Bernhard Hommel
Stephen B.R.E. Brown
Dieter Nattkemper

Human Action Control

From Intentions to Movements

Springer

Human Action Control

Bernhard Hommel • Stephen B.R.E. Brown
Dieter Nattkemper

Human Action Control

From Intentions to Movements

Springer

Bernhard Hommel
Institute of Psychology
Leiden University
Leiden, The Netherlands

Stephen B.R.E. Brown
Institute of Psychology
Leiden University
Leiden, The Netherlands

Dieter Nattkemper
Institute of Psychology
Humboldt University of Berlin
Berlin, Germany

ISBN 978-3-319-09243-0 ISBN 978-3-319-09244-7 (eBook)
DOI 10.1007/978-3-319-09244-7

Library of Congress Control Number: 2016932402

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG Switzerland

Preface

The purpose of this textbook is to provide insight into the areas of action planning and action control. These two areas are still relatively separated, and many of the topics we will be discussing are studied in different disciplines and subdisciplines, such as cognitive vs. motivational psychology, movement and sport sciences, neurophysiology, and biology. Unfortunately, this diversity has generated different kinds of scientific jargon and theorizing, which is difficult to relate to one another. Despite the varied studies throughout multiple disciplines, action still plays an underprivileged role in psychology. It does not feature as a relevant topic in 99 % of psychological introductory textbooks. We believe that every effort has to be taken to better integrate action into psychology's canon of knowