of an irrelevant colour singleton. Importantly (Mounts 2000) showed that the identification of the letter was slowest when it was located next to the irrelevant colour singleton (see also for a similar result Mathot et al. 2010). Kim and Cave (1999) employed the additional singleton search task in combination with a probe detection task and showed that at a 60 ms SOA the probe RT at the location of the distractor singleton was about 20 ms faster than at the target singleton location. At the 150 ms SOA however this pattern was reversed: the probe RT at the target location was about 15 ms faster than at the distractor location. It was concluded that early on at 60 ms after display onset–attention was first captured by uniquely coloured distractor, while soon thereafter (at 150 ms) the probe at the target singleton received more attentional activation. Geyer et al. (2008) showed that capture depended on the frequency of occurrence of the distractor singleton. It was shown that the less frequent a distractor the larger the interference effect Lu and Han (2009) showed that when the search task becomes more difficult (more serial) attentional capture is reduced or even absent (see also Proulx and Egeth 2006). Schübo (2009) showed large RT interference effects when observers searched for a shape singleton while a colour singleton was presented, but not the reversed. Dalton and Lavie (2007) reported an equivalent effect of attentional capture in the auditory domain: Irrelevant high intensity singletons interfered with an auditory search task when the target itself was also a feature singleton. Others have demonstrated similar effects using measures related to d-prime (Theeuwes and Chen 2005; Theeuwes et al. 2004), saccadic eye movements (Godijn and Theeuwes 2002b; Ludwig and Gilchrist 2002; Mulckhuyse et al. 2009; Theeuwes et al. 1998; Theeuwes et al. 1999); and hand movements (Hunt et al. 2007).

Automatic Attentional Selection? On the basis of the additional singleton paradigm, Theeuwes (1994a, 1992, 1994b) claimed that spatial attention was captured in a bottom-up way to the location of the item having the highest salience. Even though some have challenged these claims (Bacon and Egeth 1994; Leber and Egeth 2006) there is evidence from a whole host of behavioural (Theeuwes 1995; Theeuwes et al. 2000; van Zoest et al. 2004b) and ERP studies (Hickey et al. 2006) that strongly suggest that attention is in fact that captured spatially to the location of the salient singleton (for an overview see Theeuwes 2010a, 2010b). The question for the current discussion is whether this selection is in fact automatic.

Several aspects of the task suggest that attentional capture is indeed automatic. Consistent with Shiffrin and Schneider's theory, we do find that search in the additional singleton task is fast and occurs in parallel as evidenced by basically flat search functions (see Fig. 2). We also claim that under these circumstances search is indeed completely stimulus-driven, and that the activation by the irrelevant colour singleton cannot be avoided resulting in a shift of spatial attention to the location that generates the highest activation.

It is important to realize that according to Shiffrin and Schneider (1977) controlled attentional processes are slow, serial and effortful. This implies that in circumstances where search is slow and serial, a salient singleton may not capture attention. Research conducted in the late 1980s by Yantis and colleagues (Jonides and Yantis 1988; Yantis and Egeth 1999; Yantis and Jonides 1984) is consistent with this notion (see also Theeuwes 1990). Yantis and colleagues adopted a visual search task, such that the target of search was a non-singleton letter (search for a target letter among other letters). This type of search was slow, effortful and serial as search times increased linearly with the number of elements present in the display. In each display, there was always one salient element (for example an element with a unique colour) and the question was whether observers would start searching at the salient element (i.e., the element with the unique colour). Since the salient element was the target at chance level, there was no incentive to deliberately start searching at the salient singleton. Jonides and Yantis (1988) showed that observers did not start searching at the salient element in the display. When the unique element happened to be the target (e.g., an element with a unique colour or unique luminance), the search slopes were basically the same as in the condition in which a non-unique element was the target. It was concluded that salient static singletons are treated in the same way as other non-salient elements in the visual field. Uniqueness in colour or luminance is not sufficient to capture attention when it is irrelevant to the top-down goal.

The findings of Jonides and Yantis (1988) fit perfectly with the distinction between automatic and controlled processes as described by Shiffrin and Schneider (1977). Only if search is slow and serial there appears to be attentional control in the sense that attention is not captured by irrelevant salient signals. Note however that Jonides and Yantis (1988) showed that elements appearing with an abrupt onset grab attention even when search is slow and serial (see also Theeuwes 1990 for motion). Even though it is important to establish that when search is slow and serial, static singletons do not capture attention, one has to explain why in some circumstances static singletons do capture attention (as in Theeuwes 1991, 1992) while in other circumstances, a static singleton does not capture attention (as in Jonides and Yantis 1988).

A possible solution for this apparent contradiction was first offered by Theeuwes (1994a) who argued that *"top-down control over visual selection can be accomplished by endogenously varying the spatial attentional window"* (p. 436) (see also Theeuwes 2004, 2010a). The idea is that an attentional window adopted by observers could be one of the factors explaining why salient colour singletons fail to capture attention in some studies (as in Jonides and Yantis 1988) while in other studies they do capture attention (as in Theeuwes 1992). Belopolsky et al. (2007) directly tested this idea in a visual search task which resembled the original Jonides and Yantis (1988) paradigm. As in Jonides and Yantis, participants had to serially search for a target letter, which had a unique colour at chance level. In this particular study, the size of the attentional window was manipulated by asking participants to detect either a global (diffuse attention) or a local shape (focused attention) before starting the search for a non-singleton target. The results showed that when attention was initially focused at the centre (focused attention condition) the salient colour singleton was examined just as frequently as the other elements in the display. This result was similar to the classic finding of Jonides and Yantis (1988). However, when attention was initially diffused over the global stimulus arrangement (diffuse attention condition), attention was captured frequently which was evidenced by faster responses and a significantly reduced search slope when the coloured element happened to be the target. It was concluded that the size of the attentional window plays a crucial role in attentional capture: when the window is wide salient stimuli capture attention, but when it is small salient stimuli falling outside of the window can be ignored (see also Hernandez et al. 2010).

In a more recent study Belopolsky and Theeuwes (2010) tested the idea of the attentional window using the classic additional singleton task of Theeuwes (1992). They found that when observers were in a diffuse attentional state, the classic attentional capture interference effect was found: the presence of an irrelevant colour singleton slowed search for a shape singleton (as in Fig. 2; left panels). However, this very same capture effect was abolished when just before the presentation of the display attention was in a focused state. If attention was not spread over the display, but focused in the centre, the presence of an irrelevant singleton no longer captured attention. Belopolsky and Theeuwes (2010) concluded that the attentional window is a determining factor in the occurrence of attentional capture.

Given these studies, one can ask the question whether the capture of attention by salient singletons is an automatic process. If one adheres a very stringent definition of automaticity one can argue that it is not an automatic process. Indeed, it appears that attentional control is possible by endogenously directing attention to a restricted area within the visual field. In this sense attentional capture does not fulfil what is known as the "intentionality criterion of automaticity" which states that *"automatic processes are under the control of stimulation rather than under control of the intentions (strategies, expectancies, plans) of the person"* (Neumann 1984, p. 258). So in the very strict sense of automaticity, attentional capture is not an automatic process because some attentional control seems to be possible. Consistent with Shiffrin and Schneider (1977) the actual initiation of the task (and whether one spreads attention or not) is very much under volitional control. However, we claim that once the process has started (once a person is engaged in visual search), the task runs to a completion without any further top-down guidance. It is feasible that there is no control because once a task is set into motion there is no conscious feedback regarding the selection processes. One possible way to gain control is by slowing down the task execution dramatically or by restricting the attentional window.

As argued, the direction of spatial attention to a restricted area in visual space is the only top-down control that can prevent automatic attentional capture. This implies that when attention is spread across the display there is no top-down control implying that the calculation of a local feature difference (cf. salience) occurs in a bottom-up fashion. This calculation is not penetrable in a top-down, volitional way. Once the most salient singleton is selected, its identity becomes available and then only then top-down knowledge (such as the fact that the observer is looking for a red target) will play a crucial role. If the automatically selected feature difference signal is the target singleton, a response can be given. If it is not the target singleton, top-down down processing (that occurs after the item has been selected) allows a quick and fast disengagement of attention from the location having the highest salience signal.

Our claim is that while the size of the attentional window is under top-down control, within the attentional window processes runs off automatically.

Our notion that the distribution of spatial attention across the visual field is one of the top-down mechanism that can control the occurrence of automatic processes fits very well with results from the Stroop effect. The Stroop effect is generally considered to be the prime example of automatic processing. Many researchers claim that the word processing involved in Stroop is automatic and occurs without intent and cannot be prevented (Macleod 1991). However, it has been shown that a narrow attentional focus reduces or even eliminates the Stroop effect. For example, Besner and Stoltz (1999) precued a single letter position to narrow the focus of attention before the word was presented. The results showed that this focusing of attention prevented the activation of the word recognition system. These findings suggest that Stroop just like attentional capture is an automatic process that runs off automatic unless the input is restricted by focusing spatial attention to a limited spatial area.

To sum up: when attention is divided across the visual field, attentional selection runs off in an automatic way, selecting in turn the items according to their raw salience. When attention is focused on a restricted area, there is no pre-attentive extraction of the pop-out features outside the attended area. When engaged in this type of controlled processing (Schneider and Shiffrin 1977), there is no automatic attentional capture anymore but search is relatively slow, serial and effortful.

Oculomotor Capture

Background Theeuwes et al. (1998; 1999) developed the so-called oculomotor capture task, a task which is comparable to the additional singleton task. Instead of inferring capture on the basis of a manual RT, capture is reflected by erroneous eye movements toward the irrelevant distractor item. In this task, observers view displays containing six equi-spaced grey circles presented on an imaginary circle around a central fixation point. After 1 s all of the circles but one change their colour to red. Participants have the explicit instruction to make a saccade towards the only grey element in the display. On some trials, an irrelevant red circle, presented with an abrupt onset, is added to the display. In Theeuwes et al. (1999) a control condition was used in which an additional non-onset distractor was added to the display at the beginning of the trial. In Theeuwes et al. (1998) there was no additional non-onset distractor on trials without an onset. Both studies showed that when no item was added to the display, observers generated correct saccades that went directly towards the uniquely coloured circle. However, on those trials on which an item was added to the display, the eyes went in the direction of the onsetting item in about 30–40 % of the trials, stopped briefly, and then went on to the target. Figure 3 shows the results. The graphs on the left side depict the control condition without the onset; the graphs on the right side depict the condition in which an onset was presented. Note that in the condition with the onset, the eyes often went to the onset. This occurred even when the onset appeared on the opposite side of the target circle.

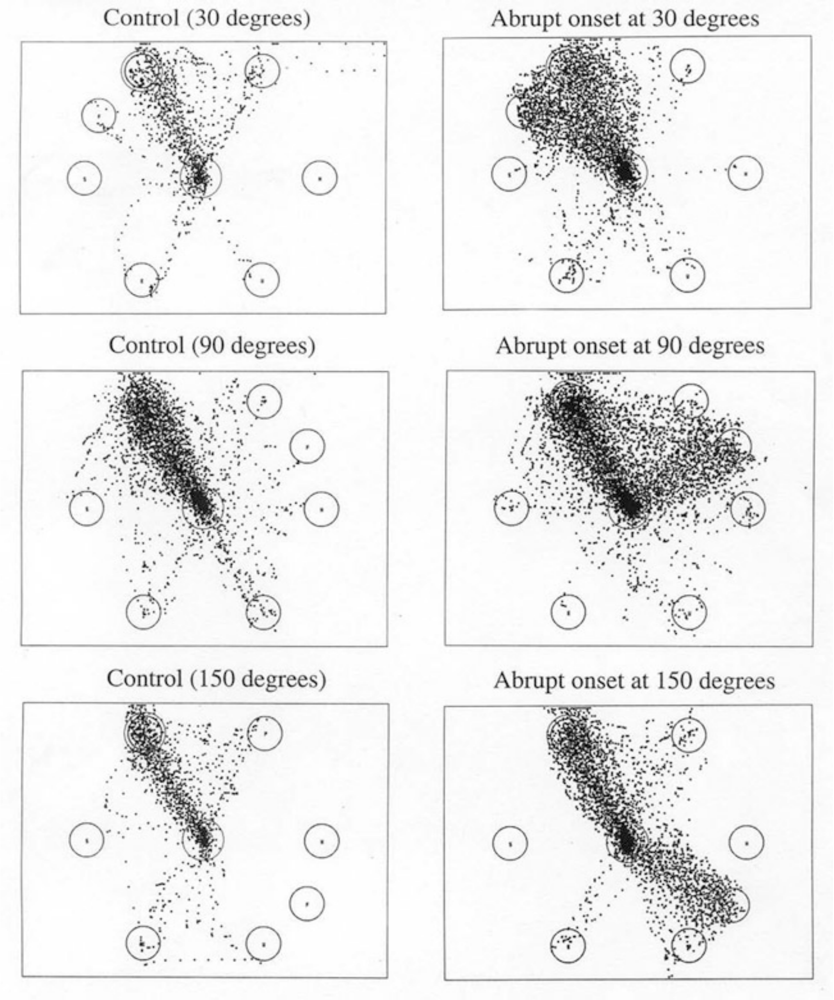


Fig. 3 Oculomotor capture. Eye movement behaviour in the condition in which an abrupt onset distractor was presented simultaneously with the target. The results are collapsed over all eight participants and normalized with respect to the position of target and onset. Sample points (every 4 ms) are only taken from the first saccade. *Left panels*: Eye movement behaviour in the control condition in which no abrupt onset was presented. *Right panels*: Eye movement behaviour in the condition in which an abrupt onset was presented; Either close to the target (*top*) somewhat away from the target (*middle*) and or at the opposite side from the target (*bottom*). (Data from Theeuwes et al. 1999)

Since participants were required to execute a saccade to the uniquely coloured elements they had a clear top-down goal. However, despite this clear top-down goal, on about 30–40 % of the trials on which an onset was presented a saccade was executed toward the abrupt onset (see Fig. 3). These saccades are considered to be genuinely bottom-up, since they are completely irrelevant for the task at hand and were executed even though there was an explicit instruction to move the eyes to another location. The results of the oculomotor capture task has been replicated in various variations many times (e.g., Belopolsky et al. 2008; Born et al. 2011; Godijn and Theeuwes 2002b, 2003; Hunt et al. 2007; Wu and Remington 2003).

It is important to note that in the oculomotor capture task, observers are typically not aware of the fact that their eyes move to the location of the distractor (Theeuwes et al. 1998). For example, in Theeuwes et al. (1998) after the experiment observers were asked whether the onset affected their eye movement behaviour. Most observers did not recall that there was a abrupt onset and none of them thought their eye movement behaviour was affected by the onset. Similarly, in Belopolsky et al. (2008) observers had to indicate after each trial whether they moved their eyes directly to the target or not. The results showed that on trials where the eyes were misguided (19 %) observers only reported this correctly in about 5 % of the trials. This study clearly shows that people have only very limited knowledge about where their eyes went even in conditions in which they know they have to report where their eyes went on each and every trial. Feedback about our eye movement behaviour turns out to be very limited.

Automatic Oculomotor Selection? An obvious question is whether oculomotor system is automatically driven by the presentation of the abrupt onset distractor. Given the fact that most observers are not aware that they actually made an erroneous saccade to the distractor there may indeed reasons to argue that oculomotor capture occurred automatically. Tse et al. (2002) directly tested a strong version of the automaticity hypothesis. They adapted the original oculomotor capture paradigm of Theeuwes et al. but instead of asking observers to make a saccade to the target, they asked them to remain fixated in the middle, at the centre fixation point. Maybe not surprising, the presentation of the abrupt onset had no effect on eye movements, as observers remained fixated at the central fixation point. Tse et al. concluded that the oculomotor system does not automatically react to stimuli presented in the environment. Only when the system has been preset to make a saccade, an abrupt onset can grab the eye and disrupt the saccade generation process. Obviously this result is not unexpected because in every-day life we need to be able to fixate and identify an object fixation without being distracted by abrupt onsets. It is concluded that abrupt onsets do affect the oculomotor system when the observer has the willingness and intention to make a saccade. When the goal is to remain fixated, observers have no trouble doing so even if there are very large abrupt onsets. The results of Tse at al. are consistent with the notion that focusing attention to a restricted area (i.e., focusing the attentional window) prevents capture of attention and the eyes by events that occur outside the restricted focused area.

The question whether such an abrupt onset capture attention (but not the eyes) was not addressed in the study of Tse et al. (2003), but it is feasible that there is attentional capture without oculomotor capture. For example, Godijn and Theeuwes (2002a) showed that even when the eyes did not go to the location of an abrupt onset, orienting back to that location was delayed (also known as Inhibition of Return, Klein 2000) suggesting that even though the eyes did not go to the location of the onset, attention did go there.

Theeuwes et al. (1998) also showed that the execution of the saccade to the abrupt onset is not fully automatic. In their experiment 2, Theeuwes et al. precued the location of the upcoming target before display onset allowing observers to focus attention to the location where the target would be presented. The results showed that precuing had a dramatic effect on eye movement behaviour: in the condition in which the location was precued, oculomotor capture was eliminated. The behaviour was virtually identical to a condition in which no abrupt onset was presented. Consistent with the earlier discussed notion of the attentional window, this study (and that of Tse et al. 2002) shows that directing spatial attention to a restricted area in visual space prevents the eyes from being captured by an abrupt onset. Again, space appears the only feature that allows control over processes that otherwise run off automatically.

It is important to note that in one aspect oculomotor capture is quite different than attentional capture. In a study by Theeuwes et al. (2003) it was shown that in some conditions one does get attentional capture (increase in RT) without a oculomotor capture. Theeuwes et al. adapted the traditional additional singleton task for eye movements such that observers had to make a saccade to a green circle and ignore the irrelevant colour singleton. The results showed that when the colour singleton remained the same across trials, there was an attentional capture effect of about 20 ms but no oculomotor capture (the eyes did not go to the location of the colour singleton). However, when the colour of the distractor changed from trial to trial (as in the Theeuwes 1991 task) there was both strong attentional (about 200 ms RT increase) and oculomotor capture (in 38 % of the trials the eyes went to the colour singleton). It is likely that there was such strong capture in this condition because selection was entirely driven by relatively automatic bottom-up priming mechanisms (see also Pinto et al. 2005). Wu and Remington (Wu and Remington 2003) also showed that a reflexive shift of attention does not necessarily initiate the execution of a saccade. Even though we never tested it directly, it is likely that a shift of spatial attention to the location of the distractor is only followed by a subsequent eye movement when attention is long enough at the location of the distractor. Only when the time that attention resides at the distractor location is longer than a critical duration will a saccade be launched. If the distractor does not look like the target at all, attention will stay only very briefly at the location of the target, possibly not long enough to launched a saccade (see Born et al. 2011, for a detailed discussion).

To sum up: the eyes do not move automatically to those things that are salient. Clearly this would make it impossible to operate, read and navigate in the world. However, when observers have the intention to move the eyes, salient events such as onsets or movement may grab our eyes in an automatic fashion. Similar to attentional capture, when spatial attention is focused on to a particular limited spatial area, there will be no eye movements to salient events outside this area. Whether a shift of attention ultimately results in a shift of the eyes depends on how long attention resides at a location after it has been shifted there. These findings suggest that just like attentional capture, oculomotor capture is an automatic process that runs off without any control unless observers focus their attention to a restricted area.

Selection and Implicit Memory (Priming)

Background Priming is a well-known phenomenon and refers to a nonconscious influence of past experience on current performance. Priming represents an example of what is referred to as implicit memory or nondeclarative memory (Graf and Schacter 1985). In classic priming experiments the prime and target are usually presented at the same spatial location. In other words, visual search and selection do not play much of a role. However, in a pioneering study of Maljkovic and Nakayama (1994) a similar effect was shown in visual search labelled as 'priming of pop-out' demonstrating that what was selected on a given trial (say a red circle) would automatically be selected on the next trial. More importantly, it was shown that this automatic selection tendency could not be counteracted by top-down volitional set. Even when a target on a given trial was 100 % predictable (i.e., target definition changed in an AABBAABBAA… manner), volitional knowledge-based expectations could not modulate feature-specific intertrial effects. In another study (Kristjánsson et al. 2002) a similar effect was shown, and Kristjansson et al. argued that "the role of priming in visual search is underestimated in current theories of visual search and that differences in search times often attributed to top-down guidance may instead reflect the benefits of priming" (p. 37).

Recently, we (Theeuwes et al. 2006; Theeuwes and Van der Burg 2007) showed a similar finding using another procedure. Instead of looking at intertrial effects, we presented a cue in de centre of the display before the start of the trial. Observers either searched for a red circle or a green diamond and responded to the orientation of the line segment inside of the singleton (as is always done in the additional singleton task). In some conditions, we presented a verbal label (e.g., we presented the word "green" in the centre of the display) which indicated with 80 % validity the feature of the upcoming target. The cue telling what to select on the upcoming target was presented in total for about 1.5 s which gave observers ample opportunity to prepare for the upcoming target. Importantly, providing this verbal information regarding the relevant feature of the upcoming target had no effect at all at selection efficiency (see also Theeuwes and Van der Burg 2007). The results indicate that whether one "knows" whether the target is green or whether it is a diamond did not improve performance.

However, when the actual object of search was used as a cue predicting the identity of the target (see Fig. 4), there was a reliable validity effect. As can be seen in Fig. 4 the cue had an overall validity of 80 % (in 80 % of the trial the cue matched the target) observers were slightly faster than when the cue did not match the target (Fig. 4; dotted line). On the face of it, this finding appears to indicate that top-down set does have an influence on selection efficiency. Such a finding is hard to reconcile with our notion that selection runs off in a more or less automatic fashion. However,

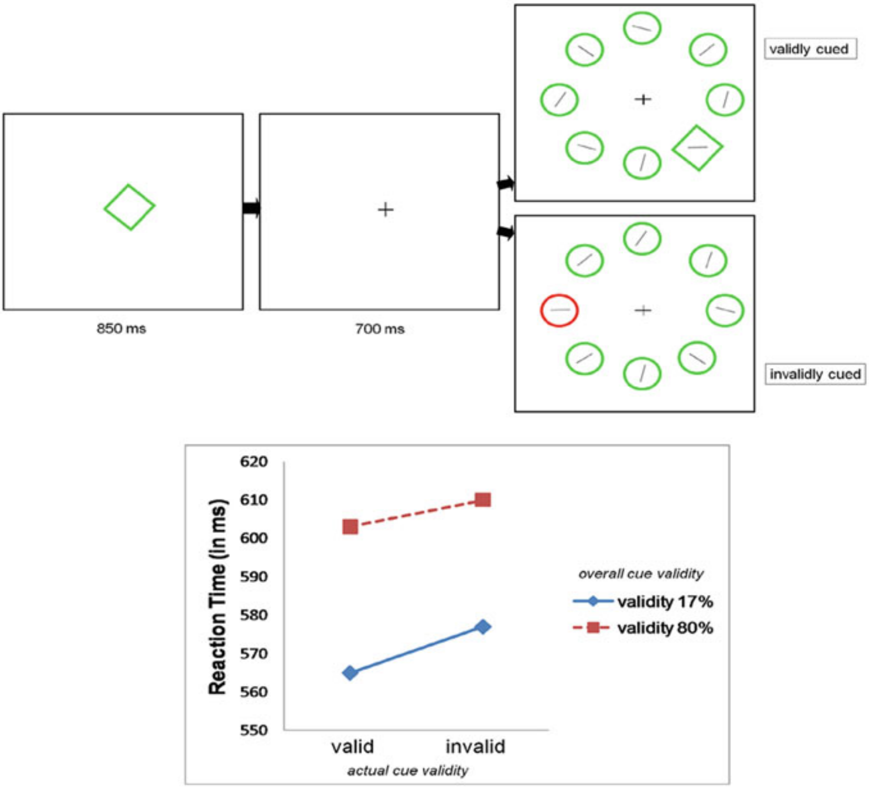


Fig. 4 Stimuli and data from Theeuwes et al. (2006). Observers responded to the orientation of the line segment inside the singleton. The singleton containing the target line segment was either a *diamond* (shape singleton) or a *red circle* (colour singleton). The cue (which was identical to the target singleton) indicated with either 80 % or 17 % validity the target on the upcoming trial. The RT data show that a valid cue speeded up responding suggesting that seeing the cue before the display improved selection efficiency. However, since the overall cue validity did not modulate this effect, the results suggest that the effects are completely due to an automatic priming which is not penetrable by volitional top-down control

we replicated the experiment with an overall cue validity of 17 %. For example, in this experiment it implied that when a green diamond was presented as a cue there was an 83 % chance that the target would be a red circle. Also, when a red circle would be presented as a cue, there was an 83 % chance that a diamond would be presented as a target. Even though the cue was *counter*-predictive, and observers were told about it, it still had a reliable effect on selection. For example, when a green diamond was presented as a cue and the target happened to be a green diamond (even though this occurred only on 17 % of the trials) observers were faster than when the cue was a green diamond and the target was a red circle (even though this occurred on the majority of the trials). The same held for when the cue was a red circle. Notably, there was no statistical difference between the 80 % and 17 % validity conditions, suggesting that knowing that the cue had predictive value, did not and possibly could not, alter selection. We explained these findings in terms of priming. Processing the cue before the presentation of the search display facilitates the processing of the target (when it matches) independent of its validity. This study shows that volitional control (i.e., preparing for a particular target) had no effect; instead selection was driven by bottom-up priming from the cue (Pinto et al. 2005) for a similar result). On the basis of these findings we suggested that priming runs off automatically and is not penetrable by volitional control (Theeuwes and Van der Burg 2007).

In another study we addressed this issue again but now there were two salient singletons simultaneously present (Theeuwes and Van der Burg 2011). Observers viewed displays consisting of seven gray circles in which one colour singleton was the target (say a red circle) while the other colour singleton was the distractor (e.g. a green circle). Before each trial, observers received a cue telling them what the target would be on the next trials. The cue was 100 % valid. For example, the word "red" would be shown telling the observer that the target singleton was the red singleton in the display. As before observers responded to the line segment inside the target singleton. The results showed that observers could not attend exclusively to the target colour singleton (the one indicated by the 100 % valid cue). The irrelevant colour singleton captured attention. Only when the colour of the target singleton happened to remain the same from one trial to the next, selection was perfect and attentional capture could be prevented. This effect was again thought to be the result of passive automatic intertrial priming.

It is important to realize that the type of priming that we discuss here has nothing to do with response priming or the buildup of automatic associations between stimuli and response tendencies. In the classic priming of pop-out paradigms (Maljkovic and Nakayama 1994) as well as in our cueing tasks (Theeuwes et al. 2006; Theeuwes and Van der Burg 2007) the response is completely different from what participants search for. Priming in our paradigm represents the speed with which the item can be selected not the speed with which the response can be emitted (see Theeuwes and Van der Burg 2007).

Automatic Selection due to (Intertrial) Priming? The question is whether the (intertrial) priming results in an automatic selection of that feature on the next trial. The data of our visual experiments (Theeuwes and Van der Burg 2007, 2011; Theeuwes et al. 2006) present a strong case for automatic selection: selection takes place even when observers actively tries to attend to the other feature. One way to explain priming in these visual search experiments is to assume that it is easier to attend to a feature that just has been attended. Such facilitation may occur because following target selection on a given trial, activation of the target feature may persist to the next trial and thereby speeds the selection of the repeated feature target.

From a theoretical level, Theeuwes (2010a, 2010b) argued that priming may change the salience of a stimulus such that for example a red stimulus that is primed appears to be more salient than a red stimulus that is not primed. Even though physically it is the same stimulus, it appears to be more salient because it was processed just before. There is in fact evidence for such a conception. Desimone (1996) suggested that repeated processing of a stimulus produces a "sharpening" of its cortical representation, possibly making it more salient within its environment. Recent evidence (Bichot and Schall 2002) showed an increased activity at a neural level as evident from increased activity in the frontal eye fields (FEF) while performing a priming of pop-out task. The results indicated an increased firing rate for repeated targets, providing evidence that the representation of salience in the FEF was increased due to target priming. The FEF is a region that has been implicated to be the neural substrate of the salience map (Thompson et al. 2005).

This brings us to the notion that the salience of a stimulus is not solely defined by the physical appearance of a stimulus in the outside world, but depends on its representation in the salience map (see Fig. 1). As noted, the processing of a stimulus leads to a change in representation of that stimulus in the salience map and this change occurs independently of top-down intentions. To appreciate what this implies: if observers process an object with a red colour on a given trial, on the next upcoming trial, observers will be biased to process red objects even when they are told to look for an object with a different colour (Theeuwes et al. 2006). Processing the colour red has sharpened the cortical representation of red such that on the next occasion (i.e., the next trial) when this colour is encountered again, it appears more salient. Because it is more salient, it is more likely to be selected and this occurs independent of the intentions of the observer. In this sense, bottom-up processing is considered to be automatic and passive, not sensitive to top-down set.

There is compelling evidence that indicates that priming takes place very early in processing. Olivers and Hickey (2010) showed that intertrial priming results in latency shifts and amplitude differences in the P1 component of the EEG signal, a signal that is seen 80–130 ms following display onset. Obviously since priming affects visual processing so early, it is unlikely the result of top-down processing. Because priming has already an effect during the first feedforward sweep of processing (<150 ms) one has to conclude that priming is automatic, passive and takes place without top-down control.

The discussion about priming is somewhat muddled because priming has been associated with top-down control. For example, Wolfe and colleagues (Wolfe et al. 2003) argued that priming is an example of implicit top-down guidance. Others associated priming with contingent capture (Ansorge et al. 2010), with motivational factors related to reward (Kristjansson 2010) or to the so-called feature search mode (Egeth et al. 2010). The argument is often that because priming is the result of prior history with a particular stimulus, it has to be the result of some top-down processing. Our claim is the opposite; the processing of the stimulus changes the representation of that stimulus above and beyond its physical appearance. The effects of priming on visual search are automatic and cannot be counteracted by top-down control.

The proposition that search runs off in an automatic fashion once it is set in motion is also consistent with recent experiments from our lab using the well-known spatial precueing paradigm of Folk and colleagues (e.g., Folk et al. 1992; Folk et al. 1994). In the original Folk et al. paradigm a cue display precedes the search display which consisted of either a colour or an onset singleton. Observers are typically required to identify the unique element. In the colour display, the target is red while the other three elements are white. In the onset display, only one element is presented, and so the target was characterized as being the only element presented with an abrupt onset. Immediately preceding the target display at an SOA of 150 ms, a cue display is presented: this cue display either consists of a colour cue (in which one location was surrounded by red dots and the other three locations were surrounded by white dots) or an onset cue (in which one location was surrounded by an abrupt onset of white dots and the remaining locations remained empty). Importantly, in all Folk et al. experiments, observers consistently search for one particular type of target (colour or onset singleton) throughout a block of trials. On the basis of these experiments, Folk et al. formulated the contingent capture hypothesis which states that selection is always under volitional top-down control of the observer.

In a recent study we (Belopolsky et al. 2010) have shown that what is known as contingent capture may in fact be the result of intertrial priming. The crucial point is that in all Folk et al.'s experiments (and related contingent capture studies) the target observers are looking for remains the same over a block of trials. Because it remains the same one will obtain strong intertrial effect which may appear and have been interpreted as being a form of top-down control. Belopolsky et al. (2010) used exactly the same spatial cueing paradigm as Folk et al. (1992). Rather than keeping the target fixed over a whole block of trials (as was originally done with contingent capture experiments), observers had to adopt a top-down set before the start of each trial. In other words, observers were cued at the beginning of each trial to either look for a unique colour or the unique onset. If, as claimed by the contingent capture hypothesis, top-down attentional set determines which property captures attention, then one would expect that only properties that match the top-down set would capture attention. Belopolsky et al. showed that even though participants knew what the target would be on the upcoming trial, both relevant and irrelevant properties captured attention. In other words, there was no sign of contingent capture, instead both the relevant cue that matched the target as well as the irrelevant cue, captured attention.

More importantly for the present discussion the Belopolsky et al. (2010) study also showed that when the target on the current trial was unknown, the target on the previous trial had a large influence on which cue property would capture attention on the current trial. For example, if the target on the previous trial was an abrupt onset, only onset cues captured attention on the following trial, whereas the colour cues were ignored. A similar trend was also found for the colour targets on the previous trial, with colour cues on the current trial capturing attention and onset cues capturing attention somewhat less. These results demonstrate that contingent capture is driven by the target properties encountered on the previous trial, instead of a top-down set for a particular target property. As noted, in basically all contingent capture experiments the target remained constant over a whole block of trials. These findings suggest that what is known as a prime example of top-down attention control over selection (contingent capture hypothesis) may in fact be completely driven by automatic priming effects which cannot be affected by top-down set.

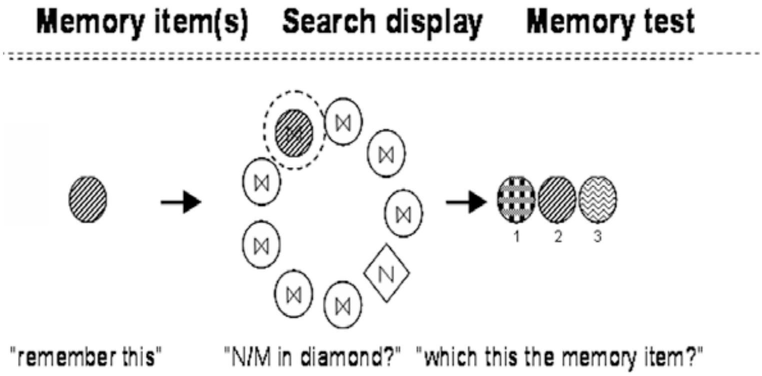


Fig. 5 The main procedure of Olivers et al. (2006). Participants had to memorize the colour of a circle. When searching for the shape singleton (the *diamond*) an irrelevant colour singleton was present. The colour of the singleton could match the colour kept in working memory. After search participants received a memory test in which they had to indicate the colour they had kept in working memory. The results provide evidence for automatic memory-driven capture: when the colour of the irrelevant singleton matched the colour held in memory there was more capture than when it did not match (from Olivers et al. 2006)

In summary: we have shown here that priming plays a key role in visual search. We argued priming sharpens the cortical representation of stimulus properties such that on the next occasion when a stimulus with the same property (e.g., colour, shape, etc.) is encountered again, it appears more salient. Because it is more salient, it is more likely to be selected and this occurs automatic, independent of the intentions of the observer. Priming is considered to be automatic and passive, not sensitive to volitional control.

Selection and Explicit Memory

Background Priming is often considered to represent implicit memory because it affects behaviour while observers are not aware of it and did not actively prepare for it. Explicit memory is a conscious act: observers actively try to keep an item active in memory. Recently, we (Olivers et al. 2006) studied explicit memory in relation to visual search. We addressed the question whether an object held in working memory would capture attention in an automatic fashion. In this study, we used the classic additional singleton paradigm (see Fig. 2) and tested whether a singleton that is kept in working memory would cause more capture than a singleton not kept in memory. In Olivers et al. (2006) observers were asked to remember a particular colour (red, green, blue or yellow). At the end of the trial, their memory was tested by asking them to choose the original colour from a set of three alternatives (see Fig. 5). We used two versions of the memory task. In what is called the “more verbal” version, the memory test consisted of easily distinguishable alternatives for which verbal labels are readily available, for example red, green, and blue. For this type of memory, one can use the verbal label (e.g., green or red) without any effort in trying to create a visual memory of the exact shade of red. In contrast, in the "more visual" version, the to-be-remembered colour had to be distinguished from highly similar colours from the same category. For example, a particular shade of red had to be distinguished from other shades of red. In this condition, we assumed that observers would use their visual working memory.

After memorizing the colour for a few seconds, participants had to search for a grey diamond among grey circles. Participants responded to the identity of the letter presented inside the diamond. On many trials, however, one of the distractors carried a unique colour. The important finding here was that the interference was stronger for distractors that matched the content of memory than for unrelated colour distractors. Another important finding was that this was only the case for the "more visual" memory condition. In the "more verbal" condition there was no effect of the relationship between the visual distractor and the contents of memory. Note that participants had no reason to attend to the distractor: It only interfered with the goal of responding to the gray diamond. Thus, these results are consistent with the idea that visual working memory and visual attention share the same content. Moreover, follow-up experiments excluded a number of alternative explanations in terms of implicit perceptual priming, perceptual encoding, strategic memory updating, and delayed attentional disengagement (see Olivers 2009). In one of the experiments, we also used an eye movement version of the task. Similar to the findings obtained with the attentional version of this task we showed that keeping an item in memory causes more eye movements towards an object relative to a condition in which that very same object was not kept in memory. The results show that "things in memory" that are not relevant for the search task at hand will generate both automatic attentional and oculomotor capture.

Automatic Selection due to Explicit Memory? The question for the current discussion is whether this type of selection is automatic. Obviously, storing information into working memory is a volitional act, and will not occur automatically. However, once the object is stored in memory, and observers engage in a visual search task in which this object (which is irrelevant for the search task) is present, attention and eye movement may be automatically drawn to this object kept in working memory. In this sense, it is feasible that keeping an object in working memory may increase its salience in a similar way as occurs in intertrial priming. Because its salience is increased, attention will be drawn to it in an automatic fashion.

In addition to our study (Olivers et al. 2006) there are several other studies that seem to provide evidence for the notion that storing an object in working memory should automatically alter the processing of that object when it appears in the external world. For example, in one study observers were asked to form a mental image of a picture just before a sequence of pictures was presented (Pashler and Shiu 1999). While keeping this mental image in memory observers performed an RSVP task looking for a target digit. When in half of the trials, the imagined picture appeared in the sequence just before the target digit, performance in detecting the target digit suffered. Pasher and Shui (1999) argued that the formation of the mental image caused the subsequent presentations of that image to capture attention.

Downing (2000) reported a related finding. Observers had to memorize an object during a period of 3.5 s. During this period, two objects were presented on either side of the fixation point, one of which matched the item in working memory. Observers responded as fast as possible to a probe, occurring on top of one of the objects. In conditions in which the probe appeared on top of the item that was stored in memory, observers were faster than when it appeared on top of the object that was not in memory. Consistent with Olivers et al. (2006), these results provide compelling evidence that when an object in the environment matches the one held in working memory, it captures attention. Theeuwes, Kramer and Irwin (2011) pushed this idea a bit further. In this experiment, observers were required to hold four distinctly coloured circles in visual working memory. Each circle was positioned at one of the corners of the display. After storing these four items in visual working memory, we asked observers whether one of the colours was present in the memory array (e.g., “was red present?”). In some trials, a visual probe dot that required an immediate response was presented on the empty computer screen at a location that previously was occupied by one of the four circles. The probe dot location could, at chance probability, coincide with the location of the coloured circle that had to be retrieved from visual working memory. We found that when we asked whether a particular colour was present in the memory array (“was red present?”) and the probe happened to be presented at the location of the colour that participants had to retrieve (the probe was presented at the location that happened to contain the red circle), probe RTs were reliably faster than when the probe was presented at any of the other locations (see also Theeuwes et al. 2009). This finding suggests that the four coloured circles as they were presented at their representative locations on the screen were stored exactly in this spatial make-up in working memory. When asked about a particular colour, attention was automatically shifted to the location as it was stored in working memory. We concluded that accessing information from memory is not much different than accessing information from the outside world. In both cases spatial visual attention plays a key role in accessing this information.

Soto and colleagues (Soto et al. 2005; Soto et al. 2008) also provided compelling evidence that stimuli held in working memory automatically drive selection priority. They demonstrated that stimuli held in working memory affects the direction of the first saccade, and the fastest reaction times in detecting pop-out targets. Crucially, when observers just had to look at the stimuli and not report for later recall Soto et al. reported no effect on search, result inconsistent with the findings discussed in the previous section regarding priming. If anything, one expects that the mere exposure to a stimulus should have some effect on the subsequent allocation of attention. It is unclear why Soto et al. did not find evidence for priming but one concern is that what they labelled as pop-out search was in fact slow, serial and effortful search with search times up to 1,600 ms (Soto et al. 2005). In conditions in which observers serially item-by-item search through a display, it is unlikely that priming can occur. By the time, serial focal attention has reached the item that was primed so much processing has taken place during the scanning of the other items that the priming effect has worn off. Clearly in our studies we always used pop-out singleton search, a condition that can reveal early priming effects of selection priority.

To make things more complicated, there are various studies that have failed to find an effect of working memory on the sequent employment of attention (Houtkamp and Roelfsema 2006; Woodman and Luck 2007). There are several clear discrepancies between the studies that do and do not find effects of working memory on the subsequent deployment of attention (for an overview see Olivers 2009). The question is then whether keeping an item in working memory automatically biases attention. On the basis of these studies it seems that it is a matter of priorities. If one keeps an object in working memory but the task requires that one searches for another item (which may change from trial to trial) then it is likely that the effect of the item in working memory is small or even negligible. The working memory item is probably replaced by an item that is more important for the task. As such the effect of working memory content is much less automatic then what we have described in the previous paragraph regarding inter-trial priming. As noted inter-trial priming occurs even when observers actively try to counteract its effect. With respect to explicit working memory one can simply override the working memory content with information that is immediately relevant for the current search task. Obviously, if the representation is working memory is relatively weak one will not find an automatic bias toward this items when encountered in the outside world.

In summary: there is quite some evidence that objects that are stored in working memory automatically affect the deployment of attention. However, if these items are removed, altered or overwritten in working memory, the effects are no longer seen. In this respect, it is automatic as long as the storage in working memory is active. We have shown such automatic deployment of attention can occur because objects and/or their location in memory matches that in the outside world.

Selection and Reward

Background It is well-known that behaviour can be is strongly modulated by previous experiences, i.e., by the consequences that have led to it in the past. For example, in studies of learning, visual exploration—much like any other behaviour—is thought to be largely guided by the prior experience of reward (cf. law of effect Thorndike 1911). In essence it may not be surprising that reward has an effect on the deployment of attention. For example, if one would receive a monetary reward for selecting say a red circle, it may be not be surprising that the person will try to select that very same red circle again on the next occasion. This may not be particularly interesting since it indicates that prior experience of reward has a strong impact on strategic preparation, and thus on the establishment of goal-driven attentional control.

Many studies have in fact shown this less interesting effect because it is difficult to separate strategic from non-strategic (automatic) effects. Typically, in these studies human or animal observers receive stimuli that predict reward outcome for the current trial. Results usually show that visual processing of the reward-predictive stimulus is facilitated (e.g., Ikeda and Hikosaka 2003; Kiss et al. 2009; Platt and Glimcher 1999). For example, in Peck et al. (2009) monkeys viewed displays that contained two placeholders (Peck et al. 2009). On each trial, monkeys had to make a saccade to a target presented at one of these two placeholder locations. A cue superimposed on one of the placeholders preceded the target and indicated the likelihood that the monkey would receive a liquid reward in that trial. This cue did not predict the location of the target and was therefore task irrelevant. The results showed that saccades to a target presented at a location at which a high-reward cue had been, were faster and more accurate than saccades to a target presented at a location where a low-reward cue had been. Peck et al. (2009) recorded single-unit activity from cells in lateral intraparietal cortex (LIP), an area assumed to represent a high-level salience map, integrating basic stimulus activation with prior knowledge regarding task confines and target characteristics (e.g., Platt and Glimcher 1999). Results from Peck et al. (2009) suggest that LIP is important area that integrates reward contingencies with representations of environmental stimuli. However, it is not clear whether the impact of reward on processing in this brain structure does in fact reflect an automatic instance of reward learning or whether it is simply a strategic effect. It is very possible that the monkeys in Peck et al. (2009) and human and animal observers in similar experiments may have strategically ’looked out’ for the high-reward cue, resulting in the establishment of a top-down attentional set for such an item. As such this study and similar other studies just demonstrate that reward may have a strategic effect on the deployment of attention.

A recent study from our lab (Hickey et al. 2010b) involving human observers was designed to circumvent these problems using a variant of the additional singleton task of Theeuwes (1991, 1992). We used basically the same paradigm as the one described earlier (Fig. 2) with one major exception: the characteristics defining the target and distractor could switch from trial to trial; the target might be a red diamond in one trial, with the distractor a green circle, but in the next trial the target might be a green diamond, and the distractor a red circle. This is a version of the additional singleton task as was originally developed by Theeuwes (1991). Figure 6 gives an overview of the procedure and the results.

In this study, given that their response was correct, observers received either 1 or 10 cents. Note that the reward was in fact randomized and therefore not tied to performance. As is clear from the Figure a high magnitude reward biased attention towards the features that were rewarded. After receiving a high magnitude reward, observers were quick when the target had the same colour as it did in the immediately preceding trial, but they were slow when the colours switched. For low magnitude reward the pattern reversed: a low-magnitude reward resulted in a relative devaluation of features that characterize a target such that attention is less likely to be deployed to objects characterized by these features in the next trial. Observers were slow when the target had the same colour as it did in the preceding trial, and now relatively quick when the colours switched between trials.

An obvious interpretation of these findings would be that observers strategically looked out for the colour that was just rewarded, even though this had no benefit for them. If this would be the case it would imply that reward only has a strategic effect

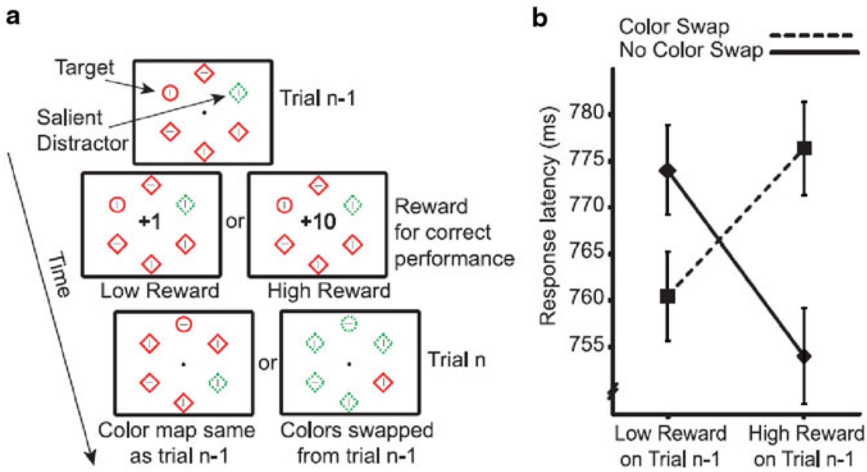


Fig. 6 Stimuli and data from Hickey et al. (2010a) **a** Observers searched for a circle and responded to the orientation of the line segment inside of it. Colours could randomly switch from trial to trial. Observers received a high or low reward (1 cent or 10 cents) which was randomly administrated (given that the response was correct). **b** Reaction times to find the target. After receiving a high magnitude reward, and the colours stayed the same observers were fast; if the colour switched they were relatively slow. After receiving a low magnitude reward the effect reversed. After a low reward and the colours switched observers were relatively fast; if it stayed the same, they were relatively slow

on the deployment of attention. In a follow-up experiment we addressed this issue. Observers were told that following a high magnitude reward the colour would switch (this did happen in 80 % of the trials). Obviously, now observers should stop looking for reward-associated visual features because they knew this was counter-productive. Importantly, however, we found exactly the same pattern of results as in the first experiment (see Fig. 6). Observers still selected the object with the same colour as the high-reward target in the last trial, even though this slowed their response, made them less accurate, and ultimately cost them money. These latter findings are crucial: it seems that the "brain" keeps selecting the features that were rewarded even though observers try to do the opposite. These findings are very similar to the processes described regarding priming: as with priming, it seems that the effects of reward on attention is beyond strategic endogenous volitional control. The reward received determines what will be selected on the next trial, independent of what observers are trying to do volitionally.

Reward and Automatic Selection? Even though nobody disputes strategic effects of reward on the attentional deployment (e.g., I got rewarded to look for red so I look for it again), our study was one of the first to show that there are automatic effects of reward that are beyond strategic control. A criticism that could be raised to our study is that in the counter-predictive condition (e.g., high reward implies a switch of colours) observers did not bother to use the information. However a control study shows that this is in fact unlikely. In a control experiment that was very similar to the one with the switching colours, we gave information about the location of the upcoming target. In a task, observers did use this information to improve their performance. Our results showed that observers were quicker when the target was presented at the cued location then when it was presented at the uncued location. So we have evidence that observers try to use this information to improve their task performance when they receive it, yet in the colour switch experiment this strategically "trying" had no effect on behaviour.

In addition to the behavioural findings that suggest that reward affects the deployment of attention in an automatic way, we also collected ERP data that confirms the idea of automatic effects of reward on the attentional deployment (Hickey et al. 2010a). We used the same paradigm as described before (Fig. 6) and looked at the early ERP component (the posterior P1 component) The P1 is assumed to reflect relatively early visual processing in extrastriate cortex (e.g., Luck and Hillyard 1994) and is not sensitive to top-down attentional set for visual features (Hopf et al. 2000). We found a clear P1 effect which suggests that the receiving the reward changed visual processing at early stages which cannot be attributed to top-down strategic effects. There was also a distractor-elicited N2pc which became evident when the distractor was defined by the colour which rewarded on the previous trial confirming the notion that attention was captured to the distractor (e.g. Hickey et al. 2006). Importantly we only observed this in the high-reward, switch colour condition and not in the low-reward, switch colour condition or any other condition. These results suggest that a high magnitude reward makes target become more salient. As noted, the increase in salience after a high reward is not a strategic effect; it is the result of an automatic enhancement of salience in early visual brain areas such as the ventral-lateral occipital lobe.

The ERP study of Hickey et al. (2010b) also reveals the underlying brain mechanism that may drive the automatic effects of reward. As a theoretical background we used the incentive salience hypothesis of Berridge and Robinson (1998), which assumes that subcortical processing in the dopamine reward system results in perceptual and attentional priming of reward associated visual features. Even though the actual midbrain dopaminergic structures are too deep in the brain to be detected by ERP, one can isolate activity from these structures at the anterior cingulate cortex (ACC, Holroyd and Coles 2002) and measure this activity known as the medial frontal negativity (MFN, Gehring et al. 1993). In our study we found that the medial frontal negativity (MFN) that was elicited by reward feedback was strongly predictive of the behavioural reward priming effect: observers that had a larger MFN to high-magnitude reward feedback were also those that were more likely to select a distractor defined by reward-associated visual features.

In summary, there is compelling evidence that reward has a direct and automatic effect on the deployment of attentional selection that is beyond strategic control. We believe that receiving a reward results in the release of dopamine in the midbrain structures (including ACC). In turn, this activity in this midbrain structure changes the representation in early visual areas such that the features that are associated with reward become more salient. This increase in salience of stimulus features that are associated with reward is an automatic effect which cannot be counteracted with top-down attentional control. It is believed that these automatic attentional biases related to reward are also seen in addicted individuals who show an attentional bias for substance-related words and pictures (e.g., Robbins and Ehrman 2004). Crucially even if the addicted individual tries to ignore cues that are related to the drug abuse, through the mechanism described above, the 'brain' decides otherwise and makes these drug-related so salient that they will be selected automatically. Addicted people cannot help attending those stimuli that are related to their addiction.

Summary

In the current paper we have tried to reduce the role of the homunculus, the 'little man in the head' that is still so prominent in many psychological theories regarding the control our behaviour. We have claimed that once engaged in a task (which is a volitional act), visual selection may run off more or less in an automatic fashion. We have claimed that once a task is set in motion, selection is beyond top-down control also because there is not much conscious feedback from what we are selecting. Even though visual selection may run off automatically, there is a way to gain attentional control. By focusing attention to a limited spatial area, automatic processes (capture) outside the focus of attention can be stopped. By focusing attention, one will gain control at the expense of become slow resulting in slow and effortful search.

We have argued that the salience map that drives automatic selection is not only determined by raw physical salience of the objects in the environment but also by the way these objects appear to the person. We have argued that priming (feature, working memory and reward priming) may sharpening the cortical representation of these objects such that these objects appear to be more salient above and beyond their physical salience. Priming itself is not under volitional control: it occurs even if observers try to volitionally prepare for something else. In other words, looking at red prepares our brain for things that are red even if we volitionally try to prepare for green.

References

Ansorge, U., Horstmann, G., & Scharlau, I. (2010). Top-down contingent attentional capture during feed-forward visual processing. *Acta Psychologica, 135*(2), 123–126.

Awh, E., Armstrong, K. M., & Moore, T. (2006). Visual and oculomotor selection: links, causes and implications for spatial attention. *Trends in Cognitive Sciences, 10*(3), 124–130.

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception & Psychophysics, 55*, 485–496.

Baldauf, D., & Deubel, H. (2008). Properties of attentional selection during the preparation of sequential saccades. *Experimental Brain Research, 184*(3), 411–425.

Bargh, J. (1992). The ecology of automaticity: Toward establishing the conditions needed to produce automatic processing effects. *American Journal of Psychology, 105*(2), 181–199.

Bargh, J. A., & Chartrand, T. L. (1999). The unbearable automaticity of being. *American Psychologist, 54*(7), 462–479.

Belopolsky, A., Kramer, A. F., & Theeuwes, J. (2008). The role of awareness in processing of oculomotor capture: evidence from event-related potentials. *Journal of Cognitive Neuroscience, 20*(12), 2285–2297.

Belopolsky, A. V., Schreij, D., & Theeuwes, J. (2010). What is top-down about contingent capture? *Attention Perception & Psychophysics, 72*(2), 326–341.

Belopolsky, A. V., & Theeuwes, J. (2009). When are attention and saccade preparation dissociated? *Psychological Science, 20*(11), 1340–1347.

Belopolsky, A. V., & Theeuwes, J. (2010). No capture outside the attentional window. *Vision Research, 50*(23), 2543–2550.

Belopolsky, A. V., Zwaan, L., Theeuwes, J., & Kramer, A. F. (2007). The size of an attentional window modulates attentional capture by color singletons. *Psychonomic Bulletin & Review, 14*(5), 934–938.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Research Review, 28*(3), 309–369.

Besner, D., & Stolz, J. A. (1999). Unconsciously controlled processing: The Stroop effect reconsidered. *Psychonomic Bulletin & Review, 6*(3), 449–455.

Bichot, N. P., & Schall, J. D. (2002). Priming in macaque frontal cortex during popout visual search: Feature-based facilitation and location-based inhibition of return. *Journal of Neuroscience, 22*, 4675–4685.

Born, S., Kerzel, D., & Theeuwes, J. (2011). Evidence for a dissociation between the control of oculomotor capture and disengagement. *Experimental Brain Research*, 208, 621–631.

Broadbent, D. E. (1958). *Perception and communication*. Oxford: Pergamon.

Burnham, B. R. (2007). Displaywide visual features associated with a search display's appearance can mediate attentional capture. *Psychonomic Bulletin & Review, 14*(3), 392–422.

Burnham, B. R., & Neely, J. H. (2007). Involuntary capture of visual-spatial attention occurs for intersections, both real and "imagined". *Psychonomic Bulletin & Review, 14*(4), 735–741.

Carrasco, M., Ling, S., & Read, S. (2004). Attention alters appearance. *Nature Neuroscience, 7*(3), 308–313.

Cavanaugh, J., & Wurtz, R. H. (2004). Subcortical modulation of attention counters change blindness. *Journal of Neuroscience, 24*(50), 11236–11243.

Corbetta, M., Kincade, J. M., & Shulman, G. L. (2002). Neural systems for visual orienting and their relationships to spatial working memory. *Journal of Cognitive Neuroscience, 14*(3), 508–523.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nature Neuroscience, 3,* 201–215.

Craighero, L., Nascimben, M., & Fadiga, L. (2004). Eye position affects orienting of visuospatial attention. *Current Biology, 14*(4), 331–333.

Dalton, P., & Lavie, N. (2007). Overriding auditory attentional capture. *Perception & Psychophysics, 69*(2), 162–171.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222.

Desimone R. (1996). Neural mechanisms for visual memory and their role in attention. *Proceedings of the National Academy of Sciences USA 93,* 13494.

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

Downing, P. E. (2000). Interactions between visual working memory and selective attention. *Psychological Science, 11*(6), 467–473.

Duncan, J. (1985). Visual search and visual attention. In M. I. Posner, & O. S. M. Marin (Eds.), *Attention and performance* (Vol. XI, pp. 85–106). Hillsdale: Erlbaum.

Egeth, H. E., Leonard, C. J., & Leber, A. B. (2010). Why salience is not enough: Reflections on top-down selection in vision. *Acta Psychologica, 135*(2), 130–132.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception and Performance, 18*(4), 1030–1044.

Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: Contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human Perception and Performance, 20,* 317–329.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error-detection and compensation. *Psychological Science, 4*(6), 385–390.

Geyer, T., Muller, H. J., & Krummenacher, J. (2008). Expectancies modulate attentional capture by salient color singletons. *Vision Research, 48*(11), 1315–1326.

Godijn, R., & Theeuwes, J. (2002a). Oculomotor capture and inhibition of return: Evidence for an oculomotor suppression account of IOR. *Psychological Research: Psychologische Forschung, 66*(4), 234–246.

Godijn, R., & Theeuwes, J. (2002b). Programming of endogenous and exogenous saccades: Evidence for a competitive integration model. *Journal of Experimental Psychology: Human Perception and Performance, 28*(5), 1039–1054.

Godijn, R., & Theeuwes, J. (2003). Parallel allocation of attention prior to the execution of saccade sequences. *Journal of Experimental Psychology: Human Perception and Performance, 29*(5), 882–896.

Graf, P., & Schacter, D. L. (1985). Implicit and explicit memory for new associations in normal and amnesic subjects. *Journal of Experimental Psychology: Learning Memory and Cognition, 11*(3), 501–518.

Hernandez, M., Costa, A., & Humphreys, G. W. (2010). The size of an attentional window affects working memory guidance. *Attention Perception & Psychophysics, 72*(4), 963–972.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010a). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience, 30*(33), 11096–11103.

Hickey, C., McDonald, J. J., & Theeuwes, J. (2006). Electrophysiological evidence of the capture of visual attention. *Journal of Cognitive Neuroscience, 18*(4), 604–613.

Hickey, C., van Zoest, W., & Theeuwes, J. (2010). The time course of exogenous and endogenous control of covert attention. *Experimental Brain Research, 201*(4), 789–796.

Hoffman, J. E. (1986). Spatial attention in vision: Evidence for early selection. *Psychological Research, 48,* 221–229.

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis. of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*(4), 679–709.

Hopf, J.-M., Luck, S. J., Girelli, M., Hagner, T., Mangun, G. R., Scheich, H., et al. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex, 10*, 1233–1241.

Houtkamp, R., & Roelfsema, P. R. (2006). The effect of items in working memory on the deployment of attention and the eyes during visual search. *Journal of Experimental Psychology: Human Perception and Performance, 32*(2), 423–442.

Hunt, A. R., & Kingstone, A. (2003). Covert and overt voluntary attention: linked or independent? *Cognitive Brain Research, 18*(1), 102–105.

Hunt, A. R., von Muhlenen, A., & Kingstone, A. (2007). The time course of attentional. and oculomotor capture reveals a common cause. *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 271–284.

Ikeda, T., & Hikosaka, O. (2003). Reward-dependent gain and bias of visual responses in primate superior colliculus. *Neuron, 39*(4), 693–700.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203.

James, W. (1890). *The principles of psychology*. London: MacMillan.

Jonides, J., & Yantis, S. (1988). uniqueness of abrupt visual onset in capturing attention. *Perception & Psychophysics, 43,* 346–354.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading—from eye fixations to comprehension. *Psychological Review, 87*(4), 329–354.

Kim, M. S., & Cave, K. R. (1999). Top-down and bottom-up attentional control: On the nature of interference from a salient distractor. *Perception & Psychophysics, 61*, 1009–1023.

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science, 20*(2), 245–251.

Klein, R. (1980). Does oculomotor readiness mediate cognitive control of visual attention? In R. S. Nickerson (Ed.), *Attention and performance VIII.* Hillsdale: Lawrence Erlbaum.

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

Klein, R. M., & Pontefract, A. (1994). Does oculomotor readiness mediate cognitive control of visual attention? In: C. Umilta & M. Moscovitch (Eds.), *Attention & Performance XV: Conscious and Unconscious Processing.* (pp. 323–350) Cambridge: MIT Press.

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

Kowler, E., Anderson, E., Dosher, B., & Blaser, E. (1995). The role of attention in the programming of saccades. *Vision Research, 35*(13), 1897–1916.

Kramer, A. F., & Hahn, S. (1995). Splitting the beam: Distribution of attention over noncontiguous regions of the visual field. *Psychological Science, 6,* 381–386.

Kristjansson, A. (2010). Priming in visual search: A spanner in the works for Theeuwes's bottom-up attention sweeps? *Acta Psychologica, 135*(2), 114–116.

Kristjánsson, Á., Wang, D., & Nakayama, K. (2002). The role of priming in conjunctive visual search. *Cognition, 85,* 37–52.

Kumada, T. (1999). Limitations in attending to a feature value for overriding stimulus-driven interference. *Perception & Psychophysics, 61,* 61–79.

Leber, A. B., & Egeth, H. E. (2006). It's under control: Top-down search strategies can override attentional capture. *Psychonomic Bulletin & Review, 13*(1), 132–138.

Li, Z. (2002). A saliency map in primary visual cortex. *Trends in Cognitive Sciences, 6*(1), 9–16.

Libet, B., Gleason, C. A., Wright, E. W., & Pearl, D. K. (1983). Time of conscious intention to act in relation to onset of cerebral-activity (readiness-potential)—the unconscious initiation of a freely voluntary act. *Brain, 106*(Sep.), 623–642.

Logan, G. D. (1985). Executive control of thought and action. *Acta Psychologica, 60*, 193–210.

Lu, S. N., & Han, S. H. (2009). Attentional capture is contingent on the interaction between task demand and stimulus salience. *Attention Perception & Psychophysics, 71*(5), 1015–1026.

Luck, S. J., & Hillyard, S. A. (1994). electrophysiological correlates of feature analysis during visual-search. *Psychophysiology, 31*(3), 291–308.

Ludwig, C. J. H., & Gilchrist, I. D. (2002). Stimulus-driven and goal-driven control over visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 902–912.

Macleod, C. M. (1991). Half a century of research on the stroop effect—an integrative review. *Psychological Bulletin, 109*(2), 163–203.

Maljkovic, V., & Nakayama, K. (1994). Priming of pop-out: I. Role of features. *Memory & Cognition, 22*(6), 657–672.

Mathot, S., Hickey, C., & Theeuwes, J. (2010). From reorienting of attention to biased competition: Evidence from hemifield effects. *Attention Perception & Psychophysics, 72*(3), 651–657.

Moore, T., & Fallah, H. (2001). Control of eye movements and spatial attention (vol. 98, pg 1273, 2001). *Proceedings of the National Academy of Sciences of the United States of America, 98*(7), 4276.

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132*(2), 297–326.

Mounts, J. R. W. (2000). Evidence for suppressive mechanisms in attentional selection: Feature singletons produce inhibitory surrounds. *Perception & Psychophysics, 62*, 969–983.

Mulckhuyse, M., Van Der Stigchel, S., & Theeuwes, J. (2009). Early and late modulation of saccade deviations by target distractor similarity. *Journal of Neurophysiology, 102*(3), 1451–1458.

Müller, H. J., Heller, D., & Ziegler, J. (1995). Visual search for singleton feature targets within and across feature dimensions. *Perception & Psychophysics, 57*, 1–17.

Müller, H. J., Reimann, B., & Krummenacher, J. (2003). Visual search for singleton feature targets across dimensions: Stimulus and expectancy-driven effects in dimensional weighing. *Journal of Experimental Psychology: Human Perception and Performance, 29*, 1021–1035.

Muller, J. R., Philiastides, M. G., & Newsome, W. T. (2005). Microstimulation of the superior colliculus focuses attention without moving the eyes. *Proceedings of the National Academy of Sciences of the United States of America, 102*(3), 524–529.

Neisser, U. (1967). *Cognitive psychology*. New York: Appleton Century Crofts.

Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of "popout" under anesthesia. *Visual Neuroscience, 16*(1), 15–34.

Olivers, C. N., & Hickey, C. (2010). Priming resolves perceptual ambiguity in visual search: Evidence from behaviour and electrophysiology. *Vision Research, 50*(14), 1362–1371.

Olivers, C. N. L. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1275–1291.

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1243–1265.

Pashler, H. E., & Shiu, L. P. (1999). Do images involuntarily trigger search? A test of Pillsbury's hypothesis. *Psychonomic Bulletin & Review, 6*(3), 445–448.

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *Journal of Neuroscience, 29*(36), 11182–11191.

Pinto, Y., Olivers, C. N. L., & Theeuwes, J. (2005). Target uncertainty does not lead to more distraction by singletons: Intertrial priming does. *Perception & Psychophysics, 67*(8), 1354–1361.

Platt, M. L., & Glimcher, P. W. (1999). Effects of expected reward on LIP neuronal activity when monkeys choose to shift gaze away from the neuronal response field. *Investigative Ophthalmology & Visual Science, 40*(4), S60.

Posner, M. I. (1978). *Chronometric explorations of mind*. Hillsdale: Erlbaum.

Posner, M. I. (1980). Orienting of attention. The VIIth Sir Frederic Bartlett lecture. *Quarterly Journal of Experimental Psychology, 32*, 3–25.

Posner, M. I., Davidson, B. J., & Snyder, C. R. R. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General, 109*, 160–174.

Posner, M. I., & Petersen, S. E. (1990). The Attention System of the Human Brain. *Annual Review of Neuroscience, 13*, 25–42.

Proulx, M. J., & Egeth, H. E. (2006). Target-nontarget similarity modulates stimulus-driven control in visual search. *Psychonomic Bulletin & Review, 13*(3), 524–529.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience, 27*, 611–647.

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*, 31–40.

Rizzolatti, G., Riggio, L., & Sheliga, B. M. (1994). Space and selective attention. In C. Umilta, & M. Moscovitch (Eds.), *Attention and performance XIV*. Cambridge: MIT Press.

Robbins, S. J., & Ehrman, R. N. (2004). The role of attentional bias in substance abuse. *Behavioral & Cognitive Neuroscience Reviews, 3*(4), 243–260.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictable switch between simple cognitive tasks. *Journal of Experimental psychology: General, 124*(2), 207–231.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: i. detection, search, and attention. *Psychological Review, 84*, 1–66.

Schubo, A. (2009). Salience detection and attentional capture. *Psychological Research-Psychologische Forschung, 73*(2), 233–243.

Sheliga, B. M., Riggio, L., & Rizzolatti, G. (1994). Orienting of attention and eye movements. *Experimental Brain Research, 98*, 507- 522.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*(3), 127–190.

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 248–261.

Soto, D., Hodsoll, J., Rotshtein, P., & Humphreys, G. W. (2008). Automatic guidance of attention from working memory. *Trends in Cognitive Sciences, 12*(9), 342–348.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Theeuwes, J. (1990). Perceptual selectivity is task dependent—evidence from selective search. *Acta Psychologica, 74*(1), 81–99.

Theeuwes, J. (1991). Cross-dimensional perceptual selectivity. *Perception & Psychophysics, 50*(2), 184–193.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*(6), 599–606.

Theeuwes, J. (1994a). Endogenous and exogenous control of visual selection. *Perception, 23*(4), 429–440.

Theeuwes, J. (1994b). Stimulus-driven capture and attentional set—selective search for color and Visual Abrupt Onsets. *Journal of Experimental Psychology: Human Perception and Performance, 20*(4), 799–806.

Theeuwes, J. (1995). Perceptual selectivity for color and form: On the nature of the interference effect. In A. F. Kramer, M. G. H. Coles, & G. D. Logan (Eds.), *Converging operations in the study of visual attention* (pp. 297–314). Washington DC: American Psychological Association.

Theeuwes, J. (2004). Top-down search strategies cannot override attentional capture. *Psychonomic Bulletin & Review, 11*(1), 65–70.

Theeuwes, J. (2010a). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*(2), 77–99.

Theeuwes, J. (2010b). Top-down and bottom-up control of visual selection: Reply to commentaries. *Acta Psychologica, 135*(2), 133–139.

Theeuwes, J., Atchley, P., & Kramer, A. F. (2000). On the time course of top-down and bottom-up control of visual attention. In S. M. J. Driver (Ed.), *Attention & performance* (Vol. 18, pp. 105 125): Cambridge: MIT Press.

Theeuwes, J., Belopolsky, A., & Olivers, C. N. L. (2009). Interactions between working memory, attention and eye movements. *Acta Psychologica, 132*(2), 106–114.

Theeuwes, J., & Chen, C. Y. D. (2005). Attentional capture and inhibition (of return): The effect on perceptual sensitivity. *Perception & Psychophysics, 67*(8), 1305–1312.

Theeuwes, J., De Vries, G. J., & Godjin, R. (2003). Attentional and oculomotor capture with static singletons. *Perception & Psychophysics, 65*(5), 735–746.

Theeuwes, J., & Hagenzieker, M. P. (1993). Visual search of traffic scenes: On the effect of location expectations. In A. Gale (Ed.), *Vision in vehicles* (Vol. 4, pp. 149–158). Amsterdam: Elsevier.

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.

Theeuwes, J., Kramer, A. F., Hahn, S., Irwin, D. E., & Zelinsky, G. J. (1999). Influence of attentional capture on oculomotor control. *Journal of Experimental Psychology: Human Perception and Performance, 25*(6), 1595–1608.

Theeuwes, J., Kramer, A. F., & Irwin, D.E. (2011). Attention on our mind: the role of spatial attention in visual working memory. *Acta Psychologica, 137*(2), 248–251.

Theeuwes, J., Kramer, A. F., & Kingstone, A. (2004). Attentional capture modulates perceptual sensitivity. *Psychonomic Bulletin & Review, 11*(3), 551–554.

Theeuwes, J., Reimann, B., & Mortier, K. (2006). Visual search for featural singletons: No top-down modulation, only bottom-up priming. *Visual Cognition, 14*(4–8), 466–489.

Theeuwes, J., & Van Der Burg, E. (2007). The role of spatial and nonspatial information in visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 33*(6), 1335–1351.

Theeuwes, J., Van Der Burg, E., & Belopolsky, A. (2008). Detecting the presence of a singleton involves focal attention. *Psychonomic Bulletin & Review, 15*(3), 555–560.

Theeuwes, J., & Van Der Burg, E. (2011). On the limits of top-down control. *Attention, Perception & Psychophysics*, *73*, 2092–2103

Thompson, K. G., Bichot, N. P., & Sato, T. R. (2005). Frontal eye field activity before visual search errors reveals the integration of bottom-up and top-down salience. *Journal of Neurophysiology, 93*, 337–351.

Thorndike, E. L. (1911). *Animal intelligence*. New York: Macmillan.

Treisman, A., & Kahneman, D. (1981). An early interference effect in visual-perception. *Bulletin of the Psychonomic Society, 18*(2), 68–68.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*, 97- 136.

Tse, P. U., Sheinberg, D. L., & Logothetis, N. K. (2002). Fixational eye movements are not affected by abrupt onsets that capture attention. *Vision Research, 42*(13), 1663–1669.

Van Der Stigchel, S., Meeter, M., & Theeuwes, J. (2007). The spatial coding of the inhibition evoked by distractors. *Vision Research, 47*(2), 210–218.

Van Der Stigchel, S., & Theeuwes, J. (2007). The relationship between covert and overt attention in endogenous cuing. *Perception & Psychophysics, 69*(5), 719–731.

van Zoest, W., Donk, M., & Theeuwes, J. (2004a). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746–759.

van Zoest, W., Donk, M., & Theeuwes, J. (2004b). Stimulus-driven and goal-directed control in saccadic target selection. *Perception, 33,* 146–146.

Wolfe, J. M. (1994). Guided Search 2.0. A revised model of visual search. *Psychonomic Bulletin & Review, 1,* 202–238.

Wolfe, J. M., Butcher, S. J., Lee, C., & Hyle, M. (2003). Changing your mind: on the contributions of top-down and bottom-up guidance in visual search for feature singeltons. *Journal of Experimental Psychology, 29*(2), 483–502.

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 363–377.

Wu, S. C., & Remington, R. W. (2003). Characteristics of covert and overt visual orienting: Evidence from attentional and oculomotor capture. *Journal of Experimental Psychology: Human Perception and Performance, 29*(5), 1050–1067.

Wundt, W. (1887). *Grundzüge der physiologischen Psychologie [Foundations of physiological psychology].* Leipzig: Engelmann.

Yantis, S., & Egeth, H. E. (1999). On the distinction between visual salience and stimulus-driven attentional capture. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 661–676.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 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*, 121–134.

Guidance of Visual Search by Memory and Knowledge

Andrew Hollingworth

Abstract To behave intelligently in the world, humans must be able to find objects efficiently within the complex environments they inhabit. A growing proportion of the literature on visual search is devoted to understanding this type of natural search. In the present chapter, I review the literature on visual search through natural scenes, focusing on the role of memory and knowledge in guiding attention to task-relevant objects.

Keywords Visual search · Memory · Goal-directed vision · Natural scenes

The laboratory study of visual search began as a means to an end. Early research used visual search as an experimental paradigm for probing perceptual efficiency in identification and selection (e.g., Schneider and Shiffrin 1977; Treisman and Gelade 1980). The purpose was not necessarily to understand how we find objects in natural scenes (where on earth did I leave my keys?) but rather to isolate the mechanisms that enable humans to select goal-relevant items in the presence of perceptual competition. The visual search paradigm has become indispensable to the study of attention and has been applied in research ranging from the single unit study of neurons in the superior colliculus (e.g., McPeek and Keller 2002) to the study of individual differences in psychopathology (see Weierich et al. 2008). Over the last 10–15 years, however, a large and growing proportion of the research on visual search has been devoted to understanding how we find goal-relevant objects within the types of complex environments that comprise daily life. Several of the chapters in the present volume are examples of this trend.

How we find objects in scenes is one of the central topics a science of cognition must address, because finding objects efficiently is essential for everyday behavior. Consider the activity of making tea discussed by Land et al. (1999). Making tea requires the sequential selection and use of several different objects, and they must be selected in a particular sequence defined by the task. The tea kettle must be found to fill it with water. A cup must be located, and then the tea bags are needed to place one in the cup. Milk is required later in the task, as is sugar, and a spoon

---

A. Hollingworth (✉)
Department of Psychology, University of Iowa, Iowa City, IA 52242, USA
e-mail: andrew-hollingworth@uiowa.edu

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_4,

must be found to measure the sugar and stir it into the liquid. The objects in a scene cannot be perceived simultaneously, so each of these component operations requires visual search for a particular target object. Moreover, the goal of search and the representation of the relevant target must evolve as the sub-goals of the task change. For example, the template specifying the relevant target object must switch at some point from representing features of the sugar bowl to features of a spoon. The efficiency by which we find each object controls the overall efficiency of the task, as anyone who has attempted to make tea in an unfamiliar kitchen can attest. Similar sequential search demands are imposed by many everyday tasks, from brushing one's teeth to retrieving a letter from the mailbox. Thus, efficient human behavior is dependent on numerous visual searches that unfold, one after another, as goals change.

To what extent can research using the traditional visual search paradigm be applied to understanding object search in scenes? Many researchers have identified a direct relationship between search arrays and natural environments (both consist of numerous objects), and many have identified a relationship between the task in visual search (find the "T") and real-world tasks (find my keys). Hundreds of papers and talks using traditional visual search paradigms have started with a real-world example of search, indicating that the authors saw their results as relevant to understanding how we find objects in natural environments. However, the traditional visual search task is, in many key respects, ill-suited to this endeavor. Below, I discuss the properties of real-world scenes that are important for understanding natural search, most of which are poorly captured by traditional search paradigms (see Fig. 1). This discussion is not a criticism of existing approaches to visual search. In fact, the properties that make the visual search paradigm ill-suited to understanding natural search are often precisely those properties that have given researchers sufficient control to draw inferences about basic mechanisms of visual perception and attention.

Scenes and the Objects within them are Meaningful Visual scenes are meaningful environments, such as a kitchen or office (Henderson and Hollingworth 1999). We know a great deal about scene categories that could be brought to bear on search operations. For example, mixers tend to be found in kitchens, whereas staplers tend to be found in offices. One's search for a stapler will proceed much differently if one is searching for the stapler within an office versus within a kitchen (Henderson et al. 1999; Neider and Zelinsky 2006; Torralba et al. 2006).

Objects are Complex, and their Identities and Visual Forms are Highly Variable Most traditional search arrays are composed of highly similar (often identical) items. In contrast, natural environments are composed of objects that vary widely in their identities and visual properties. In addition, objects typically have complex shapes composed of multiple parts and contain several values on any given feature dimension (e.g., multiple colors). It is not clear that targets in a natural scenes can be segregated efficiently from distractors on the basis of simple visual features, such as those studied in the traditional search literature (see Wolfe et al. 2011).

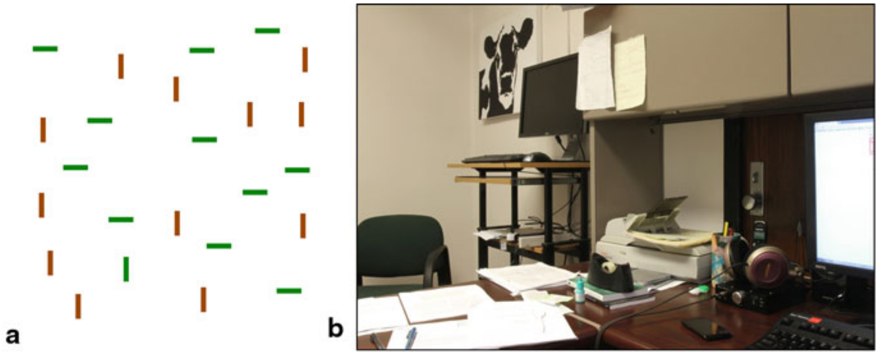


Fig. 1 **a** Typical stimulus in the traditional visual search paradigm (modeled after Treisman and Gelade 1980). The array consists of simple, meaningless stimuli arranged randomly. Across trials, locations are generated anew, precluding any possibility of spatial learning. Inter-item similarity is very high, including multiple item repetitions. Objects can be discriminated from each other on the basis of a small set of well-defined feature values. Object identification does not necessarily require movements of the eyes to fixate the elements of the array. **b** Typical real-world scene (the author's office). The scene is clearly identifiable as belonging to a particular category, as are the individual objects that comprise the scene. Objects are constrained to appear on surfaces. In addition, their locations are related to their functions (the keyboard appears near the monitor) and are often consistent across multiple searches (the monitor tends to stay put). Inter-object similarity is very low compared with the traditional search array. Objects are complex, composed of multiple values on several feature dimensions. It would be difficult to discriminate some of the objects from distractors on the basis of a single (or even several) feature values. The photograph is static, but under dynamic conditions characteristic of natural vision, changes in illumination, occlusion, perspective, viewing distance, and so forth would alter the appearance of the objects significantly. Within such a large, complex, crowded environment, object identification depends on movements of the eyes to obtain foveal information from individual items

Search Through Scenes Depends on the Retrieval of Target Features from Memory When searching for a stapler in an office, one is not provided with a visual image of the stapler immediately before commencing search (If one saw the stapler before searching, there would be no need to search for it). Therefore, the representation of the target object (i.e., the target template) must be retrieved from memory. If one knows something about the features of the particular stapler, then the template could be specific to that object. Alternatively, one could draw upon categorical knowledge to establish a search template that is general to the class of stapler (Yang and Zelinsky 2009).

Scenes have Coherent Spatial Structure Unlike standard search arrays, the locations of objects in scenes are not randomly determined. Spatial constraints arise from several sources. The most obvious is that objects are constrained by gravity to be located on surfaces (or to be supported in some other way, such as hanging from the wall) (see, e.g., Biederman et al. 1982). If one seeks an object in a scene, one can generally limit the search to surfaces that plausibly support the object. In addition, object locations are related to object function. Staplers often appear on desks because that location provides efficient access in the course of stapling. Similar functional considerations apply across exemplars of a particular scene type, enabling categorical inferences about the probable locations of objects.

Scenes are Experienced Repeatedly The vast majority of environments we inhabit are familiar to us, and searches are repeated many times for common tasks. This raises the possibility that search through scenes is strongly influenced by learning (Chun 2000).

Selection in Scenes Depends on Overt Shifts of Gaze Traditional work on visual search has often minimized the contribution of eye movements. Yet, the size and complexity of natural scenes requires eye movements (and head and body movements) to obtain high-resolution, foveal information from objects. Moreover, gaze fixation is used to link objects to the motor programs that operate over them (Ballard et al. 1997; Hayhoe 2000; Land and Hayhoe 2001; Land et al. 1999). These two roles mesh seamlessly. In the course of search, the eyes are directed sequentially to possible targets until the required object is fixated. Once fixated, gaze supports motor interaction with that object so that it can be used in the current task. (For discussion of the importance of studying eye nmovements in search, see Eckstein 2011; Findlay and Gilchrist 2003; Henderson 2003; Zelinsky 1996, 2008; Zelinsky et al. 1997).

Weaving through this analysis of the properties of scenes and search within them is the idea that natural search is likely to be strongly guided by knowledge and memory. We have extended experience with scene exemplars and scene types, we search repeatedly through particular environments, and the nature of the search operation changes dynamically as goals change. In fact, we construct and organize our environments with the intention of maximizing the efficient use of memory and knowledge in visual search. The remote control is kept in a little basket on the side table so that one need only inspect a single location to find it. After using a pen, one tends to place it back in the same location on the desk so that memory for its location can guide attention back to it. Progress toward understanding natural search will therefore depend on understanding how memory, knowledge, and goals guide the search process and how these mechanisms interact with the visual processing of the scene.

Before discussing the roles of memory and knowledge in natural search, let's first consider the possibility that search is instead dominated by the *visual* properties of the scene. In an influential model, Itti and Koch (2000) developed a "salience map" approach to explain search through scenes (and saccade target selection in scenes more generally). The model computes the relative salience of local scene regions as a function of various low-level visual properties (such as luminance, contrast, and contour density). These regions are then ranked within a saccade "priority map" and fixated sequentially, with initial fixations on the scene directed to the most salient regions. The approach is attractive in that it does not demand any role at all for memory, knowledge, or goals, is relatively straightforward to implement in computer systems, and incorporates properties of neural systems involved in visual selection.

However, such a model would appear to have difficulty accounting for natural search behavior. Returning to the example of making tea, consider the possibility that as each object becomes relevant to the task (kettle, then cup, then tea bag), gaze is directed to the most visually conspicuous object in the scene. This would obviously be a complete disaster; the tea would not get made. The visual salience of an object in a scene is not closely related to it functional relevance. In addition, the visual salience of objects does not change systematically as goals change, providing no means to adapt search to changing task demands.

A large body of research supports this analysis. Low-level visual salience accounts for only a small proportion of the variance in selection during goal-directed search through natural scenes. Selection is much more strongly influenced by memory and knowledge (Einhauser et al. 2008a, b; Foulsham et al. 2011; Foulsham and Underwood 2007, 2008; Henderson et al. 2007; Henderson et al. 2009; Itti 2005; Tatler et al. 2011; Torralba et al. 2006; Underwood et al. 2009; Zelinsky et al. 2006). When given the task to search for a target object in a novel scene, participants tend to direct their gaze to regions of the scene where an object of that type is likely to be found (Eckstein et al. 2006; Ehinger et al. 2009; Henderson et al. 1999; Neider and Zelinsky 2006; Torralba et al. 2006), independently of the visual salience of those regions (Henderson et al. 2009). In addition, attention is guided toward regions of the scene that match the features of the search target, regardless of whether the target itself is physically salient (Hwang et al. 2009; Kanan et al. 2009; Malcolm and Henderson 2009, 2010; Pomplun 2006; Zelinsky 2008). Although the visual properties of the scene over which search occurs are obviously central to the search operation, the selective mechanisms that determine where attention and gaze will be allocated within the scene are driven primarily by the observer's goals combined with knowledge about the scene context and knowledge of the visual properties of the object the observer is trying to find.

Memory in Traditional Visual Search Tasks

There is no question that memory and knowledge play a central role in visual search through natural scenes. However, the role of memory in traditional search tasks has been more controversial. Horowitz and Wolfe (1998) argued that search through random arrays of symbols has no memory. In their experiments, the positions of search items were static or were scrambled every 111 ms. If search depends on using memory to avoid previously attended locations (e.g., Klein 1988) then scrambling should impair search efficiency. Yet, the search slopes did not differ between the two conditions. Converging evidence came from a repeated search paradigm, in which multiple searches over the same static array did not generate a major benefit in performance (Wolfe et al. 2000). Wolfe (1999) proposed that after attention is withdrawn from an object, the visual object representation dissolves into its elementary features, leaving no memory that could facilitate search either within a trial or across trials.

This claim has been subject to significant debate (e.g., Shore and Klein 2000), and subsequent research has established a central role for memory in traditional visual search paradigms. If targets in the scrambled condition of the Horowitz and Wolfe (1998) paradigm are placed in previous distractor locations, search efficiency is impaired relative to the static condition (Kristjánsson 2000). Larger set sizes also generate slope differences between static and scrambled conditions (Kristjánsson 2000) as does search for multiple targets (Gibson et al. 2000). von Mühlenen et al. (2003) showed that the original Horowitz and Wolfe finding was probably caused by an idiosyncratic strategy: Participants kept their attention in a single location and waited for the target to appear nearby. Other experiments demonstrated memory in several components of the search operation. Refixations during search were observed less frequently than predicted by a no-memory model (McCarley et al. 2003; Peterson et al. 2001), and previous demonstrations of inhibition of return during search (Klein 1988) were confirmed (Müller and Von Mühlenen 2000; Takeda and Yagi 2000). Memory for distractors during search has been found to facilitate subsequent searches when the distractor becomes a target (Körner and Gilchrist 2007), and repeated search across multiple trials generates reliable savings (e.g., Chun and Jiang 1998). Thus, the balance of evidence indicates that memory processes plays multiple important roles in traditional visual search tasks.

Guidance by Categorical Knowledge of Scene Types

Our extensive experience with scene types allows us to predict where particular objects are likely to be found. Milk tends to be found in the refrigerator and not in the pantry. Staplers tend to be found on top of desks rather than beneath them. Indeed, search for objects whose locations are predictable within a scene is more efficient than search for objects whose locations are not (Eckstein et al. 2006; Henderson et al. 1999; Mack and Eckstein 2011; Neider and Zelinsky 2006). This type of knowledge can be applied rapidly within a search operation to guide attention and gaze toward regions of the scene that have a high probability of containing the target. For example, Torralba et al. (2006) found that eye movements during search for pedestrians were rapidly directed to sidewalks and other regions of the scene where pedestrians were likely to be found. Torralba et al. developed a *contextual guidance model* to account for these effects of scene and object knowledge on visual search. The model learns the associations between large-scale image features within a scene and the locations of particular object types. During a search operation, global image features and local salience are computed in parallel. These two sources of guidance are combined within a priority map that governs the order of scene regions fixated. The inclusion of contextual guidance allows the model outperform a model based solely on visual salience. In addition, if one adds a second top-down guidance component to the model that uses knowledge of the target appearance to prioritize candidate regions, the combined model can account for an impressive proportion of the variance in human eye movements during visual search (Ehinger et al. 2009).

Guidance from the Learning of Scene Exemplars and Repeated Search

We experience scenes repeatedly. We conduct multiple searches through the same scene for different objects and search for the same object in a scene multiple times. We conduct all kinds of non-search activities through those same scenes. Humans have ample opportunity to encode information from particular scene exemplars, such as the spatial structure of the environment and the locations of individual objects. And humans have the *capability* to remember these features of environments: Long-term memory (LTM) for objects and scenes can be quite precise (Hollingworth and Henderson 2002) and has a remarkably large capacity (Brady et al. 2008; Hollingworth 2004, 2005; Konkle et al. 2010; Standing 1973; Standing et al. 1970). The role of scene memory in search has been addressed in studies using scene previews and repeated search and has been focused on several key questions. Does scene memory facilitate visual search? If so, what types of scene information are functional? How much experience is necessary? Is learning task-general or specific to visual search?

A preview of the scene facilitates later search (Becker and Rasmussen 2008; Castelhano and Heaven 2011; Castelhano and Henderson 2007; Hillstrom et al. 2012; Hollingworth 2009, in press; Võ and Henderson 2010) and repeated search leads to reliable savings (Brockmole et al. 2006; Brockmole and Henderson 2006b; Brooks et al. 2010; Ehinger and Brockmole 2008; Võ and Wolfe 2012). Thus, there is no doubt that memory for a scene exemplar is indeed employed to facilitate visual search. Using a scene preview design, (Hollingworth 2009; see also Hollingworth 2006) examined two forms of scene memory that are likely to control the allocation of attention and gaze: memory for object locations (which could guide attention directly to the target location) and memory for the spatial structure of a scene (which could guide attention to the locations where the target object is likely to be found). Participants viewed a scene preview for 10 s. Then, a single search target object was presented. Next, participants searched through the scene and reported whether the target in the scene had the same or different left-right orientation. Two preview conditions were compared with a no-preview control condition. A preview that was identical to the search scene except for the presence of the target improved search, as measured both by reaction time and the elapsed time to the first fixation on the target. Thus, memory for the spatial structure of the scene guides attention during search. A preview that contained the target led to further savings, demonstrating that participants had encoded the locations of specific objects from the preview and had used this to guide attention to the remembered location.

Guidance from scene memory can be observed even with an extremely brief glimpse of a scene before search. Castelhano and Henderson (2007) presented a preview of a scene for 250 ms, followed by a label indicating the search target, followed by search through the scene. [In this and related experiments, Castelhano, Henderson and colleagues used a moving window technique to limit visible information during search to a small region around the current fixation point; however, this aspect of the design does not appear to be necessary, as similar effects are found for search through fully visible scenes (Hillstrom et al. 2012)]. A brief glimpse facilitated search relative to a no-preview baseline, and subsequent work has found that a masked preview of only 50 ms is sufficient to generate a search benefit (Võ and Henderson 2010). Thus, scene information can be extracted very rapidly from an image in a manner that can guide attention and gaze to plausible target locations. The effect is not due to simply identifying the *type* of scene over which search will occur, because a preview using a different exemplar from the same scene category provided no benefit at all (Castelhano and Henderson 2007), and preview effects are observed even for target objects that are semantically inconsistent with the scene in which they appear (Castelhano and Heaven 2011). Instead, the preview allows participants to encode the spatial structure of the scene exemplar and locate the plausible surfaces on which an object could appear. Note that this spatial representation is likely to be quite abstract; a change in the absolute size of the scene from preview to search had no effect on the magnitude of the preview benefit (Castelhano and Henderson 2007).

In preview studies, participants do not conduct explicit search during the preview itself, but the preview is embedded within the context of a search experiment, and participants know that they will later search through the previewed scene. To what extent is the scene learning that facilitates search specific to the context of a visual search task? Evidence from Võ and Wolfe (2012) suggested that learning is highly task-specific. They included an initial preview session in which each scene was viewed for 30 s. Participants did not know, during the preview session, that they would later search through the scenes. The preview task was either closely related to the informational demands of search (memorize the locations of the objects in the scenes) or not closely related (decide whether the room is inhabited by a man or a woman). Relative to a no-preview control, there was no evidence at all that a scene preview facilitated search, even in the case that participants were instructed to memorize object locations. However, repeated *searches* generated significant savings. Thus, the transfer of memory to visual search appeared to be limited to memory representations formed during previous searches for a particular object. Võ and Wolfe reasoned that in order to transfer, learning must occur within the context of visual search. This finding is broadly consistent with theoretical approaches holding that visual memory encoding and application is strongly constrained by task (Ballard et al. 1995; Droll et al. 2005).

The interpretation of the Võ and Wolfe (2012) results is limited by the fact that their preview and no-preview trials were assigned to different participant groups, and they had relatively little power to detect an effect of preview. Hollingworth (in press) replicated the method but in a more sensitive within-subjects design. Participants viewed half the scenes items in a preview session followed by a search session containing all scene items. A preview task of memorizing object locations led to a robust benefit in later search. Further, a preview task that did not involve remembering object locations (decide which object is the least likely to be present in a scene of this type) also generated a robust search benefit, and the magnitude of the effect was similar to that found for the location-memorization preview. Contrary to the evidence of Võ and Wolfe, our results establish that the transfer of learning is not strongly constrained by task. Object and scene information acquired outside of the context of visual search transfers reliably to the task of finding objects in scenes. Scene memory appears to be applied flexibly.

Consistent with this conclusion is the finding that when distractors are fixated during search (i.e., when they are not directly task relevant), a memory trace is established (Castelhano and Henderson 2005; Williams et al. 2005) that facilitates later search when that distractor becomes a target (Howard et al. 2011; Võ and Wolfe 2012). In three recent studies (Hollingworth in press; Howard et al. 2011; Võ and Wolfe 2012), participants searched for a series of different objects within the same scene image. For example, in a static bedroom scene a participant would first search for an alarm clock, then for a cell phone, then for a candle, and so on. As multiple searches progressed, participants became faster as they searched for objects that had been distractors during previous searches (Hollingworth in press; Võ and Wolfe 2012) and this benefit was related to whether the distractor had been fixated during a previous search (Howard et al. 2011). Again, scene memory appears to be quite flexible. Information acquired incidentally from an object when it was not task relevant is stored and used later to facilitate search for that object.

Further insights into the nature of scene learning across repeated searches comes from the literature on contextual cuing (for a review, see Chun and Turk-Browne 2008; Chap. 6). Initial studies used relatively simple arrays of randomly organized symbols. In the basic task, participants search for a target symbol. The spatial configuration of a subset of search arrays is repeated throughout the experiment. Search is facilitated as participants learn the spatial relationship between repeated search contexts and target locations. In the standard contextual cuing paradigm, the learning of target locations is incremental, implicit, and local. Arrays require several repetitions before an advantage emerges over the novel condition, and RTs for repeated arrays do not reach asymptote until at least 10–15 repetitions. Once established, the magnitude of the contextual cuing effect is modest (typically 40–70 ms). At the end of an experiment, participants are at chance when asked to discriminate between repeated arrays and novel arrays, indicating that the memory representation responsible for facilitating search is not consciously accessible. Finally, the learning in contextual cuing studies appears to be highly local, with the effect dependent on the objects closest to the target (Brady and Chun 2007; Olson and Chun 2002). For example, Brady and Chun (2007) found that repetition limited to the two distractors closest to the target was sufficient to produce contextual cuing equivalent to that produced by repetition of the entire array of 11 distractors.

But the story is, in fact, much more complicated. Studies of repeated search through *naturalistic* scenes have found learning to be neither incremental, implicit, nor local; instead, the learning of target location is extremely rapid, is explicitly available, and is strongly dependent on global properties of the scene context (Brockmole et al. 2006; Brockmole and Henderson 2006a, b; Brooks et al. 2010). In Brockmole and Henderson (2006b) participants searched for small letter targets embedded within images of natural scenes. The very first repetition of a particular scene context led to search savings, and search reached floor-level RTs following only 3–4 repetitions. In addition, the magnitude of the repetition benefit was extremely large. After a few repetitions, search through repeated scenes was approximately 2 s faster than search through novel scenes.

Differences between paradigms are not limited to the rate of learning or to the magnitude of the effect; the nature of the learning process itself appears to differ. At the end of the session in Brockmole and Henderson (2006b), participants reliably discriminated repeated scenes from novel scenes, indicating that the learning of target locations in scenes depends on memory processes that are available to conscious report. More importantly, object location was coded relative to the global structure and/or identity of the scene, rather than locally. Brockmole et al. (2006) included a transfer session after the main search session. For a repeated scene, either the local context around the target was disrupted (e.g., replacing the table on which the target appeared with a different table) or the global context was disrupted (e.g., replacing all of the rest of the scene except the table on which the target appeared). The former did not have any observable influence on the magnitude of the repetition effect, whereas the latter eliminated it completely. For real-world scenes, then, it is clear that target learning and/or the expression of learning is strongly dependent on global processing of the scene. In the experiments of Brockmole, Henderson, and colleagues, memory for the location of the target was likely to have been associated with a representation of the identity of the scene. Brockmole and Henderson (2006a) included a transfer session in which they mirror-reversed repeated scene items. Mirror reversal did not radically impair scene identification, but it disrupted the spatial structure of the scene and the absolute location of the target. Initial saccades during search were very frequently directed to the original *screen* location of the target rather than to the original *scene* location of the target. Participants appeared to associate an absolute target location directly with a scene identity, independently—at least to some extent—of the internal spatial structure of the scene.

The preceding discussion depends on a distinction between global and local properties of search contexts, but these are difficult to define for natural scenes. The functional context involved in search for a desk in an office (the office) will be different from the context involved in the search for a stapler on the desk (the desk). Brooks et al. (2010) proposed that memory for scenes is organized hierarchically and is accessed hierarchically in search. It would highly inefficient to code the location of the stapler relative to the coordinate system established by the large-scale features of the office (or office building). The locations of local objects are more likely to be coded relative to the immediate context for search, such as the desk on which the stapler usually appears. This localist coding assumption is consistent with evidence from the traditional contextual cuing literature, reviewed above. However, Brooks et al. proposed that that retrieval of the appropriate search context will depend on scene recognition. That is, retrieval of the desk context will be strongly dependent on

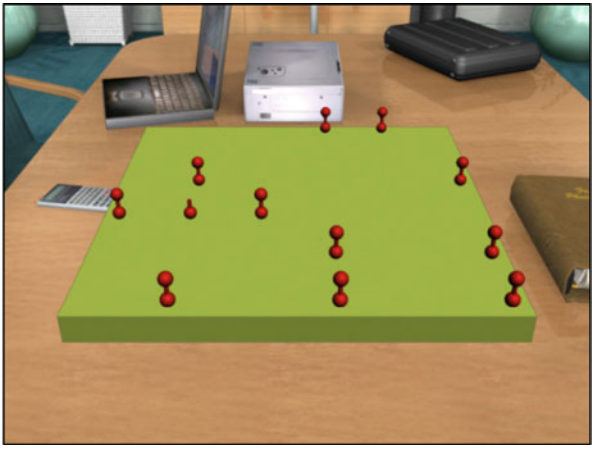


Fig. 2 Example of a hybrid scene stimulus in which a search array is embedded within a scene background. (From Brooks et al. 2010)

identifying the office in which it appears, consistent with the results of Brockmole et al. (2006).

To test this proposed resolution of the conflicting theoretical accounts of contextual cuing, Brooks et al. (2010) embedded search arrays within images of natural scenes (see Fig. 2). During several search blocks, a particular target location was consistently paired with a particular array configuration and with a particular scene background. Participants showed reliable contextual cuing. In a transfer session, we disrupted either the search array (by scrambling the distractor locations) or the scene background (by replacing it with a different background). A purely localist view predicts transfer in the latter condition but not the former. A purely globalist view predicts the reverse. The hierarchical hypothesis predicts transfer in neither: disrupting the array eliminates the functional context within which the target location was learned; disrupting the background blocks the identification of the scene and the retrieval of the appropriate local context. The data confirmed the hierarchical prediction. Evidence of local learning and effects of scene identity can therefore be integrated within a larger view holding that learning is relative to a particular functional search context, but that the retrieval of this contextual information depends on scene recognition. Brooks et al. found that search arrays could be learned in the absence of a consistent background, suggesting that the search context and scene identity are integrated only when there is a consistent pairing between the two.

Guidance from a Target Template

Thus far, we have discussed the guidance of attention and gaze by categorical knowledge and by exemplar memory. Guidance also derives from knowledge of the visual properties of the target object; searches will be more efficient when we know something specific about the appearance of what we are looking for (e.g., Wolfe et al.

2004). Most theories of attentional selection depend on a mechanism that can specify and keep active the features of the to-be-selected item (Bundesen 1990; Desimone and Duncan 1995; Duncan and Humphreys 1989; Wolfe 1994), providing a set of parameters for the guidance of selection. The mechanism of guidance has traditionally been conceived as a *search template* specifying the relevant features that distinguish the target from distractors (e.g., Wolfe 1994). The search template properties must be maintained during the course of search, which can extend over multiple shifts of attention and gaze to individual objects. Clearly, such extended maintenance requires a relatively robust form of memory. VWM is the natural candidate, and several theories of attention assume that VWM is indeed the substrate of the search template (Bundesen et al. 2005; Desimone and Duncan 1995). VWM also provides the type of flexibility needed to configure search in the course of real-world tasks. As goals change, new target information can be rapidly loaded into VWM (Vogel et al. 2006), providing a means to re-set the parameters of search efficiently (e.g., from those specifying the features of the sugar bowl to those specifying the features of the spoon).

Recent research has provided strong confirmatory evidence that VWM plays precisely this role. First, a concurrent VWM load interferes with search efficiency (Woodman et al. 2007), except when the target features repeat over many trials and do not themselves require active maintenance in VWM (Hollingworth and Maxcey-Richard in press; Woodman and Luck 2004). Second, the capability to store multiple objects in VWM is reflected in the ability to establish multiple templates that guide attention simultaneously to different sets of objects (Beck et al. 2012). Third, an electrophysiological correlate of VWM maintenance, the contralateral delay activity (CDA) event-related potential (ERP) component, is sustained across the course of visual search, and the amplitude of the CDA predicts search success (Woodman and Arita 2011).

Finally, and most directly, VWM can automatically bias selection during visual search. Several studies have found that attention is oriented to distractors that match the content of VWM (Han and Kim 2009; Hollingworth et al. (in press); Hollingworth and Luck 2009; Olivers 2009; Olivers et al. 2006; Soto et al. 2005; Soto et al. 2006; Soto et al. 2010). In these experiments, participants maintain a VWM load (typically a color) during a visual search task. The search array is made up of colored items, but the target feature is orthogonal to color. On a subset of trials, a distractor is rendered in the remembered color. Participants know that this object will never contain the target value, yet its presence attracts gaze and slows search. There are some circumstances in which participants can avoid or even reverse this *memory-driven capture*, but these appear to be the exception rather than the rule, occurring when the search task is extremely difficult and when participants have sufficient time to configure a template that excludes the remembered feature value (Downing and Dodds 2004; Han and Kim 2009; Woodman and Luck 2007). The basic memory-driven capture effect suggests that VWM and attentional selection are tightly coupled. VWM modulates the competition among multiple objects for selection (Desimone and Duncan 1995), with attention biased toward items in a scene that match VWM content. In capture studies, this bias impairs performance, but under normative conditions, when the content of VWM is a set of features that define the desired target, the same bias would guide attention toward relevant objects.

VWM guidance plays a central role in a particular form of visual search generated by the simple act of making saccades (Hollingworth et al. 2008). Saccades are notoriously inaccurate. In cluttered, natural environments, when the eyes fail to land on the intended saccade target, there may be multiple objects visible near the landing position. This creates a classic visual search problem: The original saccade target must be located among distractors so as to generate an appropriate corrective saccade. We make tens of thousands of saccades each day, and a significant proportion of these fail to land on the intended object. Search for the original saccade target is almost certainly the most common form of search behavior in which humans engage, and the success of correction will determine the efficiency by which the eyes are ultimately directed to each task-relevant object.

Hollingworth et al. (2008) had participants execute a saccade to one target object in a circular array of colored disks. During the saccade, the array was rotated so that the eyes landed between the target and an equidistant distractor. Gaze correction to the original target required VWM for the target color, because the rotation during the saccade was not directly visible. Gaze correction was accurate, rapid, and automatic, suggesting a fundamental, and potentially low-level, interaction between the content of VWM and saccadic orienting. In addition, gaze correction was impaired by loading VWM with task-irrelevant information, just as found in more traditional search tasks (Woodman et al. 2007). Moreover, if the distractor object matched the content of a secondary VWM load, correction was also impaired (Hollingworth and Luck 2009), consistent with the finding of memory-driven capture in standard search tasks (Olivers et al. 2006; Soto et al. 2005). Thus, a similar guidance mechanism appears to be functional at multiple levels of the search process. VWM guides attention and the eyes to the ultimate target of the search. VWM also guides gaze correction for the individual eye movements that are embedded within the larger search task.

What is the locus of the interaction between VWM and attentional orienting? Recent studies indicate that one form of interaction occurs during the initial sensory processing of objects. This is plausible given evidence that VWM maintenance generates feature-specific activation in early sensory regions of the brain (Harrison and Tong 2009; Serences et al. 2009). Two studies have demonstrated that VWM directly influences perceptual experience: Holding a particular motion direction in VWM modifies motion perception (Kang et al. 2011); holding a particular orientation in VWM modulates the competition between stimuli in binocular rivalry, biasing perception toward the matching orientation (Pearson et al. 2008). In addition, attentional selection on the basis of color (i.e., feature-based selection, which is likely to depend on a VWM template) modulates the P1 ERP component at latencies of $\sim$100 ms post stimulus (Zhang and Luck 2009), indicating an effect on the initial feed-forward sweep of sensory processing. An early sensory interaction is also consistent with the fact that VWM modulates capture by abrupt onsets during search (Mannan et al. 2010).

The effects of VWM on perceptual processing and orienting are observed even in the absence of stimulus competition. In Hollingworth et al. (in press), orienting saccades to single onset targets were influenced by whether the object did or did not match the content of a task-irrelevant VWM load: Saccades to matching targets were generated more rapidly and were more accurate (i.e., landed closer to the center of the target). These effects were observed on saccades with latencies near the limit of human capabilities, between 100 and 150 ms post stimulus (a result that contrasts with theories claiming that rapid saccades are influenced solely by physical salience, Ludwig and Gilchrist 2002; van Zoest et al. 2004). Further, VWM modulated the competition between the target and a task-irrelevant distractor, with the landing position of the saccade strongly biased toward whichever object matched VWM. Thus, even the most elementary operations of the oculomotor system are influenced by the interaction between VWM and perceptual processing, providing a plausible mechanism by which VWM can guide selection during visual search.

These recent findings provide a means to revive theories that explain oculomotor selection during natural search in terms of low-level visual salience (Itti and Koch 2000, 2001; Parkhurst et al. 2002). Visual salience has been modeled on the basis of the physical properties of local scene regions. Physical salience does not correlate strongly with task relevance, nor does it change systematically as goals change. If, as suggested above, visual salience is instead a joint property of an object's physical properties and their match with perceptual features maintained in VWM, then visual salience would be directly related to the features of task-relevant objects (maintained in VWM), and visual salience would vary systematically as goals change (and the content of VWM is modified to reflect the properties of newly relevant objects). This possibility is broadly consistent with models of search in which salience is modulated by the match between local scene regions and the target features maintained in a search template (Kanan et al. 2009). Of course, a modulating effect of VWM on visual salience would not necessarily provide a comprehensive explanation of guidance, as categorical knowledge and scene memory will nonetheless place strong constraints on selection during search through scenes.

Is template-based guidance plausible when searching for objects in natural environments? Natural search introduces significant hurdles to the generation and implementation of a search template. In the traditional visual search literature, the trial usually begins with the presentation of an image of the target, which simplifies the problem of guidance significantly: the target image typically matches the target exactly, containing all of the features of the target, and the perceptual representation of the target image can be loaded efficiently into VWM. But when we are looking for a particular object in a scene, rarely do we get a chance to see what we are looking for before we start. Instead, target features must be retrieved from LTM, which may limit both their precision and the extent to which those features are represented robustly in VWM. In addition, the visual features of an actual object in a scene will vary significantly as a result of idiosyncrasies in orientation, lighting, occlusion, distance, and so forth. It would be an extraordinary coincidence to obtain an exact visual match between the internal representation of the target (retrieved from memory) and any particular object in the scene. Moreover, if one is searching broadly for any object matching a particular category, such as "sugar bowl", variation in the perceptual properties of the exemplars that make up that category will also limit the precision of the search template (Yang and Zelinsky 2009). Wolfe et al. (2011) identified several additional constraints on template use in natural search: (1) natural target objects have multiple feature values on multiple dimensions, and it is not clear that guidance can be implemented simultaneously for a complex set of features; (2) these values are often shared by many other objects in the scene, limiting their discriminative power; and (3) effects of template guidance may be quite small relative to guidance from scene knowledge (e.g., by knowledge of the where the object is likely to be found in the scene).

The literature on guidance in scenes suggests that, despite these limitations, template-based guidance is possible, is not necessarily overshadowed by the effects of scene knowledge, and occurs even when the template must be retrieved from LTM (Bravo and Farid 2009, 2012; Hollingworth and Matsukura 2011; Hwang et al. 2009; Malcolm and Henderson 2009, 2010; Pomplun 2006; Wolfe et al. 2011; Zelinsky 2008; Zelinsky et al. 1997). Malcolm and Henderson (2009) had participants search for objects in photographs of natural scenes, such as a kettle in a kitchen. They were provided either a verbal description of the target or a picture of the target immediately before scene onset (see Schmidt and Zelinsky 2009; Vickery et al. 2005; Wolfe et al. 2004). The availability of a precise visual template in the picture condition reduced the time taken to scan the scene for the target and reduced the time needed to verify a target match. The effect of visual guidance was observed above and beyond any guidance attributable to scene knowledge, as knowledge of the plausible location of a kettle in a kitchen was available in both conditions. Indeed, manipulation of both template precision and contextual constraint (i.e., the extent to which the target location was predictable on the basis of scene knowledge) indicated that the two sources of guidance combine additively (Malcolm and Henderson 2010).

Template-based guidance during natural tasks depends on the retrieval of features from LTM. Bravo and Farid (2009) trained participants to associate five photographs of tropical fish with five different species labels. They then had participants determine whether scenes of coral reefs did or did not contain a fish. Before each search, participants received a specific verbal cue (the label of one of the five species) or a non-specific cue ("fish"). In addition, the actual fish in the scene was either identical to a studied fish, a transformed version of the studied fish (rescaled, rotated, and mirror reversed), or a different exemplar from the same species. The use of verbal labels meant that participants had to retrieve visual details of the studied images from LTM. Relative to a non-specific cue, the specific cue facilitated search equally in the identical and transformed conditions but did not facilitate search at all in the different exemplar condition. That a specific verbal cue facilitated search at all demonstrates guidance from a visual template that was retrieved from LTM. The absence of facilitation in the different exemplar condition suggests that the retrieved template was specific to the particular fish image that had been studied. Because different exemplars of the same species had similar visual properties, the template must have been quite precise, retaining particular visual details of the studied image. However, the equivalence of facilitation in the identical and transformed conditions indicates that template guidance was not necessarily based on low-level image matching. The template generalized over metric differences in size and orientation. As discussed by

Bravo and Farid, such generalization makes a good deal of sense in dynamic, natural search, as the viewer's distance from the target and orientation with respect to the target are subject to change both within a search episode and across search episodes.

Converging evidence that LTM supports visual template guidance comes from Wolfe et al. (2011). In a repeated search paradigm, participants located natural objects in scenes on the basis of a verbal label. The second search for a particular object was much faster than the first search. Much of this improvement was attributable to learning the specific visual features of the target object. On the first search, participants had only a verbal label and did not know anything specific about the visual form of the object exemplar in the scene. Once the target had been found the first time, however, participants were able to encode its visual form into LTM. On the second search, retrieval of the target's features from LTM provided substantial guidance, improving search efficiency. This effect was observed above that attributable to learning the locations of the target objects.

Does retrieval from LTM lead to active maintenance of the search template in VWM? If so, then we can establish a bridge between the literature on VWM guidance in traditional search tasks and template guidance while searching for natural objects in scenes. In the memory-driven capture literature, capture is observed even if the memory item is specified by a label, such as "red" (Mannan et al. 2010; Soto and Humphreys 2007). In this case, activation in VWM is presumably driven by retrieval from LTM. Similarly, Olivers (2011) found that long-term perceptual associations for an item maintained in VWM influence search. Participants were given a verbal description of a canonically colored item (such as "stop sign"). They then searched for a grayscale version this item within an object array. Significantly greater distraction was generated by a colored distractor that matched the canonical color of the target (in this case, red) versus a colored distractor that did not, despite the fact that the colored distractor was never the target. Again, the feature value associated with the target must have been retrieved from LTM, and the capture effect indicates an interaction between the sustained activation of that feature in VWM and perceptual processing of the search display. Thus, it seems quite plausible that template features retrieved from LTM in the course of natural search would be represented in VWM and would interact with selection to increase the perceptual salience of scene regions matching VWM content.

Formal theories of search through naturalistic environments differ in their dependence on a search template. The original salience map models (Itti and Koch 2000; Parkhurst et al. 2002) contain no target representation and thus did not implement guidance on the basis of target features (but see Navalpakkam and Itti 2005). The contextual guidance model (Torralba et al. 2006) adds guidance by knowledge of plausible target locations to the salience-map account, but it likewise does not implement guidance by a search template (but see Ehinger et al. 2009). In contrast, several models of natural search depend centrally on a search template without introducing constraints from scene knowledge (Hwang et al. 2009; Zelinsky 2008). For example, in Zelinsky's target acquisition model, the search template, maintained in memory, is a multidimensional feature vector derived from an exact image representation of the region of the scene containing the target. During search, this feature vector is correlated with feature vectors computed at each pixel location within the search image (the scene vectors are strongly influenced by fixation position within the scene and thus vary with each eye movement). The "salience" of each location in the search image is then a function of the extent to which features present in the template are present at that location and features absent from the template are absent from that location. Saccades are directed to regions of high correlation.

The target acquisition model (Zelinsky 2008) has been highly adept at capturing key features of human eye movements during search through scenes. Its assumption that a memory representation of the target features is combined with perceptual information to bias selection toward matching regions of the scene is consistent with the general claims of VWM-based guidance discussed above. However, the type of template specified in the target acquisition model and in other image-based models (Hwang et al. 2009; Pomplun 2006) would be difficult to integrate with this view. VWM representations are not imagistic; rather, they are abstracted away from the precise, metric structure of early vision (Irwin 1991; Phillips 1974). The only visual memory system capable of maintaining the type of low-level image representation specified in these models is sensory persistence (Averbach and Coriell 1961; Coltheart 1980; Sperling 1960), but sensory persistence decays within less than 500 ms of a perceptual event, so it could not plausibly support template maintenance during natural search. In addition, templates will rarely be derived from precise image representations, as they will need to be retrieved from LTM, and LTM representations, like VWM representations, are not imagistic (for a review, see Hollingworth 2008). Instead, template-guidance is likely to depend on visual representations that are abstracted away from the precise metric properties of early vision, allowing generalization over image-level variation within and across searches (e.g., Bravo and Farid 2009).

Online Memory for the Search: Inhibition of Return

The efficiency of search for objects in scenes would be improved by a mechanism that prevents attention and gaze from returning to previously fixated objects. Posner and Cohen (1984) discovered that target detection at a peripherally cued location is impaired starting approximately 300 ms after a neutral cue. This inhibition of return (IOR) effect has been taken to suggest that attention is biased against returning to a previously attended location (but see Berlucchi 2006; Posner and Cohen 1984 for alternativbe accounts). Salience models of visual search make this assumption explicit, with an IOR mechanism that deprioritizes previously fixated regions so that gaze does not cycle continuously between the most visually salient regions of the scene (Itti and Koch 2001). More generally, IOR has been proposed to act as a foraging facilitator in complex environments, ensuring that attention is consistently oriented to new locations that might contain the target (for a review, see Wang and Klein 2010).

The most direct evidence that IOR plays a functional role in the task of finding objects in scenes comes from a study by Klein and MacInnes (1999). Participants completed a very difficult visual search task: Finding well-camouflaged Waldos in “Where’s Waldo?” images. At a pre-specified point during search, a dot onset was introduced, and participants executed a saccade to the onset dot as rapidly as possible. Oculomotor IOR was observed as increased saccade latency to targets that appeared in the reverse direction from the preceding saccade (i.e., back to a previously fixated location) relative to targets that appeared in the same direction as the preceding saccade (i.e., toward a new fixation location). This basic finding has been replicated several times (Dodd et al. 2009; Smith and Henderson 2011a, b). In addition, the duration of a fixation before a saccade to a new location tends to be shorter than the duration preceding a saccade that returns gaze to a just-fixated location (Hooge et al. 2005).

A functional account of IOR is bolstered by two additional lines of evidence. First, oculomotor IOR is strongly tied to the particular scene context in which the search occurs. If the scene background is removed immediately before the onset dot, the latency deficit for saccades to previously fixated locations is eliminated (Klein and MacInnes 1999; see also Müller and Von Mühlenen 2000; Takeda and Yagi 2000). Second, the presence of IOR is strongly tied to the demands of the task. Dodd et al. (2009) modified the Klein and MacInnes task with a manipulation of viewing instructions. Participants either searched for a small letter target within scenes, memorized the scenes, assessed the aesthetics of the scenes, or viewed freely. Orienting latency to onset targets was measured at previously fixated locations and new locations. Only the search task placed a premium on inhibiting orienting to previously fixated locations, and only in the search task was IOR observed. Saccades to old locations were actually facilitated under the other task conditions, a phenomenon of *facilitation of return*. Thus, IOR appears to be tied directly to the demands of visual search.

The phenomenon of oculomotor IOR—increased saccadic latency to targets at previously fixated locations—is well established, but its functional relevance to natural search depends on whether this effect reduces the probability that a scene location will be refixated under free viewing conditions (Hooge et al. 2005). That is, the latency deficit must be indicative of a spatial bias against directing saccades back to previously fixated regions (i.e., making refixations). Such biases certainly exist. Refixations during visual search through scenes are generally rare, much lower than would be expected by chance (Gilchrist et al. 2001; for similar results using abstract arrays, see Peterson et al. 2001). Moreover, refixations are less frequent during visual search than during non-search tasks (Dodd et al. 2009). However, it has not been demonstrated that these large-scale biases against refixation during search are caused by the same IOR mechanism that produces latency differences, and several recent studies suggest that they may not be.

At question is whether the increased latency of return saccades (that reverse a saccade just executed) is accompanied by a reduction in the probability that a return saccade will be executed. Several studies have now found that return saccades are generated more often than would be expected by chance, indicating a bias to direct gaze *back to*, rather than *away from*, the most recently fixated item, and researchers have interpreted this finding as evidence against the idea that IOR interacts significantly with the selection of saccade locations to facilitate foraging (Hooge et al. 2005; Smith and Henderson 2009, 2011a, b).[1] The relatively high frequency of return saccades eliminates the possibility that IOR dominates other sources of constraint to ensure that the eyes are *always* directed to new objects. It is, however, possible that an immediate return saccade is a special case. Return saccades are quite likely to reflect incomplete processing of the target during the initial fixation. Gaze is returned relatively rapidly (Hooge et al. 2005) to the object so as to compete the acquisition of information. This does not violate the general goal that saccades should be biased away from previously examined objects, because the immediate return can be considered as completion of the process of initial examination. Diagnostic evidence would come from objects fixated two or more fixations earlier, when the processing of the fixated object has clearly been completed. Smith and Henderson (2011b) found that the probability of return to a 2-back location was not reduced relative to other locations within the scene, and thus there is currently no evidence that IOR facilitates foraging by reducing the probability of refixation.

It is still the case that over the entire course of search, refixations are relatively rare. There must be *some* memory mechanism that implements this behavioral bias. Although IOR might contribute, it is unlikely to be the only mechanism. Oculomotor IOR is observed for, at most, four preceding fixations (Dodd et al. 2009) and appears to depend on the VWM system (Castel et al. 2003), which has a severely limited capacity. If IOR is found, ultimately, to cause reduced refixation probability, that bias could be operational for only a few of objects over a couple of seconds. Yet, real-world searches often unfold over many seconds of exploration. A bias to avoid refixation across extended search cannot plausibly depend on IOR alone. Instead, it is likely that there are longer-term learning mechanisms that serve this purpose. The memory mechanisms supporting longer-term biases during search through scenes have received no systematic study whatsoever.

In summary, it still remains possible that IOR plays a functional role in the task of finding objects in scenes. But there is currently no positive evidence that IOR is associated with a reduced probability of refixation, and tests of the immediate return of gaze suggest that return is actually more likely than expected by chance. Without evidence of a bias against refixations, there is currently no direct link between the IOR phenomenon and the facilitation of search within scenes, and thus no direct support for the foraging facilitator hypothesis. Moreover, any influence of IOR on the selection of saccade targets must be limited, as it does not appear to prevent immediate returns, and the temporal range of IOR is quite brief (four fixations at most). Biases against refixation during extended search events must therefore depend on longer-term memory mechanisms.

---

[1] Smith and Henderson (2009) use the term "facilitation of return" to refer to the finding that return saccades are observed more frequently than expected by chance. Dodd et al. (2009) use "facilitation of return" to refer to the finding that, in non-search tasks, saccades are generated more quickly to abrupt onsets at previously fixated locations than at new locations.

Conclusions

Visual search is one of the most common behaviors in which humans engage. When looking for objects in scenes, multiple sources of guidance are used to direct attention and gaze to the target location efficiently, including categorical knowledge of plausible object positions in scenes, memory for the particular environment in which search occurs, and knowledge of the perceptual features that characterize the desired object. In addition, online memory for inspected locations is likely to improve efficiency by biasing attention and gaze to new locations. These various sources of guidance appear to be complementary (Castelhano and Heaven 2010; Ehinger et al. 2009; Malcolm and Henderson 2010; Wolfe et al. 2011), and future work will need to account for the manner in which they jointly contribute to search efficiency.

References

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal, 40*(1), 309–328.

Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience, 7*(1), 66–80. doi: 10.1162/jocn.1995.7.1.66.

Ballard, D. H., Hayhoe, M. M., Pook, P. K., & Rao, R. P. (1997). Deictic codes for the embodiment of cognition. *Behavioral and Brain Sciences, 20*(4), 723–767.

Beck, V. M., Hollingworth, A., & Luck, S. J. (2012). Simultaneous control of attention by multiple working memory representations. *Psychological Science*, *23*(8), 887–898. doi: 10.1177/0956797612439068

Becker, M. W., & Rasmussen, I. P. (2008). Guidance of attention to objects and locations by long-term memory of natural scenes. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(6), 1325–1338. doi: 10.1037/a0013650.

Berlucchi, G. (2006). Inhibition of return: A phenomenon in search of a mechanism and a better name. *Cognitive Neuropsychology, 23*(7), 1065–1074. doi: 10.1080/02643290600588426.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*(2), 143–177. doi: 10.1016/0010-0285(82)90007-X.

Brady, T. F., & Chun, M. M. (2007). Spatial constraints on learning in visual search: Modeling contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 798–815. doi: 10.1037/0096-1523.33.4.798.

Brady, T. F., Konkle, T., Alvarez, G. A., & Oliva, A. (2008). Visual long-term memory has a massive storage capacity for object details. *Proceedings of the National Academy of Sciences of the United States of America, 105*(38), 14325–14329. doi: 10.1073/pnas.0803390105.

Bravo, M. J., & Farid, H. (2009). The specificity of the search template. *Journal of Vision, 9*(1), 34, 31–39. doi: 10.1167/9.1.34.

Bravo, M. J., & Farid, H. (2012). Task demands determine the specificity of the search template. *Attention, Perception, & Psychophysics, 74*(1), 124–131. doi: 10.3758/s13414-011-0224-5.

Brockmole, J. R., Castelhano, M. S., & Henderson, J. M. (2006). Contextual cueing in naturalistic scenes: Global and local contexts. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*(4), 699–706. doi: 10.1037/0278-7393.32.4.699.

Brockmole, J. R., & Henderson, J. M. (2006a). Recognition and attention guidance during contextual cueing in real-world scenes: Evidence from eye movements. *Quarterly Journal of Experimental Psychology, 59*(7), 1177–1187. doi: 10.1080/17470210600665996.

Brockmole, J. R., & Henderson, J. M. (2006b). Using real-world scenes as contextual cues for search. *Visual Cognition, 13*(1), 99–108. doi: 10.1080/13506280500165188.

Brooks, D. I., Rasmussen, I. P., & Hollingworth, A. (2010). The nesting of search contexts within natural scenes: Evidence from contextual cuing. *Journal of Experimental Psychology: Human Perception and Performance, 36*(6), 1406–1418. doi: 10.1037/a0019257.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97*(4), 523–547. doi: 10.1037/0033-295X.97.4.523.

Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention: Bridging cognition and neurophysiology. *Psychological Review, 112*(2), 291–328. doi: 10.1037/0033-295x.112.2.291.

Castel, A. D., Pratt, J., & Craik, F. I. M. (2003). The role of spatial working memory in inhibition of return: Evidence from divided attention tasks. *Perception & Psychophysics, 65*(6), 970–981. doi: 10.3758/BF03194827.

Castelhano, M. S., & Heaven, C. (2010). The relative contribution of scene context and target features to visual search in scenes. *Attention, Perception, & Psychophysics, 72*(5), 1283–1297. doi: 10.3758/app.72.5.1283.

Castelhano, M. S., & Heaven, C. (2011). Scene context influences without scene gist: Eye movements guided by spatial associations in visual search. *Psychonomic Bulletin & Review, 18*(5), 890–896. doi: 10.3758/s13423-011-0107-8.

Castelhano, M. S., & Henderson, J. M. (2005). Incidental visual memory for objects in scenes. *Visual Cognition, 12*(6), 1017–1040. doi: 10.1080/13506280444000634.

Castelhano, M. S., & Henderson, J. M. (2007). Initial scene representations facilitate eye movement guidance in visual search. *Journal of Experimental Psychology: Human Perception and Performance, 33*(4), 753–763. doi: 10.1037/0096-1523.33.4.753.

Chun, M. M. (2000). Contextual cueing of visual attention. *Trends in Cognitive Sciences, 4*(5), 170–178. doi: 10.1016/S1364-6613(00)01476-5.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36*(1), 28–71. doi: 10.1006/cogp.1998.0681.

Chun, M. M., & Turk-Browne, N. B. (2008). Associative learning mechanisms in vision. In S. J. Luck, & A. Hollingworth (Eds.), *Visual memory* (pp. 209–246). New York: Oxford University Press.

Coltheart, M. (1980). The persistences of vision. *Philosophical Transactions of the Royal Society B: Biological Sciences, 290*(1038), 269–294 (1980). doi: 10.1098/rstb.1980.0082.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18*, 193–222. doi: 10.1146/annurev.ne.18.030195.001205.

Dodd, M. D., Van der Stigchel, S., & Hollingworth, A. (2009). Novelty is not always the best policy: Inhibition of return and facilitation of return as a function of visual task. *Psychological Science, 20*(3), 333–339. doi: 10.1111/j.1467-9280.2009.02294.x.

Downing, P. E., & Dodds, C. M. (2004). Competition in visual working memory for control of search. *Visual Cognition, 11*(6), 689–703. doi: 10.1080/13506280344000446.

Droll, J. A., Hayhoe, M. M., Triesch, J., & Sullivan, B. T. (2005). Task demands control acquisition and storage of visual information. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1416–1438. doi: 10.1037/0096-1523.31.6.1416.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96*(3), 433–458. doi: 10.1037//0033-295X.96.3.433.

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision, 11*(5), 14: 11–36. doi: 10.1167/11.5.14.

Eckstein, M. P., Drescher, B. A., & Shimozaki, S. S. (2006). Attentional cues in real scenes, saccadic targeting, and Bayesian priors. *Psychological Science, 17*(11), 973–980. doi: 10.1111/j.1467-9280.2006.01815.x.

Ehinger, K. A., & Brockmole, J. R. (2008). The role of color in visual search in real-world scenes: Evidence from contextual cuing. *Perception & Psychophysics, 70*(7), 1366–1378. doi: 10.3758/pp.70.7.1366.

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. doi: 10.1080/13506280902834720.

Einhauser, 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: 1–19. doi: 10.1167/8.2.2.

Einhauser, W., Spain, M., & Perona, P. (2008b). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14), 18: 11–26. doi: 10.1167/8.14.18.

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

Foulsham, T., Alan, R., & Kingstone, A. (2011). Scrambled eyes? Disrupting scene structure impedes focal processing and increases bottom-up guidance. *Attention, Perception, & Psychophysics, 73*(7), 2008–2025. doi: 10.3758/s13414-011-0158-y.

Foulsham, T., & Underwood, G. (2007). How does the purpose of inspection influence the potency of visual salience in scene perception? *Perception, 36*(8), 1123–1138. doi: 10.1068/p5659.

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. doi: 10.1167/8.2.6.

Gibson, B. S., Li, L., Skow, E., Brown, K., & Cooke, L. (2000). Searching for one versus two identical targets: When visual search has a memory. *Psychological Science, 11*(4), 324–327. doi: 10.1111/1467-9280.00264.

Gilchrist, I. D., North, A., & Hood, B. (2001). Is visual search really like foraging? *Perception, 30*(12), 1459–1464. doi: 10.1068/p3249.

Han, S. W., & Kim, M. S. (2009). Do the contents of working memory capture attention? Yes, but cognitive control matters. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1292–1302. doi: 10.1037/a0016452.

Harrison, S. A., & Tong, F. (2009). Decoding reveals the contents of visual working memory in early visual areas. *Nature, 458*(7238), 632–635. doi: 10.1038/nature07832.

Hayhoe, M. (2000). Vision using routines: A functional account of vision. *Visual Cognition, 7*(1–3), 43–64. doi: 1080/135062800394676.

Henderson, J. M. (2003). Human gaze control during real-world scene perception. *Trends in Cognitive Sciences, 7*(11), 498–504. doi: 10.1016/j.tics.2003.09.006.

Henderson, J. M., Brockmole, J. R., Castelhano, M. S., & Mack, M. (2007). Visual saliency does not account for eye movements during search in real-world scenes. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 537–562). Oxford: Elsevier.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50,* 243–271. doi: 10.1146/annurev.psych.50.1.243.

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. doi: 10.3758/pbr.16.5.850.

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25*(1), 210–228. doi: 10.1037//0096-1523.25.1.210.

Hillstrom, A. P., Scholey, H., Liversedge, S. P., & Benson, V. (2012). The effect of the first glimpse at a scene on eye movements during search. *Psychonomic Bulletin & Review, 19*(2), 204–210. doi: 10.3758/s13423-011-0205-7.

Hollingworth, A. (2004). Constructing visual representations of natural scenes: The roles of short- and long-term visual memory. *Journal of Experimental Psychology: Human Perception and Performance, 30*(3), 519–537. doi: 10.1037/0096-1523.30.3.519.

Hollingworth, A. (2005). The relationship between online visual representation of a scene and long-term scene memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*(3), 396–411. doi: 10.1037/0278-7393.31.3.396.

Hollingworth, A. (2006). Visual memory for natural scenes: Evidence from change detection and visual search. *Visual Cognition, 14*(4–8), 781–807. doi: 10.1080/13506280500193818.

Hollingworth, A. (2008). Visual memory for natural scenes. In S. J. Luck, & A. Hollingworth (Eds.), *Visual memory* (pp. 123–162). New York: Oxford University Press.

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition, 17*(1–2), 273–291. doi: 10.1080/13506280802193367.

Hollingworth, A. (in press). Task specificity and the influence of memory on visual search: Commentary on Võ and Wolfe (2012). *Journal of Experimental Psychology: Human Perception and Performance*.

Hollingworth, A., & Henderson, J. M. (2002). Accurate visual memory for previously attended objects in natural scenes. *Journal of Experimental Psychology: Human Perception and Performance, 28*(1), 113–136. doi: 10.1037//0096-1523.28.1.113.

Hollingworth, A., & Luck, S. J. (2009). The role of visual working memory (VWM) in the control of gaze during visual search. *Attention, Perception, & Psychophysics, 71*(4), 936–949. doi: 10.3758/APP.71.4.936.

Hollingworth, A., & Matsukura, M. (2011). *Visual working memory content biases the allocation of gaze during search through natural scenes*. Paper presented at the Annual Meeting of the Psychonomic Society, Seattle, WA.

Hollingworth, A., Matsukura, M., & Luck, S. J. (in press). Visual working memory modulates rapid eye movements to simple onset targets. *Psychological Science*.

Hollingworth, A., & Maxcey-Richard, A. M. (in press). Selective maintenance in visual working memory does not require sustained visual attention. *Journal of Experimental Psychology: Human Perception and Performance*.

Hollingworth, A., Richard, A. M., & Luck, S. J. (2008). Understanding the function of visual short-term memory: Transsaccadic memory, object correspondence, and gaze correction. *Journal of Experimental Psychology: General, 137*(1), 163–181. doi: 10.1037/0096-3445.137.1.163.

Hooge, I. T. C., Over, E. A. B., van Wezel, R. J. A., & Frens, M. A. (2005). Inhibition of return is not a foraging facilitator in saccadic search and free viewing. *Vision Research, 45*(14), 1901–1908. doi: 10.1016/j.visres.2005.01.030.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(6693), 575–577. doi: 10.1038/29068.

Howard, C. J., Pharaon, R. G., Körner, C., Smith, A. D., & Gilchrist, I. D. (2011). Visual search in the real world: Evidence for the formation of distractor representations. *Perception, 40*(10), 1143–1153. doi: 10.1068/p7088.

Hwang, A. D., Higgins, E. C., & Pomplun, M. (2009). A model of top-down attentional control during visual search in complex scenes. *Journal of Vision, 9*(5), 25: 21–18. doi: 10.1167/9.5.25.

Irwin, D. E. (1991). Information integration across saccadic eye movements. *Cognitive Psychology, 23*(3), 420–456. doi: 10.1016/0010-0285(91)90015-G.

Itti, L. (2005). Quantifying the contribution of low-level saliency to human eye movements in dynamic scenes. *Visual Cognition, 12*(6), 1093–1123. doi: 10.1080/13506280444000661.

Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research, 40*(10–12), 1489–1506. doi: 10.1016/S0042-6989(99)00163-7.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2*(3), 194–203. doi: 10.1038/35058500.

Kanan, C., Tong, M. H., Zhang, L., & Cottrell, G. W. (2009). SUN: Top-down saliency using natural statistics. *Visual Cognition, 17*(6–7), 979–1003. doi: 10.1080/13506280902771138.

Kang, M. S., Hong, S. W., Blake, R., & Woodman, G. F. (2011). Visual working memory contaminates perception. *Psychonomic Bulletin & Review, 18*(5), 860–869. doi: 10.3758/s13423-011-0126-5.

Klein, R. M. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334*(6181), 430–431. doi: 10.1038/334430a0.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science, 10*(4), 346–352. doi: 10.1111/1467-9280.00166.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21*(11), 1551–1556. doi: 10.1177/0956797610385359.

Körner, C., & Gilchrist, I. D. (2007). Finding a new target in an old display: Evidence for a memory recency effect in visual search. *Psychonomic Bulletin & Review, 14*(5), 846–851. doi: 10.3758/bf03194110.

Kristjánsson, A. (2000). In search of remembrance: Evidence for memory in visual search. *Psychological Science, 11*(4), 328–332. doi: 10.1111/1467-9280.00265.

Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Research, 41*(25–26), 3559–3565. doi: 10.1016/S0042-6989(01)00102-X.

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. doi: 10.1068/p2935.

Ludwig, C. J. H., & Gilchrist, I. D. (2002). Stimulus-driven and goal-driven control over visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 28*(4), 902–912. doi: 10.1037//0096-1523.28.4.902.

Mack, S. C., & Eckstein, M. P. (2011). Object co-occurrence serves as a contextual cue to guide and facilitate visual search in a natural viewing environment. *Journal of Vision, 11*(9), 9: 1–16. doi: 10.1167/11.9.9.

Malcolm, G. L., & Henderson, J. M. (2009). The effects of target template specificity on visual search in real-world scenes: Evidence from eye movements. *Journal of Vision, 9*(11), 8: 1–13. doi: 10.1167/9.11.8.

Malcolm, G. L., & Henderson, J. M. (2010). Combining top-down processes to guide eye movements during real-world scene search. *Journal of Vision, 10*(2), 4: 1–11. doi: 10.1167/10.2.4.

Mannan, S. K., Kennard, C., Potter, D., Pan, Y., & Soto, D. (2010). Early oculomotor capture by new onsets driven by the contents of working memory. *Vision Research, 50*(16), 1590–1597. doi: 10.1016/j.visres.2010.05.015.

McCarley, J. S., Wang, R. X. F., Kramer, A. F., Irwin, D. E., & Peterson, M. S. (2003). How much memory does oculomotor search have? *Psychological Science, 14*(5), 422–426. doi: 10.1111/1467-9280.01457.

McPeek, R. M., & Keller, E. L. (2002). Superior colliculus activity related to concurrent processing of saccade goals in a visual search task. *Journal of Neurophysiology, 87*(4), 1805–1815. doi: 10.1152/jn.00501.2001.

Müller, H. J., & Von Mühlenen, A. (2000). Probing distractor inhibition in visual search: Inhibition of return. *Journal of Experimental Psychology: Human Perception and Performance, 26*(5), 1591–1605. doi: 10.1037/0096-1523.26.5.1591.

Navalpakkam, V., & Itti, L. (2005). Modeling the influence of task on attention. *Vision Research, 45*(2), 205–231. doi: 10.1016/j.visres.2004.07.042.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research, 46*(5), 614–621. doi: 10.1016/j.visres.2005.08.025.

Olivers, C. N. L. (2009). What drives memory-driven attentional capture? The effects of memory type, display type, and search type. *Journal of Experimental Psychology: Human Perception and Performance, 35*(5), 1275–1291. doi: 10.1037/a0013896.

Olivers, C. N. L. (2011). Long-term visual associations affect attentional guidance. *Acta Psychologica, 137*(2), 243–247. doi: 10.1016/j.actpsy.2010.07.001.

Olivers, C. N. L., Meijer, F., & Theeuwes, J. (2006). Feature-based memory-driven attentional capture: Visual working memory content affects visual attention. *Journal of Experimental Psychology: Human Perception and Performance, 32*(5), 1243–1265. doi: 10.1037/0096-1523.32.5.1243.

Olson, I. R., & Chun, M. M. (2002). Perceptual constraints on implicit learning of spatial context. *Visual Cognition, 9*(3), 273–302. doi: 10.1080/13506280042000162.

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. doi: 10.1016/S0042-6989(01)00250-4.

Pearson, J., Clifford, C. W. G., & Tong, F. (2008). The functional impact of mental imagery on conscious perception. *Current Biology, 18*(13), 982–986. doi: 10.1016/j.cub.2008.05.048.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12*(4), 287–292. doi: 10.1111/1467-9280.00353.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception & Psychophysics, 16*(2), 283–290. doi: 10.3758/BF03203943.

Pomplun, M. (2006). Saccadic selectivity in complex visual search displays. *Vision Research, 46*(12), 1886–1900. doi: 10.1016/j.visres.2005.12.003.

Posner, M. I., & Cohen, Y. (1984). Components of visual orienting. In H. Bouma, & D. G. Bouwhuis (Eds.), *Attention and performance X* (pp. 531–556). Hillsdale: Erlbaum.

Schmidt, J., & Zelinsky, G. J. (2009). Search guidance is proportional to the categorical specificity of a target cue. *Quarterly Journal of Experimental Psychology, 62*(10), 1904–1914. doi: 10.1080/17470210902853530.

Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information-processing: 1. Detection, search, and attention. *Psychological Review, 84*(1), 1–66. doi: 10.1037/0033-295x.84.1.1.

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. doi: 10.1111/j.1467-9280.2009.02276.x.

Shore, D. I., & Klein, R. M. (2000). On the manifestations of memory in visual search. *Spatial Vision, 14*(1), 59–75. doi: 10.1163/156856801741369.

Smith, T. J., & Henderson, J. M. (2009). Facilitation of return during scene viewing. *Visual Cognition, 17*(6–7), 1083–1108. doi: 10.1080/13506280802678557.

Smith, T. J., & Henderson, J. M. (2011a). Does oculomotor inhibition of return influence fixation probability during scene search? *Attention, Perception, & Psychophysics, 73*(8), 2384–2398 (2011a). doi: 10.3758/s13414-011-0191-x.

Smith, T. J., & Henderson, J. M. (2011b). Looking back at Waldo: Oculomotor inhibition of return does not prevent return fixations. *Journal of Vision, 11*(1), 3: 1–11 (2011b). doi: 10.1167/11.1.3.

Soto, D., Heinke, D., Humphreys, G. W., & Blanco, M. J. (2005). Early, involuntary top-down guidance of attention from working memory. *Journal of Experimental Psychology: Human Perception and Performance, 31*(2), 248–261. doi: 10.1037/0096-1523.31.2.248.

Soto, D., & Humphreys, G. W. (2007). Automatic guidance of visual attention from verbal working memory. *Journal of Experimental Psychology: Human Perception and Performance, 33*(3), 730–737. doi: 10.1037/0096-1523.33.3.730.

Soto, D., Humphreys, G. W., & Heinke, D. (2006). Working memory can guide pop-out search. *Vision Research, 46*(6–7), 1010–1018. doi: 10.1016/j.visres.2005.09.008.

Soto, D., Wriglesworth, A., Bahrami-Balani, A., & Humphreys, G. W. (2010). Working memory enhances visual perception: Evidence from signal detection analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*(2), 441–456. doi: 10.1037/a0018686.

Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs, 74*, (11, Whole no. 498).

Standing, L. (1973). Learning 10,000 pictures. *Quarterly Journal of Experimental Psychology, 25*(2), 207–222. doi: 10.1080/14640747308400340.

Standing, L., Conezio, J., & Haber, R. N. (1970). Perception and memory for pictures: Single-trial learning of 2500 visual stimuli. *Psychonomic Science, 19*(2), 73–74.

Takeda, Y., & Yagi, A. (2000). Inhibitory tagging in visual search can be found if search stimuli remain visible. *Perception & Psychophysics, 62*(5), 927–934 (2000). doi: 10.3758/BF03212078.

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: 1–23. doi: 10.1167/11.5.5.

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, 113*(4), 766–786. doi: 10.1037/0033-295X.113.4.766.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12*(1), 97–136. doi: 10.1016/0010-0285(80)90005-5.

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*(6–7), 812–834. doi: 10.1080/13506280902771278.

van Zoest, W., Donk, M., & Theeuwes, J. (2004). The role of stimulus-driven and goal-driven control in saccadic visual selection. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 746–759. doi: 10.1037/0096-1523.30.4.746.

Vickery, T. J., King, L. W., & Jiang, Y. (2005). Setting up the target template in visual search. *Journal of Vision, 5*(1), 81–92. doi: 10.1167/5.1.8.

Võ, M. L. H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision, 10*(3), 14: 11–13. doi: 10.1167/10.3.14.

Võ, M. L. H., & Wolfe, J. M. (2012). When Does Repeated Search in Scenes Involve Memory? Looking At Versus Looking For Objects in Scenes. *Journal of Experimental Psychology: Human Perception and Performance, 38*(1), 23–41. doi: 10.1037/a0024147.

Vogel, E. K., Woodman, G. F., & Luck, S. J. (2006). The time course of consolidation in visual working memory. *Journal of Experimental Psychology: Human Perception and Performance, 32*(6), 1436–1451. doi: 10.1037/0096-1523.32.6.1436.

von Mühlenen, A., Müller, H. J., & Müller, D. (2003). Sit-and-wait strategies in dynamic visual search. *Psychological Science, 14*(4), 309–314. doi: 10.1111/1467-9280.14441.

Wang, Z., & Klein, R. M. (2010). Searching for inhibition of return in visual search: A review. *Vision Research, 50*(2), 220–228. doi: 10.1016/j.visres.2009.11.013.

Weierich, M. R., Treat, T. A., & Hollingworth, A. (2008). Theories and measurement of visual attentional processing in anxiety. *Cognition & Emotion, 22*(6), 985–1018. doi: 10.1080/02699930701597601.

Williams, C. C., Henderson, J. M., & Zacks, R. T. (2005). Incidental visual memory for targets and distractors in visual search. *Perception & Psychophysics, 67*(5), 816–827. doi: 10.3758/BF03193535.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238. doi: 10.3758/bf03200774.

Wolfe, J. M. (1999). Inattentional amnesia. In V. Coltheart (Ed.), *Fleeting memories* (pp. 71–94). Cambridge: MIT Press.

Wolfe, J. M., Alvarez, G. A., Rosenholtz, R., Kuzmova, Y. I., & Sherman, A. M. (2011). Visual search for arbitrary objects in real scenes. *Attention, Perception, & Psychophysics, 73*(6), 1650–1671 (2011). doi: 10.3758/s13414-011-0153-3.

Wolfe, J. M., Horowitz, T. S., Kenner, N., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*(12), 1411–1426. doi: 10.1016/j.visres.2003.11.024.

Wolfe, J. M., Klempen, N., & Dahlen, K. (2000). Postattentive vision. *Journal of Experimental Psychology: Human Perception and Performance, 26*(2), 693–716. doi: 10.1037/0096-1523.26.2.693.

Wolfe, J. M., Võ, M. L., Evans, K. K., & Greene, M. R. (2011). Visual search in scenes involves selective and nonselective pathways. *Trends in Cognitive Sciences, 15*(2), 77–84. doi: 10.1016/j.tics.2010.12.001.

Woodman, G. F., & Arita, J. T. (2011). Direct electrophysiological measurement of attentional templates in visual working memory. *Psychological Science, 22*(2), 212–215. doi: 10.1177/0956797610395395.

Woodman, G. F., & Luck, S. J. (2004). Visual search is slowed when visuospatial working memory is occupied. *Psychonomic Bulletin & Review, 11*(2), 269–274. doi: 10.3758/BF03196569.

Woodman, G. F., & Luck, S. J. (2007). Do the contents of visual working memory automatically influence attentional selection during visual search? *Journal of Experimental Psychology: Human Perception and Performance, 33*(2), 363–377. doi: 10.1037/0096-1523.33.2.363.

Woodman, G. F., Luck, S. J., & Schall, J. D. (2007). The role of working memory representations in the control of attention. *Cerebral Cortex, 17*, 118–124. doi: 10.1093/cercor/bhm065.

Yang, H., & Zelinsky, G. J. (2009). Visual search is guided to categorically-defined targets. *Vision Research, 49*(16), 2095–2103. doi: 10.1016/j.visres.2009.05.017.

Zelinsky, G. J. (1996). Using eye saccades to assess the selectivity of search movements. *Vision Research, 36*(14), 2177–2187. doi: 10.1016/0042-6989(95)00300-2.

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*(4), 787–835. doi: 10.1037/a0013118.

Zelinsky, G. J., Rao, R. P. N., Hayhoe, M. M., & Ballard, D. H. (1997). Eye movements reveal the spatiotemporal dynamics of visual search. *Psychological Science, 8*(6), 448–453. doi: 10.1111/j.1467-9280.1997.tb00459.x.

Zelinsky, G. J., Zhang, W., Yu, B., Chen, X., & Samaras, D. (2006). The role of top-down and bottom-up processes in guiding eye movements during visual search. In Y. Weiss, B. Scholkopf, & J. Platt (Eds.), *Advances in neural information processing systems* (Vol. 18, pp. 1569–1576). Cambridge: MIT Press.

Zhang, W., & Luck, S. J. (2009). Feature-based attention modulates feedforward visual processing. *Nature Neuroscience, 12*(1), 24–25 (2009). doi: 10.1038/nn.2223.

Reward and Attentional Control in Visual Search

Steven Yantis, Brian A. Anderson, Emma K. Wampler
and Patryk A. Laurent

Abstract It has long been known that the control of attention in visual search depends both on voluntary, top-down deployment according to context-specific goals, and on involuntary, stimulus-driven capture based on the physical conspicuity of perceptual objects. Recent evidence suggests that pairing target stimuli with reward can modulate the voluntary deployment of attention, but there is little evidence that reward modulates the involuntary deployment of attention to task-irrelevant distractors. We report several experiments that investigate the role of reward learning on attentional control. Each experiment involved a training phase and a test phase. In the training phase, different colors were associated with different amounts of monetary reward. In the test phase, color was not task-relevant and participants searched for a shape singleton; in most experiments no reward was delivered in the test phase. We first show that attentional capture by physically salient distractors is magnified by a previous association with reward. In subsequent experiments we demonstrate that physically inconspicuous stimuli previously associated with reward capture attention persistently during extinction—even several days after training. Furthermore, vulnerability to attentional capture by high-value stimuli is negatively correlated across individuals with working memory capacity and positively correlated with trait impulsivity. An analysis of intertrial effects reveals that value-driven attentional capture is spatially specific. Finally, when reward is delivered at test contingent on the task-relevant shape feature, recent reward history modulates value-driven attentional capture by the irrelevant color feature. The influence of learned value on attention may provide a useful model of clinical syndromes characterized by similar failures of cognitive control, including addiction, attention-deficit/hyperactivity disorder, and obesity.

Keywords Attentional capture · Reward · Incentive salience · Visual search

Selective attention gates access to awareness. Attentional control therefore determines the contents of awareness and the starting point for almost any behavioral or cognitive act—perceiving, remembering, learning, or behaving. Attentional control

---

S. Yantis (✉) · B. A. Anderson · E. K. Wampler · P. A. Laurent
Department of Psychological & Brain Sciences, Johns Hopkins University,
3400 N. Charles St., Baltimore, MD 21218-2686, USA
e-mail: yantis@jhu.edu

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_5, © Springer Science+Business Media New York 2012

has long been a core issue in cognitive psychology and continues to drive a great deal of empirical and theoretical research.

Two broad domains of control have long been known to determine deployments of attention (Corbetta and Shulman 2002; Egeth and Yantis 1997; Yantis 2000). Voluntary or top-down attentional control is driven by current perceptual goals. When an individual is searching for a particular object or feature, or searching in a particular location, they can voluntarily direct overt attention (eye movements) or covert attention (without eye movements) to the task-relevant object, feature, or location. Such deployments of attention increase the speed and accuracy of behavioral responses (e.g., Pashler 1998) and evoke strong modulation of neural activity in the brain (e.g., Moran and Desimone 1985; Yantis 2008). A vast literature has documented the spatial and temporal properties of voluntary, goal-directed attentional control (see Egeth and Yantis 1997 and Pashler 1998 for reviews).

A second form of control is often referred to as bottom-up or stimulus-driven control (Itti and Koch 2001; Parkhurst et al. 2002; Theeuwes 1992, 2010;Yantis and Jonides 1984; Yantis 1993, 2000). When a salient, unexpected event occurs (e.g., the appearance of a new object, Christ and Abrams 2006; Yantis and Hillstrom 1994; or looming motion, Lin et al. 2009) the perceiver will often orient to that event, even if it may interfere with other ongoing cognitive operations. Here, too, a large literature has explored the limits of stimulus-driven attentional capture and its interactions with voluntary attentional control. Yantis and Jonides (1984) reported that the abrupt onset of a new perceptual object captures attention in visual search even when the onset does not reliably predict the target location (but top-down deployment of attention elsewhere can override this automatic attention response; Yantis and Jonides 1990). Theeuwes (1992) showed that when people search for a shape singleton target (that is, a unique shape in an array of otherwise identical shapes, e.g., a diamond among many circles), then the presence of an irrelevant color singleton (e.g., a red nontarget item when all the remaining items in the display are white) significantly slows search, an indication that it captured attention involuntarily.

A variety of intermediate or hybrid cases of attentional control have also been described. These are situations in which the deployment of attention is not entirely voluntary, but the effect appears to be at least partly a function of information held by the perceiver, either in the form of explicit goals or recent perceptual history. One example of this hybrid category is contingent attentional capture (Folk et al. 1992; Anderson and Folk 2010). In this scenario, subjects are instructed to search for a target defined by a salient feature (e.g., identify the red target in an array of otherwise white items, or identify the object that has an abrupt onset). Immediately preceding the array, a to-be-ignored "cue" appears (this is probably better labeled a "distractor"). The distractor consists of a set of salient items that appear surrounding one of the potential target locations; it occurs shortly before the search array appears and should always be ignored (by virtue of its location—surrounding a possible target location—and time of onset—before the main search array). The main finding is that if the distractor carries a feature that matches the target-defining feature (e.g., the target is red and the distractor is red), then the distractor captures attention (indexed by a faster RT when the distractor appears at the target location than when it appears in a nontarget location). When the distractor does not match the target defining feature (e.g., the target is red and the distractor is an abrupt onset) then the distractor does not capture attention. So in this case, the content of the search set (e.g., red) guides attention to matching (or similar) features, even ones that are known to be irrelevant (because of their location, shape, or timing; Anderson and Folk 2010; Folk and Remington 1998).

In this chapter, we will explore an influence on attentional control that has received increasing scrutiny in recent years: attention to reward-related stimuli. Reward has long been known to play a key role in learning and cognition (Pessoa and Engelmann 2010; Schultz et al. 1997; Platt and Glimcher 1999; Sugrue et al. 2005). This is hardly surprising: organisms must procure food and water to survive; these primary rewards evoke powerful responses in the brain that lead to learning. For example, "where was I when I found this source of food, so I can find it again?" Saharan camels no doubt learn to associate a clump of palm trees with vital and rewarding water. In order to reproduce, animals must engage in sexual behavior, and sex is among the most powerfully motivating rewards we know. Here again, reward leads to learning, which can in turn increase the probability that the reward can be obtained again in the future.

Powerful learning mechanisms in the brain provide a way for stimulus-reward associations to be learned rapidly and persistently. For example, an animal may learn that a certain kind of tasty berry has a particular color, size, and shape, and so whenever they see that berry, they orient to it and consume it immediately. They may further learn that this kind of berry can be found near a particular species of tree, and so they learn to notice that tree (which is perhaps easier to spot in the dense jungle) and indeed when they see that tree they begin to anticipate the berries they soon will be eating.

A number of recent studies have investigated the role of reward on attention. They have shown that reward plays an important role in voluntary, deliberate deployments of attention in a variety of contexts. Raymond and O'Brien (2009) showed participants several novel faces during a training phase, and consistently followed each face with different amounts of positive or negative monetary reward. Following the training phase, participants carried out an attentional blink task in which two targets, each followed by a mask, are shown in rapid succession. The second target was one of the faces that had appeared during the training phase. In attentional blink tasks, the identification of the second target (termed T2) is typically impaired at short delays between the two targets. The probability of correctly recognizing a face was much greater when the face had been associated with large positive or negative rewards during training than if it had been associated with low or no reward. Indeed, faces associated with a large positive reward showed no evidence of an attentional blink. The authors concluded that learning to associate stimuli with high positive reward during learning enhances the degree to which those task-relevant stimuli can compete for attention.

Della Libera and Chelazzi (2009) had participants learn associations between complex shapes and reward in a matching task. Participants were presented with two superimposed colored shapes on one side of fixation, and a black shape on the other. They compared one of the two superimposed colored shapes to the black shape and indicated whether they were the same or different. A color cue at the beginning of each trial informed participants which colored shape to use in the comparison. Monetary rewards were delivered that varied with the identity of the colored shapes; regardless of color, some shapes tended to predict more reward than others. Following performance of this task, which lasted several thousand trials over multiple days, participants were tested without rewards (i.e., in extinction) both in the same task and in a simple visual search task. The results showed that formerly reward-predictive shapes impaired performance as the to-be-ignored shape in the matching task, but did not impair performance as irrelevant distractors in the visual search task. However, visual search was facilitated for goal-relevant targets that were formerly predictive of high reward. This was taken as evidence that reward learning can have persistent effects on attentional selection.

Krebs et al. (2010) employed a Stroop task to investigate the effect of reward association on attention; certain color words were associated with the delivery of monetary reward when the color was correctly named. Participants were to name the ink color of color words; in some cases, the color words were different than the ink color, and these incompatible trials often slow responding due to competition between the required ink-color response and the automatic word-reading response. Overall, responses to rewarded ink colors were faster than those to unrewarded colors, consistent with increased attentional priority to the rewarded feature. When the to-be-ignored color name was reward-related (e.g., “blue” when the ink color blue predicted reward), it tended to magnify the usual slowing caused by color-word conflict (compared to when the color word was not reward-related). This study shows that stimuli associated with high reward in the current task context tend to draw attention even when those stimuli should be ignored.

Peck et al. (2009) measured neural responses in monkey area LIP, which is thought to contain a map-like representation of attentional priority (Bisley and Goldberg 2010). On each trial of the experiment, a target appeared in the display and the monkey had to make a rapid eye movement to the target’s location. Almost a second before the target appeared, one of two cue shapes appeared in the display; the cue’s location did not reliably predict the subsequent target location—in other words, it was not informative about the upcoming required eye movement. One of the two cue shapes indicated that the trial would be followed by a juice reward; the other shape indicated that no reward would be delivered. When the cue indicated that a reward was forthcoming, it evoked a much stronger neural response than if it indicated no reward would be delivered. Furthermore, behavioral responses were faster and more accurate when the subsequent target stimulus appeared in the location of a reward-predicting cue compared to a cue predicting no reward (indeed, the no-reward cues exhibited evidence of attentional repulsion). This behavioral effect persisted in second task in which well-learned cues no longer predicted reward, but only when the monkeys knew reward would not be available on the current trial. The latter finding suggests a persisting attentional effect of the former reward learning, but one that can be largely overcome by current reward-related goals.

To account for how physical salience and reward value might be combined to give rise to a representation of attentional priority, Navalpakkam et al. (2010) presented visual search displays to human participants; each display contained multiple nontarget bars all with the same tilt and two targets that differed from the nontargets—one with a tilt that differed only slightly from the nontargets (low physical salience) and one with a very different tilt (high physical salience). The reward value of the two targets was varied across blocks so that various combinations of physical salience and reward value could be tested. They found that performance depended on both physical salience and relative amount of reward in a pattern that was best accounted for by a Bayesian ideal observer model in which expected value is maximized.

Several studies have shown that reward delivery gives rise to involuntary deployment of attention on the very next trial, a phenomenon that can be termed "reward priming" (Della Libera and Chelazzi et al. 2006 ; Hickey et al. 2010a, b). For example, Hickey et al. (2010a) conducted a visual search task based on the additional singleton paradigm introduced by Theeuwes (1992). Participants searched for a shape singleton (e.g., a diamond in an array of circles) and reported the orientation of a small line segment contained in the target. On some trials, all the shapes were rendered in the same color (red or green). On many trials, however, one of the nontarget shapes had a unique color (red among green or vice-versa). Each trial was followed by a feedback display containing the amount of reward that was received on that trial, either 1 point or 10 points, which translated into monetary reward at the end of the experiment. Reward was delivered randomly.

The key manipulation was whether the colors of targets and distractors swapped from one trial to the next. For example, if on trial N the target and most of the nontargets were red (and the color singleton was green) then on trial $N + 1$ the color assignment could be the same, or it could swap so that now the target and most of the distractors were green (and the color singleton was red). The main result was that on trials in which the colors did not swap, a high reward on trial N yielded faster responses on trial $N + 1$ than did a low reward on trial N. However, when the colors swapped, this pattern reversed: a high reward on trial N gave rise to slower responses on the following trial than did a low reward on trial N. This outcome indicates that when a particular color is rewarded, that color appears to draw attention to itself on the immediately following trial—even though color is not relevant to this task. Della Libera and Chelazzi (2006) reported a similar result in the context of a global/local judgment task.

Serences (2008) used a choice paradigm to study the effect of reward history on both the representation and selection of visual stimuli. Human participants selected one of two colored circles via a button press, and their selection either was or was not followed by the delivery of monetary reward. Throughout the course of the experiment, the relative probably of receiving a reward for selecting one color over the other was either 1:1, 1:3, or 3:1, which changed periodically during the experiment. The results showed that the recent reward history of each color predicted both stimulus selection and stimulus-evoked response in early visual areas as measured by functional magnetic resonance imaging (fMRI). Using a similar paradigm, Serences and Saproo (2010) extended this finding by showing that oriented gratings associated with larger rewards are represented with greater precision in early visual areas of the human brain. Additionally, Shuler and Bear (2006) found that when light flashes predicted reward, responses in rat area V1 reflected temporal expectations concerning reward delivery, with activity being either maximal or minimal at the time of expected reward. Collectively, these results argue that current stimulus-reward associations bias perception, consistent with attentional priority to high-value features.

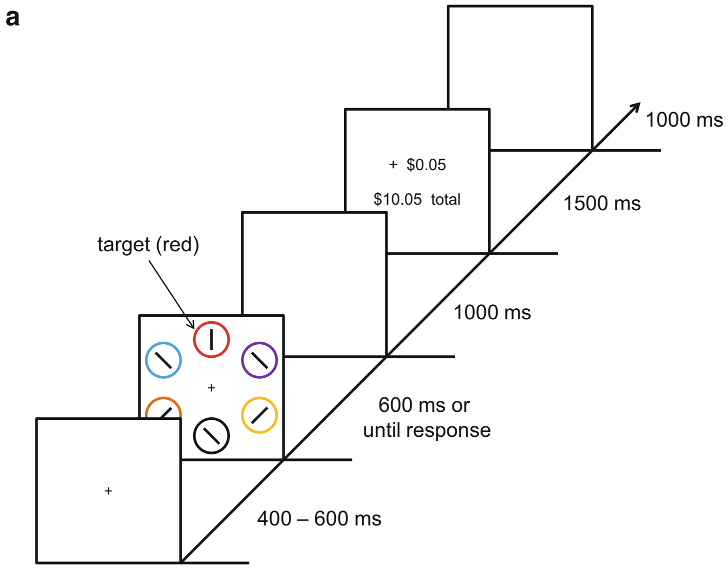
In the studies summarized above, positive effects of reward—usually, faster and/or more accurate responses, accompanied by increased brain activity, to high-reward stimuli—were observed when the task in question involved currently rewarded stimuli and/or stimuli that are currently task-relevant (e.g., they were the targets of search or they predicted reward). However, reward-related effects of task-irrelevant stimuli have only been observed in trials immediately following reward. Although investigators have in a few cases examined the effect of a to-be-ignored stimulus previously associated with reward, in most cases the previously reward-related stimuli failed to capture attention (Della Libera and Chelazzi 2009, Experiment 2; Krebs et al. 2010, Experiment 2; Raymond and O'Brien 2009, Experiment 2; for an exception, see Peck et al. 2009, Fig. 8). In these experiments, however, the stimuli in question were typically complex multi-feature or multidimensional objects such as complex shapes, words, or faces. This aspect of the stimuli may have precluded them from exerting a significant, persistent effect on observable behavior.

In this chapter, we describe several experiments that explore the degree to which otherwise neutral and elementary visual features (e.g., colors) can be associated with different amounts of monetary reward through training and feedback, and later come to capture attention involuntarily when they appear in contexts in which those stimuli should be ignored. The results of the experiments discussed in this chapter were originally reported in Anderson et al. (2011a, b, 2012).

Learned Value Modulates Salience-Based Attentional Capture

As noted earlier, when a person searches for a salient visual stimulus defined as a feature singleton in one dimension (e.g., a unique shape in an array of other shapes—diamond among circles or vice-versa), then the presence of a more salient feature singleton in another dimension captures attention and slows search (Theeuwes 1992). In the first experiment, we asked whether this form of stimulus-driven attentional capture can be modulated by reward associations.

Experiment 1 has several features in common with most of the experiments reported in this chapter. The experiment had two phases, a training phase and a test phase. In both phases, participants engaged in visual search for a target among five nontargets (Fig. 1). In the training phase, subjects searched for a red or a green target circle in an array of six circles, each rendered in a different color, presented on a black background. Exactly one item in each display was red or green (each color appeared as the target equally often in random order). Inside every nontarget circle was a white line segment that was oriented 45° to the left or right (randomly selected in each nontarget circle). The line segment inside the (red or green) target circle was either vertical or horizontal. The subject's task was to press one button if the line



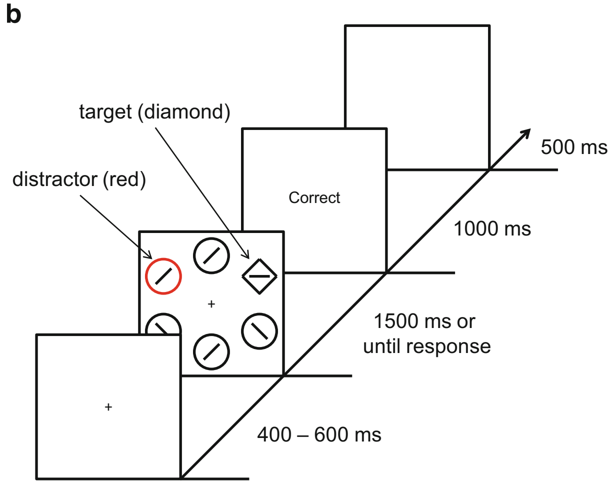


Fig. 1 Sequence of trial events in Experiment 1. **a** During training, an array of six *circles* each rendered in a different color appeared on a *black* background; one of these was either *red* or *green*. The participant determined the orientation of the line segment within the *red* or *green circle* and pressed a corresponding button. A subsequent feedback display indicated the amount of reward obtained on that trial and the cumulative reward earned so far in the experiment. The high-reward target color was followed on 80 % of the trials by a high reward (5¢) and on 20 % of the trials by a low reward (1¢); this mapping was reversed for the other target color. **b** At test, participants were to determine the orientation of the line segment within the shape singleton (*diamond* among *circles* or *circle* among *diamonds*, unpredictably). On half the trials, all the items were *white*. On one-quarter of the trials one of the nontargets was *red* and on one-quarter one of the nontargets was green. No reward was delivered during the test phase

segment contained within the target circle was vertical and another button if it was horizontal. Following each correct response, a feedback display indicated that the participant had received a small monetary reward. The cumulative value of all rewards earned so far was also displayed. Following an incorrect response, a feedback display indicated that no money had been earned on the trial.

The magnitude of the reward following correct responses was either large or small (in this experiment, 5¢ or 1¢). One color had a high probability ($p = 0.8$) of yielding a large reward and a low probability ($p = 0.2$) of yielding a small reward; this mapping was reversed for the other color. This partial reinforcement schedule, which is more resistant to extinction (Rescorla 1999), was modeled after Della Libera and Chelazzi 2009). The high- and low-reward colors were counterbalanced across subjects. The training phase was designed so that participants came to associate one color with high reward and the other color with low reward. In this experiment, the training phase included 1008 trials.

After a short break, the test phase began. In the test phase, subjects searched for a unique shape (diamond among circles or circle among diamonds, unpredictably). Once again, each shape contained a small line segment that was oriented at ±45°, but the target shape contained either a vertically or horizontally oriented line segment (Fig. 1b). The subject was to identify the orientation of the line segment within the unique shape and press the corresponding button as rapidly as possible. The test phase consisted of 480 trials.

On most trials in the test phase, all the items were rendered in white on a black background. On a randomly-selected one-quarter of the trials, one of the nontarget shapes was rendered in red, and on one-quarter in green. Subjects were told to ignore color—the target shape was never the color singleton. Each response was followed by feedback about whether the response was correct or not—no reward feedback was provided during the test phase.

Many previous studies have demonstrated that the presence of a color singleton (red or green) captures attention and slows responses in the shape-search task (e.g., Theeuwes 1992). Thus, in this experiment, we expected to find that responses on trials containing either a red or a green distractor were significantly slowed compared to responses on trials without a color singleton distractor. The main question was whether the reward association established in the training phase affects the magnitude of this slowing. Eighteen participants completed the experiment in a single 2 h session, and earned between US$ 21 and 28 (mean = US$ 25.22).

We first checked the mapping of color to reward (i.e. red vs. green as the high-reward color) and found no interaction with the effect of reward on distraction ($F < 1$), so we collapsed across color in the remaining analyses. RT differed significantly in the three distractor conditions [Fig. 2a, $F(2,34) = 48.6$, $p < 0.001$]. Both the high-value and low-value distractor significantly slowed RT compared to the no-distractor condition [$t(17) = 8.45$ and 6.31, respectively, both $p < 0.001$], which replicates the standard effect of a salient color singleton on search time.

More importantly, we found that the presence of a high-value distractor slowed search more than the presence of a low-value distractor [$t(17) = 3.37$, $p = 0.004$]. The difference in slowing due to reward history was not due to differences in physical

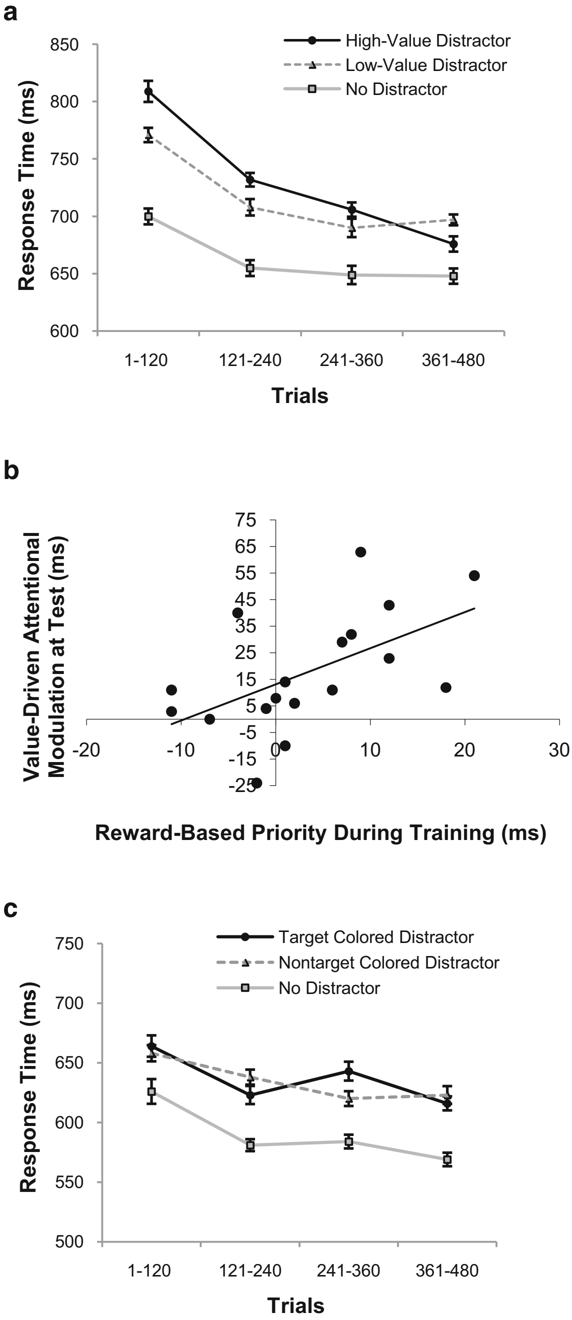


Fig. 2 Results for experiments 1 and 2. **a** Mean response time (in milliseconds) for each condition in the test phase of Experiment 1. **b** Correlation between reward-based performance during training (defined as the mean RT difference to low- and high-reward targets), and the magnitude of value-driven attentional modulation at test (defined as the mean difference in RT on trials containing high- and low-value distractors, respectively) in Experiment 1. **c** Mean response time (in milliseconds) for each condition in the test phase of Experiment 2. Error *bars* represent ± within-subjects SEM

salience (red slowed search more than green when it had been the high-reward color, and green slowed search more than red when it was the high-reward color). The effect of reward history on RT declined over the course of the test phase, as revealed by a linear trend in the difference between RTs for high- and low-value distractor trials over trial epoch [Fig. 2a, $F(1{,}17) = 17.22$, $p = 0.001$].

Several previous studies had shown that reward influences both behavioral and neural responses (Platt and Glimcher 1999; Simen et al. 2009; Sugrue et al. 2005). In this experiment, there was not a robust difference in RT to high- and low-reward targets during the training phase, when rewards were being delivered in the form of monetary feedback (the mean difference in RT to the high-reward and low-reward color targets during the training phase was just 3 ms). However, there were substantial individual differences in this effect—some participants responded as much as 20 ms faster to the high-reward color than to the low-reward color on average, while others exhibited a negligible or even slightly negative difference. We wondered whether these individual differences could reflect the degree to which people differed in their sensitivity to reward. Such differences have been investigated in a variety of contexts and are thought to underlie differences in self-control in those contexts (Braver et al. 2010).

To examine this, we plotted the magnitude of value-driven attentional capture (that is, the mean difference in RT to high-value vs. low-value distractors) in the test phase as a function of reward-based speed of responding (that is the mean difference in RT to the high-reward vs. low-reward targets) in the training phase (Fig. 2b). The Pearson's correlation coefficient was significant ($r = 0.553$, $p = 0.017$), showing that individuals who responded more rapidly to high-reward targets during the training phase also responded more slowly in the presence of high-value distractors in the test phase.

We divided the 18 participants into two groups according to the mean difference in RT to high- and low-reward targets in the training phase (median split). Although the magnitude of value-driven attentional capture in the test phase task declines for the group as a whole (Fig. 2a), the slowing persisted longer for the subset of participants who exhibited greater reward sensitivity during the training phase (Fig. 3).

These results suggest that a salient distractor that was previously associated with high reward slows search more than a salient distractor previously associated with low reward. It is possible that this effect was merely the result of a persisting search set for the two color targets, and particularly—and for some participants—for the highly rewarded target, during the test phase, even though in the test phase color was not task relevant and was unrewarded. This possibility appears unlikely given recent evidence that people can adjust their deliberate search set rapidly and effectively: Lien et al. (2010) showed that participants could rapidly and flexibly adjust their search set on a trial-by-trial basis according to task requirements. However, some studies have shown that unrewarded former targets can capture attention under certain circumstances (Kyllingsbaek et al. 2001; Shiffrin and Schneider 1977). To definitively exclude the possibility of a persisting search set for the color distractors, eighteen new participants completed a control experiment that was very similar to Experiment 1, except that no reward feedback was provided during the training

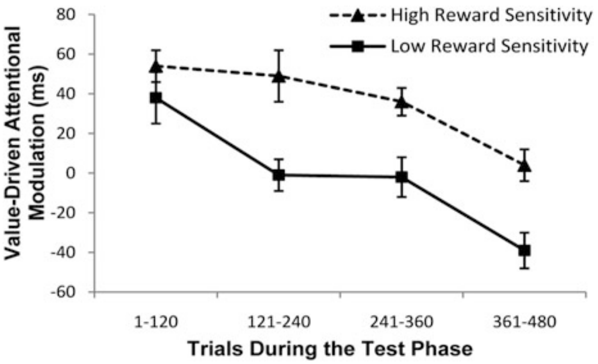


Fig. 3 Mean difference in response time (in milliseconds) for high- and low-value distractors over the course of the test phase of Experiment 1, plotted separately for two subgroups of participants: those who exhibited the largest difference in RT to high- and low-reward targets in the training phase and those who exhibited the smallest difference (median split). The effect of reward at test is larger and persists longer for participants who were more sensitive to reward during training

phase. Each participant in Experiment 2 was given a flat compensation of US$ 25, approximately the mean amount earned by participants in Experiment 1.

In the training phase of Experiment 2, half the participants searched for red and blue targets in a multicolored array, and green was among the nontargets on half the trials (equally often with each color target); the other half of the participants searched for green and blue targets and red was among the nontargets half the time. The test phase was identical to that in Experiment 1: subjects searched for a unique shape (diamond among circles or vice-versa, unpredictably). All items were white on a black background, except that on one-quarter of all trials, one of the nontargets was red and on one-quarter of the trials one of the nontargets was green. Because participants had just spent 1008 training trials searching for red and always ignoring green (or vice-versa for half the participants), any residual tendency to continue to search for the previously prioritized target color should be evident in comparing trials in which a former target was the distractor vs. trials in which a former nontarget was the distractor.

The presence of a color singleton distractor significantly slowed search, as expected. However, there was no difference in RT on trials containing a former target vs. those containing a former nontarget [Fig. 2c, $t(17) = 0.34$, *n.s.*]. Furthermore, the amount of slowing caused by a high-value distractor in Experiment 1 was greater than that caused by a former target-colored distractor in Experiment 2 [mean difference = 27 ms, $t(34) = 2.29$, $p = 0.025$], confirming that learned associations between stimuli and reward indeed increase distraction beyond that produced by previously prioritized stimuli that are not associated with reward. We take this as evidence that slowing in the test phase in Experiment 1 was not due to a persisting search set for a former target.

Experiment 1 shows that when a salient but irrelevant item appears in a search array, the degree to which it slows search depends on its reward history. There are at least two possible mechanisms for this effect that can be considered. One is that reward history makes an item "more salient"—that is, the physical properties of the item that make it perceptually salient (e.g., local color contrast) are combined with the reward associations to magnify the conspicuity or pertinence (Bundesen 1990) of the item, increasing the probability that the item will be selected during search. In effect, the item competes more effectively for attention with the shape target.

A second possibility is that reward history has an effect only *after* the item has captured attention by virtue of its physical salience, by prolonging dwell time or increasing disengagement costs (Belopolsky et al. 2010; Duncan et al. 1994). On this account, the increased reward association causes an attended item to continue to "hold" attention when that item had previously been highly rewarded.

The present experiments cannot distinguish between these accounts. However, in the remaining experiments we ask whether reward history can evoke an involuntary shift of attention even when the item in question is not physically salient. If it can, then we may be able to conclude that reward history has an effect that is not strictly a consequence of increasing dwell time or disengagement costs.

Physically Inconspicuous Items Previously Associated With Reward Capture Attention

In the remainder of the chapter, we describe several experiments that investigate how otherwise inconspicuous and task-irrelevant stimuli associated with reward during training capture visual attention. Experiment 3 was similar in design to Experiment 1, except that the items in each trial of the test phase were rendered in multiple different colors, thereby ensuring that the previous target colors were not themselves physically salient (Fig. 4). The training phase was identical to that in Experiment 1: 1008 trials of visual search for a red or a green circle containing a vertical or horizontal line segment that specified the correct response. For half the participants, red targets were followed by high reward with probability $p = 0.8$ and by a low reward with $p = 0.2$, and green targets had the complementary contingencies; for the remaining participants, this mapping was reversed (Table 1).

In this experiment, because the red and green distractors are not physically salient, any slowing of RT they produce must be attributed to the value with which they were imbued in the training phase. Responses were indeed slowed by value: RT was significantly slower on trials containing a high-value distractor than when no value-related distractor was present; the mean RT in the presence of a low-value distractor was intermediate [ANOVA: $F(2{,}50) = 6.07$, $p = 0.004$; linear trend: $F(1{,}25) = 12.19$, $p = 0.002$; see Table 2). This result shows that an otherwise neutral stimulus feature captures attention when that feature has previously been associated with reward.

There were individual differences in susceptibility to value-driven attentional capture. We wondered if those differences were systematically related to other cognitive abilities or personality traits that have been shown in other studies to covary with

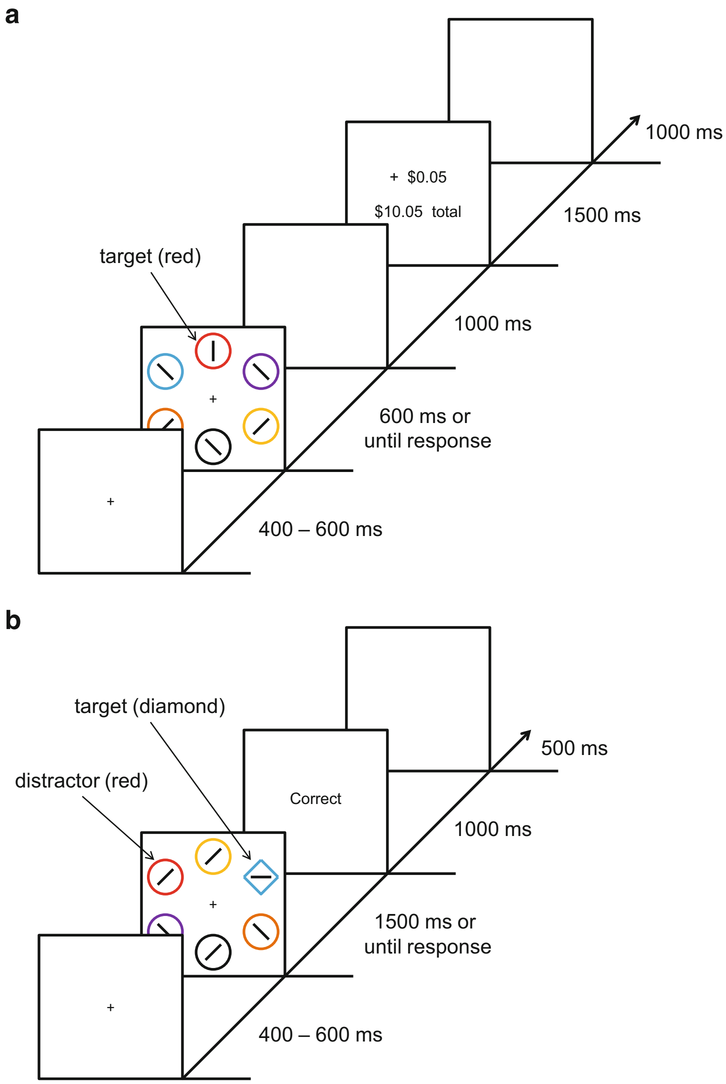


Fig. 4 Sequence of trial events in experiment 3. **a** During training, an array of six *circles* each rendered in a different color appeared on each trial; one of these was either *red* or *green*. The participant determined the orientation of the line segment within the *red* or *green circle* and pressed a corresponding button.Text indicating monetary reward feedback was presented after each correct response. **b** During the test phase, participants searched for a unique shape (*diamond* among *circles* or *circle* among *diamonds*, unpredictably); no reward feedback was provided. The items were rendered in six different colors. On 25 % of the trials, one of the nontarget items was red and on 25 % of the trials, one of the nontarget items was green; on the remaining trails none of the items were *red* or *green*

Table 1 Error rates by distractor condition for Experiments 1 and 2

Distractor condition in Experiment 1			Distractor condition in Experiment 2		
None	Low-value	High-value	None	Nontarget colored	Target colored
0.09 (0.003)	0.10 (0.004)	0.10 (0.005)	0.11 (0.003)	0.13 (0.005)	0.13 (0.005)

The error terms, in parentheses, reflect the within-subjects SEM

Table 2 Mean response time (in milliseconds) and error rate, respectively, in the test phase of Experiments 3 and 5 in which reward was delivered for each of three training conditions: Experiment 3, long training (1008 trials) with low and high reward of 2¢ or 10¢ per trial; Experiment 5, brief training (240 trials) with rewards of 2¢ and 10¢ per trial, and Experiment 5, brief training followed by a delay of 4–21 days

Training phase	Distractor condition in the test phase		
	None	Low-value	High-value
1008 trials	665 (2.8)	673 (2.8)	681 (2.6)
	0.11 (0.004)	0.10 (0.004)	0.11 (0.004)
240 trials	667 (2.0)	675 (3.0)	682 (2.9)
	0.12 (0.005)	0.12 (0.006)	0.12 (0.006)
4–21 days ago	614 (1.8)	624 (2.7)	630 (3.3)
	0.06 (0.004)	0.07 (0.006)	0.08 (0.005)

The error terms, in parentheses, reflect the within-subjects SEM

attentional control. One such ability is visual working memory capacity. Fukuda and Vogel (2009) measured working memory capacity using a visual change detection task in which a memory array of colored squares is presented briefly, followed by a 900 ms retention interval, and then a test display containing a single probe square that is either the same or different in color than the square previously presented at the probed location. Performance generally declines with the number of items to be remembered during the delay interval. Using a simple formula, a parameter K, representing the individual's working memory capacity, can be estimated.

Fukuda and Vogel estimated visual working memory capacity, and then asked subjects to carry out a task requiring them to focus attention at a spatially cued location (without moving their eyes) to perform a visual discrimination task. Immediately following the visual display, a potentially distracting dot appeared either at the target location or at a nontarget location. The investigators reasoned that a person who was able to focus attention very well at the cued location should exhibit a much larger neural response to the dot when it appeared at the cued location than when it appeared elsewhere. A person who has less effective control over attention should exhibit a smaller difference in the magnitude of the neural response to probe dots appearing in the cued and uncued locations, respectively—in particular, the response to the dot should be larger when it appears in a to-be-ignored location particularly for low-capacity individuals compared to high-capacity individuals. These two measures (effectiveness of focused attention and working memory capacity) were strongly correlated across individuals ($r = 0.73$). The authors argued that working memory requires effective attentional control in order to maintain information over a delay interval by minimizing the degree to which distracting information displaces to-be-remembered items. If this is correct, then we might expect that individuals with high working memory capacity should also be less vulnerable to value-driven attentional capture.

Individual differences in impulsivity and reward sensitivity have also been shown to be related to attentional control. Hickey et al. (2010b) showed that reward sensitivity (measured by a normed questionnaire that queries the degree to which the individual seeks reward in everyday life) is correlated with the degree to which a color distractor that was associated with high reward on the preceding trial of a search task would slow search on the current trial. Dickman and Meyer (1988) found that individual differences in impulsivity are associated with individual differences in the speed and accuracy in the performance of visual tasks. Impulsivity is thought to contribute to the comorbidity of ADHD and predisposition to substance abuse (Groman et al. 2008). We therefore measured trait impulsivity using the Barratt Impulsiveness Scale (Patton et al. 1995) as another covariate. Here again, we would expect individuals who are impulsive by self-report should also be more vulnerable to value-driven attentional capture.

We used visual working memory capacity and trait impulsivity as predictor variables in a simultaneous regression model of value-driven attentional capture. The model accounted for a significant proportion of the variance in value-driven capture ($R^2 = 0.355$, $p = 0.006$); the regression weights for both predictor variables were significantly greater than zero (est. $\beta = 0.378$, $p = 0.038$ for impulsivity; est. $\beta = -0.554$, $p = 0.004$ for WM capacity). Thus, individuals exhibiting greater impulsivity and lower visual WM capacity were more vulnerable to value-driven attentional capture. Figure 5 (diamonds and grey line) shows a scatter plot of WM capacity vs. value-driven capture in Experiment 3.

We performed a control experiment, as before, to examine the possibility that the effects observed here reflected a continuing top-down set for former targets, rather than an effect that depended critically upon the receipt of reward during the training phase. As noted in the description of Experiment 2, there is evidence that search sets can be adjusted rapidly in visual search, but we wished to confirm this in the present context. Experiment 4 was identical to Experiment 3, with ten naïve participants, except that no reward feedback was provided during the training phase. All participants were given a flat US$ 25 payment for their participation, an amount that matched the average reward earned in Experiment 3. The absence of reward during training eliminated any slowing by former targets in the test phase [$t(9) = -0.39$, *n.s.*; see Table 3]. There was also no difference in error rate among the red, green, and no-distractor conditions [$F(2,18) = 2.30$, *n.s.*].

In Experiment 5, we reduced the amount of training and increased the delay between training and test to examine the robustness and persistence of value-driven attentional capture. The training and test phases of the experiment were identical to those in Experiment 3, with two exceptions: first, the number of trials in the training and test phases were reduced to 240 each (compared to 1008 and 480, respectively, in Experiment 3); second, the magnitude of high and low rewards delivered in the

Fig. 5 Scatter plot shows the magnitude of value-driven attentional capture (RT in the presence of a high-value distractor minus RT when neither distractor was present) vs. visual working memory capacity in Experiment 3 (long training, *diamonds*, *grey line*) and Experiment 5 (short training, *circles*, *black line*). Best-fitting regression lines are shown

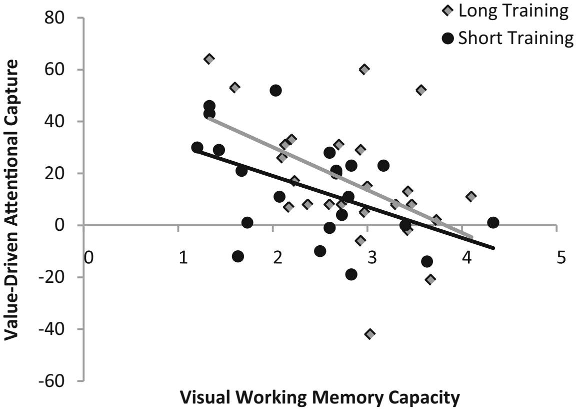


Table 3 Mean response time (in milliseconds) and error rate, respectively, in the test phase of Experiment 4 in which no reward was delivered

Training phase	Distractor condition in the test phase		
	None	Red	Green
1008 trials	602 (3.9)	606 (2.1)	593 (3.9)
(Unrewarded)	0.14 (0.004)	0.17 (0.006)	0.15 (0.005)

The error terms, in parentheses, reflect the within-subjects SEM

training phase were increased to 10¢ and 2¢, respectively (compared to 5¢ and 1¢ in Experiment 3).

We observed significant value-driven attentional capture with this reduced training regimen that was nearly as large in magnitude as that observed in Experiment 3 [$F(2,46) = 5.17$, $p = 0.009$; see Table 2). We also observed a significant negative correlation between visual working memory capacity and value-driven capture ($r = -0.468$, $p = 0.021$; Fig. 5, circles and black line) but the correlation with trait impulsivity, while positive, was not significantly greater than zero ($r = 0.093$, *n.s.*). We noted that the variance of impulsivity in Experiment 5 was quite low compared both to that observed in Experiment 3 and to that reported by Patton et al. (1995); either this restriction of range in this college-student population or the reduced number of training trials could have led to the absence of a significant correlation with impulsivity.

We invited the participants in this experiment back to the lab after several days had elapsed since training (4–21 days, mean = 8.8 days); 19 of the 24 participants agreed to return. These participants only completed the test phase of the experiment with no further training and no reward delivered. We continued to observe significant slowing due to formerly rewarded color distractors [$F(2,36) = 5.81$, $p = 0.007$; see Table 2].

The slowing caused by formerly rewarded distractor items in the test phase could have one of two possible sources. The first possibility is that the presence of a formerly highly rewarded item in the display caused a general slowing or filtering cost (Folk et al. 2009). The second possibility is that the formerly rewarded distractor caused a spatially specific, involuntary shift of attention to its location. To examine these possibilities, we took advantage of the spatially-specific effects of inhibition of return (IOR; Theeuwes and Godijn 2002). When attention is directed to an item and subsequently withdrawn, a residual inhibition of subsequent targets appearing in that location is reliably observed.

We examined response times on the subset of all trials that met two criteria: (a) they did not contain a low- or high-value distractor item, and (b) a high-value distractor appeared on the previous trial. We then separated these into those trials in which the target on trial N appeared in the same location as the high-value distractor on trial N–1, and those in which the target appeared in a different location than the preceding high-value distractor. We found RTs were 66 ms slower when the target appeared in the location of a distractor on the previous trial than when it appeared elsewhere [$t(23) = 3.13, p = 0.005$]. This shows that value-driven capture is a spatially-specific deployment of attention.

In Experiment 6, we tested whether learned associations between a stimulus feature (color) and reward can have an involuntary influence on attentional priority that extends to different stimuli and different task contexts. To this end, we had 21 new participants engage in a training phase similar to that employed in Experiment 5, with the exception that three different colors were used for the critical items (red, green, and blue). Two colors served as targets during training: one predicted high reward with probability $p = 0.8$ while the other predicted low reward with probability $p = 0.8$. The third color was always a nontarget (among other things, this manipulation allowed us to extend our findings to colors other than red and green). In the 480-trial test phase, participants engaged in a flankers task (Eriksen and Eriksen 1974) in which they reported the identity of a centrally presented white letter while ignoring colored flanking letters on the left and right. These flanking letters could either be associated with the same response as the centrally presented letter (response compatible) or a different response (response incompatible), although the center and flanking letters were never physically identical. The compatibility effect, that is, the degree to which RTs are slowed on incompatible vs. compatible trials, can be taken as an index of the degree to which the irrelevant flankers have been processed. We observed larger compatibility effects for flankers that were the color of a formerly high-reward target compared to those that were the color of a formerly low-reward target [mean difference = 11.6 ms, $t(20) = 2.59$, $p = 0.017$], demonstrating value-driven attentional capture that transferred across stimuli (colored outline shapes to colored letters) and across tasks (visual search to a flankers task).

Reward at Test can Magnify Value-Driven Attentional Capture

In all the experiments described to this point, the search for shape in the test phase was unrewarded, except in the usual sense that participants experience an amorphous sense of satisfaction from carrying out the task they have been asked to do. In everyday life, however, multiple items may be associated with different amounts of reward, and these typically compete for attention—for example, when you open your refrigerator in order to find something non-specific to snack on, each food item competes for attention according to how rewarding its consumption has been in the past and according to your current motivational state.

In Experiment 7, we used a similar training regimen as in Experiment 5 (240 trials, 6¢ and 2¢, respectively, for high and low reward). The test phase lasted 240 trials, but now the two shape targets (diamond among circles and vice-versa) were each associated with reward: one shape was followed with high reward (3¢) on 80 % of the trials and low reward (1¢) on 20 % of the trials; the other shape had the complementary mapping. The mapping of reward to shape was as always counterbalanced across subjects.

The first, simplest, question was whether the delivery of reward at test that is driven by the currently relevant target shapes would overpower and abolish value-driven capture by the formerly rewarded color distractors. The answer is no: value-driven capture by the color distractors was robust in the face of competition from currently rewarded shapes [$F(2,30) = 16.63$, $p < 0.001$], and indeed larger in magnitude than in the previous experiments, although the difference between high- and low-value distractors was no longer evident.

This experiment offers an opportunity to examine recent reward history on the magnitude of value-driven attentional capture. This is because, unlike all the previous experiments, rewards were delivered on each trial in the test phase, and prior research shows that participants are sensitive to the amount of reward received in the last few trials (e.g., Serences 2008). Thus, we can examine the degree to which the delivery of monetary reward on recent trials affects how strongly a formerly reward-associated color slows search on the current trial. One sensible possibility is that if a particular shape target has received high reward on recent trials, search for the rewarded shape would be facilitated on the current trial, and thereby reduce the effect of value-associated color distractors.

We observed exactly the opposite outcome. We examined whether the magnitude of value-driven attentional capture (i.e., the difference in RT on trials containing a distractor vs. trials containing neither of the value-associated colors) depended on recent reward history for the current shape target. Figure 6a shows value-driven capture as a function of the mean reward obtained over the last five trials on which the current target shape was the target. The larger the reward the current target had received, the greater the magnitude of value-driven capture by a formerly rewarded color distractor [$F(1,15) = 6.97$, $p = 0.019$]. This strongly suggests that when a particular shape appears that has recently been associated with high reward, it causes the visual system to be particularly susceptible to capture by *any* stimulus that has been associated with reward. Of course, given the reward schedule, the shape associated with high reward was very often the high reward shape, and the shape associated with low reward was very often the low reward shape. Thus, this analysis reveals an effect of reward association.

A more interesting analysis involves examining not reward associated with the stimuli, but instead the effect of reward-prediction error on the previous trial. Reward prediction error is an extremely important driver of learning to associate particular

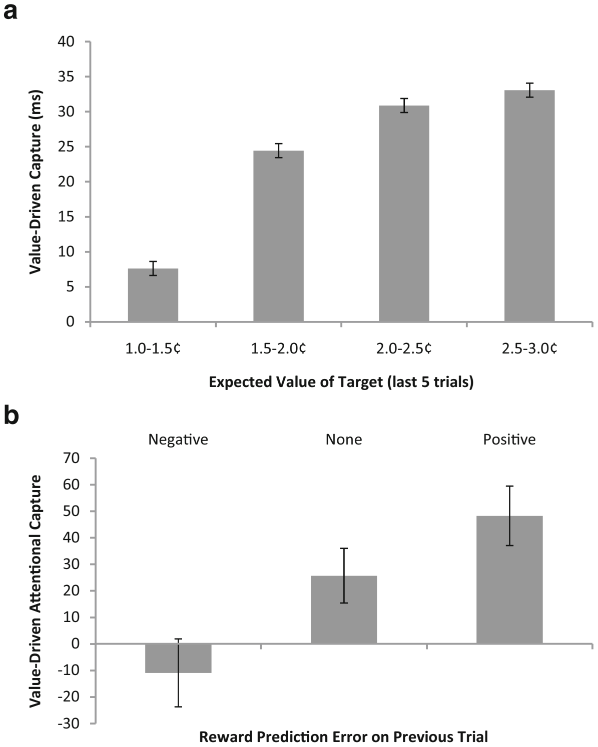


Fig. 6 Results for Experiment 7. **a** Value-driven attentional capture (defined here as the mean RT difference between distractor present and distractor absent trials) as a function of the average reward that the current target has received over the last five trials on which it was followed by reward. **b** Value-driven attentional capture as a function of the reward prediction error on the previous trial

stimuli with reward (Schultz et al. 1997) . When a neutral stimulus appears, the animal expects no reward; if this low reward-prediction is then followed by the delivery of an unexpected reward, a strong reward-prediction error response can easily be measured in the ventral striatum following the reward (e.g., McClure et al. 2003; O'Doherty et al. 2003; Schultz et al. 1997). As learning proceeds, the appearance of the reward-predicting stimulus itself begins to evoke a reward prediction response, and the reward-prediction error signal at the delivery of the expected reward declines, because the reward is no longer unexpected. Similarly, if a reward-predicting stimulus appears but no reward is delivered, the negative reward-prediction error causes a reduction in the response of these neurons when the expected reward should have been delivered.

In the current situation, we can look at the magnitude of value-driven attentional capture when the immediately preceding trial involved different reward prediction errors. In Experiment 7, three types of reward-prediction error were possible: a target with a low probability of high reward followed by high reward (this is a positive reward-prediction error—the participant is pleasantly surprised); a target with a high probability of a high reward followed by low reward (this is a negative reward-prediction error—disappointment); and the reward following a given stimulus is as expected (zero reward-prediction error). When we examined the magnitude of value-driven capture following each of these three amounts of reward-prediction error, we found the greatest value-driven capture following positive reward-prediction error on trial N–1, and no significant value-driven capture following negative reward-prediction error on trial N–1 [Fig. 6b, ANOVA: $F(2,30)=4.63$, $p=0.018$]. This finding, which is not entirely independent of the finding reported earlier, also suggests that the degree to which a perceiver is susceptible to value-driven capture depends in part on recent reward learning. Large positive reward-prediction error produces strong value-driven capture, as if the visual system is temporarily hyper-responsive to high-value stimuli.

Discussion

In this chapter we have described a series of experiments that provide new evidence concerning the role of reward learning in the deployment of attention during visual search. The experiments used well-validated visual search tasks to show that a stimulus feature associated with increased reward during training slows visual search more than a feature associated with lower or no reward.

In the first experiment, physically salient color distractors prolonged visual search more during a shape singleton search task when it was previously associated with high reward than when it had been associated with low reward. Experiment 2 showed that this effect depended upon the delivery of reward during the training phase, and was not merely a persisting top-down set for former targets. Experiment 6, in which value-driven attentional capture generalized to a flankers task involving letters (rather than the circles that were used in training), showed that value-driven capture results from associations between prior reward and predictive stimulus features (in this case color), and does not reflect attentional capture that is specific to previously rewarded objects.

The slowing observed in Experiment 1 has at least two different interpretations. It could be that a color singleton formerly associated with high reward has a greater probability of capturing attention than one associated with low reward. Alternatively, it could be that the initial capture of attention by the color singleton was unaffected by reward history, but that the time required to disengage from the color distractor depended on whether it had previously been associated with high or low reward. Of course, both factors could play a role.

In order to determine whether reward history directly causes attentional capture, the remaining experiments used color distractors that were not physically salient—that is, they were always presented within arrays of items that varied in color. Experiment 3 showed that a physically inconspicuous nontarget that is rendered in a color that had been associated with high reward during the training phase slowed search more than one formerly associated with low reward or when neither color was present among the nontargets. This shows that the reward association itself gives rise to a unique mode of attentional capture, one that differs from the well-established stimulus-driven and goal-directed modes. Experiment 4 confirmed that the rewards delivered during training are essential to these observed attention effects.

Value-driven attentional capture can be evoked with even brief training, as shown in Experiment 5, and it persists for at least a week—indeed, in later experiments we have observed significant value-driven attentional capture when participants came back to the lab as long as 8 months after initial training.

We also found that individuals vary in the degree to which they are susceptible to value-driven attentional capture. These individual differences are correlated with visual working memory capacity (low-capacity individuals exhibit more prolonged slowing due to value-driven capture) and with trait impulsivity (high impulsive individuals exhibit stronger value-driven capture). Individual differences in visual working memory capacity are thought to reflect variation in a general ability to resist distraction, such that individuals with high working memory capacity are better able to restrict selection to relevant items both in vision and in working memory (Fukuda and Vogel 2009, 2011). Our results support this idea by showing a consistently negative correlation with visual working memory capacity and susceptibility to distraction by valuable but task-irrelevant stimuli. Trait impulsivity is thought to reflect, in part, a measure of the degree to which an individual is able to inhibit a prepotent response (Dickman and Meyer 1988; Groman et al. 2008). Our results provide evidence that, at least with longer training regimens, more impulsive people are less able to avoid the tendency to select the valuable stimulus rather than the current target of visual search.

The effect of value-driven attentional capture is spatially specific: RT to targets appearing in a location occupied on the previous trial by a high-value distractor are especially slow—a manifestation of inhibition of return (Theeuwes and Godijn 2002). When a salient distractor captures attention, the subsequent active suppression of that item in order to withdraw and direct attention to the target of search gives rise to a persisting inhibitory signal at that location. Subsequent voluntary deployment of attention is slowed by this inhibition, and it is this slowing that we observed. This IOR-based signature provides strong evidence for a spatially-specific instance of involuntary attentional deployment.

The present findings may provide a model for impaired attentional capture in a variety of clinical syndromes that involve failures of cognitive control, including, for example, substance abuse. Consumption of alcohol, nicotine, cocaine, and other drugs of abuse cause the release of dopamine in the nucleus accumbens in the ventral striatum of the basal ganglia, and through repeated use can come to usurp the brain's reward circuitry, leading to compulsive craving (Robinson and Berridge 2003). Different accounts of the role of dopamine in addiction have been proposed. According to the hedonic account, dopamine is directly involved in the pleasurable experience evoked by reward (e.g., Koob and Le Moal 1997), and as tolerance to the drug increases, the homeostatic response to the drug leads to an unpleasant withdrawal state. However, drug–seeking persists well after the pleasurable effects of using have subsided; this undermines a purely hedonic account of substance abuse.

Another account is dopamine's role in learning associations between predictive cues, actions, and reward delivery. According to reinforcement learning theory, learned reward predictions are used for optimal action selection (Sutton and Barto 1998). Attended stimuli associated with reward during a training phase may therefore continue to capture attention in the test phase because the act of attending to them reliably preceded reward delivery. Everitt et al. (2001) suggested that exposure to addictive drugs can result in a transition from action-outcome learning in the ventral striatum to the formation of automatic stimulus-response habits mediated by the dorsal striatum. Robinson and Berridge (2003) argue, however, that over learned, automatic habits (e.g., tying your shoes) do not generally give rise to compulsive motivation. Some additional mechanism is needed.

Berridge and Robinson (1998); Robinson and Berridge (2003) review evidence for their proposal that the release of dopamine that accompanies the receipt of reward increases the *incentive salience* of reward-related stimuli. Incentive salience produces a motivation to want the reward-associated stimuli; in addiction, wanting persists even when a drug no longer produces pleasure. This mechanism of motivated behavior is typically adaptive: animals are motivated to seek rewarding stimuli. However, when incentive salience overrides top-down intentions, this outcome can become debilitating.

Several of the studies reviewed above provide evidence that task-relevant stimuli that predict reward elicit enhanced behavioral and neural responses (Hickey et al. 2010a, b; Navalpakkam et al. 2010; Peck et al. 2009; Raymond and O'Brien 2009; Serences 2008). The present experiments demonstrate that reward learning can imbue stimuli with value that can override top-down intention and give rise to suboptimal behavior. Together the results of these experiments extend a growing understanding of how reward associations strongly influence attention. The attentional priority accorded to reward-related stimuli expands the landscape of attentional control beyond the well known stimulus-driven (Theeuwes 1992; Yantis and Jonides 1984) and goal-directed (Folk et al. 1992) modes of attentional capture.

The results reported in this chapter are broadly consistent with the notion that the representation of value and attentional priority are critically linked. Studies by Serences (2008), Serences and Saproo (2010), Shuler and Bear (2006) demonstrated that reward-associated features are represented more robustly in early visual cortex. It is also well established that stimuli that predict reward come to evoke the striatal response formerly associated with the reward itself (e.g., Schultz et al. 1997; Hollerman et al. 1998). Our findings suggest that such activity biases attention to reward-related stimulus features which have been imbued with incentive salience. Experiment 7 demonstrates that recent reward history can potentiate value-driven capture, which is consistent with this notion.

Because attention determines the content of perceptual experience and the resulting awareness of one's surroundings, it contributes to optimal behavior in all spheres of life. Disorders of attention and cognitive control accompany a variety of clinical syndromes. It is possible that some form of value-driven attentional capture plays a role in these syndromes, including drug addiction (Field and Cox 2008; Garavan and Hester 2007; Robinson and Berridge 2008), attention-deficit/hyperactivity disorder (Bush 2010), obsessive-compulsive disorder (Sheppard et al. 2010), and obesity (Davis 2010). These conditions tend to co-occur (Davis 2010; Sheppard et al. 2010), and correlations with individual differences in working memory capacity and impulsivity suggest that there may be common underlying mechanisms that make some individuals more susceptible to value-driven attentional capture and the disorders to which it may contribute.

Acknowledgments We thank H. Egeth, J. Flombaum, L. Gmeindl, P. Holland, D. E. Meyer, and J. Serences for fruitful discussions and suggestions. The experiments reported here were supported by US National Institutes of Health grant R01-DA013165 to S.Y.

References

Anderson, B. A., & Folk, C. L. (2010). Variations in the magnitude of attentional capture: Testing a two-process model. *Attention, Perception, & Psychophysics, 72,* 342–352.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011a). Value-driven attentional capture. *Proceedings of the National Academy of Sciences U S A, 108,* 10367–10371.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011b). Learned value magnifies salience-based attentional capture. *PLoS One, 6*(11), e27926.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2012). Generalization of value-based attentional priority. *Visual Cognition, 20,* 647–658.

Belopolsky, A. V., Schreij, D., & Theeuwes, J. (2010). What is top-down about contingent capture? *Attention, Perception, & Psychophysics, 72,* 326–341.

Berridge, K. C., & Robinson, T. E. (1998). What is the role of dopamine in reward: Hedonic impact, reward learning, or incentive salience? *Brain Research Reviews*, *28,* 309–369.

Bisley, J. W., & Goldberg, M. E. (2010). Attention, intention, and priority in the parietal lobe. *Annual Review of Neuroscience, 33,* 1–21.

Braver, T. S., Cole, M. W., & Yarkoni, T. (2010). Vive les differences! Individual variation in neural mechanisms of executive control. *Current Opinion in Neurobiology, 20,* 242–250.

Bundesen, C. (1990). A theory of visual attention. *Psychological Review, 97,* 523–547.

Bush, G. (2010). Attention-deficit/hyperactivity disorder and attention networks. *Neuropsychopharmacology, 35,* 278–300.

Christ, S. E., & Abrams, R. A. (2006). Abrupt onsets cannot be ignored. *Psychonomic Bulletin & Review, 13*, 875–880.

Corbetta, M., & Shulman, G. L. (2002). Control of goal-directed and stimulus driven attention in the brain. *Nature Reviews Neuroscience, 3*, 201–215.

Davis, C. (2010). Attention-deficit/hyperactivity disorder: associations with overeating and obesity. *Current Psychiatry Reports, 12,* 389–395.

Della Libera, C., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary reward. *Psychological Science, 17,* 222–227.

Della Libera, C., & Chelazzi, L. (2009). Learning to attend and to ignore is a matter of gains and losses. *Psychological Science, 20,* 778–784.

Dickman, S. J., & Meyer, D. E. (1988) Impulsivity and speed-accuracy tradeoffs in information processing. *Journal of Personality & Social Psychology, 54,* 274–290.

Duncan, J., Ward, R., & Shapiro, K. (1994). Direct measurement of attentional dwell time in human vision. *Nature, 369,* 313–315.

Egeth, H. E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology, 48*, 269–297.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics, 16,* 143–149.

Everitt, B. J., Dickinson, A., & Robbins, T. W. (2001).The neuropsychological basis of addictive behaviour. *Brain Research Reviews, 36,* 129–138.

Field, M., & Cox, W. M. (2008). Attentional bias in addictive behaviors: A review of its development, causes, and consequences. *Drug and Alcohol Dependence, 97,* 1–20.

Folk, C. L., & Remington, R. W. (1998). Selectivity in distraction by irrelevant featural singletons: Evidence for two forms of attentional capture. *Journal of Experimental Psychology: Human Perception & Performance, 24,* 847–858.

Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance, 18,* 1030–1044.

Folk, C. L., Remington, R. W., & Wu, S. C. (2009). Additivity of abrupt onset effects supports nonspatial distraction, not the capture of spatial attention. *Attention Perception & Psychophysics, 71*, 308–313.

Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience, 29,* 8726–8733.

Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science, 22,* 361–368.

Garavan, H., & Hester, R. (2007). The role of cognitive control in cocaine dependence. *Neuropsychological Review, 17,* 337–345.

Groman, S. M., James, A. S., Jentsch, J. D. (2008). Poor response inhibition: at the nexus between substance abuse and attention deficit/hyperactivity disorder. *Neuroscience & Biobehavioral Reviews, 33*, 690–698.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010a). Reward changes salience in human vision via the anterior cingulate. *Journal of Neuroscience, 30,* 11096–11103.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2010b). Reward guides vision when it's your thing: Trait reward-seeking in reward-mediated visual priming. *PLOS One, 5,* e14087.

Hollerman J. R., Tremblay L., & Schultz W. (1998). Influence of reward expectation on behavior-related neuronal activity in primate striatum. *Journal of Neurophysiology, 80,* 947–963.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience, 2,* 194–203.

Koob, G. F., & Le Moal, M. (1997). Drug abuse: Hedonic homeostatic dysregulation. *Science, 278*, 52–58.

Krebs, R. M., Boehler, C. N., & Woldorff, M. G. (2010). The influence of reward associations on conflict processing in the Stroop task. *Cognition, 117,* 341–347.

Kyllingsbaek, S., Schneider, W. X., & Bundesen, C. (2001). Automatic attraction of attention to former targets in visual displays of letters. *Perception & Psychophysics, 63,* 85–98.

Lien, M.-C., Ruthruff, E., & Johnston, J. V. (2010). Attentional capture with rapidly changing attentional control settings. *Journal of Experimental Psychology: Human Perception & Performance, 36,* 1–16.

Lin, J. Y., Murray, S. O., & Boynton, G. M. (2009). Capture of attention to threatening stimuli without perceptual awareness. *Current Biology, 19,* 1118–1122.

Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science, 229*, 782–784.

McClure, S. M., Berns, G. S., & Montague, P. R. (2003). Temporal prediction errors in a passive learning task activate human striatum. *Neuron*, *38,* 339–346.

Navalpakkam, V., Koch, C., Rangel, A., & Perona, P. (2010). Optimal reward harvesting in complex perceptual environments. *Proceedings of the National Academy of Sciences U S A, 107,* 5232–5237.

O'Doherty, J. P., Dayan, P., Friston, K., Critchley, H., & Dolan, R. J. (2003). Temporal difference models and reward-related learning in the human brain. *Neuron, 38,* 329–337.

Parkhurst, D., Law, K., & Niebur, E. (2002). Modeling the role of salience in the allocation of overt visual attention. *Vision Research, 42,* 107–123.

Pashler, H. (Ed.). (1998). *Attention*. London: Psychology Press.

Patton, J. H., Stanford, M. S., & Barratt, E. S. (1995). Factor structure of the barratt impulsiveness scale. *Journal of Clinical Psychology, 51,* 768–774.

Peck, C. J., Jangraw, D. C., Suzuki, M., Efem, R., & Gottlieb, J. (2009). Reward modulates attention independently of action value in posterior parietal cortex. *Journal of Neuroscience, 29*, 11182–11191.

Pessoa, L., & Engelmann, J. B. (2010). Embedding reward signals into perception and cognition. *Frontiers in Neuroscience, 4*(17). doi: 10.3389/fnins.2010.00017

Platt, M. L., & Glimcher, P. W. (1999). Neural correlates of decision variables in parietal cortex. *Nature, 400*, 233–238.

Raymond, J. E., & O'Brien, J. L. (2009). Selective visual attention and motivation: The consequences of value learning in an attentional blink task. *Psychological Science, 20,* 981–988.

Rescorla, R. A. (1999). Partial reinforcement reduces the associative change produced by nonreinforcement. *Journal of Experimental Psychology: Animal Behavior Processes, 25,* 403–414.

Robinson, T. E., & Berridge, K. C. (2003). Addiction. *Annual Review of Psychology, 54,* 25–53.

Robinson, T. E., & Berridge, K. C. (2008). The incentive sensitization theory of addiction: some current issues. *Philosophical Transactions of the Royal Society: B Biological Sciences, 363,* 3137–3146.

Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Science, 275,* 1593–1599.

Serences, J. T. (2008). Value-based modulations in human visual cortex. *Neuron, 60,* 1169–1181.

Serences, J. T., & Saproo, S. (2010). Population response profiles in early visual cortex are biased in favor of more valuable stimuli. *Journal of Neurophysiology, 104,* 76–87.

Sheppard, B., Chavira, D., Azzam, A., Grados, M. A., Umaa, P., Garrido, P., & Mathews, C. A. (2010). ADHD prevalence and association with hoarding behaviors in childhood onset OCD. *Depression & Anxiety, 27,* 667–674.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing II: Perceptual learning, automatic attending, and general theory. *Psychological Review, 84,* 127–190.

Shuler, M. G., & Bear, M. F. (2006). Reward timing in the primary visual cortex. *Science, 311,* 1606–1609.

Simen, P., Contreras, D., Buck, C., Hu, P., Holmes, P., & Cohen, J. D. (2009). Reward rate optimization in two-alternative decision making: empirical tests of theoretical predictions. *Journal of Experimental Psychology: Human Perception & Performance, 35,* 1865–1897.

Sugrue, L. P., Corrado, G. S., & Newsome, W. T. (2005). Choosing the greater of two goods: neural currencies for valuation and decision making. *Nature Reviews Neuroscience, 6,* 363–375.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement learning: An introduction*. Cambridge: MIT Press.

Theeuwes, J. (1992). Perceptual selectivity for color and form. *Perception & Psychophysics, 51*, 599–606.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135,* 77–99.

Theeuwes, J., & Godijn, R. (2002). Irrelevant singletons capture attention: Evidence from inhibition of return. *Perception & Psychophysics, 64,* 764–770.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science, 2,* 156–161.

Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. In S. Monsell, & J. Driver (Eds.), *Attention and performance* (Vol. 18, pp. 73–103). Cambridge: MIT Press.

Yantis, S. (2008). Neural basis of selective attention: Cortical sources and targets of attentional modulation. *Current Directions in Psychological Science, 17*, 86–90.

Yantis, S., & Hillstrom, A. P. (1994). Stimulus-driven attentional capture: Evidence from equiluminant visual objects. *Journal of Experimental Psychology: Human Perception & Performance, 20,* 95–107.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception & Performance, 10,* 350–374.

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

Statistical Learning and Its Consequences

Nicholas B. Turk-Browne

Abstract Statistical learning refers to an unconscious cognitive process in which repeated patterns, or regularities, are extracted from the sensory environment. In this chapter, I describe what is currently known about statistical learning. First, I classify types of regularities that exist in the visual environment. Second, I introduce a family of experimental paradigms that have been used to study statistical learning in the laboratory. Third, I review a series of behavioral and functional neuroimaging studies that seek to uncover the underlying nature of statistical learning. Finally, I consider ways in which statistical learning may be important for perception, attention, and visual search. The goals of this chapter are thus to highlight the prevalence of regularities, to explain how they are extracted by the mind and brain, and to suggest that the resulting knowledge has widespread consequences for other aspects of cognition.

Keywords Regularities · Memory systems · Perception · Selective attention · Generalization · fMRI

Introduction

Human behavior is often geared towards one object at a time, as in picking up a coffee mug, recognizing a friend's face, or noticing a car's age. This fact is even more apparent in visual search, where we typically seek one target object among other distracting objects: looking for my coffee mug among many others in the office lounge; trying to track down a particular friend at a cocktail party; or, searching for my car in an airport parking garage. How we succeed (and fail) in these kinds of searches is the topic of the 59th Nebraska Symposium, including critical factors such as attention, memory, reward, and real-world complexities. The purpose of this chapter is to highlight another important factor in visual search, 'statistical learning'.

N. B. Turk-Browne (✉)
Department of Psychology, Princeton University, Green Hall,
Princeton, NJ 08540, USA
e-mail: ntb@princeton.edu

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_6,

Although the goal of visual search is to find a target object, we rarely need to start from scratch. Rather, we can use knowledge about when and where this object appears relative to other objects to find what we're looking for. In the searches above, for example, I may know from prior experience that my coffee mug sits on top of a shelf rather than floating in air; that my friend hangs out with certain people who may also be at the party; and, that I tend to park near the elevator in parking garages. Indeed, we repeatedly come across the same people, places, and things, and over time they tend to appear in similar spatial configurations and temporal sequences. Statistical learning is an unconscious process by which we extract these patterns (or 'regularities') in how objects appear relative to each other in the visual environment.

Statistical Regularities in the Visual Environment

Regularities are aspects of the environment that repeat over time, such as the fact that beaches tend to look the same, that football players tend to appear on football fields, or that my office is across the street from a pharmacy and an ice cream shop. These regularities can be roughly classified along two dimensions: the timescale of learning (the interval over which learning happens), and the domain of knowledge (what kind of information is learned). Three varieties of regularities that differ on these dimensions are reviewed below: (1) regularities that have molded the visual system over evolutionary time and extensive training to basic physical properties of the environment ('physical regularities'), (2) regularities that are acquired throughout the lifespan about kinds of objects grouped as visual concepts/categories ('semantic regularities'), and (3) regularities that are learned in minutes about happenstance relationships between particular features and objects ('token regularities'). The boundaries between these types are inherently fuzzy, and the separation below is not intended to reify strict distinctions. Rather, this separation is used rhetorically to provide intuitions about the prevalence of regularities in many aspects of perception.

Physical Regularities

The natural environment has been stable for a long time, and so the range of input received by the visual system is constrained. For example, natural landscapes have horizontal but not vertical horizons, and natural light comes from above but not below. Over phylogenetic time, our brains have adapted to such regularities: if a large set of natural images is decomposed into independent basis functions with a constraint on efficient coding, the resulting components match the receptive field properties of V1 neurons (Olshausen and Field 1996). In other words, the visual system has been tuned during evolution to the statistics of the natural world (Simoncelli and Olshausen 2001).

Beyond regularities in the general appearance of natural images, there are also physical constraints on how objects in the world interact. For example, objects move continuously through space and time, objects only rest on horizontal surfaces unless affixed, and objects 'touch' their shadows only where they rest on the obscured surface. These and other constraints may be built into the visual system, as evidenced by studies of infant cognition. For example, 3–4 month old infants know that objects continue to exist after being occluded and cannot pass through each other (Baillargeon 1987). While present early on, knowledge of physical regularities becomes more sophisticated throughout development, growing from knowledge about simple occlusion to, for example, knowledge about containment and transparency (Baillargeon 2008).

Along with changes over phylogenetic time, extensive experience with specific input can alter the visual system over ontogenetic time. For example, repeated subliminal exposure to one direction of motion improves the detection threshold for that direction but not other directions (Watanabe et al. 2001). Such 'perceptual learning' reflects long-lasting changes in visual cortex, including enlarged and refined representations of trained stimuli (Goldstone 1998; Fahle and Poggio 2002; Sasaki et al. 2010). This kind of learning may supplement hard-wired general assumptions about the visual environment (as described above) by further tuning the visual system to the natural and artificial environments that we inhabit during our lifetime.

Semantic Regularities

In addition to physical properties and object interactions, regularities exist in how *types* of objects appear in the visual environment. For example, regardless of color and shape, all fire hydrants appear on sidewalks not on top of mailboxes, and all toilets appear in bathrooms and not kitchens. The visual system relies on these canonical locations and scene contexts for object recognition, such that objects violating semantic regularities are recognized more poorly (Palmer 1975; Biederman et al. 1982; Davenport and Potter 2004; cf. Hollingworth and Henderson 1998). In addition, there are semantic regularities in terms of whether and how objects can change over time. For example, stoplights can change color but stop signs can't, and humans can move locations but not shrink in size (at least not quickly). These kinds of regularities influence visual awareness: improbable object changes have a higher incidence of change blindness (Beck et al. 2004).

The distinction between semantic and physical regularities is not rigid. For example, images representing the same scene category (e.g., a beach) have similar global physical properties. Despite differences in the local details of particular exemplars (e.g., the number of bathers, color of sand, presence of beach chairs, etc.), such properties allow us to quickly and effortlessly categorize novel scenes (Oliva and Torralba 2006). However, the objects belonging to a semantic regularity can *but need not be* especially physically similar. For example, the exemplars of many categories (e.g., buildings, trees, dogs) are physically heterogeneous. More generally, while physical properties can denote particular semantic classes, this is just one type of linking cue: objects may share semantic relations in spite of physical dissimilarity because, for example, they accomplish a similar function (e.g., microwave and BBQ grill) or are required components of a broader concept (e.g., the equipment, players, and field in football).

Token Regularities

In addition to regularities in what semantic types of objects co-occur, the visual environment is littered with regularities between object *tokens*. For example, when learning a new navigation route, the sequence of landmarks that one comes across is highly regular (e.g., a religious billboard, then a fast food restaurant, then an open field, etc.). Such regularities can be spatial in addition to temporal. For example, beyond physical and semantic constraints, there are regularities in the configuration of objects in each kitchen that are unique with respect to all other kitchens. Finally, some token regularities are not inherently spatial or temporal, such as sets of faces that tend to co-occur (e.g., in groups of friends).

There is no necessary semantic relationship between objects in token regularities (e.g., a billboard and a restaurant, two facial identities), and indeed many such regularities are composed of conceptually distant objects (e.g., the curtains, electronics, and art in a room). More importantly, semantic classes often provide too coarse a grain of description (e.g., moving around my living room in the dark or finding a particular utensil in the kitchen does not benefit much from my general knowledge of living rooms or kitchens). Token regularities also differ from the kinds of physical regularities described earlier that persist over very long timescales: particular sequences and configurations of objects are happenstance and can change frequently. Moreover, I will emphasize the *relational* nature of token regularities—that regularities exist in the statistical relationships (e.g., conditional probabilities) between two or more objects—whereas perceptual learning of physical regularities is often viewed as imprinting of discrete stimuli (Goldstone 1998). In sum, statistical learning about token regularities occurs over faster timescales (e.g., when moving to a new city, or using a new computer operating system) and requires tracking particular object exemplars rather than semantic classes or physical properties per se.

Studying Statistical Learning

Unlike physical and semantic regularities, which are built into the visual system or learned over the course of development, token regularities can be introduced and tested in a laboratory setting. In addition to introducing new regularities, experiments about statistical learning typically use novel objects without preexisting semantic associations. Thus, subjects begin *de novo* when entering these experiments, with

Fig. 1 Auditory statistical learning. Subjects are briefly exposed to a structured speech stream, and subsequently express greater familiarity with words than part-words and non-words (see text for details). (Based on Saffran et al. 1996a)

minimal prior knowledge and biases relevant to the regularities (or to the objects from which they are generated). This experimental control allows us to characterize the operation of statistical learning precisely. Given the space of possible regularities in the real world, statistical learning is generally studied in an artificial environment containing a small number of relatively simple regularities. Whether these environments provide a suitable proxy for how statistical learning operates in more naturalistic contexts is an important and open question. Later, I will consider evidence that statistical learning can handle some of the complexities of natural environments.

Where It All Started: Auditory Statistical Learning

This chapter is predominantly focused on *visual* statistical learning. But the impetus for current research on visual statistical learning (including the author's own initial interest), comes from the study of auditory statistical learning; in particular, from studies of how auditory statistical learning may be important for word learning and segmentation in speech streams, and for language acquisition more generally (Saffran et al. 1996a,b).

In a typical study of this type (Fig. 1; Saffran et al. 1996a), subjects (infants, kids, or adults) are exposed to a brief speech stream composed of syllables. Unbeknownst to subjects, the stream of syllables has been constructed to contain statistical regularities in terms of which syllables follow each other. In particular, the 12 total syllables (e.g., tu, pi, ro, bi, da, ku, go, la, bu, pa, do, ti) have been assigned to four trisyllabic 'words' (e.g., tupiro, bidaku, golabu, padoti), such that the first syllable is always followed by the second, and the second always by the third (transitional probabilities of 1.0). What occurs after the third syllable of each word is any of the first syllables from the three other words (back-to-back word repetitions are prevented; transitional probabilities of 0.33). Insofar as subjects pick up on the statistical discrepancy between within- and between-word transitions, they can use the transitional probabilities to segment the words. Critically, other than the differential transitional probabilities, there are no prosodic or other cues that would indicate word boundaries.

After only 2 min of exposure, 8-month-old infants expressed knowledge of the words by dishabituating to new words that were composed of the same syllables, but in combinations with lower (or zero) transitional probabilities based on what they had heard (Saffran et al. 1996a). While not discussed further here, these results have been extended significantly to incorporate, among other things, prosody (e.g., Thiessen and Saffran 2003), word meaning (e.g., Graf Estes et al. 2007), and multiple languages (Gebhart et al. 2009).

Temporal Visual Statistical Learning

As highlighted earlier, statistical regularities are ubiquitous in the visual environment. Regularities exist in both how objects are arrayed in space and how they appear over time. Tasks have been developed to study statistical learning in both of these dimensions. Visual statistical learning about temporal regularities is considered first, because it is studied in a similar way to auditory statistical learning. Temporal information in vision arises in one of at least two ways:

First, the world is dynamic and objects move over time. The way that objects move and change over time contains regularities, such as the trajectory of a tennis ball, the different views of a car as it passes, or the sequence of movements underlying an action (Baldwin et al. 2008). These kinds of temporal regularities have an intrinsic structure or order, and may underlie our representations of events (Avrahami and Kareev 1994; Zacks and Tversky 2001).

Second, the way that we sample the world (both static and dynamic parts) guarantees that visual input will be distributed over time. Specifically, we acquire information about one part of the visual environment at a time through eye movements and covert shifts of attention. We must continually sample the environment due to the capacity limitation of visual short-term memory—i.e., the small number of items that can be held simultaneously in memory and the brief durations of such representations (e.g., Henderson and Hollingworth 2003; Zhang and Luck 2009). Objects captured by each eye fixation or attentional shift receive enhanced processing, such that scanning creates a parade of objects through the visual system. Temporal regularities of this type may not be inherently ordered or structured, but can exist in terms of which objects appear in a broader context (e.g., the set of objects in one room), or can be driven by the likelihood that two or more objects will be fixated sequentially (e.g., because of close spatial proximity or similar salience). Thus, temporal regularities are prevalent in vision as well as in audition, due to the existence of inherently temporal events and actions, and to our serial sampling of spatial environments.

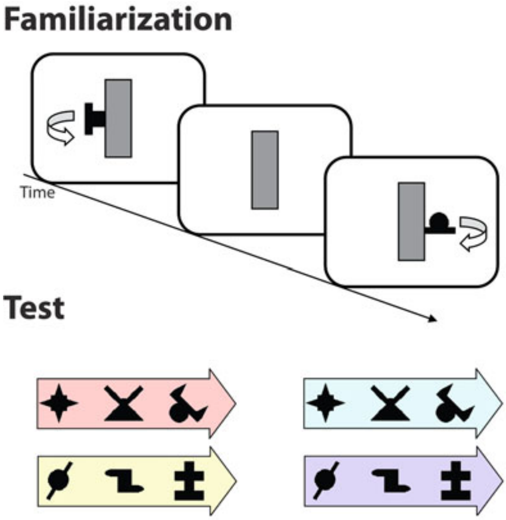


Fig. 2 Temporal visual statistical learning. Subjects are briefly exposed to a structured stream of shapes, and subsequently express greater familiarity with triplets than foils (see text for details). (Based on Fiser and Aslin 2002)

Studies of temporal visual statistical learning often rely on an experimental design adapted from the original auditory statistical learning experiments (Fig. 2; Fiser and Aslin 2002; see also Olson and Chun 2001). A set of novel nonsense shapes (e.g., A, B, C, D, E, F, G, H, I, J, K, L) is divided without replacement into temporal pairs or triplets (e.g., ABC, DEF, GHI, JKL). During an initial phase, subjects are exposed to a continuous stream constructed from these triplets, with shapes appearing one at a time (e.g., DEFJKLABCDEFGHI. . .). Critically, subjects are not oriented to the presence of triplets. The shapes can either appear as the only thing on the screen, or cycle back and forth behind an occluder changing identity each time when occluded.

After several minutes of exposure, subjects are then given a surprise familiarity test. On each trial they are presented with two three-item sequences: (1) a triplet from familiarization, and (2) a foil generated from the same shapes but rearranged into new groupings (e.g., AEI, DHL, GKC, JBF). Each triplet is tested against each foil, to ensure equal frequency of the alternatives at test. Thus, triplets can be discriminated from foils only based on the higher transitional probabilities within triplets vs. within foils. Subjects perform very well in this task, despite the fact that learning is incidental and that subjects often express low confidence in their test judgments. Such findings have been used to suggest that statistical learning happens automatically as a result of mere exposure to regularities (Saffran et al. 1999; Fiser and Aslin 2002).

Spatial Visual Statistical Learning

Temporal regularities derive partly from repeated sampling of structured spatial environments. Indeed, objects and their parts do not appear in random locations in scenes, but rather in predictable locations based on the locations of other objects and

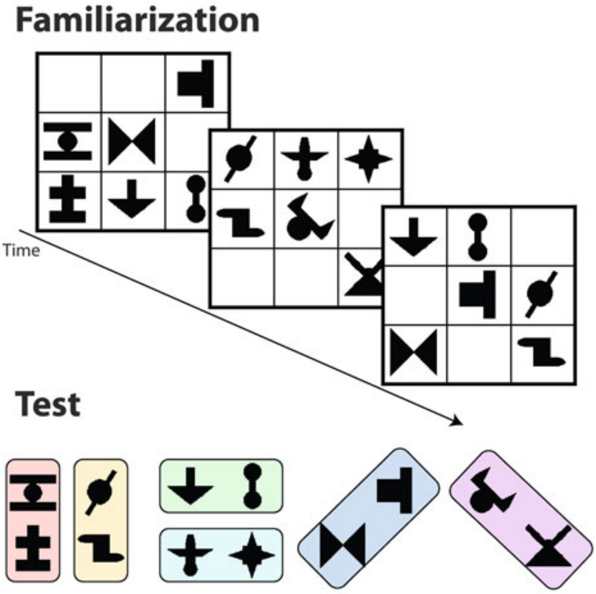


Fig. 3 Spatial visual statistical learning. Subjects are briefly exposed to a series of grids containing spatial pairs, and subsequently express greater familiarity with pairs than foils (see text for details). (Based on Fiser and Aslin 2001)

parts (Biederman et al. 1982). These spatial regularities are partly due to physical constraints (sofas cannot float in the sky) and partly due to semantic knowledge (hood fans appear above rather than below stoves). However, additional regularities exist in the particular configuration of parts and objects. Sensitivity to these regularities may be important for learning about the composition of novel objects and about the layout of novel environments. Indeed, space is the dominant dimension in the visual modality (as time is in the auditory modality; Kubovy 1988), and thus it is important to characterize how statistical learning operates over spatial regularities.

Studies of spatial visual statistical learning employ a design that is somewhat different from temporal statistical learning studies. However, at its core, this design replicates the essential property of temporal statistical learning experiments: that regularities can only be segmented on the basis of statistics. In other words, much like the continuous temporal stream from which shape triplets must be extracted, our visual system is confronted with an undifferentiated spatial image and must segment the meaningful chunks based solely on spatial probabilities.

In such studies (Fig. 3; Fiser and Aslin 2001; see also Chun and Jiang 1999), subjects are presented with simplified visual scenes generated from a 3 × 3 grid. Six shapes are shown in each grid and—unbeknownst to subjects—the shapes appear in spatial pairs. Specifically, a set of 12 shapes (e.g., A, B, C, D, E, F, G, H, I, J, K, L) is randomly assigned without replacement to six pairs (e.g., AB, CD, EF, GH, IJ, KL). Two pairs are assigned to each of three orientation types: horizontal (e.g., AB, CD), vertical (e.g., EF, GH), and diagonal (e.g., IJ, KL). Each scene is generated by selecting one pair of each orientation (e.g., AB, EF, IJ) and placing them on the grid so that all shapes are abutted by at least one shape beyond the paired shape. With these constraints, a total of 144 grids can be generated. Subjects are briefly presented with each scene and the entire stream lasts 5 min. Critically, to learn the pairs, subjects need to extract the joint probabilities between adjacent shapes. To measure learning, subjects then complete a familiarity test in which they are repeatedly presented with a pair and a foil (e.g., AD, CB, EH, GF, IL, KJ) and must choose which alternative is most familiar. Performance at discriminating pairs from foils is excellent in this task.

A modified design provides a test of *which* statistics subjects can extract. In the basic task, joint and conditional probabilities are confounded (pairs have higher joint and conditional probabilities than foils): i.e., $P(AB) = 0.5 > P(AD) = 0$ and $P(B|A) = 1.0 > P(D|A) = 0$. To examine whether subjects could pick up conditional probabilities in isolation, Fiser and Aslin (2001) doubled the frequency of a subset of the pairs (e.g., AB, IJ) such that the joint probability of two adjacent shapes that crossed pairs (e.g., BJ) was equal to the joint probability of infrequent original pairs (e.g., CD). Importantly, these two pair types differed in conditional probability: $P(D|C) = 1 > P(J|B) \sim 0.5$. Subjects were still able to successfully discriminate these pairs, suggesting that they had extracted conditional probabilities. In contrast to joint probabilities, conditional probabilities may be especially important for prediction (see ‘Anticipation’ section).

Aside: Other Related Kinds of Learning

Visual statistical learning as a cognitive process may occur in many task contexts. Most prominently, a form of visual statistical learning may underlie contextual cueing during visual search (Chun and Jiang 1998). In such tasks, subjects are presented with visual search arrays composed of a T target and L distractors. Critically, several configurations of targets and distractors are repeated during the experiment, and search performance gets faster and faster as a result of learning. This task provides an elegant online measure of learning, allowing the timecourse of learning to be assessed, unlike most statistical learning tasks. However, the spatial regularities in contextual cueing are qualitatively different from statistical learning: subjects learn configurations of locations *per se*, rather than configurations of object identities. Thus, contextual cueing may help in locating a target during visual search given the known locations of other objects (cf. Kunar et al. 2007), while spatial statistical learning results in stimulus-specific knowledge of relative locations (e.g., that A is above B). Variations on contextual cueing in which the identities of distractors predict the location and/or identity of the target (Chun and Jiang 1999; Endo and Takeda 2004) provide a bridge to the spatial statistical learning literature. There are many other related forms of learning, including in the serial reaction time task where sequences of spatial locations (rather than object identities) can be readily learned (e.g., Mayr 1996).

The Nature of Statistical Learning

Statistical regularities are everywhere in the world, and evidence of statistical learning is almost equally ubiquitous. Statistical learning occurs in every sensory modality in which it has been tested (e.g., Conway and Christiansen 2005), in time and in space (e.g., Fiser and Aslin 2001, 2002), for many types of stimuli (e.g., Saffran et al. 1999; Baldwin et al. 2008), and in many subject populations (e.g., Kirkham et al. 2002; Toro and Trobalón 2005). Statistical learning seems to be a powerful and fundamental part of cognition... but how does it work? Over the past seven years, my collaborators and I have attempted to uncover the nature of statistical learning. Here I review findings from this research program as an update on what is known about statistical learning.

When Does Learning Take Place?

An important challenge for statistical learning is that the huge set of regularities we experience is a small subset of all possible regularities that could exist. For example, consider walking through your home airport: you may repeatedly encounter the same restaurants, ticket agents, and gates, and they can even appear in fixed temporal orders and spatial layouts; but on any given trip, you encounter many other objects, such as random tourists, food stands, bestsellers in the bookstore, etc., and these objects may be gone by your next trip. How do we extract the meaningful and stable relationships between certain objects, while discounting transient intrusions from other objects? The problem is one of 'combinatorial explosion': to learn about which relationships are regular, one must in principle represent the co-occurrence of all possible groupings of objects. In other words, how can statistical learning determine *a priori* which relationships will reappear in the future? Indeed, if such advance knowledge were possible, learning would have already taken place! This is especially problematic for statistical learning since it occurs in an unsupervised manner (Fiser and Aslin 2002), without feedback about when and what to learn.

In a series of behavioral studies (Turk-Browne et al. 2005), we explored this issue by examining the automaticity of visual statistical learning. Namely, we asked whether statistical learning occurs whenever the visual system is confronted with regularities, or whether selective attention can determine which regularities are learned. The latter possibility would place an important constraint on when statistical learning takes place. In our studies, subjects were presented with a temporal stream of nonsense shapes. This stream was itself composed of two separate streams that had been interleaved (Fig. 4). Each stream appeared in a different color, and contained a unique set of shapes. Unbeknownst to subjects, each of the two colored streams was generated using triplets of shapes as in the standard temporal visual statistical learning task.

Critically, subjects were instructed to attend to one color of shapes (e.g., green), and to perform a one-back task only for shapes appearing in that color (i.e., detecting

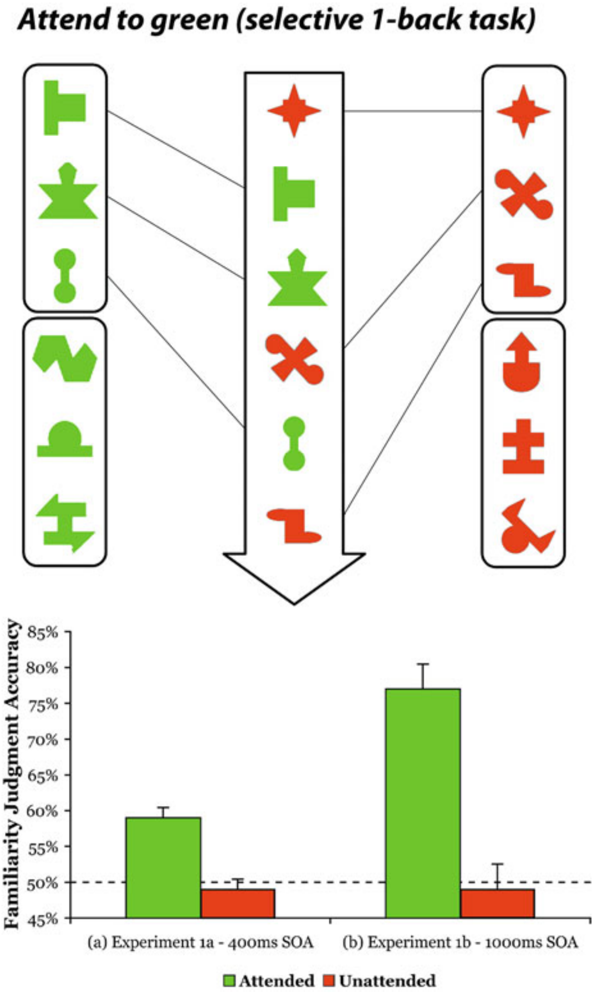


Fig. 4 Selective attention. Familiarization stream constructed from two sets of regularities, one attended and the other unattended. Statistical learning only occurs for attended regularities. (Based on Turk-Browne et al. 2005)

whether the current green shape was the same as the last green shape). Shapes were presented one at a time, and were thus all attended spatially. However, only the attended shapes were task-relevant. Insofar as learning occurs by mere exposure to regularities (e.g., Saffran et al. 1999), regularities in both streams may be learned. If selective attention instead gates statistical learning, we would expect better learning of the attended shapes.

Across five experiments, we not only found better learning of the attended regularities, but no learning whatsoever of the unattended regularities. This was even true when we used an implicit response time (RT) measure that may be more sensitive to unattended learning. Thus, selective attention determines the input to statistical learning. These results were recently used as a case study of how to prove null hypotheses (Gallistel 2009)—in our case, the null hypothesis was that *no* learning occurs without attention (confirmed by way of Bayesian analysis).

While statistical learning of object relationships occurred only when the objects were task-relevant, this learning happened without conscious awareness. Indeed, statistical learning was robust despite the fact: (1) that subjects were not informed about the presence of regularities, (2) that they performed a distracting cover task (one-back) during familiarization, (3) that the shapes were presented quickly, (4) that regularities from the two streams were interleaved, adding noise to the transitional probabilities, and (5) that learning was evident in an implicit RT measure. Moreover, during careful debriefing in the RT experiment, no subjects expressed awareness of the structure in the displays. These findings suggest that statistical learning is and is not automatic: selective attention to objects is required for their relationships to be learned, but once this input has been selected, learning takes place without conscious intent or effort.

Does Task-Relevance Guarantee Learning?

The previous section described evidence that statistical learning is constrained by top-down selective attention, i.e. that statistical learning only occurs for task-relevant objects. However, the simple working memory task used in those experiments is only one of a large number of cognitive tasks that we routinely engage in. Does statistical learning take place whenever objects are task-relevant, or does the nature of the task matter? In the latter case, finding that some tasks are better or worse for statistical learning may help uncover the component processes at work during statistical learning.

In a recent behavioral study (Zhao et al. 2011), we examined one particular kind of task that bears an interesting resemblance to statistical learning: statistical summary perception. The ability to perceive summary statistics has received much consideration recently (e.g., Ariely 2001; Chong and Treisman 2003; Alvarez and Oliva 2008). In statistical summary perception tasks, subjects are presented with a set of objects, and are instructed to make subsequent judgments about some statistical property of the set (e.g., mean size). Critically, subjects can often extract summary statistics from a set of objects without being able to identity the constituent members of the set (e.g., Ariely 2001; Alvarez and Oliva 2008).

On the surface, statistical learning and statistical summary perception are quite different: statistical learning involves extracting regularities over repeated experience, while statistical summary perception involves extracting statistics from a single display; and statistical learning involves acquiring stimulus-specific relationships (i.e. that this particular object co-occurs with another particular object), while statistical summary perception (by definition) involves representing the general properties of a collection. Despite these surface differences, however, these two processes are both inherently statistical: they involve aggregating a sample, and distilling this sample

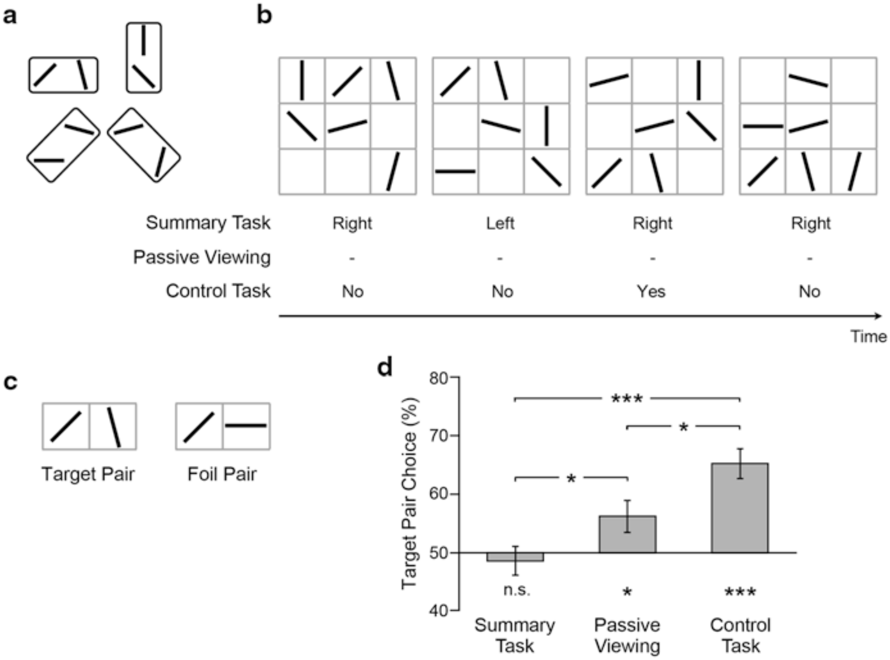


Fig. 5 Statistical summary perception. While being exposed to spatial line pairs, different groups of subjects performed a statistical summary task (*Is the mean line orientation to the left or right of the vertical meridian?*), dual-task control (*Are there any duplicate lines?*), or passive viewing. Statistical learning was blocked by the summary task. (Based on Zhao et al. 2011)

to statistics (e.g., transitional probability matrix, or mean). Thus, statistical learning and statistical summary perception may interact in meaningful ways. We tested what happens to statistical learning when objects are attended during a statistical summary task.

The design of this study most closely matches a spatial visual statistical learning task. Subjects were presented with grids of lines (rather than shapes) of different orientations (Fig. 5). There were eight possible orientations, and unbeknownst to subjects, the orientations were arranged into four spatial pairs. On each grid, three of the four pairs were selected and placed such that at least one other pair appeared adjacently. Thus, the only cue to pair structure was the co-occurrence of particular orientations.

The critical manipulation concerned which task subjects performed during familiarization. One group of subjects was instructed to passively watch the grids, as has been done in prior studies (Fiser and Aslin 2001). A second group of subjects performed a statistical summary perception task on the grids: they were instructed to compute the mean orientation of the lines, and to discriminate this mean as being to the left or right of the vertical meridian. A third group of subjects performed a control task to ensure that differences between the passive viewing and statistical summary tasks could not be attributed to merely needing to perform a second task (which could perhaps enhance attention in a generic way). The control task group determined whether each grid contained a duplicate line. The displays were identical in all conditions, and thus any consequences of the three tasks must reflect the impact of different task sets on statistical learning.

Statistical learning was robust in the passive viewing and control task conditions, but was weaker—actually, non-existent—in the summary task condition. This suggests that computing summary statistics may interfere with statistical learning, possibly because of a reliance on shared statistical computations, or because of a reliance on different spatial scales of attention. That is, global attention helps summary performance (Chong and Treisman 2005), while local attention may be necessary for stimulus-specific learning. In any event, these results provide a further constraint on when statistical learning takes place, and suggest an unforeseen connection between two varieties of statistical processing that had been studied separately.

Aside: The Quest for a Great Cover Task

The findings presented so far demonstrate that what you are doing determines whether you will learn. This fact shows the importance of an often-overlooked methodological aspect of studies on statistical learning: the task that subjects perform during familiarization. The earliest studies of statistical learning employed no task at all during familiarization (Saffran et al. 1996b; Fiser and Aslin 2001, 2002). Passive listening or viewing is not without merit: as we have already seen, certain tasks may be detrimental to statistical learning (Zhao et al. 2011). At the same time, anybody who has run a statistical learning experiment with passive viewing knows the inherent awkwardness in asking subjects to "simply watch". Subjects seem puzzled by such instructions ("how can I do nothing"), and often suspect that the experimenter has ulterior motives. This may lead subjects to implement idiosyncratic strategies in searching out the meaning behind the displays. While such searches rarely turn up the true structure, they may nevertheless be detrimental to learning. In other words, not giving subjects a task does not mean that they won't impose their own task, and, without careful debriefing, such tasks are unknowable and uncontrolled from the experimenter's perspective.

By analogy, the study of functional connectivity in the brain is typically conducted while subjects rest in a functional magnetic resonance imaging (fMRI) scanner with no task (Fox and Raichle 2007). Because no task has been imposed, it is assumed that any resulting patterns of brain activity reflect the stable intrinsic functional architecture of the brain. Yet, recent tasks seep into resting state brain activity (e.g., Stevens et al. 2010), suggesting that subjects are not performing no task (or a default task), just that the experimenters typically do not know what it is.

Thus, a certain amount of experimental control is gained by asking subjects to perform a task during familiarization. It is worth clarifying that such tasks are never to learn the structure of the display (i.e., intentional learning). Not only would this kind of explicit task violate the unsupervised spirit of statistical learning, intentional learning tasks are not always beneficial to learning: in a classic artificial grammar learning study, intentional learning instructions *hurt* learning (Reber 1976). Instead, tasks during familiarization can be considered 'cover' tasks in that they do not themselves disclose the presence of regularities.

When designing studies of statistical learning for which the nature of the cover task does not matter *per se*, one often simply wants to ensure that statistical learning takes place. From many successful and unsuccessful experiments, it seems that the most conducive tasks to learning are those that emphasize the identity of the objects being perceived. For example, the interleaved one-back task described above requires maintaining each shape in working memory, and judging whether the next shape has the same identity. Other tasks that emphasize object identity work well too as shown below, such as detecting motion jitter (requires some shape processing to resolve motion correspondence) and classifying objects into semantic categories. Tasks that do not require any object recognition are less conducive to learning, such as requiring detection of fixation luminance changes. Whether particular tasks are conducive (boost statistical learning with respect to some baseline) or permissive (prevent the blocking of learning that other intrinsic tasks might cause) remains to be further clarified. However, the big picture of why some tasks are good and others are bad has more than methodological implications: humans may be able to control when statistical learning takes place simply by engaging in certain behaviors, without even realizing that they possess this power.

What Happens During Learning?

Conventional statistical learning designs all have something in common: they require two parts. In particular, statistical learning occurs during the first part (familiarization), and then is tested in the second part (test). A separate test is needed because of the canonical use of passive viewing during familiarization. The test is often a two-alternative forced choice familiarity task in which subjects must discriminate between a regularity from familiarization and a foil constructed from the same elements arranged into a new sequence or configuration. While this kind of test has been used successfully in many studies (including many of our own), this design has two drawbacks: First, supposedly incidental and unconscious statistical learning is being tested with an explicit familiarity judgment. While familiarity can be informed by implicit processes, a familiarity test may not be the most sensitive measure of implicit learning. Second, testing learning after the fact means that information about what happens during learning (e.g., about the timecourse of learning) is lost. Indeed, by using two parts, one makes assumptions about how much exposure is needed for learning since familiarization must end before the test can begin—familiarity cannot be measured at multiple intervals, since this would explicitly cue subjects to the presence of regularities.

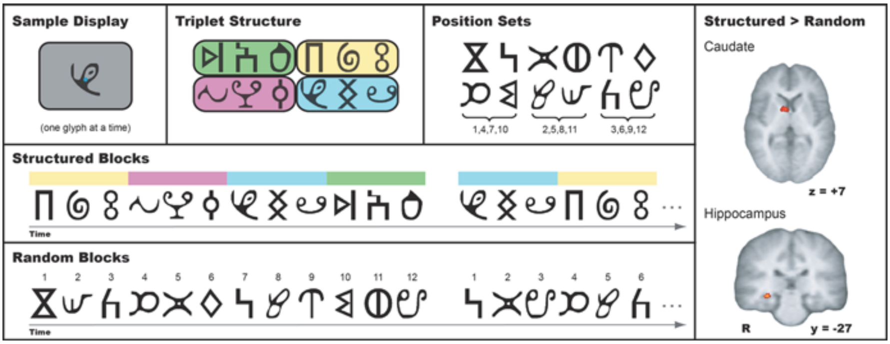


Fig. 6 Neural evidence. Subjects were exposed to blocks of glyphs constructed from triplets or from otherwise matched but random sequences. Both the caudate and posterior hippocampus responded more strongly to structured blocks, providing evidence of statistical learning after only a handful of triplet repetitions. (Based on Turk-Browne et al. 2009)

To resolve these issues, we conducted an fMRI study of what happens in the brain during familiarization (Turk-Browne et al. 2009). We sought to answer three questions about the process of statistical learning in this study: What is the relationship between statistical learning and other forms of learning? How efficient is statistical learning? What is the relationship between our online measure of statistical learning in the brain and subsequent familiarity? Using fMRI, we monitored changes in the brain related to statistical learning without requiring an online behavioral measure (e.g., Hunt and Aslin 2001), and perhaps before these changes manifest in behavior.

Subjects were presented with blocks of nonsense glyphs from ancient alphabets, where each block contained 12 glyphs (Fig. 6). Their task was to detect whenever one of the glyphs jiggled on the screen (a subtle horizontal motion). The blocks alternated between two sets of 12 unique glyphs. One of the sets, which was used to generate the Structured blocks, was composed of four triplets of glyphs. In generating each Structured block, the triplets were randomly sequenced in an order that had never previously been seen. The glyphs appeared sequentially, one at a time, and thus triplets could only be discriminated based on the higher transitional probabilities for glyphs within vs. between triplets. As a comparison, the other set of glyphs, which was used to generate the Random blocks, was composed of three 'position sets'. Namely, four of the glyphs could appear in the same serial position as the first item in each triplet, four other glyphs could appear in the same serial position as the second item in each triplet, etc. Other than these constraints, the order of glyphs was randomized in each Random block. The overall block sequence alternated between Structured and Random blocks, and we contrasted neural responses to the two block types as a measure of learning. The Structured and Random blocks were identical in terms of item frequency and serial position frequency, and thus any neural differences must reflect sensitivity to the stronger transitional probabilities within triplets in the Structured blocks.

We were especially interested in whether statistical learning would engage well-known memory systems involved in other types of learning. Two systems of particular interest included the medial temporal lobe and the striatum. The medial temporal lobe, and the hippocampus in particular, is the primary brain system for declarative memory (Cohen and Eichenbaum 1993), and has been implicated in some implicit forms of relational (Ryan et al. 2000), configural (Chun and Phelps 1999), and sequence learning (Schendan et al. 2003), which may all be related to statistical learning. The striatum, and the caudate in particular, is often linked to non-declarative or procedural learning, and has been implicated in many forms of implicit learning, including category learning (Seger and Cincotta 2005), artificial grammar learning (Lieberman et al. 2004), and motor sequence learning (Toni et al. 1998). After a few minutes of exposure, the hippocampus and caudate responded more strongly to the Structured blocks than to the Random blocks. These findings suggest potential connections between statistical learning and other forms of learning and memory. We are actively exploring the role of the MTL in statistical learning, for example, examining whether representations in MTL cortex are tuned based on regularities (see Miyashita 1993).

To examine the timecourse of learning, we explored when the difference between Structured and Random blocks emerged during learning using smaller windows of time. In the caudate and hippocampus, as well as other regions, we found initial evidence that statistical learning can occur very quickly, after only 2–3 triplet repetitions. To examine the relationship between this neural evidence of statistical learning and the conventional familiarity measure used in previous studies, we repeated our analyses including only those subjects who performed at or below chance on the familiarity test. Despite not exhibiting any statistical learning by conventional standards, the caudate in these subjects nevertheless discriminated Structured vs. Random blocks. This result suggests that neural evidence of statistical learning can exist without explicit familiarity, perhaps preceding behavioral expressions of learning. In sum, this study uncovered some of the dynamics of learning and identified the neural systems involved.

How Does Learning Handle Real-World Complexity?

We have so far considered when statistical learning takes place and what happens during statistical learning. In this and the next section, I describe what is represented in memory as a result of statistical learning. This question becomes salient when we consider the types of real-world objects over which statistical learning operates. In contrast to the monochromatic shapes used in the prior studies, real-world objects are complex, or ‘multidimensional’. For example, an object can have many features, including color, texture, and shape, and these features can vary over the different parts of an object. How does statistical learning scale up to more complex stimuli? More specifically, when confronted with regularities of complex objects such as colorful shapes, what is learned: patterns of bound objects (e.g., sequences or configurations

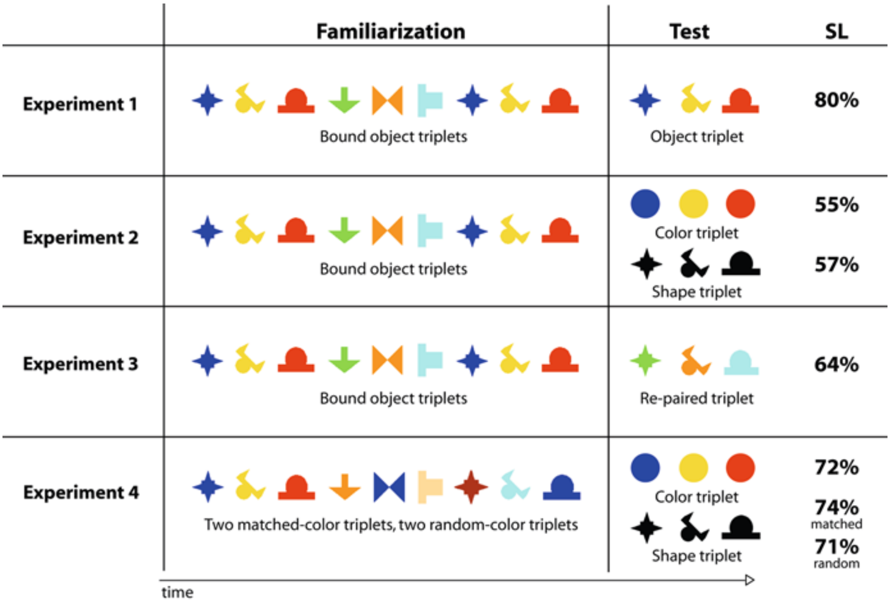


Fig. 7 Multidimensional objects. When exposed to a stream of shape-color objects (each shape has a unique color), statistical learning is better expressed to the bound object (Experiment 1) than to separated features (Experiment 2), or a recombination of features (Experiment 3). When mapping between shape and color is variable, learning is robust for features (Experiment 4). (Based on Turk-Browne et al. 2008)

of shape-color conjunctions) or patterns of the constituent features (e.g., sequences or configurations of shapes or colors)? The answer is not obvious. On one hand, objects provide a strong organizing system for features, and object-based effects have been observed in many domains (e.g., Luck and Vogel 1997; Scholl 2001). On the other hand, some features are not stable over time (e.g., luminance, shadows) and thus learning over conjunctions may be inefficient.

We examined these issues in a series of behavioral studies that looked at statistical learning for multidimensional objects (Turk-Browne et al. 2005). In a temporal visual statistical learning task, subjects were shown a continuous stream of shape-color objects (Fig. 7). Each of twelve shapes was assigned a unique color, and these shape-color conjunctions were grouped into four triplets without subjects' knowledge. To examine whether statistical learning operates over bound objects or separated features, we used a transfer logic: if triplets of bound objects are learned, then familiarity for object triplets should be high at test and familiarity for triplets of the separated shape or color features should be low. Instead, if triplets of separated features are learned, then familiarity for the feature triplets should be robust. The result was that statistical learning extracted the regularities between objects: familiarity was much higher for triplets of bound objects than for triplets of separated features, or even triplets created at test from new combinations of feature triplets.

We further hypothesized that non-diagnostic features should be discounted in learning about objects. For example, certain objects have canonical colors (e.g., fruits and vegetables) while others do not (e.g., furniture, books, hats). We examined whether feature triplets would be better learned when combined into an object with variable or non-diagnostic features from another dimension. We assigned two shape triplets to have fixed and unique colors (as before), and the remaining two shape triplets to appear in randomized colors. Statistical learning now occurred for the shape triplets that were paired with variable colors. Surprisingly, the shape triplets paired with fixed colors were also now better learned, as were the color triplets that appeared with these shape triplets. These findings suggest that the general covariance between feature dimensions determines whether statistical learning outputs knowledge about object or feature regularities: when covariance is high, regularities are learned at the level of objects; when covariance is low, regularities are learned at the level of features.

Another way to think about these results is that statistical learning may always be object-based, but that feature diagnosticity determines what counts as an object: when colors are diagnostic of shapes (and vice versa), objects consist of both a shape and color; when colors are not diagnostic of shapes, objects consist of a single shape or color feature. This interpretation suggests that statistical learning can operate within objects—in addition to between objects—extracting relationships between features (and feature dimensions) to determine which conjunctions are reliable. An object-based bias, coupled with a sensitivity to feature covariance, may allow statistical learning to operate in natural contexts containing regularities among complex objects.

How Flexible Is Learning?

The research above raises a broader question: after learning regularities, how flexible are the acquired representations to changes in the appearance of objects and/or their relationships? In other words, what kinds of changes in the environment can be tolerated when expressing knowledge about regularities? One type of change that has already been discussed is color: we found that statistical learning of colored shapes can be expressed despite the removal of color, but only when colors are non-diagnostic. Another common type of change relates not to the features of individual objects, but rather to the spatiotemporal patterns in which multiple objects appear. Indeed, the visual environment is highly dynamic, meaning that sequences and configurations change over time. For example, the set of people we encounter at the office is relatively stable, but the particular sequence and locations in which we encounter them may vary from day to day. Moreover, because a major source of temporal information in vision comes from eye movements, changes in how we fixate in a given environment (e.g., entering through the back door of a house instead of the front door) will lead to very different sequences. If statistical learning is highly specific to the spatiotemporal details of experience (e.g., Jiang and Song 2005), such variability would be disastrous to the expression of learning. Indeed, given the inherent variability of the world, statistical learning may not be of much use if it cannot cope with some changes between when regularities are acquired and expressed.

We examined these issues in a series of behavioral studies (Turk-Browne and Scholl 2009) where changes were made between the familiarization and test phases of visual statistical learning tasks. Specifically, we examined the flexibility of statistical learning (i.e., generalization) using a similar 'transfer' logic to above: if statistical learning can be expressed despite some change at test, then the changed aspect of the displays is not an integral part of the learned knowledge.

We first examined whether statistical learning could be expressed at test for triplets that appeared in a new temporal order. Subjects watched a continuous stream containing four triplets in the familiarization phase. We then tested whether temporal order was an important part of what they had learned by including both the original (forwards) triplets and, on other trials, reversed (backwards) versions of the same triplets. Insofar as temporal visual statistical learning can generalize across temporal order, subjects should exhibit some familiarity with the backwards triplets. Surprisingly, they expressed equally strong familiarity with the forwards and backwards triplets. This suggests that the knowledge generated by statistical learning is invariant to order.

Since forwards and backwards triplets elicited the same amount of familiarity, were they even distinguishable? Was all order information lost? To test this, we ran an additional study in which subjects were asked to discriminate forwards vs. backwards triplets directly (rather than forwards and backwards triplets vs. nonword foils). Subjects now expressed greater familiarity with forward triplets. These results suggest that statistical learning generalizes across temporal order when order is not needed to discriminate triplets vs. foils, but that this information is represented nonetheless and can be accessed when necessary (when two alternatives can only be discriminated based on order).

Since temporal order information was easily discarded during statistical learning, we then tested whether temporal information was necessary at all for learning to be expressed at test. Following the standard temporal statistical learning task, subjects were tested on triplets vs. foils, but now the objects in the triplets were presented simultaneously in a spatial configuration (Fig. 8). If statistical learning can generalize over the time dimension entirely, subjects should be able to discriminate spatial triplets vs. foils. This is what we found, even when the test displays were presented too quickly for eye movements. These results suggest that temporal visual statistical learning may be useful in building up spatial representations.

Finally, we also examined the flexibility of spatial visual statistical learning. We asked whether learning of spatial pairs would transfer to temporal displays, just like temporal triplets transferred to spatial displays. Saving the details for a little later, the basic result was that spatial learning led to temporal processing benefits. These findings suggest that statistical learning produces knowledge that can be applied flexibly in new contexts. In sum, the studies reported in this section show the power of statistical learning, and its fit to the constraints and properties of the natural environment.

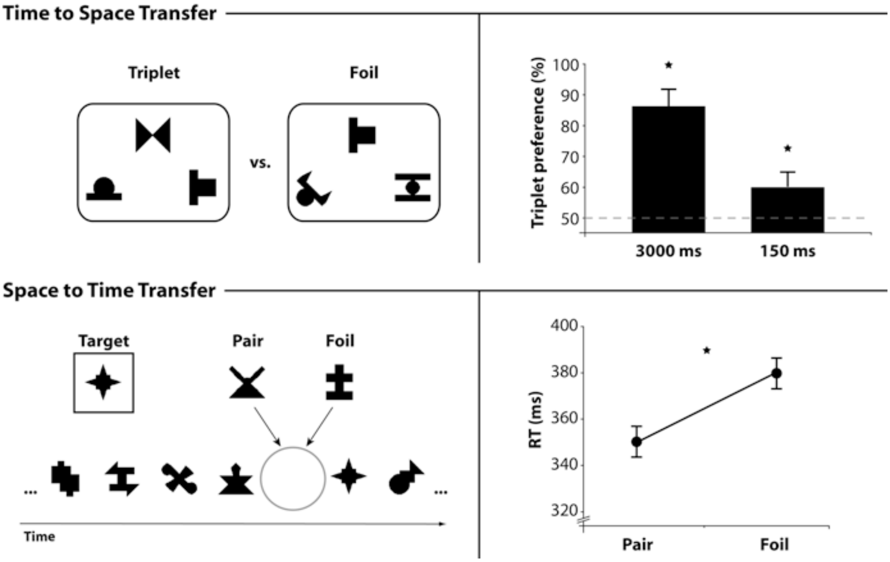


Fig. 8 Spatiotemporal transfer. After temporal visual statistical learning, triplets can be recognized in spatial configurations, even when they are presented too briefly for eye movements. Moreover, after spatial visual statistical learning, pairs lead to RT priming effects in time. (Based on Turk-Browne and Scholl 2009)

Consequences of Statistical Learning

What is statistical learning good for? Many of the studies discussed thus far have relied on familiarity as a measure of statistical learning. Does statistical learning only result in increased familiarity with regularities? This seems somewhat epiphenomenal, in the sense that fleeting familiarity signals may not have much functional significance for ongoing behavior. In this section, I present three other possible consequences of statistical learning, and consider how they might relate to visual search. While there has not been much research specifically testing how statistical learning and visual search interact, the results presented here suggest that such an investigation could be fruitful.

Anticipation

One potential consequence of statistical learning that is often emphasized in the application of statistical learning to language is ‘segmentation’—the use of regularities to parse the world into meaningful lexical units or chunks (Giroux and Rey 2009). Similar approaches have been used to model visual statistical learning (Orbán et al.

2008). However, both familiarity and segmentation are *retrospective*, requiring entire patterns to be experienced before they can occur. That is, the relevant kind of familiarity and the ability to place a boundary occur only after an entire regularity has been perceived. Indeed, a hallmark of chunking models is that chunks are indivisible into their component parts (Orbán et al. 2008). Such rigidity may help us recognize higher-order chunks, such as scenes, events, and words, but is not well suited to the dynamic nature of perception. We have thus become interested in whether statistical learning can be used *prospectively*.

In a recent fMRI study (Turk-Browne et al. 2010), we examined what happens at the beginning of temporal regularities—can the brain use knowledge of regularities to anticipate upcoming, predictable objects? For example, imagine repeatedly meeting a particular host when entering a restaurant before being led into the dining room; over time, does perceiving the host's face trigger automatic anticipation of the layout and content of the dining room? In other words, can we rely on learned regularities to live with one foot in the perceptual future? Such anticipation could have important consequences for perception, allowing us to more quickly recognize objects and cope with occluded or degraded visual input.

Subjects in our study were presented with a continuous stream of face and scene images, appearing one at a time and separated by a jittered inter-trial interval (for fMRI analysis purposes). Subjects' task was to decide for each image whether it depicted a face or a scene. Unbeknownst to subjects, eight of the 12 images in each scanning run were divided into four cross-category pairs: two face → scene pairs and two scene → face pairs. The remaining four images (two faces, two scenes) were unpaired, and served as a baseline. We were interested in three conditions: the First image in each of the pairs, the Second image in each of the pairs, and the Unpaired images. Insofar as statistical learning affords anticipation, we predicted that: (a) First images would elicit anticipatory responses compared to Unpaired images, and (b) this anticipation would lead to facilitated processing of Second images.

This latter prediction was apparent in RTs: subjects were faster to categorize the Second vs. Unpaired images, suggesting that the First image had caused priming. Such associative priming effects, resulting from statistical learning, have been observed in two of our other studies as well (Turk-Browne et al. 2005; Turk-Browne and Scholl 2009). In both of these cases, the task on each trial required detecting a pre-specified target embedded in a rapid serial visual presentation stream. We manipulated where the target object appeared with respect to preceding items. For example, in studying whether spatial visual statistical learning transfers to the temporal dimension, the target object was sometimes preceded by the object that it was paired with in space during familiarization and sometimes by an equally familiar foil object. RTs were faster when the target was preceded by its spatial pair, providing evidence that spatial learning can produce temporal cueing benefits. Such effects of statistical learning on object detection and discrimination highlight an important behavioral consequence of statistical learning for object recognition.

Back to the fMRI study of anticipation. Interestingly, RTs to the First images were slower than to Unpaired images. We interpreted this as evidence of anticipation: when perceiving an object that affords predictions about what will come next, this

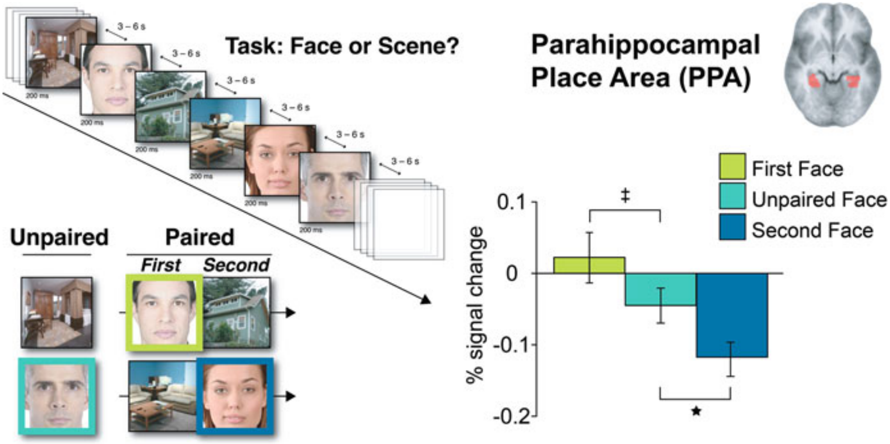


Fig. 9 Implicit perceptual anticipation. Objects that are predictive of what will appear next in the trial sequence produced robust hippocampal activity (not shown). Faces, which otherwise do not activate the scene-selective PPA (see Unpaired condition), elicited enhanced PPA responses when they predicted a scene, and suppressed PPA responses when they could be predicted as non-scenes. (Based on Turk-Browne et al. 2010)

predictive response may act as an implicit dual task and interfere with the current task of categorizing the image in front of you. We are following up on this effect, to determine whether it reflects a generic cost of anticipation (i.e., that performance is slowed whenever anticipation is possible), or prospective response conflict (i.e., that the response for the anticipated image interfered with the required response for the current image). The current study was equivocal: because pairs always contained objects of different categories (for reasons to be described shortly), the response to the First image always conflicted with the response to the anticipated Second image.

Along with these behavioral results, we observed a robust and selective fMRI response in the right hippocampus to the First images compared to the Unpaired images. It is worth emphasizing that First and Unpaired images are identical in all respects but one: they are presented an equal number of times and neither image type is itself predictable based on which image came before; but critically, First images are predictive and allow anticipation of the Second images, while Unpaired images afford no such predictions. Surprisingly, the hippocampal response was evident after fewer than six repetitions of each pair, again providing evidence for the remarkable speed of statistical learning. Finally, although the hippocampus has been traditionally linked to declarative or explicit forms of memory (Cohen and Eichenbaum 1993), subjects reported no awareness that pairs even existed in the stream, suggesting that these anticipatory effects reflect *implicit* perceptual anticipation.

We were especially interested in whether implicit perceptual anticipation could influence visual cortex; in particular, whether regularities could be exploited to potentiate visual processing (Fig. 9). To examine this possibility, we localized category-selective parts of ventral temporal cortex that responded selectively to faces and scenes. If regularities can modulate the visual system, then faces that are predictive of scenes may prospectively elicit scene-related activity (and vice versa). Indeed, scene-selective cortex (the parahippocampal place area, PPA; Aguirre et al. 1998; Epstein and Kanwisher 1998) provides a clean test of this hypothesis since it does not otherwise respond to faces. This can be seen in the response of the PPA to Unpaired faces, which was slightly below baseline. Critically, First faces, which reliably predicted that a particular scene would appear next, elicited an enhanced PPA response. Along with this enhancement, the PPA response to Second faces, which were predictable as non-scenes, was suppressed. These findings provide evidence that anticipation based on statistical regularities can prospectively alter visual cortex.

The idea that statistical learning can prime the detection and recognition of predictable or probable objects may have important consequences for visual search. Indeed, search is speeded when the configuration or identities of distractors predict the identity of the target during visual search (Chun and Jiang 1999; Endo and Takeda 2004). Such effects may result from an effect of regularities on biased competition (Desimone and Duncan 1995), whereby perceiving one object may prioritize processing for associated objects. A related alternative is that regularities could transiently increase familiarity for associated objects, with increased target familiarity leading to more efficient search (Flowers and Lohr 1985).

Relatedly, semantic regularities influence visual search by guiding eye movements to locations in scenes where the target is semantically licensed or probable (Henderson et al. 1999; Neider and Zelinsky 2006; Torralba et al. 2006). For example, receiving a scene preview without a search target facilitates subsequent search when the target is introduced, ostensibly because the scene context in the preview could be leveraged to restrict search to likely target locations (Hollingworth 2009). The brain seems keen on generating predictions, and if those predictions are grounded in regularities, searching for a target object may be aided by perceiving and exploiting associated distractors.

Shifting Modes of Attention

The efficiency of visual search depends on the extra time needed to find a target item for each additional distractor added to a search set, quantified as the slope of RT as a function of set size. Efficiency varies a lot depending on the nature of the target and distractors (Wolfe 2001). Search is most efficient (shallow slopes) when the target differs from all distractors in terms of the presence of one simple feature (Treisman and Gelade 1980). Such targets may be detected pre-attentively, reflecting parallel processing of all items in the display—a ‘distributed mode’ of attention (Chong and Treisman 2005). In contrast, search is least efficient (steep slopes) when the target and distractors share features, and the target is defined by the conjunction of features (Duncan and Humphreys 1989). Such targets require serial processing of each item in the display—a ‘focused mode’ of attention (Treisman and Souther 1985). Search

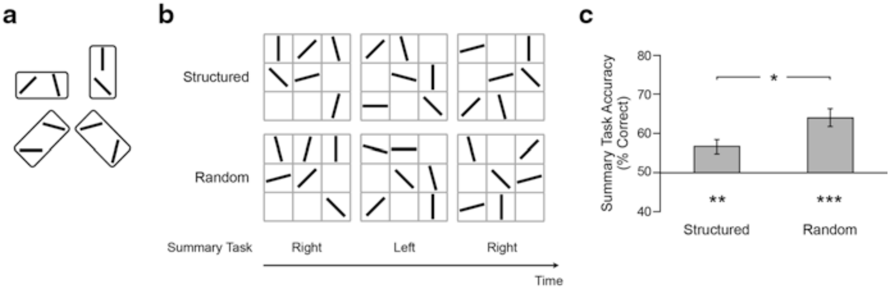


Fig. 10 Regularities and summary perception. Subjects performed a summary task (judging the mean line orientation) over grids that contained spatial regularities (Structured) vs. grids that were otherwise matched but did not contain spatial regularities (Random). Summary perception improved when regularities were removed. (Based on Zhao et al. 2011)

efficiency is thus better when a target can be found with distributed attention, and worse when focused attention is required.

The need for distributed vs. focused attention during visual search is typically attributed to properties of targets and distractors. However, statistical regularities themselves may also control whether attention can be allocated in a distributed or focused manner. Specifically, we are exploring the hypothesis that regularities among local objects draw attention into a more focused mode. While current evidence for this hypothesis is sparse, the idea grew out of a recent behavioral study (Zhao et al. 2011). In that study, we examined how statistical summary perception, a process that depends on distributed attention (Chong and Treisman 2005), is affected by statistical learning. This is the counterpart to an experiment reported earlier on the reverse—how statistical learning is affected by statistical summary perception. The study is described below before I return to the question of how regularities affect attention.

Subjects were presented with grids of lines, and judged whether the mean line orientation was to the left or right of the vertical meridian (Fig. 10). One group of subjects (Structured) received grids that contained spatial pairs. Another group of subjects (Random) received the same grids, but where the locations of lines were shuffled on each trial to destroy the spatial regularities. We predicted that the mere presence of regularities might prompt the visual system to attempt to learn, which may in turn interfere with summary perception. This is exactly what we found: Judgments of mean line orientation were less accurate in the Structured condition.

We conducted a follow-up study to verify that statistical learning *per se* was interfering with summary perception, and not that it was more difficult to summarize the Structured vs. Random displays because they differed in some unintended way. In this new study, one group of subjects was pre-exposed to Structured displays during the duplicate detection control task described earlier (allowing for pre-learning of the line pairs), while another group performed the control task over Random displays (resulting in the same item familiarity, but no knowledge of the line pairs). Both groups then performed the summary task over Structured displays. The group that had pre-learned the pairs gave more accurate summary judgments than the group that was being exposed to the pairs for the first time, presumably because the latter group was engaged in statistical learning while the former group was not. Critically, the displays were identical during this phase, eliminating the possibility of stimulus differences, and supporting the claim that statistical learning interferes with summary perception.

Given that distributed attention is necessary for summary perception (Chong and Treisman 2005), one interpretation of our results is that regularities (even if not learned successfully) shift attention to a focused mode. There are other potential interpretations too, such as that statistical learning and summary perception rely on shared statistical computations, and that engaging in one process interferes with the other by blocking necessary resources. It will therefore be important for future experiments to test the idea that local regularities attract focused attention. Nevertheless, these findings demonstrate that statistical learning can have *costs* for other cognitive processes, along with the benefits for familiarity, segmentation, and anticipation described earlier.

Biasing Locations in Space

In typical visual search tasks, the target location on any given trial is random. When target locations are not random, search performance improves. For example, when target location is predictable from past experience with a specific configuration of distractors, the target can be found more quickly (e.g., Chun and Jiang 1998). Moreover, independently of distractor locations, the probability of targets appearing in specific locations also influences search. For example, when targets appear in one general location on 75 % of trials, target discrimination is facilitated at that location and inhibited at others (Geng and Behrmann 2005; see also Umemoto et al. 2010).

All of these findings suggest that regularities in the locations of targets and distractors can improve search by biasing the allocation of spatial attention. Besides regularities related to the search task, statistical learning may also independently bias the allocation of spatial attention in a way that could influence search. For example, the presence of statistical regularities at one location may draw spatial attention. Consider the natural environment: at any given moment, we are confronted with too many potential sources of regularities and must decide what to learn. We have previously seen that top-down or goal-directed attention to a subset of visual input limits statistical learning to that input. Without such task goals, however, how does the visual system decide what to learn? One possibility is that, like many basic visual cues (e.g., abrupt onsets), regularities themselves act as a cue for spatial attention. Indeed, we have obtained preliminary support for this possibility: targets are detected more quickly when they appear at a spatial location containing temporal regularities, even when the identity, timing, and location of the target are orthogonal to the regularities (Zhao et al., in press).

Another way that statistical learning could influence attention is after learning has occurred: knowing the structure at one location may allow us to divert limited attentional and working memory resources elsewhere in space. We have tested this possibility in a recent study (Al-Aidroos and Turk-Browne, in prep). Subjects first completed a temporal visual statistical learning task in which shapes were presented one at a time at central fixation and subjects detected one-back repetitions. In one condition, the stream was constructed from triplets; in the other condition, the stream was randomized. After some initial exposure to the stream, subjects were instructed to continue detecting repetitions, but also to determine the orientation of rare low-contrast Gabor probes that appeared in the periphery. Probe discrimination was more accurate when the probes were presented during the triplet vs. random streams. Having acquired the triplets during the initial exposure, subjects may have been released from the burden of statistical learning at fixation and better able to monitor the peripheral locations where probes appeared. Statistical learning can thus have diverse effects on spatial attention, and as a consequence, affect other processes that are modulated by attention.

Conclusions

The purpose of this chapter was to describe where regularities exist in the visual environment, to consider ways of studying how regularities are learned, to review recent studies about when statistical learning occurs, what happens during learning, and what is represented as a result, and finally to consider some potential consequences of statistical learning. All of this research highlights the reciprocal connection between perception and memory: statistical learning is an important mechanism for recording visual experience into memory; in turn, learned regularities influence ongoing perception, whether it be object recognition, statistical summary perception, or spatial attention. Given our robust ability to extract and use regularities from the visual environment, statistical learning may play a broad and fundamental role in many cognitive processes, including visual search.

References

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. *Neuron, 21,* 373–383.

Alvarez, G. A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science, 19,* 392–398.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12,* 157–162.

Avrahami, J., & Kareev, Y. (1994). The emergence of events. *Cognition, 53,* 239–261.

Baillargeon, R. (1987). Object permanence in 3 ½ - and 4 ½ -month-old infants. *Developmental Psychology, 23,* 655–664.

Baillargeon, R. (2008). Innate ideas revisited: For a principle of persistence in infants' physical reasoning. *Perspectives on Psychological Science, 3,* 2–13.

Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action via statistical structure. *Cognition, 106,* 1382–1407.

Beck, M. R., Angelone, B. L., & Levin, D.T. (2004). Knowledge about the probability of change affects change detection performance. *Journal of Experimental Psychology: Human Perception and Performance, 30,* 778–791.

Biederman, I., Mezzanotte, R. J., & Rabinowitz, J. C. (1982). Scene perception: Detecting and judging objects undergoing relational violations. *Cognitive Psychology, 14*, 143–177.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research, 43,* 393–404.

Chong, S. C., & Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics, 67,* 1–13.

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology, 36,* 28–71.

Chun, M. M., & Jiang, Y. (1999). Top-down attentional guidance based on implicit learning of visual covariation. *Psychological Science, 10,* 360–365.

Chun, M. M., & Phelps, E. A. (1999). Memory deficits for implicit contextual information in amnesic subjects with hippocampal damage. *Nature Neuroscience, 2*, 844–847.

Cohen, N. J., & Eichenbaum, H. (1993). *Memory, Amnesia, and the Hippocampal System.* Cambridge: MIT Press.

Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition B31,* 24–39.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15,* 559–564.

Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience, 18,* 193–222.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Endo, N., & Takeda, Y. (2004). Selective learning of spatial configuration and object identity in visual search. *Perception & Psychophysics, 66,* 293–302.

Epstein, R., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature, 392,* 598–601.

Graf Estes, K. M., Evans, J. L., Alibali, M. W., & Saffran, J. R. (2007). Can infants map meaning to newly segmented words? *Psychological Science, 18,* 254–260.

Fahle, M., & Poggio, T. (2002). *Perceptual Learning*. Cambridge: MIT Press.

Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12,* 499–504.

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences USA, 99,* 15822–15826.

Flowers, J. H., & Lohr, D. J. (1985). How does familiarity affect visual search for letter strings? *Perception & Psychophysics, 37*, 557–567.

Fox, M. D., & Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience, 8,* 700–711.

Gallistel, C. R. (2009). The importance of proving the null. *Psychological Review, 116,* 439–453.

Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing structures in midstream: Learning along the statistical garden path. *Cognitive Science, 33,* 1087–1116.

Geng, J. J., & Behrmann, M. (2005). Spatial probability as an attentional cue in visual search. *Perception & Psychophysics, 67*, 1252–1268.

Giroux, I., & Rey, A. (2009). Lexical and sub-lexical units in speech perception. *Cognitive Science, 33,* 260–272.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49,* 585–612.

Henderson, J. M., & Hollingworth, A. (2003). Global transsaccadic change blindness during scene perception. *Psychological Science, 14,* 493–497.

Henderson, J. M., Weeks, P. A., & Hollingworth, A. (1999). The effects of semantic consistency on eye movements during complex scene viewing. *Journal of Experimental Psychology: Human Perception and Performance, 25,* 210–228.

Hollingworth, A. (2009). Two forms of scene memory guide visual search: Memory for scene context and memory for the binding of target object to scene location. *Visual Cognition, 17,* 273–291.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127,* 398–415.

Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General, 130,* 658–680.

Jiang, Y., & Song, J.-H. (2005). Hyperspecificity in visual implicit learning: Learning of spatial layout is contingent on item identity. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 1439–1448.

Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83,* B35–B42.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy. *Journal of Experimental Psychology: Human Perception and Performance, 14,* 318–320.

Kunar, M. A., Flusberg, S., Horowitz, T. S., & Wolfe, J. M. (2007). Does contextual cuing guide the deployment of attention? *Journal of Experimental Psychology: Human Perception and Performance, 33,* 816–828.

Lieberman, M. D., Chang, G. Y., Chiao, J., Bookheimer, S. Y., & Knowlton, B. J. (2004). An event-related fMRI study of artificial grammar learning in a balanced chunk strength design. *Journal of Cognitive Neuroscience, 16,* 427–438.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390,* 279–280.

Mayr, U. (1996). Spatial attention and implicit sequence learning: Evidence for independent learning of spatial and nonspatial sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 350–364.

Miyashita, Y. (1993). Inferior temporal cortex: Where visual perception meets memory. *Annual Review of Neuroscience, 16,* 245–263.

Neider, M. B., & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision Research, 46,* 614–621.

Oliva, A., & Torralba, A. (2006). Building the gist of a scene: The role of global image features in recognition. *Progress in Brain Research, 155,* 23–36.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381,* 607–609.

Olson, I.R., & Chun, M.M. (2001). Temporal contextual cuing of visual attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27,* 1299–1313.

Orbán, G., Fiser, J., Aslin, R. N., & Lengyel, M. (2008). Bayesian learning of visual chunks by human observers. *Proceedings of the National Academy of Sciences USA, 105*, 2745–2750.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*, 519–526.

Reber, A. S. (1976). Implicit learning of synthetic languages: The role of instructional set. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 2,* 88–94.

Ryan, J. D., Althoff, R. R., Whitlow, S., & Cohen, N. J. (2000). Amnesia is a deficit in relational memory. *Psychological Science, 11,* 454–461.

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996a). Statistical learning by 8-month-old infants. *Science, 274,* 1926–1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996b). Word segmentation: The role of distributional cues. *Journal of Memory & Language, 35,* 606–621.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition, 70,* 27–52.

Sasaki, Y., Nanez, J. E., & Watanabe, T. (2010). Advances in visual perceptual learning and plasticity. *Nature Reviews Neuroscience, 11,* 53–60.

Schendan, H. E., Searl, M. M., Melrose, R. J., & Stern, C. E. (2003). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. *Neuron, 37,* 1013–1025.

Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition, 80,* 1–46.

Seger, C. A., & Cincotta, C. M. (2005). The roles of the caudate nucleus in human classification learning. *Journal of Neuroscience, 25,* 2941–2951.

Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience, 24,* 1193–1216.

Stevens, W. D., Buckner, R. L., & Schacter, D. L. (2010). Correlated low-frequency BOLD fluctuations in the resting human brain are modulated by recent experience in category-preferential visual regions. *Cerebral Cortex, 20,* 1997–2006.

Thiessen, E. D., & Saffran, J. R. (2003). When cues collide: Use of stress and statistical cues to word boundaries by 7-to 9-month-old infants. *Developmental Psychology, 39,* 706–716.

Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during motor sequence learning: A whole-brain fMRI study. *NeuroImage, 8,* 50–61.

Toro, J. M., & Trobalón, J. B. (2005). Statistical computations over a speech stream in a rodent. *Perception & Psychophysics, 67*, 867–875.

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, 113,* 766–786.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General, 114,* 285–310.

Turk-Browne, N. B., & Scholl, B. J. (2009). Flexible visual statistical learning: Transfer across space and time. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 195–202.

Turk-Browne, N. B., Jungé, J., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134,* 552–564.

Turk-Browne, N. B., Isola, P. J., Scholl, B. J., & Treat, T. A. (2008). Multidimensional visual statistical learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 399–407.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21,* 1934–1945.

Turk-Browne, N. B., Scholl, B. J., Johnson, M. K., & Chun, M. M. (2010). Implicit perceptual anticipation triggered by statistical learning. *Journal of Neuroscience, 30,* 11177–11187.

Umemoto, A., Scolari, M., Vogel, E. K., & Awh, E. (2010). Statistical learning induces discrete shifts in the allocation of working memory resources. *Journal of Experimental Psychology: Human Perception and Performance, 36,* 1419–1429.

Watanabe, T., Náez, J. E., & Sasaki, Y. (2001). Perceptual learning without perception. *Nature, 413,* 844–848.

Wolfe, J. M. (2001). Asymmetries in visual search: An introduction. *Perception & Psychophysics, 63,* 381–389.

Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin, 127,* 3–21.

Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science, 20,* 423–428.

Zhao, J., Ngo, N., McKendrick, R., & Turk-Browne, N. B. (2011). Mutual interference between statistical summary perception and statistical learning. *Psychological Science, 22,* 1212–1219.

Zhao, J., Al-Aidroos, N., & Turk-Browne, N. B. (in press). Attention is spontaneously biased towards regularities. *Psychological Science*.

Overcoming Hurdles in Translating Visual Search Research Between the Lab and the Field

Kait Clark, Matthew S. Cain, Stephen H. Adamo and Stephen R. Mitroff

Abstract Research in visual search can be vital to improving performance in careers such as radiology and airport security screening. In these applied, or "field," searches, accuracy is critical, and misses are potentially fatal; however, despite the importance of performing optimally, radiological and airport security searches are nevertheless flawed. Extensive basic research in visual search has revealed cognitive mechanisms responsible for successful visual search as well as a variety of factors that tend to inhibit or improve performance. Ideally, the knowledge gained from such laboratory-based research could be directly applied to field searches, but several obstacles stand in the way of straightforward translation; the tightly controlled visual searches performed in the lab can be drastically different from field searches. For example, they can differ in terms of the nature of the stimuli, the environment in which the search is taking place, and the experience and characteristics of the searchers themselves. The goal of this chapter is to discuss these differences and how they can present hurdles to translating lab-based research to field-based searches. Specifically, most search tasks in the lab entail searching for only one target per trial, and the targets occur relatively frequently, but field searches may contain an unknown and unlimited number of targets, and the occurrence of targets can be rare. Additionally, participants in lab-based search experiments often perform under neutral conditions and have no formal training or experience in search tasks; conversely, career searchers may be influenced by the motivation to perform well or anxiety about missing a target, and they have undergone formal training and accumulated significant experience searching. This chapter discusses recent work that has investigated the impacts of these differences to determine how each factor can influence search performance. Knowledge gained from the scientific exploration of search can be applied to field searches but only when considering and controlling for the differences between lab and field.

Keywords Applied visual search · Rare-target visual search · Multiple-target visual search · Satisfaction of search · Motivation · Anxiety · Expertise · Training · Learning · Radiology

---

K. Clark (✉) · M. S. Cain · S. H. Adamo · S. R. Mitroff
Center for Cognitive Neuroscience, Department of Psychology and Neuroscience,
Duke University, LSRC Building, 450 Research Drive, Durham, NC 27708, USA
e-mail: kait.clark@duke.edu

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_7,

General Introduction

Imagine two hypothetical individuals, Tyler and Olivia. Tyler is an undergraduate at Duke University. He is taking an Introductory Psychology course that requires him to participate in three experimental psychology studies. For one of the required experiments, Tyler selects a "visual search" study in a cognitive psychology laboratory and signs up at a time immediately after his morning Linear Algebra class. Tyler arrives at the lab at 11 a.m. and fills out the necessary consent forms before being led into a dimly lit testing room where he is instructed to search for 'T'-shaped targets on a computer screen among 'L'-shaped distractor items and to indicate whether a 'T' is present or absent on each trial. Tyler is bored by the time the practice segment is complete and begins to muddle through the task, exerting the minimal effort required. While he is sure some graduate student really cares about how he performs, he is not too concerned. After checking his email on his phone for the 25th time, he finally nears the end of this exercise. He speeds up as the end of the hour approaches; his stomach is beginning to rumble, and he's getting hungry. He knows that the faster he can get through the remainder of the trials, the sooner he'll be able to eat lunch. Tyler completes the task and heads to a university café while the experimenters in the lab examine his data, pooled with data from other participants just like Tyler, with the intentions of drawing conclusions about the nature of human visual search processes.

Olivia is an X-ray operator at the Raleigh-Durham Airport in North Carolina. She has worked with the Transportation Security Administration for 6 years and is currently a full-time employee who works 5 days a week. She starts her normal shift at 5 a.m., and during each shift she works several 30 min stints at the X-ray machine, searching for contraband that may be hidden in passengers' luggage. Olivia has completed an initial training on optimal search strategies and numerous refresher courses that are designed to make her a better searcher. As her supervisors monitor her performance, she is cognizant of maintaining a certain level of accuracy in order to keep her job. Additionally, Olivia is keenly aware of the consequences of letting a bag with a bomb slip through the cracks. Very few, if any, of the bags Olivia searches contain any actual lethal items, but her job is to remain vigilant and conduct thorough searches on every bag, regardless of the improbability of finding a bomb, a gun, or a knife.

Clearly, our hypothetical individuals, Tyler and Olivia, are conducting very different visual searches, with very different motives, in very different environments. Moreover, the differences highlighted above only scratch the surface of the variability between the nature of inexperienced and expert searchers. Realistically, how can search performance between undergraduates and career searchers be compared, given the drastic differences between these scenarios? Can we conclude anything from Tyler's performance about how Olivia should be performing her job? Does understanding Olivia's performance inform cognitive theories of visual search?

An eventual goal of basic research is translating findings from the **lab** to the **field**;[1] however, researchers often struggle to overcome the inherent differences between the sterile, controlled environment of a research lab and the complex, messy environment of the real world. The study of visual search—finding a target amongst distractors—is an excellent paradigm to illustrate this relationship. On the one hand, visual search is a powerful research method for psychologists, as it encompasses several aspects of cognition (e.g., memory, perception, attention). As such, researchers have extensively studied and theorized about the nature of visual search (see Nakayama and Martini 2010; Eckstein 2011, for recent reviews). On the other hand, visual searches are regularly conducted in everyday tasks outside of a laboratory setting.

The goal of this chapter is to discuss several of the hurdles encountered when moving between the lab and the field and how they might be overcome. We will begin with a general overview of visual search followed by a brief review of the research history and theories. We will then introduce some relevant applied visual search findings before detailing four primary hurdles that stand in the way of translating search findings between the lab and the field:

1. **Target prevalence**: Is search performance affected by the relative likelihood of a target being present (e.g., do searchers perform worse if targets are rarely present)?
2. **Number of targets and target categories**: Does search performance decline if a searcher is required to search for more than one target in the same image (e.g., multiple fractures in a medical X-ray) or more than one possible target type (e.g., a gun *or* a bomb in a luggage X-ray)?
3. **Motivation and anxiety**: How does the context within which a search is conducted affect performance? Is search performance helped or hindered by added motivation or anxiety?
4. **Level of experience**: Is search performance altered by expertise? How might years of experience alter visual search strategies or abilities?

Overview of Visual Search

Visual search is the process of finding specific target items within an environment based on particular visual features or semantic information. In its simplest form, visual search could operate via basic pattern matching; for example, detecting a red vertical line in a field of green horizontal lines would solely require invoking a red and/or vertical pattern template. However, even this easy visual search depends on many attentional and perceptual factors, and more complex searches move well beyond basic pattern matching. More common visual searches, both in the lab

---

[1] For the purposes of this chapter, we will use "lab" to refer to visual search experiments conducted by cognitive psychologists with inexperienced searchers in a laboratory setting, and we will use "field" to refer to visual searches conducted as part of normal activities in naturalistic settings that are often done by highly trained "expert" searchers.

(e.g., finding a target 'T' amongst distractor 'L's, finding a particular shape amongst variable distractors) and in the field (e.g., finding a tumor in a radiograph, finding keys in a purse), involve an array of cognitive processes. Search involves perception (i.e., processing and interpreting visual features), attention (i.e., allocating resources to the relevant areas of a visual area), and memory (i.e., storing a representation of the target item or items). Thus, search represents a fruitful and exciting area of research.

In the lab, visual search has been used extensively to learn about cognition. For example, search studies have informed theories of basic perception (e.g., Wolfe et al. 2005), the structure of visual short-term memory (e.g., Alvarez and Cavanagh 2004), and attentional capture (e.g., Yantis and Jonides 1996; Franconeri et al. 2005), to name just a few. Beyond using visual search as a powerful tool for understanding cognitive processing, researchers have also focused on search as an experimental paradigm with the goal of understanding how searches are conducted. Over the past several decades, psychological research has made tremendous headway in understanding the processes responsible for performing visual search tasks and the mechanisms that allow for the successful identification of target items. The findings from visual search research have been extensive, and, in turn, the contributions to the scientific community have been invaluable.

Given the relevance of visual search to real-world environments as well, ideally, what is learned from studying search processes in the lab can be applied to searches in the field. Beyond the vital function of search in navigating our everyday lives, the search performance of radiologists, X-ray operators, and many others can be life-or-death critical. As recent technological advancements have allowed for the improvement of screening techniques, additional key advancements lie in understanding the cognitive processes of the searchers themselves, identifying common search errors, and improving the manner in which searches are conducted.

A Brief History of Visual Search Research

Early Evidence from Non-Human Visual Searchers

While the current era of visual search research is largely laboratory-based, the first investigations of search were focused on its primary goal in the world—survival. Animals engage in survival activities that require visual search, such as finding food, avoiding predators, detecting a potential mate's signs, and locating appropriate shelter. Search was perhaps first scientifically investigated in 1890 by Edward Poulton, a zoologist who was interested in how animals elude predators. Poulton noted that a single species tends to evolve many different appearances, making it more difficult to be detected by predators, a phenomenon known as *cryptic pattern polymorphism*. For instance, a single species of forest moth appears with many different wing patterns, and Poulton noticed that it is more difficult for a bird to search for a multiple kinds of targets simultaneously than to search for a single type. The added difficulty that accompanies a search for multiple kinds of targets is now a well-documented idea in cognitive psychology (e.g., Menneer et al. 2007) and is especially relevant to current-day X-ray baggage screening at airports, as security officers must search for a number of potentially hazardous items.

Expanding on Poulton's observations, Tinbergen (1960) found that insectivorous birds maximized their rate of detection by confining their searches to only a few prey types at a given time and by focusing on either the most common prey available or those that had been seen most recently. In effect, this research demonstrated that non-human animals are sensitive to the statistics of their environments and are able to quickly adapt to maximize search efficiency, and contemporary work with human searchers has found similar results (e.g., Cain et al., in press). Pigeon studies have illustrated that search is specialized for ecologically relevant tasks, as pigeons demonstrate a fantastic ability to find food (e.g., Bond 1983) and effectively optimize their rate of food discovery. These early studies of search with non-human species have served to both establish the evolutionary basis of search processes and demonstrate the practical nature of visual search. For the remainder of this chapter, we focus on human visual search research that has built upon, and complements, these and other non-human search findings.

Early Evidence from Human Visual Searchers

Speculations about the nature of human visual search—also from an applied angle—began with Bernard Koopman in the 1950s, when he explored theories of search in the context of radar operators locating enemy ships (Koopman 1956a,b). Tasked by the US Navy to systematically determine the location of enemy ships and lost personnel, he revealed many basic theoretical properties of visual search, such as the distribution of attention and the criteria for termination (Koopman 1957), that remain fundamentally important for current theories of search (e.g., Chun and Wolfe 1996).

Cognitive psychologists entered the visual search research arena in the 1960s and 1970s and have played a primary role ever since. Early work (e.g., Neisser 1963; Schneider and Shiffrin 1977) laid the groundwork for two influential theories (see Palmer et al. 2000, for a review): the *feature integration theory* (Treisman and Gelade 1980) and the *guided search model* (Wolfe 1998).

While seeking to isolate the fundamental elements of vision, Treisman developed the feature-integration theory (FIT; Treisman and Gelade 1980), which served as a driving force of the surge of research in visual search that was soon to follow. The basic idea behind FIT is rooted in Neisser's (1967) original division of visual processing into two distinct stages, but Treisman expanded on the meaning of these stages dramatically. According to the theory, the basic features of items (color, shape, orientation, etc.) are first processed effortlessly and automatically in the early stages of vision, in separate, spatially organized maps. Next, directed attention is required during the "attention" stage in order to successfully bind the separate features into integrated object percepts (Treisman 1998). Finally, a subset of these items is selected for further processing.

FIT allows for the dissociation of two types of searches, often categorized as "parallel" and "serial." *Parallel* search occurs when all items in a search array are assessed simultaneously, as the target item is different from all distractor items on at least one dimension, so it simply "pops out" at the observer (e.g., a red vertical line amongst green horizontal lines). In parallel searches, increasing the set size (number of objects in the search display) has little effect on response time because individual processing of each item is not required. *Serial* search occurs when the individual items within an array need to be searched one-by-one (or small group by small group) because the target item does not immediately pop out at the observer. Serial searching is needed when the target is only separable from the distractors by a conjunction of multiple features, such that it shares some features with the distractors (e.g., a target red vertical line amongst red horizontal lines and green vertical lines). In these cases, response time increases as the number of items in an array increases because more items need to be searched successively.

This strictly dichotomous view of serial vs. parallel search is no longer considered an accurate characterization of search processes (e.g., Townsend 1990; Wolfe 1998), but it continues to offer a useful framework for understanding the variation in processing between simple and complex searches. Because parallel and serial searches are thought to rely on different cognitive processes, it is typically necessary to differentiate between the two when examining an effect, as many conditions may only modulate performance for one of these two types of search.

Feature-integration theory is useful in understanding a simple two-stage concept of the preattentive and focused stages of search, but preattentive processing is more complex than Treisman's original model captures (Wolfe 1998; Wolfe and Horowitz 2004). Wolfe's "Guided Search" theory (2007) has a similar, but less linear, model of the stages involved in search. In Guided Search, the basic features serve as guiding attributes to direct the deployment of attention. Both basic sensory processes and selective attention are used in tandem, as basic perception identifies relevant features and guides the observer's attention appropriately. The many versions of the Guided Search model (Wolfe et al. 1989; Wolfe 1994; Wolfe and Gancarz 1996; Wolfe 2007) offer a more comprehensive understanding of visual search. These theories of the basic mechanisms of search are important for understanding the underlying processes of visual cognition and allow for analysis of how more complex searches occur.

Bridging the Gap Between the Lab and the Field

Historical studies and key cognitive theories of search have built a solid framework for further exploration of exactly what guides visual search performance. A recent trend has been to build upon this framework to approximate critical differences between lab and field searches. In doing so, researchers have purposefully deviated from standard parameters employed in typical lab-based search tasks to introduce factors usually found in field-based searches. For instance, in a typical visual search task in the lab, only one target is possible on any given trial, and half of the trials have a target present. Such standards offer ideal experimental control and statistical power; however, they are not necessarily representative of the nature of field searches where more than one target may be present, and targets may be infrequent. Efforts to better approximate conditions in the field have begun including multiple targets and target categories and decreasing the relative frequency of the targets (e.g., in radiology, Berbaum et al. 1998; Samuel et al. 1995; Franken et al. 1994; in cognitive psychology, Menneer et al. 2007; Wolfe et al. 2005; Fleck and Mitroff 2007; Fleck et al. 2010).

In addition to modifying lab-based search tasks to more directly approximate field conditions, considerable effort has also been dedicated towards advancing technological aids for field searches. Consider, for example, the nature of airport baggage screening; searching X-rays in airports is particularly difficult because of the wide range of potential targets, variability of distractor items, clutter, and potential for purposefully hidden or obscured objects in the search array. When presented with this difficult, but critical, search scenario, it is important to pursue all available means by which to improve performance in both the technology and in the searchers themselves. Key insights have already been offered in terms of how technological advances may help or hurt the human operator by examining interactions between human factors and technology changes (e.g., Bolfing et al. 2008; Schwaninger 2006a,b,c; Schwaninger and Hofer 2004; Schwaninger and Wales 2009; von Bastian et al. 2008; Wiegmann et al. 2006).

Technological advances can improve field searches along several fronts, but search accuracy still relies on the performance of individual X-ray operators. As such, it is important to study the searchers themselves to find additional ways to increase accuracy. Several research projects have brought the lab and the field together to address this by assessing factors that may both positively and negatively affect search performance (e.g., McCarley and Steelman 2006; Mitroff and Hariri 2010; Neider et al. 2010; Schwaninger 2003a,b; Schwaninger et al. 2005). Contextual and situational factors potentially present during field searches, such as motivation and anxiety, can impact search processes and performance. Some recent work has examined the effects of motivated and anxious conditions on a variety of cognitive processes (e.g., declarative memory, Murty et al. 2011), but few studies have investigated the interplay of these factors with visual search specifically. Given the numerous cognitive mechanisms underlying successful search and the complexity of many searches in the field, career searchers may be significantly influenced by situational factors that may induce anxiety or increase motivation. Thus, research has begun to explore the impact of context on performance (e.g., Cain et al. 2011) and has determined that such factors can significantly influence search accuracy.

Finally, because of the differences in experience between undergraduates in the lab and career searchers, another technique that has been instrumental in bridging the lab and the field is to test search experts (e.g., radiologists and airport security officers) in a laboratory setting. By controlling for many of the other differences between lab and field, directly comparing the performance between inexperienced searchers and experts on the same task in the same environment allows for the assessment of the effects of expertise on search (e.g., Clark et al. 2011a; Mitroff et al. 2012). Preliminary results of such studies demonstrate an increase in accuracy with expertise and suggest that the experts employ different strategies.

Many open questions stand in the way of directly translating findings from the lab to the field, and vice versa, but there are good reasons to be hopeful. The current state of visual search research suggests that it is possible to successfully bring a result from one realm to the other, and the current goal is to make this process more and more robust. In the 'Target Prevalence' section, we discuss four hurdles that present potential problems and strategies for how to overcome them. Specifically, we discuss target prevalence, target number and target category, motivation and anxiety, and level of experience.

Target Prevalence

When Tyler, our hypothetical undergraduate, begrudgingly sat through a visual search experiment in a dark room, as described in the beginning of this chapter, a target was present on half of the trials he viewed. He did not find every target, but the frequency of targets kept him alert. He was not exerting a significant amount of effort, but the fact that he was able to find a target so frequently may have helped to keep him on task. If a much smaller percentage of the trials had contained targets, perhaps Tyler would have been more likely to miss those targets.

Olivia, our hypothetical X-ray operator, has rarely encountered actual harmful items in the bags she inspects. Threatening items are, in fact, so rare that the Transportation Security Administration (TSA) has devised a method of inserting images of harmful items onto the viewing screens, superimposed over real luggage. These images, called Threat Image Projections (TIP) (Schwaninger 2006a–c), are designed to appear as real, dangerous contraband. When the X-ray operator sees a threatening item, he or she pushes the appropriate button on the console. If the item was a TIP image, rather than a legitimately harmful item, the screener is provided with immediate feedback (to avoid detaining an innocent passenger). TIP images are used to keep X-ray operators like Olivia alert, to maintain an index of performance, and to counteract potentially negative effects of the actual low prevalence of targets.

The difference in target prevalence between Tyler's and Olivia's searches highlights a potentially major hurdle: lab-based searches typically have a target present on half the trials, and field-based searches rarely have a target present so often. Lab-based searches use 50% target prevalence levels to maximize statistical power. However, as Olivia experiences, many visual searches conducted in the field do not have this nice balance of 50% target-present and 50% target-absent displays. A termite inspector seldom finds pests on routine inspection; a lifeguard, thankfully, rarely deals with a drowning swimmer; border patrol agents do not routinely see individuals trying to illegally cross a border; and pilots almost never detect a physical defect on their routine pre-flight structural inspections.

The same case follows for searches in radiology and airport security screening; there is not an abnormality in half of the X-ray images viewed by a radiologist, and there is not a dangerous item in half of the bags viewed by a TSA X-ray operator. The numbers are difficult to calculate for airport security screening, but the rate of truly hazardous items is well below 1% (e.g., Rubenstein 2001). The prevalence is a bit easier to determine in radiological screening; it is estimated that only around 0.3% of routine mammograms contain an abnormality (Gür et al. 2004). These values deviate substantially from the typical 50% used in the lab, and a critical question is whether the factor of *target prevalence* actually has a functional role in visual search performance. While laboratory search performance is usually quite good, an estimated 30% of malignancies are missed in radiological exams (e.g., Berlin 1994; Kundel 1989; Renfrew et al. 1992). Might target prevalence factors account for some of this disturbingly high miss rate? This question is critically important, as failure to identify targets in rare-target search could be potentially disastrous.

Vigilance tasks bear a strong relationship to rare-target search, as they typically consist of a monitoring task in which events occur at rare and unknown intervals, in contrast to a visual search study in which each trial demands a separate response of absent or present. Early studies with vigilance tasks found that performance declines over time while performing a monotonous task (e.g., Mackworth 1950; Parasuraman and Davies 1976; Davies et al. 1983). Because rare-target visual searches resemble vigilance tasks in the monotonous response of "no target," it is a reasonable assumption that visual search accuracy could decline over the course of time; for example, as screeners repeatedly determine that X-rays do not contain tumors.

Radiological examinations of target prevalence effects have found conflicting results. One study varied target prevalence from 20 to 60% and found a much higher accuracy rate in higher prevalence conditions (Egglin and Feinstein 1996). However, another study found no difference in performance related to prevalence rates varying from 2 to 20% (Gur et al. 2003); this lower prevalence rate better maps onto the actual rates of screenings and routine examinations. An additional study by the same group (Gur et al. 2007) demonstrated an influential effect of prevalence expectations on confidence ratings following target identification, in which decreasing prevalence tended to increase confidence ratings; yet again, the data indicated no detrimental effect on accuracy.

Given the complexities of the radiological environment, it not easy to directly assess the role of prevalence with radiologists as the participants and radiographs as the search arrays (e.g., Gur et al. 2003). Likewise, it is not easy to test such questions with X-ray operators and luggage X-rays. Prevalence is, however, possible to address in the lab with inexperienced searchers. Using simplified displays and untrained participants, Wolfe et al. (2005) found a robust prevalence effect. Participants searched arrays of line-drawn objects and were to find "tools" amongst distractor shapes drawn from other categories. Each participant completed searches in which a target "tool" could appear on 50% of the trials (high prevalence), 10% of the trials (low prevalence) and 1% of the trials (very low prevalence). Visual search accuracy significantly declined as the target prevalence decreased, suggesting that target prevalence, *per se*, may have affected performance (Wolfe et al. 2005). In the 1% target prevalence condition (where in participants searched 2,000 individual trials with only 20 actually containing a tool), participants missed 30% of the targets. While it may just be coincidental, it is nonetheless striking that this number mirrors the probable miss rate from radiology (e.g., Gür et al. 2004). For additional discussion of this point, see Wolfe (2012).

In a typical vigilance task, participants slow down over the course of the experiment (e.g., Buck 1966). However, the Wolfe et al. (2005) participants were found to speed up over the course of the 2,000 rare target trials. It is proposed that as participants repeatedly and continuously correctly reject most target-absent trials, the time taken to reject decreases dramatically. In effect, participants may become so accustomed to saying that no target is present that they stop performing a sufficient search to actually find a target, thus causing a high miss rate on the few target-present trials.

The Wolfe et al. (2005) finding of a target-prevalence effect with simple displays and inexperienced searchers has the potential to be highly relevant to visual searches in the field. If this effect has been properly modeled in the lab, then manipulations can be tested that might improve accuracy (e.g., motivation; Navalakkam et al. 2009) and more precise methods can be used to better assess *why* misses occur (e.g., eye tracking; Rich et al. 2008). However, before this lab-to-field link can directly inform visual searches conducted in the field, it is critical to ensure that the link is valid. Does the underlying mechanism of the prevalence effect found in the lab match those of possible prevalence effects in the field?

Follow-up studies have raised concerns about whether the initial prevalence effect found in the lab sufficiently matches prevalence effects in the field (Fleck and Mitroff 2007; Li et al. 2011; Madden et al., in press). Fleck and Mitroff (2007) and Li et al. (2011) suggest that the prevalence effect found in the lab may be an error of response execution rather than a perceptual or identification error. When Fleck and Mitroff (2007) offered participants an option to "correct" their responses on a previous trial, this alone removed a previously found prevalence effect. At least in these studies, participants were able to correct such errors, indicating that they were not actually "missing" the targets perceptually; they were simply responding quickly out of habit. In effect, participants fell victim to a classic "oops" problem—they were quickly responding "no" trial after trial, until suddenly they hit the "no" key when, in fact, they had not intended to do so. Such a physical perseveration or inhibition problem is not likely to underlie a prevalence effect in the field. Fast-paced responding is not a common aspect of radiology or baggage screening, and such searches offer the option to correct mistakes.

The results of Fleck and Mitroff (2007) suggest that lab-based visual searches with simple stimuli may not be able to adequately translate to the complex searches conducted in the field, given that the purported mechanism (a response-based error) is not a part of most field searches. An additional study, however, offered a different conclusion. Support for a prevalence effect was found in a study that employed realistic X-ray luggage images (Van Wert et al. 2009), even when participants were offered the option to correct their responses. This suggests that a prevalence effect can be observed in the lab, with the option to correct, as long as the stimuli are sufficiently complex. Moreover, this suggests that prevalence effects are not driven entirely by response-based errors.

A potential problem remains, however, with extrapolating prevalence effects from the lab to the field. While Van Wert et al. (2009) clearly involved more complexity than Fleck and Mitroff (2007) and Wolfe et al. (2005), the locus of the complexity is not as clear. Fleck and Mitroff (2007) and Wolfe et al. (2005) used a set of six possible target "tools" and showed pictures of the targets to the participants before the experiment. Van Wert et al. (2009) used 100 images of knives and 100 images of guns and only showed a small subset to the participants prior to the experiment. While this added desired complexity to the stimuli, it also, unfortunately, added complexity to the participant's task. On 94% of the occasions when participants used the "correction" option in Fleck and Mitroff (2007) to report that they had pressed the wrong response key by accident, they changed misses (responding "no" when a target was present) into hits (responding "yes" when a target is present). In contrast, when the participants in Van Wert et al. (2009) used the correction option in the low prevalence condition, they primarily (81% of uses) changed correct rejections (responding "no" when no target was present) into false alarms (responding "yes" when no target was present). This suggests that the participants in Van Wert et al. (2009) did not understand what was and was not a target and did not have a sufficient grasp of their task. Ultimately, it is not clear what this means for relating prevalence effects from the lab to the field.

While the effects of a correction option on rare-target search performance remain debatable, additional studies have suggested alternative mechanistic accounts of the prevalence effect that suggest viable connections between lab findings and the field. Further work by Wolfe and Van Wert (2010) demonstrated that not only did searchers' decision criteria shift toward increasing misses at low prevalence, the reverse criterion shift also occurred with very high target prevalence leading to an increase in false alarms. They also found that target prevalence not only influences the criterion shift, but also the decision of when to stop searching in target-absent trials. Another nuanced study examined the prevalence effect in older adults, who typically exert greater top-down attentional control and more cautious approaches when completing search tasks (e.g., Madden 2007). The older adults not only exhibited less severe prevalence effects but also benefited even more greatly from the ability to correct responses (Madden et al., in press). The prevalence effect was also found to vary with the number of response alternatives, as the effect was eliminated in a four-alternative forced-choice task but remained intact in the standard two-alternative forced-choice task (Rich et al. 2008). Finally, Lau and Huang (2010) varied instructions given to participants regarding whether there were a high or low number of targets and found that this sort of instruction did not affect performance but that the prevalence effect was driven by the actual distribution of the targets encountered. Furthermore, participants showed the prevalence effect in conditions with a consistent prevalence level, but the miss rate did not increase when the prevalence level varied throughout a block of trials.

These conflicting results highlight the complexities of interpreting visual search performance data. Participants fall victim to the prevalence effect in some cases but not in others. The fact that the prevalence effect differs between younger and older adults (Madden et al., in press) demonstrates that inherent differences between participants affects search performance and provides a note of caution when trying to translate results from undergraduate searchers in the lab to older, professional searchers in the field. When these findings are taken together, it becomes apparent why interpreting visual search data is almost never straightforward. Both the characteristics of the experiment and the characteristics of the participants can dramatically alter performance results, such that isolating an effect to one specific cause is often impossible. Using a rare-target search task is a far more comparable means by which to consider search performance in the field, but it must be done in an informed way while taking all nuances discussed here into consideration.

Number of Targets and Number of Target Categories

Tyler, our hypothetical undergraduate, is aware there is never going to be more than one target-'T' shape within any display he views. His task is rather simple—he searches for a single target of a single category, and once he finds the target, his search is complete. Tyler does not need to concern himself with additional targets nor additional target types. He is not required to maintain two (or more) separate templates in memory while searching, and once he finds a target, he knows no further searching is required.

Olivia, our hypothetical X-ray operator, is tasked with searching for multiple kinds of items at all times. Not only does she need to identify guns, knives, and bombs, but she also needs to search for other items such as laptops, shoes, and liquids within each bag she examines. Furthermore, when she finds one target in a bag, she cannot terminate her search after the identification of this single target because there is no limit to the number of harmful items potentially present.

The possibility of multiple targets highlights another critical question: does it matter if someone is searching for more than one thing at the same time? The majority of lab-based visual search tasks present participants with well-defined stimuli and ask them to search arrays that contain either zero or one target. However, searches in the field can often contain more than one target type (e.g., either a tumor *or* a broken bone in a single radiograph) and/or more than one target (e.g., a tumor *and* a broken bone in a single radiograph). These types of searches, which we will refer to as *multiple-category* and *multiple-target* search, respectively, are rarely employed in the lab but are frequently present in the field. Is it possible to generalize from lab-based single-target research to multiple-category and multiple-target field-based searches? What is the cognitive cost of having to maintain in memory more than one target type? Is search performance worse if there may be multiple targets in the same display? We discuss these questions in this section and explore how they may present hurdles for translating research between the lab and the field.

Multiple-Category Visual Search

An X-ray operator is tasked with finding dangerous items and must simultaneously search for guns, knives, bombs, water bottles, and several other potentially dangerous items. Multiple-category visual search has a long history in cognitive psychology (e.g., Kaplan and Carvellas 1965; Krueger and Shapiro 1980; Menneer et al. 2004, 2007, 2008, 2009; Neisser et al. 1963; Vreven and Blough 1998), and several conclusions have emerged. It is clear that there is a negative impact of having to hold more than one potential target in memory (e.g., Gould and Carn 1973). When varying the number of possible target categories, larger numbers of target categories led to steeper search slopes (Kaplan and Cavellas 1965) and slower searches overall (Metlay et al. 1970).

Kyle Cave and his colleagues have convincingly shown that multiple-category search has a detrimental effect in terms of both visual search speed and accuracy (e.g., Menneer et al. 2007; Menneer et al. 2009; Godwin et al. 2010). In one study, different groups of participants searched for either one or two colors, one or two shapes, or one or two line orientations (Menneer et al. 2007). Search times were slower, and miss rates were drastically higher in the dual-category trials. In a study that was directly inspired by airport baggage screening, participants had to search X-ray images for either weapons *or* bombs in separate searches or weapons *and* bombs in the same search (Godwin et al. 2010). They found that there were dual-category search costs in both accuracy and response time and that low-prevalence targets were missed more often than high-prevalence targets, but these factors appear to be additive. This suggests that searchers in the field who are searching for rare targets in many categories may be subject to many sources of miss errors.

The aforementioned work highlights two important points for the current discussion. First, multiple-category search is considerably slower and more error-prone than single-target visual search. Second, lab-based research can inform—and be informed by—visual searches in the field. Experiments using both simplified stimuli (e.g., Menneer et al. 2007) and X-ray baggage images (e.g., Godwin et al. 2010), have revealed ways in which multiple-category search demands impact performance and have offered suggestions for how to improve real-world searches. For example, Menneer et al. (2007) suggests that because simultaneously searching for multiple types of targets (e.g., guns, knives, and bombs) produces costs for both search speed and accuracy, it may be more effective to have multiple, specialized searchers that are focused on a single target type (e.g., only guns or only bombs).

Multiple-Target Visual Search Findings from the Lab

Multiple-category visual search requires holding more than one item in memory but does not necessarily involve identifying more than one target within the same array. Once a target has been found in a single-target search, the search can immediately be terminated, but what happens when there are potentially more targets? Does the successful detection of one target make a searcher more likely to notice additional targets? Or, does it act as a distractor and impair further search? These are critically important questions, as many visual searches in the field—where misses can be disastrous—have an unconstrained number of targets.

The nature of multiple-target search has been directly examined in the lab in a few studies (e.g., Cain et al. 2011, in press; Chan and Courtney 1995; Fleck et al. 2010; Schneider and Shiffrin 1977; Wolfe et al. 2005). One series of studies examined multiple-target visual search accuracy using an array of measures: useful field of view (Chan and Courtney 1995), target discriminability (Chan et al. 2002), and the time course of search (Chan and Chan 2000). In the first of these experiments (Chan and Courtney 1995), participants were briefly (250 ms) shown a horizontal line of shapes and were to report whether any 'o' shapes were present. The majority of the shapes were 'x's, but there were a variable number of target 'o's. When there were two targets present, participants were more likely to report the target that was presented closer to center and less likely to report the target presented in the periphery than when targets in those same locations were presented as the only target on a trial (Chan and Courtney 1995). This result implies that multiple-target search effectively reduced searchers' useful field of view, compared to single-target searches. In a version of the task designed to look at the time course of multiple target search, participants again scanned an array of 'x's for '<' and '>' but with both targets present on all trials. The time taken to find a second target was much more variable than that needed for the first target (Chan and Chan 2000), suggesting that modeling a dual-target search as two, serial single-target searches would not properly reflect actual search behavior. In another accuracy-focused study, participants searched for either two hard-to-spot shapes ('<' and '>') or an easy-to-spot and a hard-to-spot shape ('o' and 'v') among 'x's. The presence of a hard-to-spot target impaired detection of an easy-to-spot target more so than the presence of an easy-to-spot target impaired detection of a hard-to-spot target (Chan et al. 2002), reinforcing similar findings from radiology (Berbaum et al. 2001).

The above studies suggest that several factors can impact multiple-target search accuracy, and an additional study has suggested that the top-down knowledge of a multiple-target search can affect search even before the first target is located. Körner and Gilchrist (2008) compared eye movements between a condition in which there were 0 or 1 targets present and a condition in which there were either 1 or 2 targets present, with participants informed of the conditions. Even before a target was found, participants made more distractor re-fixations in the 1 vs. 2 condition on the trials with just 1 target present than on physically identical 1-target stimuli in the 0 vs. 1 condition. This difference was argued to arise from participants "setting aside" memory for a possible second target before the search began, thus limiting the available memory for which locations have been searched (Körner and Gilchrist 2008). This finding suggests that, not only do physical aspects of the search array affect performance, but that searchers' expectations about the likely number of targets may also affect the efficiency of their search.

Multiple-Target Search in Radiology: "Satisfaction of Search"

The problems accompanying the presence of more than one target and different types of targets (e.g. pulmonary nodules and fractures) have been well documented in radiology (e.g., Berbaum et al. 1998; Samuel et al. 1995; Franken et al. 1994). A classic pitfall, known as "satisfaction of search" (SOS), occurs when the identification of a second target is less successful after the identification of a first target in the same display. SOS has been a topic of radiological research since the 1960s (Tuddenham 1962), but radiologists still fall victim to SOS. In fact, 28% of radiological misses have been attributed to SOS errors, which makes this a critically important problem to solve.

Radiography studies have delineated three possible types of errors contributing to SOS (Nodine and Kundel 1987): *scanning errors* (the search path never encounters the target area, Berbaum 1996, 2005; Samuel et al. 1995), *recognition errors* (scanning in the region of a possible target but failing to dwell on the correct area for further inspection, Berbaum 2000), and *decision-making errors* (fixating and dwelling on a possible target but ultimately failing to identify it as a target, Franken 1994). To date, evidence has suggested that all three likely contribute to SOS and the latter two explanations differ primarily in the amount of time spent analyzing a potential target. The time required to examine a target stems, in part, from the relatively low spatial frequency of radiographs and radiological targets (e.g., pulmonary nodules), which may require extra analysis to visually parse targets from background noise. By categorizing errors as scanning or decision-making/recognition errors, radiologists have attempted to understand whether SOS arises primarily as a function of a basic perceptual failure to properly scan an image or more of a cognitive failure in determining whether a particular item is indeed a target (Kundel et al. 1978).

Several suggestions have been proposed for the causes of multiple-target search errors, and some possibilities include a *truncated search* (finding one target leads to a non-exhaustive search) and a *perceptual set* (e.g., once a tumor is detected, the searcher engages a "tumor set" where additional tumors are likely to be spotted but other abnormalities, for example a fracture, are less likely to be spotted, Berbaum et al. 2000). However, no clear mechanism has been identified as responsible for the SOS effect. Eye-tracking data within radiological research has indicated that search is not actually terminated early (Samuel et al. 1995) and that participants continue to search after the successful identification of a first target; participants may even fixate on a second target but fail to identify it as such. This is confirmed and extended by a recent eye-tracking study done with undergraduate searchers and simplified search displays (Cain et al. 2012a). The eye-tracking data suggest that SOS errors are likely due to a combination of scanning, recognition, and decision-making errors.

The stimuli used in radiological search studies, however, are actual radiographs, which are highly complex and greatly variable. Furthermore, participants are radiologists who have extensive training and experience with the experimental tasks. In contrast, many lab-based visual search tasks do not involve a heavy decision-making component because the targets and distractors are easily distinguishable

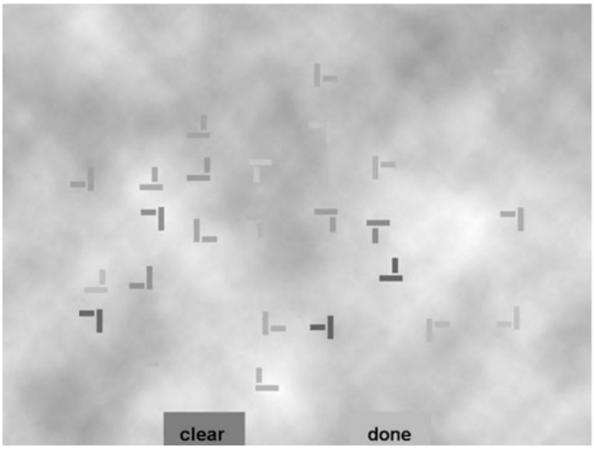


Fig. 1 Sample trial: find the "perfect" T shapes

from one another (in part so that untrained participants can perform reasonably). Recent work has begun to explore the SOS effect outside of radiology using simplified stimuli and undergraduate participants rather than radiographs and radiologists (Fleck et al. 2010) in order to generalize SOS to outside of the medical world.

In seeking to establish the scope of SOS errors in nonmedical searches, Fleck et al. (2010) aimed to understand the cognitive processes broadly involved in multiple-target search. In a series of experiments, basic 'T' and 'L' shapes were presented at varying degrees of visibility against a cloudy background. The cloudy background and the more–and less-salient targets and distractors served to approximate the noise typically present in radiographs. See Fig. 1. Each trial contained 0, 1, or 2 targets, and participants were to make a localization mouseclick on each target found and then clicked a 'DONE' button at the bottom of the screen to terminate their searches.

In previous studies, the SOS effect was typically observed when radiologists were less likely to identify a low-salience target when it was in the presence of a high-salience target than when the same low-salience target was the only target present in the array. Thus, the dual-target trials in this paradigm contained both a low–and high-salience target, and the SOS effect was calculated as the difference between a participant's accuracy in identifying low-salience targets in single-target trials and his or her accuracy in identifying low-salience targets in dual-target trials, provided the high-salience target had been successfully identified in the same trial. The SOS effect was found to be sensitive to both target prevalence and time pressure, as it was exacerbated when high-salience targets were three times as likely as low-salience targets and when participants had a time limit of 15 seconds per trial (Fleck et al. 2010).

These findings demonstrate generalized SOS errors in nonmedical searches; the inclusion of the possibility of multiple targets allows this paradigm to be applied to the field, as it is clear that search processes become infinitely more complex when multiple targets may be present. In order to gain a full understanding of search processes employed in the field, multiple-target searches should be used when exploring other issues related to lab-field differences.

Motivation and Anxiety

Tyler, our hypothetical undergraduate student, is simply not very concerned with his performance on his computer-based visual search experiment. He has little reason to care if he finds every target; regardless of how he performs, he will receive the participation credit for his Introductory Psychology class. There are also no consequences for poor accuracy. In addition to a lack of motivation, he likely has little to no anxiety about how well he does on this task since this really is just some "meaningless" experiment to him.

On the other hand, Olivia, our hypothetical X-ray operator, is strongly motivated to perform well since her job security is, at least partially, based on good performance and accurate searching. She takes great pride in her search abilities, as she has been with the TSA for quite some time and is a seasoned X-ray operator. More importantly, she is well aware of the consequences that might accompany her failure to identify harmful items in her search, and this keeps her motivated. Unfortunately, Olivia finds that she is often quite anxious while at work since passengers are always around her, and they are usually visibly (and often verbally) annoyed and hurried.

Difference in Context: Motivation

In the above scenarios, it is clear that Tyler and Olivia are faced with wildly different motivational contexts. Tyler's performance has no impact on his life, and there is little reason, beyond personal pride, to perform well. Olivia's performance can impact her livelihood (e.g., whether she has a job in the future) and others' lives (e.g., whether they are boarding a plane along with a bomb). Searches in the field are often linked to high-stakes outcomes; a radiologist or X-ray operator could save lives by identifying harmful targets in X-rays. Does performing a life-critical search cause individuals to be more motivated than when completing a lab-based task with no tangible consequences? The primary issue addressed in this section is whether differing levels of motivation affect visual search performance, and if so, how. If higher levels of motivation result in higher levels of performance, then how comparable are unmotivated, inexperienced searchers to highly motivated career searchers? This is a third fundamental hurdle for translating findings from the lab to the field.

Visual search tasks conducted in the lab often reveal a great deal of variability in performance in undergraduate participants. While some of this variability may be tied to differences in underlying search ability, some variability may also result from differences in motivation: Some participants may be intrinsically motivated to perform well regardless of a tangible outcome while others may not be motivated at all. For instance, more conscientious participants are likely to exert greater care and effort when performing the task, even though their levels of performance have no external consequences for them.

Unfortunately, it is not feasible to directly motivate laboratory participants in the same way career searchers are motivated. Participants cannot possibly believe that people's lives are in their hands nor that their careers depend on their performance in a computer-based experimental task. One reasonable approximation of motivation, however, is performance-based monetary reward. The prospect of receiving money for good performance provides an effective global incentive that, for most people, will increase their interest and effort (e.g., Camerer and Hogarth 1999).

In typical lab-based visual search experiments, monetary reward has been used to examine the impacts of motivation on attentional selection (e.g., Libera and Chelazzi 2006; Kiss et al. 2009), priming (e.g., Hickey and Theeuwes 2008; Kristjánsson et al. 2010) and attentional capture (e.g., Anderson et al. 2011). These studies have presented clear evidence that monetary rewards can improve performance; however, they have primarily focused on changes in the speed of attentional deployment. While this is a critical component of visual search performance, the majority of field-based searches place a larger emphasis on accuracy than on speed. Two recent experiments have employed monetary incentives with a focus on visual search accuracy: one with rare-target visual search (Navalakkam et al. 2009) and one with multiple-target visual search (Clark et al. 2011b).

Navalakkam et al. (2009) investigated whether the prevalence effect (that targets are missed more when they occur rarely than when they occur frequently) could be overcome when participants were sufficiently motivated. Participants searched for a target object in a cluttered scene, with the target prevalence (2, 10, and 50%) varied across blocks. A typical pattern emerged, with impairments in accuracy at low target-prevalence (e.g., Wolfe et al. 2005). However, when participants were motivated with a monetary incentive, the prevalence effect decreased significantly, restoring detection rates to near optimal levels. It was argued that fatigue, carelessness, and lack of vigilance were not responsible for the prevalence effect, but instead, the prevalence effect was caused by a shifted decision criterion, which could be modified through proper reward (Navalakkam et al. 2009).

In the 'Number of Targets and Number of Target Categories' section, we described the pitfalls of multiple-target visual search: searchers are less likely to find a target if they have already found another target in the same display (a phenomenon termed "satisfaction of search," SOS). This is a potentially dangerous problem that has been consistently observed in both lab-based and field-based visual searches (e.g., see Berbaum et al. 2010; Fleck et al. 2010). A recent series of experiments (Clark et al. 2011b) has explored whether SOS errors can be alleviated with the incentive of monetary reward. Can certain motivational frameworks lead to performance differences, and do such differences provide information on how to better structure work conditions for career searchers?

Clark et al. (2011b) employed a multiple-target search tasks that mirrored a paradigm that has previously found robust SOS errors (Fleck et al. 2010, Experiment 3, described in more detail in the 'Number of Targets and Number of Target Categories' section but manipulated the participants motivation by including a monetary incentive. Participants competed against nine other participants, and the "best" performer was awarded an additional $50. By simply adding this motivation of a performance-based reward, accuracy improved and the SOS effect was effectively eliminated.

These results raise a curious concern for the translatability of research considering that motivated, inexperienced participants show a decreased SOS effect, yet career searchers such as radiologists, presumably operating while motivated, still exhibit the SOS effect. Clearly, the incentive for the inexperienced participants differs greatly from the incentives for career searchers, but one could argue that a relatively small monetary incentive is qualitatively less than the incentive for career searchers—the chance at winning an extra $50 at some point in the next week or two is seemingly less motivating than keeping a job and preventing fatalities. How can this be reconciled? Perhaps despite the immense focus on accuracy for career searchers, the monotony of their daily routines interferes with their motivation. It is possible that the undergraduate searchers could actually be more motivated than the career searchers because they are completing a task for only an hour-long period, over which it is relatively easy to maintain a high level of motivation. Career searchers may not be not equally motivated at every hour throughout their workdays (or weeks, or months, etc.), and the SOS effect is observed may result from an inability to maintain consistently high levels of motivation.

Differences in Context: Anxiety

The dire consequences of missing a target in field searches could be potentially motivational but could also induce anxiety. Anxiety—the displeasurable psychological experience of worry or concern—is difficult to replicate in the lab, but it may be an element in many field searches. Beyond the general anxiety of knowing that missed targets could have life-threatening consequences, there is also more acute anxiety that can occur when searchers anticipate tangible stressors, such as a visit from a supervisor or a large workload. These states of heightened anxiety can be detrimental to accuracy, and anxiety has been linked to a decline in cognitive performance across species (e.g., in mice, Ohl et al. 2003; in humans, Eysenck et al. 2007).

As discussed earlier in this section, the motivation to earn rewards can significantly improve performance (e.g., Callan and Schweighofer 2008; Murayama and Kuhbandner 2011), but motivation to avoid punishments can increase anxiety and substantially diminish performance (Davis and Whalen 2001; Lang and Bradley 2009). Recent work has examined the differing effects of *approach* and *avoidance* motivation (earning rewards and avoiding punishments, respectively) on declarative memory (Murty et al. 2011). While approach motivation enhanced memory performance, avoidance motivation hindered performance, and this effect was especially amplified in participants who showed high levels of arousal. In the lab, the experience of *anticipatory anxiety* can be induced using a “threat of shock” paradigm, in which electrical shocks are administered at unpredictable intervals, unrelated to performance (e.g., Grillon et al. 2004; Rhudy and Meagher 2000). A recent study has found this type of anticipatory anxiety to be specifically detrimental to multiple-target visual search performance (Cain et al. 2011).

Cain et al. (2011) used a variant of a standard multiple-target search paradigm (Fleck et al. 2010, Experiment 5) in which an SOS effect was *not* expected. When participants were anticipating a neutral event (an innocuous tone), they did not show SOS (which replicates the previous instantiation of these particular experimental parameters). However, when those same participants were anticipating a negative event (an electrical shock) they produced SOS errors. Interestingly, the participants did not show a difference on single-target performance between the non-anxious and anxious blocks of trials—the SOS effect was due solely to poorer second-target identification (Cain et al. 2011). Moreover, this effect was modulated by the level of anxiety that participants were experiencing at the start of the experiment. Less anxious participants showed high levels of SOS when anticipating a shock but no SOS in the control condition, while more anxious participants showed mild SOS throughout the entire experiment, regardless of condition. These results suggest that both acute and generalized anxiety could negatively affect search performance in the field by inducing SOS errors. Thus, efforts should be made to shield professional searchers in the field from anticipatory anxiety in order to improve target identification in multiple-target displays. This sort of anxiety potentially poses an extra risk of misses in searchers with post-traumatic stress disorders or clinical anxiety disorders given that these individuals have been shown to be more likely to generalize specific causes of anxiety to the environment itself (e.g., Fanselow 1980; Grillon et al. 1998).

Just as the prevalence effect can be overcome with the proper motivation, multiple-target search appears to be influenced by contextual conditions. Performance on single-target searches in both motivated (Clark et al. in press) and anxious (Cain et al. 2011) conditions were unaffected by context, and influences were seen only on multiple-target conditions. The complex mechanisms responsible for the SOS effect may simply be more sensitive to contextual influences, and the motivation and anxiety inherent in career searches may work both for and against performance.

Though SOS can be eliminated in the laboratory via monetary incentive, SOS remains a problem in the field. This could be attributable to the monotony of the daily grind detracting from the value of motivation, but the anxiety associated with career searching could also contribute negatively. While motivation appears to positively affect performance, anxiety may serve as a hindrance, and it is important to take all of these factors into account when evaluating differences in searching between the lab and the field. Furthermore, enhancing motivation while decreasing anxious circumstances may be the best combination of contexts for optimal search performance.

Level of Experience

Tyler, our hypothetical undergraduate, searches for items in his everyday life—the books for his classes, the keys to his dorm room, and his cell phone. He rarely, if ever, dedicates any sort of mental effort toward improving the efficiency of these searches since they are mundane and generally completed successfully. Tyler has also never been trained on how to conduct visual searches to increase accuracy, nor does he regularly spend hours at a time conducting visual searches (e.g., he usually finds his keys within a minute or two at the most).

Olivia, our hypothetical X-ray operator, has undergone extensive training in order to improve her search skills. Additionally, she spends hours every workday actively conducting visual searches as she scans bag after bag. She has years of experience in search that have allowed her the opportunity to increase her ability to scan X-ray images for harmful items, allocate her visual attention more effectively, and utilize any superior strategies she may have developed.

The above vignettes about Tyler and Olivia highlight the last major hurdle we discuss in this chapter—experience. Many career searchers have years of training and experience on specific search tasks, and it is important to understand how this might influence their abilities. It is not clear exactly how career searchers' levels of experience may affect their performance, both on their typical job-related searches and on search tasks more generally. How might search expertise on the job translate to search performance on standard lab-based search tasks? Which conclusions drawn from inexperienced undergraduates, without extensive training, are applicable to the field?

Trained professionals are often better at visual searches related to their jobs than are novices (e.g., farmers improve their ability to sort chickens by sex with experience, Biederman and Shiffrar 1987; wine connoisseurs learn to discriminate between fine wines, Bende and Nordin 1997; bank tellers are better than the general public at detecting counterfeit currency, Klein et al. 2004; and chess players are better able to see patterns of moves on a chessboard, Chase and Simon 1973). This apparent benefit of experience leads to two key questions: What are the bases for these expertise differences? And how can the differences be accounted for when assessing the performance of inexperienced searchers in an attempt to translate from the lab to the field? Observing how expertise may alter both trained task performance specifically, and visual/cognitive abilities more generally, has the potential to inform questions about visual search as well as the general malleability of cognitive abilities.

Perceptual Training in the Lab

In most cognitive psychology studies, a participant (like our hypothetical undergraduate, Tyler) arrives in the lab, runs through a minute or so of practice, and then completes an hour-long study. They are then dismissed and may never think about the task again. The experimental results provide a useful assessment of performance but do not allow for an investigation of learning. One class of experiments, however, is focused primarily on learning effects. In *perceptual learning* experiments, a research participant may make several visits to the lab and undergo thousands of trials of the same specific task so that they ultimately receive extensive training.

Research in perceptual learning has shown that it is possible for very basic visual abilities to change with experience. If a participant is asked to make a difficult visual discrimination over many trials—often spread out over days—his or her threshold for discrimination will decrease dramatically (Westheimer and McKee 1978); these changes are often attributed to plasticity in primary visual cortex (V1). There has been evidence for a host of sensory and perceptual improvements in which basic feature discrimination improves with extensive practice of a task. Participants improve in discriminating the orientation of a line (Ramachandran and Braddick 1973; Fiorentini and Berardi 1981; Matthews and Welch 1997), identifying the direction of motion (Ball and Sekuler 1982, 1987), and show increased vernier acuity (Westheimer and McKee 1978; Saarinen and Levi 1995; Beard et al. 1995). In all of these cases and in others (e.g., Vogels and Orban 1985; Karni and Sagi 1991, 1993; Poggio et al. 1992; Fahle and Edelman 1993), learning is specific to the stimulus on which the participant was trained. In fact, a hallmark aspect of perceptual learning is that the training effects appear to be quite specific. Because the learning is believed to take place at such a basic perceptual level, improvement is only seen when examining performance on the exact trained stimulus. If, after training, participants showed an overall improvement in a task, beyond that of the trained stimulus (e.g., he/she was trained to identify rightward motion, but also improved in identification of leftward motion), more generalized training would be said to have occurred.

Generalized Training

Perceptual learning studies in the lab have produced highly specific training effects, but this does not directly inform generalized learning effects. Given the uncontrolled and variable nature of field-based searches, expertise gained through career searching likely produces more generalized benefits. In airport security screening, for example, X-ray operators never search two entirely identical suitcases, so they cannot rely on simple sensory-level template matching to successfully identify targets. The experience gained through their daily training relies on improvements that *can* transfer from bag to bag. Furthermore, visual search, even its most simplistic, laboratory form relies on the integration of both sensory perception and strategic attentional allocation.

One of the few perceptual learning studies to demonstrate generalized learning used a visual search task (Sireteanu and Rettenbach 1995). Training accumulated over the course of the experiment resulted in improved search efficiency, even on untrained stimulus sets. Perhaps because of the complex attentional processes required for effective visual search, the learning occurred in a less specialized manner; visual search involves cognitive processes more complex than basic sensory discrimination, so the improvement likely occurred at a level that can generalize beyond the perception of one specific stimulus. For example, search efficiency may improve via changes in strategies—participants may learn to better distribute their attention, disregard irrelevant cues, or react quickly to relevant ones.

Because search relies on strategies and attentional processes, it is impossible to isolate learning for visual search tasks to the sorts of low-level feature discrimination improvements seen in classic perceptual learning tasks. As such, "perceptual learning" has recently undergone a redefinition, which encompasses even strategic and attentional improvements. Using visual search paradigms to study trained improvement allows for a more comprehensive investigation of the many elements of attention that can be improved with training.

Generalized Learning via Action Video Game Playing

Basic sensory perceptual learning cannot allow for improvement in skills in the field, where search arrays consistently vary. Certain experiences, however, have been found to elicit improvement in a wide variety of skills and are far more generalized than basic perceptual learning processes. Extensive experience with specific activities can influence perceptual and attentional abilities that generalize beyond those activities, and a host of studies have shown that those who regularly play action video games (usually an average of 6 or more hours per week for at least 6 months) show improved performance on a variety of tasks. Specifically, when compared to those who did not regularly play action video games, avid action video game players respond more rapidly (Castel et al. 2005; Dye et al. 2009; Orosy-Filders and Allan 1989; Yuji 1996), have improved spatial abilities (Okagaki and Frensch 1994; Quaiser-Pohl et al. 2006; Terlecki and Newcombe 2005), have enhanced temporal abilities (Donohue et al. 2010; Green and Bavelier 2003, 2006b, 2007; West et al. 2008), can enumerate briefly displayed items more quickly (Green and Bavelier 2006b), can switch between tasks faster (Cain et al. 2012b; Karle et al. 2010), and have enhanced eye–hand coordination (Griffith et al. 1983).

Studies exploring the causal role of video game playing have trained non-gamers on action video games and shown improved performance (e.g., De Lisi and Cammarano 1996; De Lisi and Wolford 2002; Dorval and Pepin 1986; Green and Bavelier 2003, 2006a,b, 2007; however, see Boot Kramer et al. 2008 for lack of training effects; and Nelson and Strachan 2009 for more nuanced training effects). The issue of causality explores an important mechanistic explanation of gamers' benefits, but regardless of the causal nature of such benefits, differences between gamers and non-gamers have been reliably demonstrated.

However, there is a mechanistic question regarding these differences; two feasible accounts have both received support and are not mutually exclusive. The basic-sensory hypothesis suggests that action video game exposure trains better "vision" and "attention," honing basic abilities (e.g., Dye et al. 2009; Green and Bavelier 2006a, 2007; Li et al. 2009; West et al. 2008; Caplovitz and Kastner 2009). According to this hypothesis, gamers may have an increased capacity to process visual information compared to non-gamers. Alternatively, the improved-strategy hypothesis suggests that video game playing leads to the development of enhanced higher-level abilities such as attentional control (Cain et al. 2012; Chisholm et al.

2010; Hubert-Wallander et al. 2010b), shifts in attentional allocation, and improved strategy (Clark et al. 2011c) for generalized use across a variety of visually demanding tasks. In line with this account, gamers need not necessarily have an increased information-processing capacity but rather could be better able to use what resources they have to process perceptual information (e.g., Colzato et al. 2010).

Generalized Learning via Stroboscopic Training

In addition to video-game learning, stroboscopic training has been shown to improve visual cognition abilities (Appelbaum et al. 2011; Appelbaum et al., in press). Stroboscopic, or intermittent, vision is the process of presenting an individual with snapshots of the visual environment rather than a continuous visual experience. Training in such a visual environment can alter perceptual-motor abilities (e.g., Bennett et al. 2004; Mitroff et al. in press; Smith and Mitroff in press), and recent work suggests it can influence visual attention and memory as well. For example, in Appelbaum et al. (2011), participants trained on sports activities (e.g., playing catch) while either wearing transparent eyewear or stroboscopic eyewear that occluded vision at regular intervals. Before and after training, participants completed computer-based tasks without the eyewear. In one task, participants viewed patches of moving dots presented either centrally or peripherally and reported which of two sequentially-presented patched had coherent motion. Those participants who wore stroboscopic eyewear during training showed greater test-retest improvements on motion coherence sensitivity for centrally presented patches than participants who wore transparent eyewear, but no effects were seen for peripherally presented motion (Appelbaum et al. 2011). In another task, a useful field of view experiment, participants were briefly ($\sim$90 ms) shown a central letter and a dot in one of 24 peripheral locations. After a masked delay, they were asked to report the location of the dot and whether the central letter was upper or lower case. While the central task was primarily intended as a fixation control, the participants who trained with stroboscopic eyewear showed significant test-retest improvement at accurately reporting the case while the control group did not. No differences were found in peripheral performance for either group. Taken together, these results suggest that stroboscopic training may lead to generalized perceptual improvements, particularly in the center of the visual field (Appelbaum et al. 2011).

Career Training and Visual Abilities

Another form potentially generalized training occurs on the job; radiologists, for instance, spend years learning how to properly scan radiographs. Medical searches are among the most commonly studied visual searches in the field and have provided evidence for both specialized and generalized learning. Trained orthodontists are better able to detect subtle facial asymmetries than general dentists or lay people (Kokich et al. 1999), suggesting a specific enhancement in their trained skill set. Similarly, radiologists and cytologists are better able to detect abnormalities in medical images than inexperienced searchers, but they do not have enhanced memory for these abnormalities nor are they better at simple scene or object detection, suggesting specific training benefits. However, surgeons who regularly engage in video game playing were found to perform better at assessments of laparoscopic surgery (Rosser et al. 2007), suggesting a generalized benefit.

Radiologists typically have years of experience searching medical radiographs for abnormalities, but research in radiology shows they still fall victim to many of the same types of errors as inexperienced searchers. The studies of radiological visual search described in the ‘Number of Targets and Number of Target Categories’ section focused on the satisfaction-of-search (SOS) effect using radiologists as participants and real radiographs as test stimuli. Using actual radiologists and the stimuli they normally view to address research questions is entirely sensible, but it limits the ability to compare performance across different real-world expert populations. Showing an inexperienced searcher a radiograph or an X-ray of a bag may not be the best way to assess abilities since there would be a baseline difficulty that could mask their performance. One way to more directly compare abilities from the lab to the field is to use simplified displays for all participants, and a few recent studies have done so. One study used simple visual search arrays to study SOS in both undergraduate participants and training radiologists (Clark et al. 2011a), and another used simple visual search arrays to compare and contrast undergraduates and working airport baggage screeners (Mitroff et al. 2012).

To compare performance between radiologists and inexperienced searchers, Clark et al. (2011) administered a simplified multiple-target search task (e.g., Fleck et al. 2010) to both trained radiologists and to undergraduate students. A broad analysis of the data indicated that, perhaps surprisingly, radiologists and inexperienced searchers did not differ in overall search accuracy; the percentage of trials that they completed correctly (no misses, no false alarms) was not significantly different. However, the radiologists spent significantly longer per trial than did the undergraduates, and they frequently exceeded a trial time limit (15 seconds), while the undergraduates searchers hardly ever did so. The inexperienced searchers were actively deciding they had finished searching and electing to terminate their searches, while the expert searchers may have just run out of time while attempting to complete a more thorough search. When only assessing performance on trials in which participants indicated they had completed the search before the time limit, radiologists were more accurately able to detect the presence of a second target in a display, showing a reduced satisfaction of search effect relative to the inexperienced searchers.

Returning to the key hurdle discussed in the ‘Motivation and Anxiety’ section, different levels of motivation between searchers in the lab and in the field, it is worth considering whether the above differences between radiologists and inexperienced searchers might stem from radiologists simply caring more about their performance. Perhaps they took longer to respond and were more accurate because they were more motivated to perform well. This is a general concern for any such comparison, and one way to address this is to have the participants perform an additional, orthogonal task that does not tap into the specific skill in question. This was done in this study, with all participants also completing a control task, on which they made judgments about the temporal order of appearing squares. No differences between groups were found on this control task, which helps dampen the motivational concerns.

Experience appears to dramatically impact performance on cognitive tasks, but not in a simple, straightforward manner. Because of the perceptual variability from X-ray to X-ray, it is unlikely that any improvement would stem from enhanced basic sensory abilities. Instead, it appears that improvement in strategy or better attentional allocation may contribute to the improvement that comes with experience. Expert searchers are going about their searching in very different ways from inexperienced searchers and are likely more effective as a result. However, the mechanisms responsible for these differences remain largely unclear.

Discussion and Conclusions

The goal of this chapter was to explore ways in which visual search findings can translate between the lab and the field. On one side, a tremendous number of visual search experiments have been conducted in the lab, and the data have served as the basis for intricate and powerful theories of search. On the other side, career searchers conduct visual searches daily and are constantly looking for ways to improve performance. The critical question is whether each side can inform the other. Can cognitive theories and data be used in the field to guide and inform search practice? Likewise, can the nature of field-based searches be analyzed to further refine cognitive theories?

At first blush, it would be easy to say that searches from the lab and the field are not compatible given the vast differences between the manner in which search research is typically conducted in the lab and how search is performed in the field. However, such a conclusion would be both pragmatically unfortunate and empirically premature. Four significant hurdles were discussed in this chapter, and while each raises a critical concern when attempting to use lab-based findings to improve searches in the field, all show that with proper consideration, they can be overcome.

Target Prevalence The overwhelming majority of published cognitive psychology studies on visual search have employed paradigms in which targets appear on a substantial percentage of the trials. Yet, many field-based searches rarely have a target present (e.g., there is not a gun in the majority of baggage X-rays). While this difference initially presented itself as a critical hurdle for translating between the lab and the field, a number of studies have now explicitly focused on the impact of target prevalence (both in cognitive psychology searches and in radiological searches). The hope is that, with careful experimentation, the effects of target prevalence will be isolated so that researchers can explore this topic for its own sake, but can also explore field-related visual search questions without prevalence serving as a confound. Several recent studies have made significant advances along this front, and target prevalence may no longer serve as a critical hurdle.

Number of Targets and Target Categories As for target prevalence, most cognitive psychology studies have employed search arrays that only have one possible target at any given time. Likewise, most lab-based search experiments inform the searcher of their exact target prior to the start of the search. In contrast, most searches in the field can have an unconstrained number of targets and targets from multiple different categories. This difference is potentially devastating given that the possibility of multiple targets within a search array could have broad influences on search strategies and accuracy. However, recent efforts have explicitly married lab–and field-based searches (e.g., Fleck et al. 2010), finding numerous commonalities with multiple-target visual searches. This not only suggests that field-based visual search can learn from searches in the lab, but, that even more so, career searchers can partner with cognitive psychologists to take advantage of the benefits afforded by testing in the lab. It is not easy to experiment with working radiologists or X-ray operators, so anything that can be tested out in the lab, such that it will translate, can be profoundly helpful. With the knowledge that multiple-target search in the lab can translate to the field, researchers are now equipped to use these search paradigms to explore the additional differences between the lab and the field.

Motivation and Anxiety The contexts in which searches are performed in the lab and the field are extraordinarily different, given both added motivation and added anxiety when conducting life-critical searches. In examining the effects of contextual motivation and anxiety in the lab, it becomes especially apparent why the deviations from standard search paradigms must be employed in order to properly investigate the effects of these factors in the field. In a multiple-target search paradigm—in which some trials only have one target, but other trials have more than one target—there were no differences in performance on single-target trials in motivated vs. non-motivated conditions (Clark et al. 2011b) and in anxious vs. non-anxious conditions (Cain et al. 2011). In both cases, the differences were only apparent in dual-target trials. As is known from the investigation of multiple-target search in general, complex mechanisms may be interacting to cause performance differences that basic single-target searches are simply not sensitive enough to show. Searches in the field can contain more than one target and can be conducted in motivated and/or anxious contexts; by examining the effects of these factors in tandem, meaningful conclusions can be made about how these contexts may affect performance on field searches.

Levels of Experience Finally, experts have demonstrated vastly different performance on a variety of tasks when compared with inexperienced searchers. Even when the results may appear similar between the two groups, more sensitive analyses reveal that expert searchers could be approaching the tasks very differently. In these cases, one must be careful not to jump to conclusions about expert searchers from what is observed with undergraduate, inexperienced searchers. Other tasks, however, may demonstrate similar processes between the two groups. Early work used real radiographs and actual radiologists to explore performance in the lab, but since inexperienced searchers do not have the experience to identify targets in such stimuli, researchers must use simplified displays accessible to both groups. By using these tasks, researchers can test both inexperienced and expert searches, and if in some tasks, performance is similar between the groups, conduct follow-up experiments using inexperienced searchers and potentially draw conclusions about experts from these data.

Despite the tremendous differences between lab and field searches, these hurdles are not insurmountable. Experimental search research is invaluable to the applied world, but only with an acknowledgement of the differences and shortcomings. By modifying the parameters of search tasks in the lab to account for the differences in target distributions in the field, by adding contextual factors present in the field such as motivation and anxiety to tasks in the lab, and by exploring the differences in performance between inexperienced and expert searchers, researchers are able to appropriately examine visual search processes as they exist in the applied world. While these are lofty requirements, ideally all examined simultaneously, with careful experimentation, we can understand the contributions of the individual factors and how they may interact. By accounting for all of these differences, we have the ability to use the performance of Tyler, the undergraduate, to inform and improve work conditions for Olivia, the X-ray operator.

References

Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, *15*, 106–111.

Anderson, B. A., Laurent, P. A., & Yantis, S. (2011). Value-driven attentional capture. *Proceedings of the National Academy of Sciences*, *108*, 10367–10371.

Appelbaum, L. G., Schroeder, J. E., Cain, M. S., & Mitroff, S. R. (2011). Improved visual cognition through stroboscopic training. *Frontiers in Psychology, 2*, 276. doi: 10.3389/fpsyg.2011.00276.

Appelbaum, L. G., Cain, M. S., Schroeder, J. E., Darling, E. F., & Mitroff, S. R. (in press). Stroboscopic visual training improves information encoding in short-term memory.

Ball, K., & Sekuler, R. (1982). A Specific and enduring improvement in visual motion discrimination. *Science, 218,* 697–698.

Ball, K., & Sekuler, R. (1987). Direction-specific improvement in motion discrimination. *Vision Research, 27,* 953–967.

Beard, B. L., Levi, D. M., & Reich, L. N. (1995). Perceptual learning in parafoveal vision. *Vision Research, 35*(12), 1679–1690.

Bende, M., & Nordin, S. (1997). Perceptual learning in olfaction: Professional wine tasters versus controls. *Physiology & Behavior, 62,* 1065–1070.

Bennett, S., Ashford, D., Rioja, N., & Elliott, D. (2004). Intermittent vision and one-hand catching: The effect of general and specific task experience. *Journal of Motor Behavior, 36,* 442–449.

Berbaum, K. S., Franken, E. A. Jr., Dorfman, D. D., Miller, E. M., Krupinski, E. A., & Kreinbring, K. (1996). The cause of satisfaction of search effects in contrast studies of the abdomen. *Academic Radiology, 3*, 815–826.

Berbaum, K.S., Franken, E. A. Jr., Dorfman, D.D., Miller, E.M., Caldwell, R.T., Kuehn, D.M., & Berbaum, M.L. (1998). Role of faulty visual search in the satisfaction of search effect in chest radiography. *Academic Radiology, 5,* 9–19.

Berbaum, K. S., Franken, E. A. Jr., Dorfman, D. D., Caldwell, R. T., & Krupinski, E. A. (2000). Role of faulty decision making in the satisfaction of search effect in chest radiography. *Academic Radiology, 7,* 1098–1106.

Berbaum, K. S., Brandser, E. A., Franklin, E. A. J., Dorfman, D. D., Caldwell, R. T., & Krupinski, E. A. (2001). Gaze dwell times on acute trauma injuries missed because of satisfaction of search. *Academic Radiology, 8*(4), 304–314.

Berbaum, K. S., Franken, E. A. Jr., Dorfman, D. D., Caldwell, R. T., & Lu, C. H. (2005). Can order of report prevent satisfaction of search in abdominal contrast studies? *Academic Radiology, 12*, 74–84.

Berbaum, K.S., Franklin, E.A., Caldwell, R.T., & Schartz, K.M. (2010). Satisfaction of search in traditional radiographic imaging. *The handbook of medical image perception and techniques* (pp. 107–138). Cambridge: Cambridge University Press.

Berlin, L. (1994). Reporting the "missed" radiologic diagnosis: Medicolegal and ethical considerations. Radiology, *192,* 183–187.

Biederman, I., & Shiffrar, M. (1987). Sexing day-old chicks: A case study and expert systems analysis of a difficult perceptual learning task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13,* 640–645.

Bolfing, A., Halbherr, T., & Schwaninger, A. (2008). Using speed measures to predict performance in x-ray luggage screening tasks. Proceedings of the IEEE International Carnahan Conference on Security Technology, Zurich, October 5–8, 2009.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*(3), 387−398.

Buck, L (1966). Reaction time as a measure of perceptual vigilance. *Psychological Bulletin*, *65*(5), 291–304.

Bond, A. B. (1983). Visual search and selection of natural stimuli in the pigeon: The attention threshold hypothesis. *Journal of Experimental Psychology: Animal Behavior Processes, 9,* 292–306.

Cain, M. S., Vul, E., Clark, K., & Mitroff, S. R. (in press). A Bayesian optimal foraging model of human visual search. *Psychological Science.*

Cain, M. S., Dunsmoor, J. E., LaBar, K. S., & Mitroff, S. R. (2011). Anticipatory anxiety hinders detection of a second target in dual-target search. *Psychological Science, 22,* 866–871. doi: 10.1177/0956797611412393

Cain, M. S., Adamo, S. H., & Mitroff, S. R. (2012a). What eye-tracking can tell us about multiple-target visual search. Poster presented at the annual meeting of the Vision Sciences Society, Naples, FL.

Cain, M. S., Landau, A. N., & Shimamura, A. P. (2012b). Action video game experience reduces the cost of switching tasks. *Attention, Perception, & Psychophysics*. *74*(4), 641–647.

Callan, D. E., & Schweighofer, N. (2008). Positive and negative modulation of word learning by reward anticipation. *Human Brain Mapping, 29,* 237–249.

Camerer, C. F., & Hogarth, R. (1999). The effects of financial incentives in economics experiments: A review and capital-labor-production framework. *Journal of Risk and Uncertainty, 19*, 7–42.

Caplovitz, G. P., & Kastner, S. (2009). Carrot sticks or joysticks: Video games improve vision. *Nature Neuroscience, 12*(5), 527−528.

Castel, A. D., Pratt, J., & Drummond, E. (2005). The effects of action video game experience on the time course of inhibition of return and the efficiency of visual search. *Acta Psychologica, 119,* 217−230.

Chan, A. H. S., & Chan, C. Y. (2000). Validating the random search model for a double-target search task. *Theoretical Issues in Ergonomics Science*, *1*(2), 157–167. doi: 10.1080/14639220050171315.

Chan, A. H. S., Courtney, A. J., & Ma, C. W. (2002). Visual performance on detection tasks with double-targets of the same and different difficulty. *Ergonomics, 45*(13), 934–948. doi: 10.1080/00140130210166087.

Chan, H. S., & Courtney, A. J. (1995). Visual performance on detection tasks with two targets. *International Journal of Human Factors in Manufacturing*, *5*(4), 417–428.

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology 4,* 55–81.

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30,* 39–78.

Chisholm, J. D., Hickey, C., Theeuwes, J., & Kingston, A. (2010). Reduced attentional capture in action video game players. *Attention, Perception, & Psychophysics, 72*(3), 667−671.

Clark, K., Samei, E., Baker, J., & Mitroff, S. R. (2011a). Expertise in radiological screening and satisfaction of search. Poster presented at the annual Object Perception, Attention, and Memory meeting. Seattle, WA.

Clark, K., Cain, M. S. Adcock, R. A., & Mitroff, S. R. (2011b). Interactions between reward, feedback, and timing structures on dual-target search performance. Poster presented at the annual meeting of the Vision Sciences Society, Naples, FL.

Clark, K., Fleck, M. S., & Mitroff, S. R. (2011c). Enhanced change detection performance reveals improved strategy use in avid action video game players. *Acta Psychologica, 136,* 67–72.

Colzato, L. S., van Leeuwen, P. J. A., Van Den Wildenberg, W. P. M., & Hommel, B. (2010). DOOM'd to switch: Superior cognitive flexibility in players of first person shooter games. *Frontiers in Cognition, 1,* 8. doi: 10.3389/fpsyg.2010.00008.

Davis, M., & Whalen, P. J. (2001). The amygdala: Vigilance and emotion. *Molecular Psychiatry, 6,* 13–34.

Davies, D. R., Shackleton, V. J., & Parasuraman, R. (1983). Monotony and boredom. In G. R. J. Hockey (Ed.), *Stress and fatigue in human performance*. (pp. 1–32). New York: Wiley.

Donohue, S. E., Woldorff, M. G., & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics, 72,* 1120−1129.

Dorval, M., & Pepin, M. (1986). Effect of playing a video game on a measure of spatial visualization. *Perceptual and Motor Skills, 62,* 159−162.

Dye, M. W. G., Green, C. S., & Bavelier, D. (2009). Increasing speed of processing with action video games. *Current Directions in Psychological Science*, *18*(6), 321–326.

De Lisi, R., & Cammarano, D. M. (1996). Computer experience and gender differences in undergraduate mental rotation performance. *Computers in Human Behavior, 12*(3)*,* 351−361.

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Generic Psychology, 16*(3), 272−282.

Eckstein, M. (2011). Visual search: A retrospective. *Journal of Vision, 11*(5), 14, 1–36.

Egglin, T. K., & Feinstein, A. R. (1996). Context bias: A problem in diagnostic radiology. *The Journal of the American Medical Association, 276*, 1752–1755.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion, 7*(2), 336–353.

Fahle, M., & Edelman, S. (1993). Long term learning in vernier acuity: Effects of stimulus orientation, range and of feedback. *Vision Research, 33,* 397–412.

Fanselow, M. S. (1980). Conditional and unconditional components of post-shock freezing. *Pavlovian Journal Biological Science, 15,* 177.

Fiorentini, A., & Berardi, N. (1981). Learning in grating waveform discrimination: Specificity for orientation and spatial frequency. *Vision Research, 21*(7), 1149–1158.

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological Science, 18*, 943–947.

Fleck, M.S., Samei, E., & Mitroff, S.R. (2010). Generalized 'satisfaction of search': Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied, 16,* 60–71.

Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? *Psychological science, 16*(4), 275–281.

Franken, E.A. Jr., Berbaum, K.S., Lu, C.H., Kanam, Dorfman, D.D., Warnock, N.G., Simonson, T.M., & Pelsang, R.E. (1994). Satisfaction of search in the detection of plain-film abnormalities in abdominal contrast studies. *Investigative Radiology, 4,* 403–409.

Godwin, H. J., Menneer, T., Cave, K. R., & Donnelly, N. (2010). Dual-target search for high and low prevalence X-ray threat targets. *Visual Cognition, 18*(10), 1439–1463.

Gould, J. D., & Carn, R. (1973). Visual search, complex backgrounds, mental counters, and eye movements. *Perception and Psychophysics*, *14*, 125–132.

Gur, D., Rockette, H. E., Armfield, D. R., Blachar, A., Bogan, J. K., & Brancatelli, G. (2003). The prevalence effect in a laboratory environment. *Radiology, 228,* 10–14.

Gür, E., Turhan, P., Can, G., Akkus, S., Sever, L., Güzelöz, S., Çifçili, S., & Arvas, A. (2004). Enuresis: Prevalence, risk factors and urinary pathology among school children in Istanbul, Turkey. *Pediatrics International, 46,* 58–63.

Gur, D., Bandos, A. I., Fuhrman, C. R., Klym, A. H., King, J. L., & Rockette, H. E. (2007). The prevalence effect in a laboratory environment: Changing the confidence ratings. *Academic Radiology*, *14*(1), 49–53.

Green, C. S., & Bavelier, D. (2003). Action video game play modifies visual selective attention. *Nature, 423,* 534−537.

Green, C. S., & Bavelier, D. (2006a). Effects of action video games on the spatial distribution of visuospatial attention. *Journal of Experimental Psychology: Human Perception and Performance, 32,* 1465−1478.

Green, C. S., & Bavelier, D. (2006b). Enumeration versus multiple object tracking: The case of action video game players. *Cognition, 101*(1), 217−245.

Green, C. S., & Bavelier, D. (2007). Action video game experience alters the spatial resolution of vision. *Psychological Science, 18,* 88−94.

Griffith, J. L., Voloschin, P., & Gibb, G. D. (1983). Differences in eye-hand motor coordination of video-game users and non-users. *Perceptual and Motor Skills, 57*(1), 155−158.

Grillon, C., Morgan, C. A., Ill, Davis, M., & Southwick, S. M. (1998). Effects of experimental context and explicit threat cues on acoustic startle in Vietnam veterans with posttraumatic stress disorder. *Biological Psychiatry, 44*, 1027–1036.

Grillon, C., Baas, J. P., Lissek, S., Smith, K., & Milstein, J.(2004). Anxious responses to predictable and unpredictable aversive events. *Behavioral Neuroscience. 118*, 916–924.

Hickey, C., & Theeuwes, J. (2008). ERP correlates of inter-trial effects in visual search. *Journal of Vision, 8,* 6. (Abstract)

Hubert-Wallander, B., Green, C. S., & Bavelier, D. (2010). Stretching the limits of visual attention: The case of action video games. *WIREs Cognitive Science, Wiley, 1,* 1−9.

Kaplan, I. T., & Carvellas, T. (1965). Scanning for multiple targets, *Perceptual and Motor Skills, 21*, 239–243.

Karle, J. W., Watter, S., & Shedden, J. M. (2010). Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. *Acta Psychologica, 134*(1), 70−78.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Sciences USA, 88,* 4966–4970.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature, 365,* 250–252.

Kiss, M., Driver, J., & Eimer, M. (2009). Reward priority of visual target singletons modulates event-related potential signatures of attentional selection. *Psychological Science, 20,* 245–251.

Klein, R. M., Gadbois, S., & Christie, J. J. (2004). Perception and detection of counterfeit currency in Canada: Note quality, trainng and security features. *Proc. SPIE*, V. 5310. Rudolf L. van Renesse; (Ed.), *Optical Security and Counterfeit Deterrence Techniques* Vol. 1–12.

Kokich, V. O., Jr., Kiyak, H. A., & Shapiro, P. A. (1999). Comparing the perception of dentists and lay people to altered dental esthetics. *Journal of Esthetic Dentistry*, *11,* 311–324.

Koopman, B. O. (1956a). The theory of search: I. Kinematic bases. *Operations Research, 4*(3), 324–346.

Koopman, B. O. (1956b). The theory of search: II. Target detection. *Operations Research, 4*(5), 503–531.

Koopman, B. O. (1957). The theory of search: III. The optimum distribution of search effort. *Operations Research*, *5*(5), 613–626.

Körner, C., & Gilchrist, I. D. (2008). Memory processes in multiple-target visual search. *Psychological Research*, *72,* 99–105. doi: 10.1007/s00426-006-0075-1.

Kristjánsson, Á., Sigurjónsdóttir, Ó., & Driver, J. (2010). Fortune and reversals of fortune in visual search: Reward contingencies for pop-out targets affect search efficiency and target repetition effects. *Attention, Perception & Psychophysics, 72,* 1229−1236.

Krueger, L. E., & Shapiro, R. G. (1980). Repeating the target neither speeds nor slows its detection: Evidence for independent channels in letter processing. *Perception & Psychophysics, 28,* 68–76.

Kundel, H. L. (1989). Perception errors in chest radiology. *Semester Respiratory Therapy*, *10,* 203–210.

Kundel, H. L., Nodine, C. F., & Carmody, D. (1978). Visual scanning, pattern recognition, and decision-making in pulmonary nodule detection. *Investigative Radiology, 13,* 175–181.

Lang, P. J., Bradley, M. M. (2009). Emotion and the motivational brain. *Biological Psychology, 84*(3), 437–450.

Lau, J. S., Huang, L. (2010). The prevalence effect is determined by past experience, not future prospects. *Vision Research, 50*(15), 1469–1474.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience, 12,* 549−551.

Li, F-H., Cao, B-H., Xiao, F., & Li, H. (2011). The role of inhibitory control in the process of rare target detection. *Acta Psychologica, 43*(5), 509–518.

Libera, C. D., & Chelazzi, L. (2006). Visual selective attention and the effects of monetary rewards. *Psychological Science, 17*(3), 222–227.

Madden, D. J., Mitroff, S. R., Shepler, A. M., Fleck, M. S., Costello, M. C., & Voss, A. (in press). Rare target search: Diffusion model analysis and effects of adult age.

Madden, D. J. (2007). Aging and Visual Attention. *Current Directions in Psychological Science. 16*(2)*,* 70–74.

Mackworth, N. H. (1950). Researches on the measurement of human performance. *Medical Research Council Special Report*, London, Series 268.

Matthews, N., & Welch, L. (1997). Velocity-dependent improvements in single-dot direction discrimination. *Perception & Psychophysics, 59*(1), 60–72.

McCarley, J. S., & Steelman, K. S. (2006). Elements of human performance in baggage x-ray screening. Proceedings of the 4th International Aviation Security Technology Symposium, Washington, DC.

Metlay, W., Sokoloff, M., & Kaplan, I. T. (1970). Visual search for multiple targets. *Journal of Experimental Psychology, 85*, 148–150.

Mitroff, S. R., & Hariri, A. (2010). Identifying predictive markers of field performance: The potential role of individual differences in threat sensitivity. *Institute for Homeland Security Solutions Research Brief.* https://www.ihssnc.org.

Mitroff, S. R., Friesen, P., Bennett, D., Yoo, H., & Reichow, A. (in press). Enhancing ice hockey skills through stroboscopic visual training.

Mitroff, S. R., Biggs, A. T., Cain, M. S., Darling, E. F., Clark, K., Adamo, S. H., & Dowd, E. W. (2012). Visual search at the airport: Testing TSA officers. Poster presented at the annual meeting of the Vision Sciences Society, Naples, FL.

Menneer, T., Barrett, D.J.K., Phillips, L., Donnelly, N., & Cave, K.R. (2004). Search efficiency for multiple targets. *Cognitive Technology, 9,* 22–25.

Menneer, T., Barrett, D. J. K., Phillips, L., Donnelly, N., & Cave, K. R. (2007). Costs in searching for two targets: Dividing search across target types could improve airport security screening. *Applied Cognitive Psychology, 21*(7), 915–932.

Menneer, T., Stroud, M.J., Cave, K.R., Donnelly, N., & Rayner, K. (2008). Eye movements in search for multiple targets. In K. Rayner, D. Shen, X. Bai, & G. Yan (Eds.), *Cognitive and cultural influences on eye movements*. Tianjin: Tianjin People's Press

Menneer, T., Cave, K. R., & Donnelly, N. (2009). The cost of search for multiple targets: Effects of practice and target similarity. *Journal of Experimental Psychology: Applied, 15*(2), 125–139.

Navalakkam, V., Koch, C., & Perona, P. (2009). Homo Economicus in Visual Search. *Journal of Vision*, *9*(1), 31, 1–16.

Murayama, K., & Kuhbandner, C. (2011). Money enhances memory consolidation—but only for boring material. *Cognition, 119,* 120–124.

Murty, V. P., LaBar, K. S., Hamilton, D. A., Adcock, R. A. (2011). Is all motivation good for learning: Dissociable influences of approach and avoidance motivation in declarative memory. *Learning and Memory*, *18,* 712–717.

Nakayama, K., & Martini, P. (2010). Situating visual search. *Vision Research.* doi: 10.1016/j.visres.2010.09.003.

Neider, M. B., Boot, W. R., & Kramer, A. F. (2010). Visual search for real world targets under conditions of high target-background similarity: Exploring training and transfer of training in older adults. *Acta Psychologica, 134,* 29–39.

Neisser, U. (1963). Decision time without reaction time: Experiments in visual scanning. *American Journal of Psychology, 76,* 376–385.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Neisser, U., Novick, R., & Lazar, R. (1963). Searching for ten targets simultaneously. *Perceptual and Motor Skills, 17,* 955–961.

Nelson, R., & Strachan, I. (2009). Action and puzzle video games prime different speed/accuracy tradeoffs. *Perception, 38*(11), 1678−1687.

Nodine, C. F., & Kundel, H. L. (1987). The cognitive side of visual search in radiology. In J.K. O'Regan, & A. Levy-Schoen (Eds.), *Eye movements: From physiology to cognition* (pp. 573–582). North Holland: Elsevier Science Publishers B. V.

Ohl, R., Roedel, A., Binder, E., & Holsboer, F. (2003). Impact of high and low anxiety on cognitive performance in a modified hole board test in C57BL/6 & DBA/2 mice. *European Journal Neuroscience, 17*(1), 128–136.

Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology, 15*(1), 33−58.

Orosy-Filders, C., & Allan, R. W. (1989). Psychology of computer use: XII. Video-game play: Human reaction-time to visual-stimuli. *Perceptual and Motor Skills, 69*(1), 243−247.

Palmer, J., Verghese, P., & Pavel, M. (2000). The psychophysics of visual search. *Vision Research, 40*, 1227–1268.

Parasuraman, R., & Davies, D. R. (1976). Decision-theory analysis of response latencies in vigilance. *Journal of Experimental Psychology: Human Perception and Performance, 2*(4), 578–590.

Poggio, T., Fahle, M., & Edelman, S. (1992). Fast perceptual learning in visual hyperacuity. *Science, 256*(5059), 1018–1021.

Poulton, E. (1890). *The colours of animals: Their meaning of object recognition and use especially considered in the case of insects* (pp. 360). London: Kegan Paul.

Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences, 44*(3), 609−619.

Ramachandran, V. S., & Braddick, O. (1973). Orientation-specific learning in stereopsis. *Perception, 2*(3), 371–376.

Renfrew, D.L., Franken, E.A. Jr., Berbaum, K.S., Weigelt, F.H., & Abu-Yousef, M.M. (1992). Error in radiology: Classification and lessons in 182 cases presented at a problem case conference. *Radiology, 183*, 145–150.

Rhudy, J. L., & Meagher, M. W. (2000). Fear and anxiety: Divergent effects on human pain thresholds. *Pain, 84*(1), 65–75.

Rich, A. N., Kunar, M. A., Van Wert, M. J., Hidalgo-Sotelo, B., Horowitz, T. S., & Wolfe, J. M. (2008). Why do we miss rare targets? Exploring the boundaries of the low prevalence effect. *Journal of Vision, 8,* 1–17.

Rosser, J. C., Lynch, P. J. Haskamp, L., Gentile, D. A., & Yalif, A. (2007). The impact of video games in surgical training. *Archives of Surgery, 142,* 181–186.

Rubenstein, J. (2001). (Ed.) *Test and evaluation plan: X-ray image screener selection test.* Washington, DC: Office of Aviation Research.

Saarinen, J., & Levi, D. M. (1995). Perceptual learning in vernier acuity: What is learned? *Vision Research, 35*(4), 519–527.

Samuel, S., Kundel, H.L., Nodine, C.F., & Toto, L.C. (1995). Mechanisms of satisfaction of search: Eye position recordings in the reading of chest radiographs. *Radiology, 194*, 895–902.

Schneider, W., & Shriffin, R.M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention, II: Perceptual learning, automatic attending, and a general theory. *Psychological Review, 84*, 1–66.

Schwaninger, A. (2003a). Reliable measurements of threat detection. *Airport, 1,* 22–23.

Schwaninger, A. (2003b). Evaluation and selection of airport security screeners. *Airport, 2,* 14–15.

Schwaninger, A. (2006a). Threat image projection: Enhancing performance? *Aviation Security International, December 2006,* 36–41.

Schwaninger, A. (2006b). Liquid identification: Reacting to the terror threat. Analysing, controlling and adapting to meet new threats. *Airport, 5*, 30–31.

Schwaninger, A. (2006c). Airport security human factors: From the weakest to the strongest link in airport security screening. *Proceedings of the 4th International Aviation Security Technology Symposium, Washington, D.C., USA, November 27–December 1, 2006,* 265–270.

Schwaninger, A., & Hofer, F. (2004). Evaluation of CBT for increasing threat detection performance in X-ray screening. In: K. Morgan, & M. J. Spector, *The Internet Society 2004, Advances in Learning, Commerce and Security* (pp. 147–156). Wessex: WIT Press.

Schwaninger, A., & Wales, A.W.J. (2009). One year later: How screener performance improves in x-ray luggage search with computer-based training. *Proceedings of the Ergonomics Society Annual Conference 2009,* 381–389.

Schwaninger, A., Hardmeier, D., & Hofer, F. (2005). Aviation security screeners visual abilities & visual knowledge measurement. *IEEE Aerospace and Electronic Systems, 20*(6), 29–35.

Sireteanu, R., & Rettenbach, R. (1995). Perceptual learning in visual searches generalizes over tasks, locations, and eyes. *Vision Research, 40,* 2902–2949.

Smith, T. Q., & Mitroff, S. R. (in press). Stroboscopic training enhances anticipatory timing.

Terlecki, M. S., & Newcombe, N. S. (2005). How important is the digital divide? The relation of computer and videogame usage to gender differences in mental rotation ability. *Sex Roles, 53,* 433–441.

Tinbergen, N. (1960). The natural control of insects in pine woods: Vol. I. Factors influencing the intensity of predation by songbirds. *Archives Neelandaises de Zoologie*, *13,* 265–343.

Townsend, J. T. (1990). Serial and parallel processing: Sometimes they look like Tweedledum and Tweedledee but that can (and should) be distinguished. *Psychological Science, 1,* 46–54.

Treisman, A., & Gelade, G. (1980). A feature integration theory of attention. *Cognitive Psychology, 12*, 97–136.

Treisman, A. (1998). Feature binding, attention and object perception. Philosophical Transactions of the Royal Society, Series B, 353, 1295–1306.

Tuddenham, W. J. (1962). Visual search, image organization, and reader error in roentgen diagnosis: Studies of the psycho-physiology of roentgen image perception. *Radiology, 78,* 694–704.

Van Wert, M. J, Horowitz, T. S., Wolfe, J. M. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention Perception & Psychophysics, 71*(3), 541–553.

Vogels, R., & Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgments. *Vision Research, 25*(11), 1679–1687.

von Bastian, C. C., Schwaninger, A., & Michel, S. (2008). Do multiview X-ray systems improve X-ray image interpretation in airport security screening? *Zeitschrift für Arbeitswissenschaft*, *3,* 166–17.

Vreven, D., & Blough, P. M. (1998). Searching for one or many targets: Effects of extended experience on the runs advantage. *Journal of Experimental Psychology: Animal Behavior Processes*, *24,* 98–105.

West, G. L., Stevens, S. S., Pun, C., & Pratt, J. (2008). Visuospatial experience modulates attentional capture: Evidence from action video game players. *Journal of Vision, 8,* 1–9.

Westheimer, G., McKee, S. P. (1978). Stereoscopic acuity for moving retinal images. *Journal of Optical Society of America, 68*, 450–455.

Wiegmann, D., McCarley, J. S., Kramer, A. F., & Wickens, C. D. (2006). Age and automation interact to influence performance of a simulated luggage screening task. *Aviation, Space, and Environmental Medicine, 77,* 825–831.

Wolfe, J. M. (1994). Guided search 2.0: A revised model of visual search. *Psychonomic Bulletin & Review, 1*(2), 202–238.

Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention.* (pp. 13–73). East Sussex, UK: Psychology.

Wolfe, J. M. (2007). Guided search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated models of cognitive systems.* (pp. 99–119). New York: Oxford.

Wolfe, J. M., & Gancarz, G. (1996). *Guided search 3.0: Basic and clinical applications of vision science* (pp. 189–192), Dordrecht: Kluwer Academic.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15*(3), 419–433.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5,* 1–7.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature, 435,* 439–440.

Wolfe, J. M., & Van Wert, M. (2010). Varying target prevalence reveals two dissociable decision criteria in visual search. *Current Biology*, *20,* 121–124.

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1505–1513.

Yuji, H. (1996). Computer games and information-processing skills. *Perceptual and Motor Skills, 83*, 643–647.

When do I Quit? The Search Termination Problem in Visual Search

Jeremy M. Wolfe

Abstract In visual search tasks, observers look for targets in displays or scenes containing distracting, non-target items. Most of the research on this topic has concerned the finding of those targets. Search termination is a less thoroughly studied topic. When is it time to abandon the current search? The answer is fairly straight forward when the one and only target has been found (There are my keys.). The problem is more vexed if nothing has been found (When is it time to stop looking for a weapon at the airport checkpoint?) or when the number of targets is unknown (Have we found all the tumors?). This chapter reviews the development of ideas about quitting time in visual search and offers an outline of our current theory.

Keywords Visual attention · Visual search · Target prevalence · Radiology · Airport security · Guided search · Search termination · Absent trials

Visual searches, great and small, are a continuous part of our lives. As this is being written, I have just searched for Gate 22B at the Denver Airport. I then proceeded to search for an electrical outlet, my power cord, the correct port on the laptop, the link to the internet, and so on. These searches are drawn from the subset of total searches for which I have introspective awareness and some memory. We engage in search because there is too much visual information to fully process. Even if the sign for Gate B22 is in my visual field, I still need to use attentional mechanisms to select that object from the welter of other stimuli on Concourse B because attention is required to read that sign (Rayner 1983). Without worrying, for the present, about who this "I" is that is using attention, it makes some sense to imagine that I was asking my search engine to conduct these specific searches. Even if I am not engaged in what seems like deliberate search, covert attention is selecting one object after another, or maybe a few objects at a time, much as the eyes are fixating on one thing after another. The deployments of attention may be based on the bottom-up, stimulus driven salience of the stimulus (Einhauser et al. 2008; Foulsham and Underwood 2008; Koch and Ullman 1985; Masciocchi et al. 2009) (Is that a bottom-up, attention-grabbing bird flying around in Concourse B? Yes, it is! What is it doing in here?). Alternatively,

---

J. M. Wolfe (✉)
Harvard Medical School and Brigham & Women's Hospital,
Boston, MA 02115, USA
e-mail: wolfe@search.bwh.harvard.edu

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8_8,

attention might be guided by top-down task demands (Theeuwes 2010), even if those top-down demands do not usually seem to rise to conscious awareness. Consider the searches that could be involved in avoiding obstacles as you navigate down the concourse (Hamid et al. 2010; Jovancevic-Misic and Hayhoe 2009). The obstacles to be avoided might not be the most salient items but you manage to direct attention to them without introspective awareness of that search.

A vast set of research topics are present in this evocation of a trip down the airport concourse. Do we attend to objects or locations (Goldsmith 1998; Logan 1996; Roelfsema et al. 1998; Yeari and Goldsmith 2010)? What are the features that contribute to bottom-up salience (Wolfe and Horowitz 2004)? Do those features really "capture" attention (Theeuwes 1995) (Bacon and Egeth 1994)? Do new objects capture attention? (Yantis and Jonides 1996) (Franconeri et al. 2005)? How is top-down control of selection organized (Wolfe et al. 2004) (Hamker 2006) (Theeuwes 2010)? How do scene semantics guide the deployment of attention (Henderson and Ferreira 2004) (Torralba et al. 2006) (Vo and Henderson 2009)? How is this implemented in the brain (Reynolds and Chelazzi 2004) (Buschman and Miller 2009)? We could continue to list topics (for a daunting catalog from a computational viewpoint, see Tsotsos 2011), as each of these topics has generated a substantial research literature. In this chapter, however, we will focus on a different aspect of search that gets somewhat less attention. What happens when the search is unsuccessful? When is it time to abandon a search without having found a target? A moment's introspection reveals that, like successful searches, these abandoned searches occur all the time. Is there anyone I know in this airport waiting area? I can search for some period of time but, at some point, I need to give up and move on to the next task. How is that accomplished? If you find the target, there is an obvious signal that you are done. What is the signal that allows you to quit if no target is found? This problem of search termination is central to a variety of socially critical search tasks. Indeed, the airport is home to one of the signature examples of the search termination problem. Passing through security, your carry-on luggage is x-rayed and examined in a visual search for 'threats' like guns, bombs, and knives. Fortunately, most bags do not contain threats, meaning that, most of the time, the screener's task is to decide when it is time to abandon the search without finding a target. The stimulus is complex and could be examined for a long time, but, in that case, the line at the checkpoint would become unacceptably long. Of course, quitting too soon raises the possibility of missing a real threat, an error with far more consequence than a longer security line.

Similar search termination issues are raised in medical radiology. If you are screening mammograms for breast cancer, you do not want to miss any cancers but, at the same time, you need to get through all of the cases. When is it time to move to the next case? The radiology situation has some interesting characteristics that differ from the checkpoint search (beyond the obvious differences of stimulus materials). At the checkpoint, once a single gun is found, the search is done. In radiology, it is often important to find all of the signs of cancer (or whatever the radiologist may be looking for). Thus, in the medical setting, even if a 'target' has been found, there is still a search termination question. How sure are you that you have found everything that needs to be found in this image? The probability of missing a target is higher

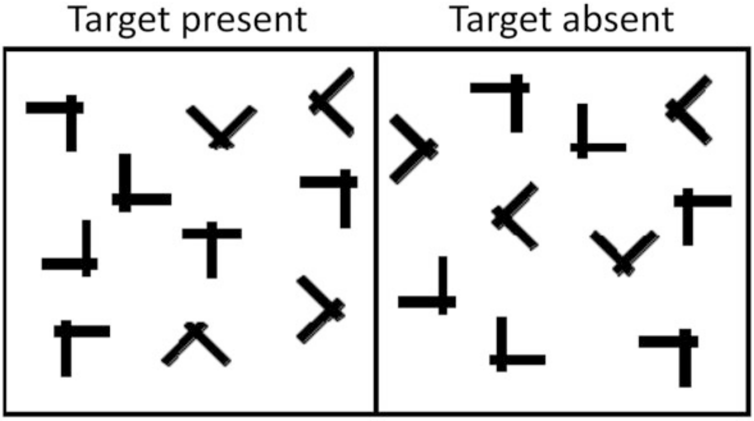


Fig. 1 A classic visual search for a T among Ls

if another target has been found; a problem known as “satisfaction of search” (Berbaum et al. 2000; Berbaum et al. 1990; Fleck et al. 2010; Nodine et al. 1992).

In this chapter, we will focus on the fundamental mechanisms of search termination with allusions to these more applied topics but without a full treatment of them. We will trace the development of ideas about search termination from early ideas about serial exhaustive search to a more plausible account and offer some pointers toward possible future progress.

How Shall We Model the Target Absent Trials?

Model 1: Serial Self-Terminating Search

Consider a basic search task, as shown in Fig. 1.

Here you are looking for a target “T” among distractor, “L”s. In an actual experiment, we would arrange for the items to be large enough so that acuity would not constrain performance. Typically, we would vary set size—the number of items in the display-and we would measure reaction time (RT) and accuracy. In cases where the display is visible until the observer responds, it is the RT data that are of most interest. An experiment of this sort would very typically produce data that look something like those shown in Fig. 2. The measure of greatest interest is the slope of the RT × set size function. In a task like this, slopes are typically in the range of 20–40 ms/item for target present trials and something more than twice that for target absent trials.

This pattern of results suggested a serial self-terminating search to Anne Treisman (Treisman and Gelade 1980) following similar ideas in memory research (Sternberg 1966). The idea, as illustrated in Fig. 3 was simple and reasonable. Items would be

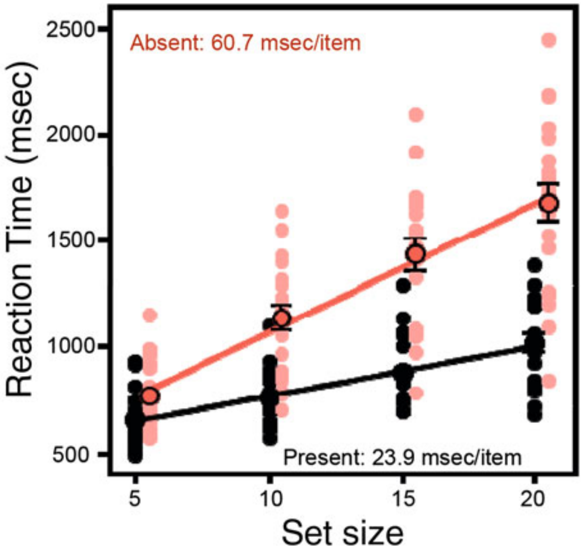


Fig. 2 Data from 20 observers performing a search for a T among Ls. *Dark black spots* show average correct target present RTs for each observer. *Light red spots* show correct target absent averages. *Larger symbols* are group averages. Error *bars* are ±1 SEM. *Lines* are best fit regression lines through the average points

selected, at random, one after another until the target was found or until all items were rejected. If there were N items, the target would be found, on average, after (N + 1)/2 selections from the display. The display could be rejected after all N items were examined. The result should be a slope ratio of close to 2:1. Treisman's data were consistent with this 2:1 prediction.

There are problems with Model 1. With more extensive data sets, it turns out that the search ratio in search tasks of this sort is typically significantly greater than 2:1 (Wolfe 1998). For the data shown in Fig. 2, for example, the hypothesis that Absent slope = 2 × (Present Slope) can be rejected (paired-*t* test, 2 × Present-Absent; $t(19) = 2.5$, $p = 0.023$). For these data, the average slope ratio is 2.5:1, very similar to what was found in Wolfe (1998). Note also that the variability of the absent trials is much higher than that of the present trials. This is also true for the RTs of individual observers contrary to what might be expected from a simple serial self-terminating model. After all, on absent trials, search always ends after all N items have been rejected while, on present trials, search could end after the first deployment of attention or the last or after any number of deployments between 1 and N.

The critical problem with a simple serial, self-terminating account is found in a classic experiment of Egeth et al. (1984). They didn't use Ts among Ls but using those stimuli as an example, imagine that half the elements were red and you were told that the target was black. You would not spend time examining red items and you would not need to search those red items in order to declare that the target was absent on blank trials. In an experiment of this sort, search slopes on target present trials would be reduced by about half of what is shown in Fig. 2. The absent slopes would be similarly reduced, suggesting that observers searched through only half the items. These and related results require a modification of the serial self-terminating model.

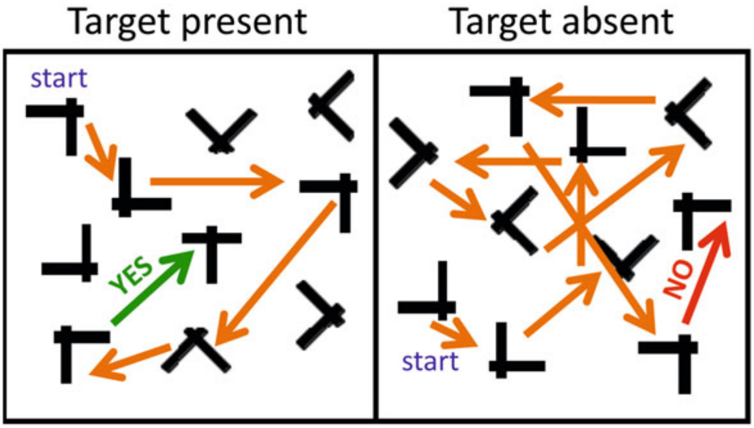


Fig. 3 Cartooned deployments of attention in a serial, self-terminating search

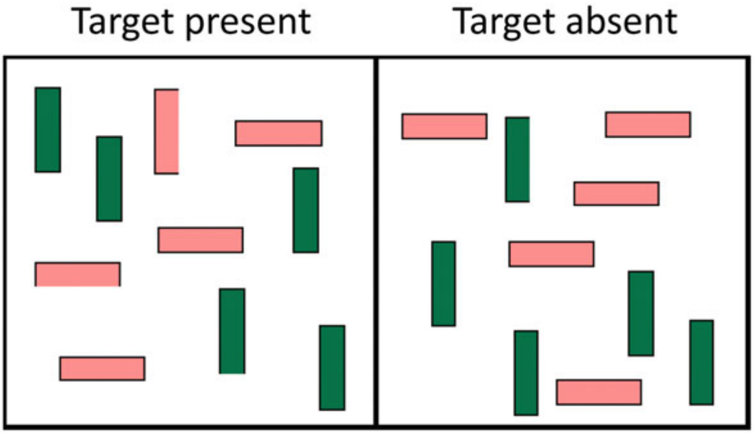


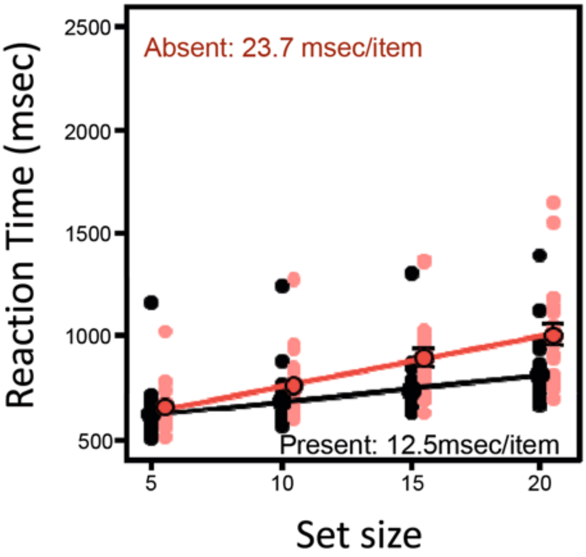
Fig. 4 A conjunction search for the *light red vertical* item among *light red horizontal* and *dark green vertical* distractors

Model 2: Serial Self-Terminating Search in a Subset

The obvious modification in the basic serial self-terminating model is to propose that the search is terminated after an exhaustive search through the relevant subset. In the example, given above, that would be the set of all red items. This model also runs into difficulties. One challenge comes from conjunction search tasks of the sort shown in Fig. 4.

In this task, Os look for a target defined by the conjunction of two features; here, the light red vertical item. Treisman had originally proposed that conjunction searches produced the same pattern of results produced by searches like the T vs L example (Treisman and Gelade 1980). However, subsequent research showed that conjunction

Fig. 5 RT × set size data for a color x orientation conjunction search. Each *dot* represents the average RT for one observer at one set size. *Light red dots* show absent trials. *Black dots* show target present trials. *Larger symbols* are group averages. Error *bars* are ±1 SEM. *Lines* are best-fit regressions for the average points. The scale is the same as in Fig. 2 for purposes of comparison



searches could be much more efficient with shallower slopes (Nakayama and Silverman 1986; Sagi 1988; Wolfe et al. 1989; Zohary and Hochstein 1989). How should observers perform on absent trials? If the relevant subset was the set of items that were either red or vertical, then the subset is the entire set—and that cannot be right. The slopes are too shallow. If the subset was the set of items that were both red and vertical, then the subset is empty on target absent trials and the slopes are too steep to support that assumption. Neither of these possibilities describes observers' behavior. They produce slopes on absent trials that are about twice the slope of target present trials. An example, from the same observers, shown in Fig. 2, is shown in Fig. 5.

One might propose that Os searched through half the items, perhaps based, on color. However, that version of a subset-search hypothesis can be rejected. When observers are forced to search through a subset based on color, performance looks very different from performance in standard conjunctions searches of the type illustrated in Fig. 4 (Friedman-Hill and Wolfe 1995). Moreover, some conjunction searches can be very efficient, with slopes near zero (Theeuwes and Kooi 1994). This is another challenge to any model that proposes that blank trials involve an exhaustive search through a feature-defined subset of items.

We will point to another challenge here and return to it later. A model based on exhaustive search through the set of items or some subset of items is plausible when the stimuli are well-isolated items on a blank background, as in typical laboratory search experiments (and as in Figs. 2 and 4). It is much more difficult to implement such a model in a real scene because it is all but impossible to decide what the set size might be (Neider and Zelinsky 2008) (Wolfe et al. 2008). Look up from this text. Examine the world in front of you and try to decide what the set size might be. Still, an exhaustive search through some subset might still be a plausible model if a way could be found to define the subset.

Model 3: Serial Self-Terminating Search up to an Activation Boundary (Guided Search 2.0)

A version of this type of subset search was proposed in Guided Search 2.0 (Wolfe 1994). In all of the incarnations of the Guided Search model, attention is guided by basic attributes of the stimulus such as color, orientation, size, etc. (Wolfe and Horowitz 2004). As noted earlier, guidance comes in two forms. Attention is guided to an item in a bottom-up, stimulus driven manner if that item differs from its neighbors in a guiding attribute (red among green, vertical among horizontal, and so forth). As discussed extensively by Duncan and Humphreys (1989), the greater the difference between target and distractors, the easier a search will be (red among green is easier than red among orange). The greater the featural heterogeneity of the distractors, the harder the search will be (red among homogeneous orange distractors is easier than red among a variety of different colors).

Guidance can also be top-down, user driven. In Fig. 4, bottom-up activity is essentially noise. Effective guidance to the light red vertical item comes from top-down guidance to red and to vertical. In Guided Search, each of these sources of guidance contributes to an overall activation map. Attention is directed to the most active item/location in that map. The map must be degraded by noise. Otherwise, a search like the conjunction search of Fig. 4 should yield a slope of 0 because the target is the only item with 2 target attributes. In the absence of noise, guidance to red and to vertical would lead directly to the one red vertical item first time, every time. Some of the noise will come from bottom-up activation. The juxtaposition of red and green or vertical and horizontal items makes those items salient. That salience is not useful.

When the sources of guidance, useful and otherwise, are summed up and some noise is added, the result will be that targets in a search like the conjunction search of Fig. 4 will have some activation drawn from a distribution and different distractor types will have activations drawn from lower but overlapping distributions. If attention is directed to the most activated item, that first item will be the target on some trials but on other trials some distractors will be examined before the target is reached. Returning to the absent trials, a reasonable approach would be to set an activation threshold to a level below which only very few targets are ever found. That threshold could define the subset on each trial and unsuccessful searches could end after an exhaustive search through that subset. If a target was not found in the set of items above the threshold activation, then it is time to quit.

Data like those shown in Fig. 3 constrain the placement of the activation threshold in this model. If the threshold is set to examine only items with high activation, then search will be abandoned on too many target-present trials before the target is attended and the miss error rate will be too high. If the threshold is set too low, few targets will be missed but the RTs will be too long. Given some assumptions about the noise in the activation values, it was possible to use one set of parameters to simulate a substantial set of search experiments in Guided Search 2.0, producing reasonable simulated target present and absent RTs (Wolfe 1994) .

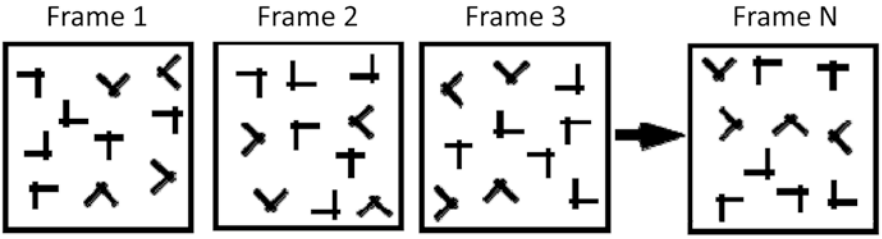


Fig. 6 Dynamic search: all items are randomly replotted on each *frame*. A target, if present, is present on every frame

There is an important assumption underlying this model and, indeed, all of the models that propose some sort of exhaustive search through some set of items on absent trials. To do an exhaustive search, one needs to know which distractors have been rejected. Put differently, search needs to sample without replacement from the display. Many models of search assumed such sampling and a mechanism, inhibition of return, had been proposed to account for this (Klein 1988). Unfortunately, the assumption does not appear to be correct.

Horowitz and Wolfe (1998) tried to test the assumption directly. They asked, what would happen if search were forced to sample with replacement? Their "dynamic search" method is illustrated in Fig. 6.

In dynamic search, observers see a sequence of frames. The items are the same on each frame but they are randomly replotted each time. A target item will be present on every frame of a target present trial and on no frames in absent trials. Dynamic search must require sampling with replacement, unless the search can be accomplished in a single frame. It can then be compared to a standard, static search condition. If, as required by the models sketched so far, rejected distractors are remembered in static search—that is, if static search is sampling without replacement—then there is a clear prediction for the relationship of slopes in the dynamic and static conditions. If standard, static search produces a slope of N ms/item, dynamic search, sampling with replacement, should produce a slope of 2N ms/item (Horowitz and Wolfe 2003).

The results rejected this hypothesis. In Horowitz and Wolfe (1998), the frame rate was 10 Hz. The slopes on target present trials were essentially the same in static and dynamic conditions. Horowitz and Wolfe reasoned that dynamic search *had* to be search with replacement. Thus, if static search produced the same result, it followed that static search was also search with replacement and they titled their paper "Visual search has no memory". Vigorous controversy ensued (Dodd et al. 2003; Gilchrist and Harvey 2006; Horowitz and Wolfe 2003; Kristjansson 2000; Peterson et al. 2001; Shore and Klein 2000). Figure 7 shows the results of a replication of the original dynamic search result from Horowitz and Wolfe (2003). This time the frame rate was 2 Hz, large set sizes were used, and in one version of the dynamic condition, targets could only appear in a few display positions, unknown to the observer. This was done to thwart "sit and wait" strategies in which the observer might pick one location and simply wait for the randomly plotted target to appear (von Muhlenen et al. 2003).

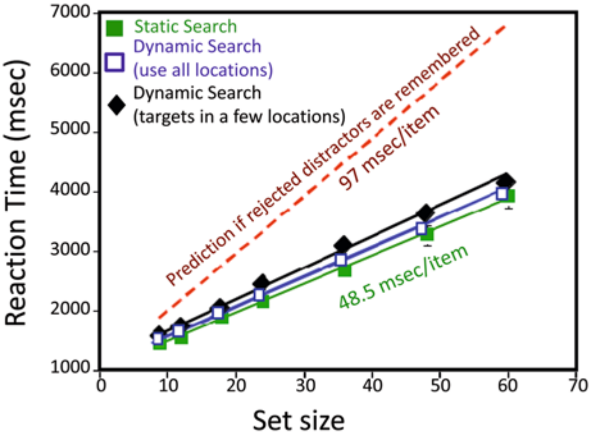


Fig. 7 Dynamic search produces similar results to static search. *Green, filled squares* show standard static search. *Blue, open squares* show dynamic search with random replotting of items. *Black filled diamonds* show a version of dynamic search with targets constrained to appear in a few locations. The *red dashed line* shows the predicted dynamic slope if static search has full memory for rejected distractors. (Replotted from Horowitz and Wolfe 2003)

The results, shown in Fig. 7, again show dynamic and static search having similar slopes.

If visual search really had no memory, one would think that perseveration would be a serious problem. Imagine that there was one salient distractor in the display. In the no memory account, what would keep attention from continuously revisiting that item? Moreover, the papers cited above that responded to the original “no memory” claim, make a case there is at least some limited memory in visual search. For methodological reasons, it is hard to differentiate between the consequences of a little memory and no memory in the dynamic search task. Perhaps the most plausible position is that “inhibition of return is a foraging facilitator in visual search” (Klein and MacInnes 1999). That is, perfect memory for the rejected distractors does not exist but there is enough inhibition to prevent perseveration and to bias attention toward new items. This seems reasonable but, returning to the problem of absent trials, the models presented thus far rely on perfect memory for rejected distractors, and that does not exist. A different type of model is needed.

Model 4: Timing or Counting Models

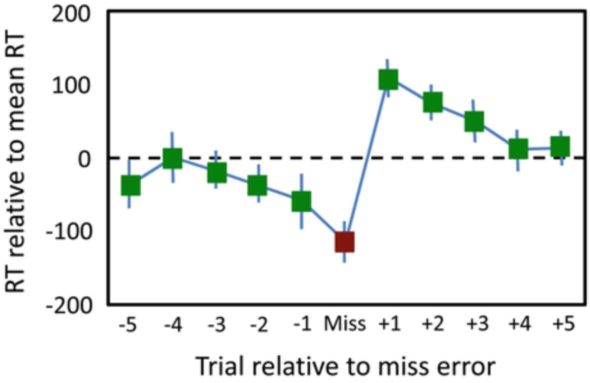
Even if the observer cannot rely on perfect memory for every deployment of attention, there is little doubt that he is accumulating some information about the ongoing searches. Suppose that an observer had information about the mean time required to find the target and the variance of that time. He could set a threshold in time rather than in activation. “If I search for N ms without finding the target, the probability that I will find a target is low enough that I might as well quit.” Observers would not, in fact, need to compute the mean and variance (implicitly or explicitly). Suppose that an adaptive process changed the quitting time on blank trials based on feedback from the ongoing sequence of trials. Threshold would be decreased and observers would quit more rapidly after correct responses and the threshold would increase after errors. If the step size on this 'staircase' is set appropriately, it would estimate a quitting threshold that would produce an acceptable error rate (Chun and Wolfe 1996). As an alternative to measuring time, the observer could count rejected distractors (sampled with or without replacement, it would not matter) and could quit after sampling some threshold number of items. As with a timing threshold, a counting threshold could be based on the number of items sampled in order to find targets on previous trials. These timing or counting models can be implemented with diffusion (Ratcliff 1978) or accumulation (Brown and Heathcote 2008) (Donkin et al. 2011) methods. In either case, search is terminated when the accumulating or diffusing signal reaches a termination threshold. That threshold, as noted, would go up in response to error and down in response to correct responses.

Chun and Wolfe (1996) looked for evidence for this adaptive mechanism. They ran observers in a triple conjunction (color X size X shape) task at a single set size of 25. Os made 3.3 % miss errors in an easy version of the task and 8.0 % errors in a harder version. Chun and Wolfe looked at RT as a function of the position of a trial, relative to a miss error. The results, replotted from the original paper, are shown in Fig. 8. It can be seen that RTs become faster after correct trials and markedly slower after a miss error.

In a function of this sort, we can see the searcher, estimating how long it should take to complete a search with an acceptable number of errors. However, there are a number of complications. First, the time for a given search is obviously dependent on how many items are present in a display. A timing or counting threshold that was established using one set size would be obviously incorrect for another set size. In practice, any such quitting threshold must be set relative to the set size on the current trial. We know this because performance on absent trials in standard search tasks does not change markedly whether set sizes are blocked or mixed. (Wolfe et al. 2010a). This raises a second problem. If observers can adjust the quitting criterion based on the set size, they must be able to derive that set size at the start of the trial. Since we know that exact counting is only possible in the subitizing range of up to about four items (Trick and Pylyshyn 1994), set size must be an estimate based on our ability to roughly enumerate larger number of items (Krueger 1984) (Dehaene 1997).

To recap, the working model would now say that the observer, at the start of a trial, makes an estimate of the set size and then sets a quitting threshold. It could be a counting threshold. In that case, the threshold would be set as some constant times the estimated set size. The constant would decrease if the search could be based on a subset (I only need to search the 25 % of objects that are green). It would also depend on whether search was sampling with or without replacement or somewhere in between. Alternatively, search termination could be based on a timing threshold, based on a calculation of the average time per item that must be devoted to the display, in order to produce a reasonable error rate. Models of this sort will run into problems when the observer is confronted with a real scene, as opposed to a display containing a countable number of items. As noted earlier, we simply have no idea

Fig. 8 Change RT relative to mean RT in a triple conjunction task, plotted with trials aligned to miss errors



how to count the number of searchable items in a real scene. We have some ideas about how to approach this problem. Observers can probably extract an "effective" set size (Neider and Zelinsky 2008) from the scene based on a variety of rapidly computed aspects of the gist of the scene (Wolfe et al. 2011b; Vo and Henderson 2010; Oliva 2005). Thus, for example, if you are looking for your thumb drive, objects the size of your computer screen probably do not enter into the calculation of effective set size. Moreover, this size constraint is probably calculated in three dimensions and not just in the image plane. Layout in depth is calculated quickly (Greene and Oliva 2009) and so the book, located across the room, that happens to subtend the same visual angle as a much closer missing thumb drive, nevertheless, is not a candidate for search because it is the wrong size in 3D even if it would be plausible in 2D (Sherman et al. 2011).

Beyond figuring out the effective set size in a scene, other properties of the scene will be important as well. Guidance by basic features like color will enter into the calculation of a quitting time. If you are looking for your car, it will make a great deal of difference if the car is an unusual shade of lime green or not. If that unusual color is not present in the visual field at all, you are likely to be able to abandon the search for the lime green car rapidly. The search for a more generic silver gray car will not be abandoned so quickly because your initial assessment of the scene will give you more hope that it is present, even if it is not. In addition, clutter and crowding become issues in real scenes (Bravo and Farid 2004; Felisberti et al. 2005; Levi 2008; Rosenholtz et al. 2009; Vlaskamp and Hooge 2006). Even if all the other factors are controlled, intuition holds that the search for a fully visible carrot peeler will be harder in a jumbled kitchen drawer than in a neat one. No one really knows how to compute clutter or crowding for these purposes, though progress is being made (Bravo and Farid 2008; Rosenholtz et al. 2009). Nevertheless, we can modify the current story to run as follows: When the scene (or an artificial search display) is presented to a viewer and a search task is defined, a quitting threshold is set based on an assessment of the gist of the scene. That gist will include an estimate of the numerosity of candidate targets. Candidate targets will be defined by their basic features and a variety of scene-based properties. The threshold will be

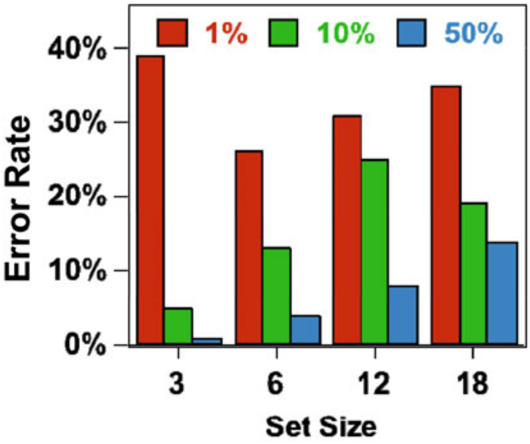


Fig. 9 Miss error rates as a function of target prevalence and set size. (Results redrawn from Wolfe et al. 2005, Nature)

further adjusted on the basis of an estimate of the ease with which candidate targets can be located and analyzed amidst whatever noise, clutter, or other obstruction is present. Diffusion or accumulator models can use a signal that measures time or that counts rejected items. Search can be terminated when that signal reaches the quitting threshold.

Target Prevalence

One factor that has not been mentioned but that has a substantial effect on search termination is the likelihood that a target is present. Returning to that carrot peeler, you should search for a longer time in the jumbled gadget drawer in your kitchen than you should search in a drawer in your office, even if we arrange for the two drawers to be visually equivalent. The prior probability of target presence is simply much higher in the former case than the latter. Moving from intuition to data, Wolfe et al. (2005) had observers search for black and white objects on a noise background. In different blocks, the targets were present on 50, 10 or 1 % of trials. These are quite laborious experiments since it takes 2,000 trials to collect a mere 20 target present trials at 1 % prevalence. Nevertheless, even with the limited statistical power imposed by the relatively small number of present trials, the results are dramatically clear, as shown in Figs. 9 and 10.

Figure 9 shows the miss ("false negative") error rates. Miss errors are much higher at low prevalence.

Figure 10 shows the RT data. The data for 50 % target prevalence show the typical RT × set size functions with absent trials being slower than present and having a slope of somewhat more than twice the present trial slope. In dramatic contrast, in this experiment, the absent RTs are actually shorter, on average than the present trial RTs. This RT result is somewhat more dramatic than is usual in prevalence studies.

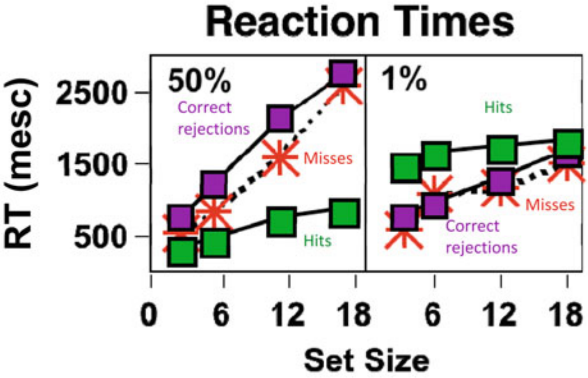


Fig. 10 RT × set size functions for 50 % and 1 % prevalence for the data shown in Fig. 9. *Green squares* show "hit" RTs, *purple-correct* absent trials, and *red asteriks* show miss error RTs

However, it is obvious that prevalence has a very substantial effect on target absent trials and that effect is not accounted for in the model sketched above. The basic prevalence result has been replicated many times (Fleck and Mitroff 2007; Godwin et al. 2010; Lau and Huang 2010; Kunar et al. 2010; Van Wert et al. 2009). There is always a rise in miss errors and a fall in target absent RTs as prevalence falls.

These effects of target frequency have been anticipated in tasks other than visual search. What has been called Hick–Hyman Law proposes RT increases with the number of alternatives (takes longer to respond with one of four keys than with one of two) and this has been taken to reflect a general relationship between stimulus frequency and RT (Hick 1952; Hyman 1953) (Maljkovic and Martini 2005). Moreover, in the vigilance literature, RT has been shown to increase as signal frequency decreases (Parasuraman and Davies 1976). The vigilance literature also documents the rise in miss errors as signal frequency decreases (Colquhoun and Baddeley 1967) (Mackworth 1970).

The prevalence effect is a potentially important phenomenon beyond the lab because a number of critical tasks are low prevalence search tasks. Clear examples include medical screening (Ethell and Manning 2001; Gur et al. 2003; Kundel 1982) and airport baggage screening. In each case, the target is very rare and in each case, miss errors are very undesirable. At the same time, the professionals doing these tasks are under time pressure to get through the workload. Is low target prevalence a source of errors in the field in any of these domains? Experiments are in progress as this is being written (and the answer is "yes", at least in breast cancer screening, even as this is being revised; Wolfe et al. 2011a). We know that expertise is not insulation against these effects. In one experiment, two groups of cytology technicians, who read Pap smear, cervical cancer tests, examined photomicrographs of cells. Each group read one set of stimuli at 50 % prevalence and another at low prevalence, either 2 or 5 %. One group simply rated slides on a 4-point normal/abnormal scale. The other group also localized apparent abnormalities. In the first group, false negatives/miss errors were 17 % at high prevalence and 30 % at low prevalence. In the second group, false negative rates rose from 27 to 42 % (One cannot make comparisons between the two groups because the stimulus sets were different). Incomplete data collection

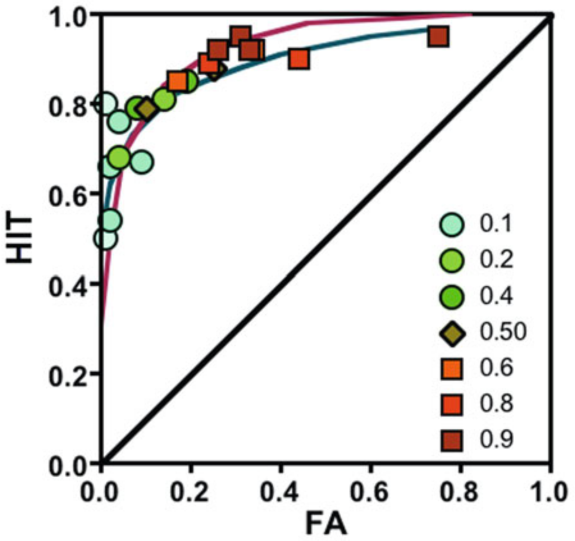


Fig. 11 ROC for an experiment with variable prevalence. *Green circles* indicate lower prevalence (< 0.5). *Red squares* indicate higher prevalence (> 0.5). Data are taken from Wolfe and Van Wert 2010. *Red ROC* assumes an equal variance. *Blue ROC* assumes unequal variance with a zROC slope of 0.6. (See text)

strongly suggests that other search experts will prove to be just as vulnerable to the prevalence effect as observers in the lab.

With miss errors going up at low prevalence and RT going down, an obvious thought is that the prevalence effect is nothing but a speed-accuracy trade-off. Fleck and Mitroff (2007) made an argument of this sort. It was based on data that showed that they could eliminate the prevalence effect if they simply allowed observers to rescind responses that they knew were in error. Everyone who has done visual search RT studies knows this phenomenon. You commit yourself to making target absent motor response. Then you find the target a moment later but it turns out to be a moment too late to recall the motor act.

There are a number of reasons to think that, while errors of this sort occur, they are not responsible for the main prevalence effect of interest. Most importantly, a speed-accuracy trade-off should represent a loss of sensitivity at low prevalence. (Note: We are using "sensitivity" to refer to what is indexed by d' or the area under an ROC curve in signal detection experiments. This is different from the usage in the medical community where sensitivity refers to the "hit" rate (P(correct|target-present)). The medical literature uses "specificity" to refer to the true negative rate (P(correct|target-absent)). In our original experiments and in the Fleck and Mitroff (2007) study, there were very few false alarm errors, making signal detection measures unreliable. When we used a simulated baggage search task that produced false alarms, it became clear that prevalence had its primary effect on response criterion, not on d'. The data shown in Fig. 11 illustrate the point. The data come from an experiment in which prevalence varied sinusoidally over the course of 1000 trials from near 1.0 to near 0 and back to 1.0 (Wolfe and Van Wert 2010). Each data point represents 50 trials from each of 13 observers. The color and shape coding of the points show the prevalence for those 50

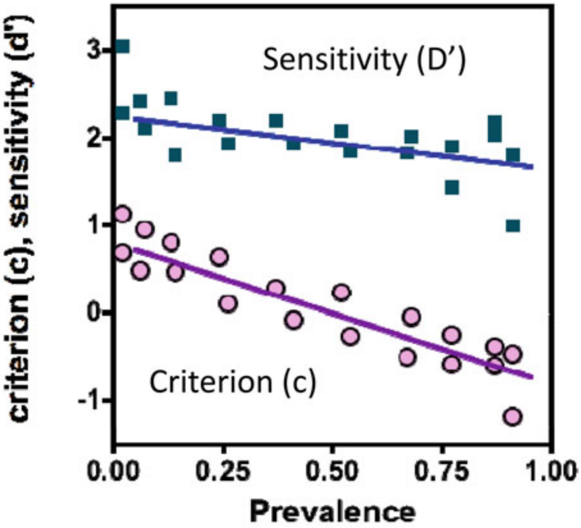


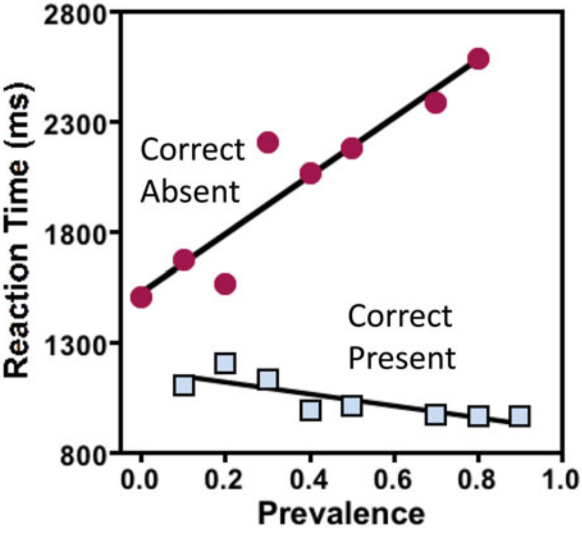
Fig. 12 Change in sensitivity (d') and criterion (c) as a function of prevalence in data taken from Wolfe and Van Wert (2010)

trials. Clearly, the data points slide along a ROC curve with low prevalence conditions characterized by high miss errors and low false alarms (conservative criterion) and high prevalence showing low miss errors and high false alarms (liberal criterion).

Figure 12 shows sensitivity (d') and criterion (c) for the data presented in Fig. 11. In a wide range of prevalence experiments, d' tends to be somewhat higher at low prevalence than at higher prevalence (Kundel 2000; Wolfe et al. 2007). As can be seen, this appears to be the case in Fig. 12. The effect is significant if all data points are included ($r^2 = 0.40$, $p = 0.0025$). If the extreme points are excluded on grounds that they are very unstable, the relationship is marginal (for Prevalence between 0.1 and 0.85, $r^2 = 0.32$, $p = 0.053$). This is probably an artifact of the underlying assumption that the "signal" and "noise" distributions are of equal variance. ROCs like the one shown in Fig. 11 become straight lines if plotted on Z-transformed axes. If the variance of the signal and the noise distributions are the same, the resulting zROC has a slope of 1. Slopes in these prevalence experiments tend to be less than 1 (0.6 for the data in Fig. 11). However, any of the various ways to deal with unequal variance preserve the strong relationship between criterion and prevalence (in various approaches, $r^2 > 0.75$, all $p < 0.0001$). The dependence of criterion on prevalence in search is anticipated in non-search tasks where changes in event frequency produce criterion shifts, rather than changes in sensitivity (Healy and Kubovy 1981; Swets and Kristofferson 1970).

The RT data provide another line of evidence suggesting that prevalence effects are not simple speed-accuracy tradeoffs. If prevalence effects were simple speed-accuracy tradeoffs, we might expect that target-present responses or, perhaps, all responses would become very fast at very high prevalence when observers could respond "present" with a good chance of being correct, no matter how quickly they pressed the key. However, as can be seen in Fig. 13, there is a fairly modest effect of

Fig. 13 Reaction time as a function of prevalence. (Average data derived from Wolfe and Van Wert 2010)



prevalence on target present trial RTs. There is a much larger effect on absent trials with very slow RTs accompanying the highest prevalence.

Prevalence is changing both speed and accuracy. Consequently, as discussed in Wolfe and Van Wert (2010), we need to think about two different criteria, each of them influenced by target prevalence. Thus far, we have been talking about a search termination threshold or criterion. When some accumulating quantity like elapsed time or number of items attended reaches that threshold, search ends, presumably with an "absent" response on most trials. While this accumulation to a termination threshold is progressing, there are other decisions that need to be made about each attended item. Is that item a target? These two alternative forced-choice decisions must have their own criterion. Prevalence alters both the search termination threshold and the target/non-target decision. As prevalence goes down, the search termination threshold comes down, meaning that observers are willing to abandon search sooner. At the same time, the 2AFC target/non-target decision criterion becomes more conservative; observers are less willing to declare that an item is a target. There are various lines of evidence to suggest that these are dissociable criteria. One of these is Experiment 2 of Wolfe et al. (2007). In an effort, to "cure" the prevalence effect, Wolfe et al. (2007) forced observers to slow down their responses. An initial block of 50 % prevalence trials established the average time required to find targets. In a subsequent low prevalence block, the computer produced a warning whenever an absent trial RT was below 1.3x of that average. This had the desired effect, at least, on RTs. Observers learned to slow their average target absent RTs by over a second (from < 1,500 to > 2,500 ms). Once behavior stabilized, they required warnings on only a few trials. Apparently, they had reset the termination threshold. However, there was no significant impact on errors. Miss errors remained much higher at low prevalence than at higher prevalence. While this was a setback in the quest to find a way to reduce miss errors, it does show that time to search and decisions about attended items are governed by dissociable responses to prevalence.

Estimating Prevalence

As noted above, the observer must estimate the set size, crowdedness, and/or clutter in a display in order to set a search termination criterion. Similarly, the observer must estimate the prevalence, if prevalence is going to have an influence on performance. One could imagine the estimate of prevalence being set by the prior history of the search. The frequency with which you find the target before the current trial would produce the estimate of the prevailing prevalence on this trial. Alternatively (or additionally), the estimate of prevalence could be based on top-down, semantic knowledge from outside of the search itself. That is, you could be told that the prevalence is 2 or 50 %. Under real world conditions, versions of both types of information are present. A radiologist knows that breast cancer is rare in a screening population and knows that she has found very few cancers in this collection of cases. Lau and Huang (2010) found no effect of explicit instruction on error rate and concluded that the prevalence effect is based entirely on past history with the task. However, Ishibashi et al. (2012) have subsequently reported a small effect of instruction on RT.

If prevalence effects are based on prior history, how much prior history is being taken into consideration. In the experiment described above (Figs. 11–13), Wolfe and Van Wert (2010) varied prevalence sinusoidally over 1,000 trials and found that error rates and RTs also varied in a roughly sinusoidal fashion. Given this variation in prevalence, each prevalence value was experienced twice, once as prevalence was falling and once as it was rising back to 1.0. Based on the difference between performance at the same prevalence value in the rising versus the falling portion of the sinusoid, Wolfe and Van Wert concluded that Os were using a prevalence estimate based on 40–50 trials. However, this may not be the best estimate of what we can call the "prevalence window" because the prevalence is changing and it is changing in a predictable manner.

How wide is the prevalence window when the prevalence for a block of trials is fixed. Even if overall prevalence is fixed, there will be local variations. Consider eight successive trials from an experiment with an overall prevalence of 50 %. Chance variation might produce 3 target present trials in one sequence, 7 in another, and so forth. Figure 14 shows the effects of just such random fluctuations in local prevalence for a window of 8 trials averaged over the data from 20 observers. The data happen to be taken from a search for a 2 among 5s (a task that will produce very few false alarm errors). Other data sets produce similar results. This particular data set includes variation in set size (5–20), which will introduce large variability into the absent RTs. Nevertheless, as can be seen, there is a substantial and significant effect of random fluctuations in local prevalence on absent trials RTs ($p = 0.04$). Target present RTs show no dependence on local prevalence ($p = 0.83$).

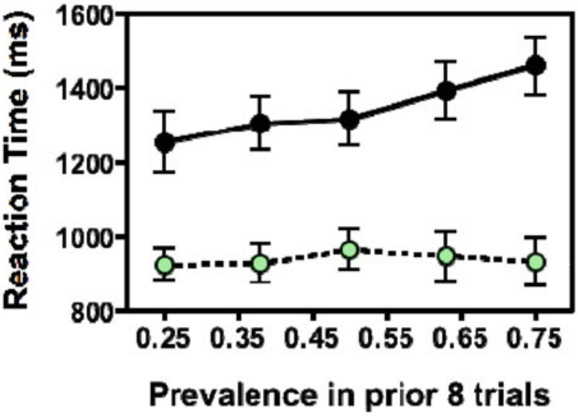


Fig. 14 RT as a function of local prevalence; in this case, the prevalence as calculated from the preceding eight trials

To estimate the prevalence window, we measured the correlation between RT and the local prevalence for windows of different sizes. Figure 15 gives the results of this analysis for a different data set; this time, a large data set where we collected 4000 trials per subject per condition in order to examine RT distributions. (Wolfe et al. 2010b). The large number of trials improves the statistical power of the analysis since, as can be seen on the y-axis of Fig. 15, while prevalence effects are reliable and quite large, they do not account for much of the variance in an experiment of this sort. Here the maximum correlation occurs in the range of 5–8 trials, suggesting a fairly small prevalence window. Interestingly, this is comparable to the range for priming of pop-out (Maljkovic and Nakayama 1994). In that paradigm, the color of items going back about eight trials into the past has an impact on the RT of a pop-out color search on the current trial.

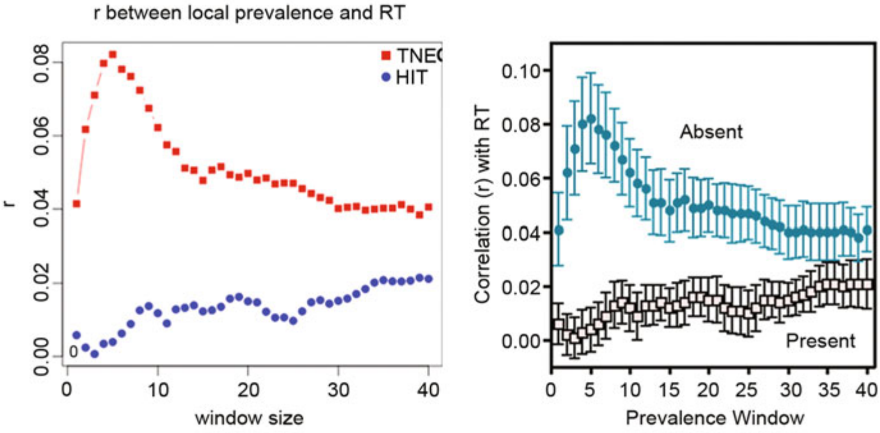


Fig. 15 Correlation of the current RT with the prevalence over the previous N trials (the "window size")

Value: One More Factor

The discussion thus far has failed to consider how badly you want to find whatever it is that you might be looking for. It seems quite obvious that this, too, will have an effect on search termination. You would search for a lost 20 dollar bill for longer than you would search for a 1 dollar bill. This example has the added virtue of making it intuitively clear that the effect of value on RT will be an effect on target-absent RTs. The 1 and 20 dollar bills are essentially the same visual stimulus. Thus, the time to find the bill, if the bill is successfully found, is unlikely to depend on its value. It is the time devoted to unsuccessful search that will be influenced by value.

There has been a recent uptick in interest in the effects of reward in visual search (Hearns and Moss 1968; Hickey et al. 2011; Hickey and Theeuwes 2008; Kristjansson et al. 2010). However, as with most other topics in search, much less attention has been devoted to the impact on target absent trials. There has been some discussion of reward in the context of the prevalence. After all, if one is concerned that low prevalence is pushing observers toward elevated miss errors, one should be able to move them the other direction on the ROC by changing the payoff. There is some evidence that prevalence effects are resistant to manipulations of payoff (Healy and Kubovy 1981; Maddox 2002) and Wolfe et al. (2007) argued that it would not work in settings like airport security and medical screening. However, more recent work shows that, if deployed correctly, payoff manipulations can affect the error rates, counteracting the prevalence effect (Navalpakkam et al. 2009). There is a need for work on reward effects on reaction time.

Moving to the Next Field

The model we have been sketching asserts that the observer in a visual search task is monitoring the time spent searching or the amount of searching that has been done. Search is terminated when the relevant quantity reaches a search termination threshold. On a given trial, that threshold is set by an estimate of the number of items in the display and an assessment of the difficulties imposed by crowding and clutter. The threshold is also influenced by the likelihood that a target is present. This estimate of prevalence seems to be based on recent search history and, perhaps, on something more like semantic knowledge. You don't need to have looked for President Obama multiple times in order to understand that he is unlikely to be in your kitchen. Finally, the termination threshold is influenced by the intrinsic value of the search target.

Now let us re-imagine the task. Suppose that each search display is a patch of a habitat in which some animal is searching for its food. In each patch, there either is or is not a food item. That assumption, convenient for 2AFC tasks, might not be entirely realistic but, if one imagines fairly sparse, evenly distributed food and patches of the right size, it is not a bad assumption. Described this way, visual search has much in common with foraging problems, as studied in Behavioral Ecology (Stephens and

Krebs 1986). The search termination problem becomes what is known as the “patch leaving” problem. When should our animal stop searching/foraging in one patch and move to the next?

Unlike visual search, where search termination is a bit of an orphan problem, patch leaving in behavioral ecology has attracted a lot of attention. Many accounts are versions of Charnov’s Marginal Value Theorem (Charnov 1976) which asserts that the animal should move when the marginal rate, the rate at which resources are being extracted from the patch, drops below the average rate of return for the environment. If you imagine picking berries from a bush, you pick at some rate. At some point, the rate begins to drop as the bush is depleted. It is time to move to another bush once the rate drops to a point below the average rate at which your berry bucket is filling up. If it takes a long time to get to the next bush, you should stay longer on the current bush because that travel time reduces the average rate of return.

Simple versions of the marginal value theorem assume that the average rate is known and uniform (Pyke et al. 1977). Realistic complications ensue if you endow the animal with an ability to sense the distribution of resource in an uneven habitat. Other variables might include how long it takes to consume an item or whether one type of item is more common than another.

It is not hard to map foraging variables to visual search variables. The various factors that influence the slope of RT × set size functions are influencing the observer’s rate of return; how many targets he eats per unit time. Endowing the observer with preattentive processes that give that observer the numerosity of a display and guide his attention to likely targets are like the processes that would allow an animal to notice that one patch appears to be more promising than another. Attentional limits have been proposed to constrain behavior in foraging as well as in the search domain (Dukas 2002; Dukas 2004). This is not to say that there is a trivial equivalence of issues in search and in foraging. However, the rich theoretical work in behavioral ecology provides a promising habitat for visual search researchers. The control and comparative ease with which visual search data can be collected represents an opportunity to test some of those theories.

A Brief Conclusion

The search termination problem is important. Many searches get terminated without success. Searches for unknown numbers of targets always face a termination problem. Search too long and you are perseverating. Quit too fast and you are leaving too many targets undiscovered. If the search is a search for a mate or food or cancer or a bomb in luggage, the costs of poor performance can be very great. The topic is under-researched in visual search but a basic model can be outlined on the basis of what we know (and can be implemented, at least in one incarnation in the supplement to Wolfe and Van Wert 2010). Figure 16 gives a final summary.

We envision a search termination decision, based on how many items have been searched or how much time has been spent in search. For any given search, there

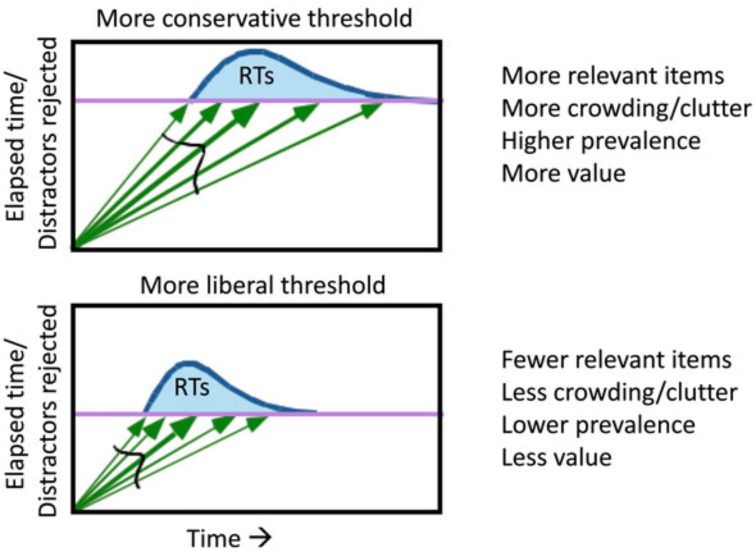


Fig. 16 Factors constraining search termination

will be a search threshold, expressed in time or item units. Information about time or items will accumulate toward that threshold in a noisy fashion (green arrows). The resulting distribution of RTs will be positively skewed (Palmer et al. 2011; Van Zandt 2002). If the threshold is more liberal, observers will quit more rapidly. Factors that will move the threshold to a more conservative position would be: (1) A greater number of relevant items (larger set size, larger number of items with the correct features, etc), (2) More crowding and clutter, making it harder to get information out of the image, (3) Higher target prevalence, and (4) Higher value.

Analogous problems exist in other domains. Here, we briefly considered the relationship to patch leaving in behavioral ecology. Ideas brought in from these neighboring fields should allow us to make progress in figuring out when to quit.

Acknowledgments This research was supported by grants from the National Institutes of Health and National Eye Institute (EY017001), the Office of Naval Research (ONR MURI N000141010278) and from the United States Department of Homeland Security (02-G-010).

References

Bacon, W. F., & Egeth, H. E. (1994). Overriding stimulus-driven attentional capture. *Perception and Psychophysics, 55*(5), 485–496.

Berbaum, K. S., Franken, E. A., Jr., Dorfman, D. D., Rooholamini, S. A., Kathol, M. H., & Barloon, T. J., et al. (1990). Satisfaction of search in diagnostic radiology. *Investigative Radiology, 25*(2), 133–140.

Berbaum, K. S., Franken, E. A., Jr., Dorfman, D. D., Caldwell, R. T., & Krupinski, E. A. (2000). Role of faulty decision making in the satisfaction of search effect in chest radiography. *Academic Radiology, 7*(12), 1098–1106.

Bravo, M. J., & Farid, H. (2004). Search for a category target in clutter. *Perception, 33*(5), 643–652.

Bravo, M. J., & Farid, H. (2008). A scale invariant measure of clutter. *Journal of Vision, 8*(1), 1–9.

Brown, S. D., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology, 57*(3), 153–178.

Buschman, T. J., & Miller, E. K. (2009). Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron, 63,* 386–396.

Charnov, E. L. (1976). Optimal foraging, the marginal value theorem. *Theoretical Population Biology, 9,* 129–136.

Chun, M. M., & Wolfe, J. M. (1996). Just say no: How are visual searches terminated when there is no target present? *Cognitive Psychology, 30,* 39–78.

Colquhoun, W. P., & Baddeley, A. D. (1967). Influence of signal probability during pretraining on vigilance decrement. *Journal of Experimental Psychology: Human Perception and Performance, 73*(1), 153–155.

Dehaene, S. (1997). *The number sense.* New York, Cambridge: Oxford University Press, Penguin press.

Dodd, M. D., Castel, A. D., & Pratt, J. (2003). Inhibition of return with rapid serial shifts of attention: Implications for memory and visual search. *Perception & Psychophysics, 65*(7), 1126–1135.

Donkin, C., Brown, S., Heathcote, A., & Wagenmakers, E.-J. (2011). Diffusion versus linear ballistic accumulation: Different models but the same conclusions about psychological processes? *Psychonomic Bulletin & Review, 18*(1), 61–69.

Dukas, R. (2002). Behavioural and ecological consequences of limited attention. *Philosophical Transaction of the Royal Society London Series B: Biolgical Science, 357*(1427), 1539–1547.

Dukas, R. (2004). Causes and consequences of limited attention. *Brain, Behavior and Evolution, 63*(4), 197–210.

Duncan, J., & Humphreys, G. W. (1989). Visual search and stimulus similarity. *Psychological Review, 96,* 433–458.

Egeth, H. E., Virzi, R. A., & Garbart, H. (1984). Searching for conjunctively defined targets. *Journal of Experimental Psychology: Human Perception and Performance, 10,* 32–39.

Einhauser, W., Spain, M., & Perona, P. (2008). Objects predict fixations better than early saliency. *Journal of Vision, 8*(14), 1–26.

Ethell, S. C., & Manning, D. (2001). Effects of prevalence on visual search and decision making in fracture detection. In E. A. Krupinski & D. P. Chakraborty (Eds.), *Image perception and performance, proceedings of spie* (Vol. 4324, pp. 249–257). SPIE.

Felisberti, F. M., Solomon, J. A., & Morgan, M. J. (2005). The role of target salience in crowding. *Perception, 34*(7), 823–833.

Fleck, M. S., & Mitroff, S. R. (2007). Rare targets are rarely missed in correctable search. *Psychological science, 18*(11), 943–947.

Fleck, M. S., Samei, E., & Mitroff, S. R. (2010). Generalized "satisfaction of search": Adverse influences on dual-target search accuracy. *Journal of Experimental Psychology: Applied, 16*(1), 60–71. doi: 10.1037/a0018629.

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.

Franconeri, S. L., Hollingworth, A., & Simons, D. J. (2005). Do new objects capture attention? *Psychological Science, 16*(4), 275–281.

Friedman-Hill, S. R., & Wolfe, J. M. (1995). Second-order parallel processing: Visual search for the odd item in a subset. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 531–551.

Gilchrist, I. D., & Harvey, M. (2006). Evidence for a systematic component within scanpaths in visual search. *Visual Cognition, 14*(5–7), 704–771.

Godwin, H. J., Menneer, T., Cave, K. R., Helman, S., Way, R. L., & Donnelly, N. (2010). The impact of relative prevalence on dual-target search for threat items from airport X-ray screening. *Acta Psychologica, 134*(1), 79–84. doi: 10.1016/j.actpsy.2009.12.009.

Goldsmith, M. (1998). What's in a location? Comparing object-based and space-based models of feature integration in visual search. *Journal of Experimental Psychology: General, 127*(2), 189–219.

Greene, M. R., & Oliva, A. (2009). The briefest of glances: The time course of natural scene understanding. *Psychological Science, 20*(4), 464–472.

Gur, D., Rockette, H. E., Armfield, D. R., Blachar, A., Bogan, J. K., Brancatelli, G., et al. (2003). Prevalence effect in a laboratory environment. *Radiology, 228*(1), 10–14.

Hamid, S. N., Stankiewicz, B., & Hayhoe, M. (2010). Gaze patterns in navigation: Encoding information in large-scale environments. *Journal of Vision, 10*(12), 28.

Hamker, F. (2006). Modeling feature-based attention as an active top-down inference process *Biosystems, 86*(1–3), 91–99.

Healy, A. F., & Kubovy, M. (1981). Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *Journal of Experimental Psychology: Human Learning and Memory, 7*(5), 344–354.

Hearns, J. F., & Moss, S. M. (1968). Effect of method of payoff on detection of targets in a visual search task. *Journal of Experimental Psychology: Human Learning and Memory, 78*(4P1), 569–573.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and action: Eye movements and the visual world* (pp. 1–58). New York: Psychology.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4,* 11–26.

Hickey, C., & Theeuwes, J. (2008). Reward primes visual search. *Perception, 37,* 46–46.

Hickey, C., Chelazzi, L., & Theeuwes, J. (2011). Reward has a residual impact on target selection in visual search, but not on the suppression of distractors. *Visual Cognition, 19*(1), 117–128.

Horowitz, T. S., & Wolfe, J. M. (1998). Visual search has no memory. *Nature, 394*(Aug 6), 575–577.

Horowitz, T. S., & Wolfe, J. M. (2003). Memory for rejected distractors in visual search? *Visual Cognition, 10*(3), 257–298.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology: Human Learning and Memory, 45*(3), 188–196.

Ishibashi, K., Kita, S., & Wolfe, J. (2012). The effects of local prevalence and explicit expectations on search termination times. *Attention, Perception, & Psychophysics, 74*(1), 115–123.

Jovancevic-Misic, J., & Hayhoe, M. (2009). Adaptive gaze control in natural environments. *Journal of Neuroscience, 29*(19), 6234–6238.

Klein, R. (1988). Inhibitory tagging system facilitates visual search. *Nature, 334,* 430–431.

Klein, R. M., & MacInnes, W. J. (1999). Inhibition of return is a foraging facilitator in visual search. *Psychological Science, 10*(July), 346–352.

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

Kristjansson, A. (2000). In search of rememberance: Evidence for memory in visual search. *Psychological Science, 11*(4), 328–332.

Kristjansson, A., Sigurjonsdottir, O., & Driver, J. (2010). Fortune and reversals of fortune in visual search: Reward contingencies for pop-out targets affect search efficiency and target repetition effects. *Attention Perception Psychophysics, 72*(5), 1229–1236.

Krueger, L. E. (1984). Perceived numerosity: A comparison of magnitude production, magnitude estimation, and discrimination judgments. *Perception & Psychophysics, 35*(6), 536–542.

Kunar, M. A., Rich, A. N., & Wolfe, J. M. (2010). Spatial and temporal separation fails to counteract the effects of low prevalence in visual search. *Visual Cognition, 18*(6), 881–897.

Kundel, H. L. (1982). Disease prevalence and radiological decision making. *Investigative Radiology, 17*(1), 107–109.

Kundel, H. L. (2000). Disease prevalence and the index of detectability: A survey of studies of lung cancer detection by chest radiography. In E. A. Krupinski (Ed.), *Medical imaging 2000: Image perception and performance* (Vol. 3981, pp. 135–144). Society of Photo Optical.

Lau, J. S., & Huang, L. (2010). The prevalence effect is determined by past experience, not future prospects. *Vision Research, 50*(15), 1469–1474.

Levi, D. M. (2008). Crowding-an essential bottleneck for object recognition: A mini-review. *Vision Research, 48*(5), 635–654.

Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103*(4), 603–649.

Mackworth, J. (1970). *Vigilance and attention*. Harmondsworth: Penguin Books.

Macmillan, N. A., & Creelman, C. D. (2005). *Detection theory*. Mahwah: Erlbaum.

Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior, 78*(3), 567–595.

Maljkovic, V., & Martini, P. (2005). Implicit short-term memory and event frequency effects in visual search. *Vision Research, 45*(21), 2831–2846.

Maljkovic, V., & Nakayama, K. (1994). Priming of popout: I. Role of features. *Memory & Cognition, 22*(6), 657–672.

Masciocchi, C. M., Mihalas, S., Parkhurst, D., & Niebur, E. (2009). Everyone knows what is interesting: Salient locations which should be fixated. *Journal of Vision, 9*(11), 1–22.

Nakayama, K., & Silverman, G. H. (1986). Serial and parallel processing of visual feature conjunctions. *Nature, 320,* 264–265.

Navalpakkam, V., Koch, C., & Perona, P. (2009). Homo economicus in visual search. *Journal of Vision, 9*(1), 31.

Neider, M. B., & Zelinsky, G. J. (2008). Exploring set size effects in scenes: Identifying the objects of search. *Visual Cognition, 16*(1), 1–10.

Nodine, C. F., Krupinski, E. A., Kundel, H. L., Toto, L., & Herman, G. T. (1992). Satisfaction of search (SOS). *Invest Radiology, 27*(7), 571–573.

Oliva, A. (2005). Gist of the scene. In L. Itti, G. Rees, & J. Tsotsos (Eds.), *Neurobiology of attention* (pp. 251–257). San Diego: Academic.

Palmer, E. M., Horowitz, T. S., Torralba, A., & Wolfe, J. M. (2011). What are the shapes of response time distributions in visual search? *Journal of Experimental Psychology: Human Perception and Performance, 37*(1), 58–71.

Parasuraman, R., & Davies, D. R. (1976). Decision theory analysis of response latencies in vigilance. *Journal of Experimental Psychology. Human Perception and Performance, 2*(4), 578–590.

Peterson, M. S., Kramer, A. F., Wang, R. F., Irwin, D. E., & McCarley, J. S. (2001). Visual search has memory. *Psychological Science, 12*(4), 287–292.

Pyke, G. H., Pulliam, H. R., & Charnov, E. L. (1977). Optimal foraging: A selective review of theory and tests. *The Quarterly Review of Biology, 52*(2), 137–154.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychology. Review, 85*(2), 59–108.

Rayner, K. (1983). Visual selection in reading, picture perception, and visual search. In H. Bouma & B. G. Bouwhuis (Eds.), *Attention and performance X* (Vol. 10, pp. 67–96). Hillside: Erlbaum.

Reynolds, J. H., & Chelazzi, L. (2004). Attentional modulation of visual processing. *Annual Review of Neuroscience, 27,* 611–647.

Roelfsema, P. R., Lammer, V. A. F., & Spekreijse, H. (1998). Object-based attention in the primary visual cortex of the macaque monkey. *Nature, 395,* 376.

Rosenholtz, R., Chan, S., & Balas, B. (2009). A crowded model of visual search. *Journal of Vision, 9*(8), 1197–1197.

Sagi, D. (1988). The combination of spatial frequency and orientation is effortlessly perceived. *Perception and Psychophysics, 43,* 601–603.

Sherman, A. M., Greene, M. R., & Wolfe, J. M. (2011). Depth and size information reduce effective set size for visual search in real-world scenes. *Poster presented at the annual meeting of the Vision Sciences Society in Naples, FL.*, May, 2011

Shore, D. I., & Klein, R. M. (2000). On the manifestations of memory in visual search. *Spatial vision, 14*(1), 59–75.

Stephens, D. W., & Krebs, J. R. (1986). *Foraging theory*. Princeton: Princeton University Press.

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153*, 652–654.

Swets, J. A., & Kristofferson, A. B. (1970). Attention. *Annual Review of Psychology, 21,* 339–366.

Theeuwes, J. (1995). Abrupt luminance change pops out; abrupt color change does not. *Perception & Psychophysics, 57*(5), 637–644.

Theeuwes, J. (2010). Top-down and bottom-up control of visual selection. *Acta Psychologica, 135*(2), 77–99. doi: 10.1016/j.actpsy.2010.02.006.

Theeuwes, J., & Kooi, J. L. (1994). Parallel search for a conjunction of shape and contrast polarity. *Vision Research, 34*(22), 3013–3016.

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 on object search. *Psychological Review, 113*(4), 766–786.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology, 12,* 97–136.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychology. Review, 101*(1), 80–102.

Tsotsos, J. (2011). *A computational perspective on visual attention*. Cambridge: MIT Press.

Van Zandt, T. (2002). Analysis of response time distributions. In H. Pashler, & J. Wixted (Eds.), *Stevens' handbook of experimental psychology* (3rd edn), *Methodology in experimental psychology* (Vol. 4, pp. 461–516). New York: Wiley.

Van Wert, M. J., Wolfe, J. M., & Horowitz, T. S. (2009). Even in correctable search, some types of rare targets are frequently missed. *Attention, Perception & Psychophysics, 71*(3), 541–553.

Vlaskamp, B. N., & Hooge, I. T. (2006). Crowding degrades saccadic search performance. *Vision Research, 46*(3), 417–425.

Vo, M. L. H., & Henderson, J. M. (2009). Does gravity matter? Effects of semantic and syntactic inconsistencies on the allocation of attention during scene perception. *Journal of Vision, 9*(3), 1–15.

Vo, M. L.-H., & Henderson, J. M. (2010). The time course of initial scene processing for eye movement guidance in natural scene search. *Journal of Vision, 10*(3), 1–13.

von Muhlenen, A., Muller, H. J., & Muller, D. (2003). Sit-and-wait strategies in dynamic visual search. *Psychol Science, 14*(4), 309–314.

Wolfe, J. M. (1994). Guided Search 2.0: A revised model of visual search. *Psychonomic Bulletin and Review, 1*(2), 202–238.

Wolfe, J. M. (1998). What do 1,000,000 trials tell us about visual search? *Psychological Science, 9*(1), 33–39.

Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience, 5*(6), 495–501.

Wolfe, J. M., & Van Wert, M. J. (2010). Varying target prevalence reveals two, dissociable decision criteria in visual search. *Current Biology, 20*(2), 121–124.

Wolfe, J. M., Cave, K. R., & Franzel, S. L. (1989). Guided Search: An alternative to the feature integration model for visual search. *Journal of Experimental Psychology: Human Perception and Performance, 15,* 419–433.

Wolfe, J., Horowitz, T., Kenner, N. M., Hyle, M., & Vasan, N. (2004). How fast can you change your mind? The speed of top-down guidance in visual search. *Vision Research, 44*(12), 1411–1426.

Wolfe, J. M., Horowitz, T. S., & Kenner, N. M. (2005). Rare items often missed in visual searches. *Nature, 435,* 439–440.

Wolfe, J. M., Horowitz, T. S., Van Wert, M. J., Kenner, N. M., Place, S. S., & Kibbi, N. (2007). Low target prevalence is a stubborn source of errors in visual search tasks. *Journal of Experimental psychology: General, 136*(4), 623–638.

Wolfe, J., Alvarez, G., Rosenholtz, R., Oliva, A., Torralba, A., Kuzmova, Y., et al. (2008). Search for arbitrary objects in natural scenes is remarkably efficient. *Journal of Vision, 8*(6), 1103–1103.

Wolfe, J. M., Horowitz, T. S., Palmer, E. M., Michod, K. O., & Van Wert, M. J. (2010a). Getting in to guided search. In V. Coltheart (Ed.), *Tutorials in visual cognition.* (pp. 93–120). Hove: Psychology.

Wolfe, J. M., Palmer, E. M., & Horowitz, T. S. (2010b). Reaction time distributions constrain models of visual search. *Vision Research, 50*(14), 1304–1311.

Wolfe, J. M., Birdwell, R. L., & Evans, K. K. (2011a). If you Don't Find it Often, You Often Don't Find It: Disease Prevalence is a Source of Miss Errors in Screening Mammography. Paper presented at the Anuual Radiological Society of North America meeting: Chicago, IL., Nov 30, 2011.

Wolfe, J. M., Vo, M. L.-H., Evans, K. K., & Greene, M. R. (2011b). Visual search in scenes involves selective and non-selective pathways. *Trends in Cognitive Sciences, 15*(2), 77–84.

Yantis, S., & Jonides, J. (1996). Attentional capture by abrupt onsets: New perceptual objects or visual masking. *Journal of Experimental Psychology: Human Perception and Performance, 22*(6), 1505–1513.

Yeari, M., & Goldsmith, M. (2010). Is object-based attention mandatory? Strategic control over mode of attention. *Journal of Experimental Psychology: Human Perception and Performance, 36*(3), 565–579.

Zohary, E., & Hochstein, S. (1989). How serial is serial processing in vision? *Perception, 18,* 191–200.

Index

M. D. Dodd, J. H. Flowers (eds.), *The Influence of Attention, Learning, and Motivation on Visual Search,* Nebraska Symposium on Motivation,
DOI 10.1007/978-1-4614-4794-8,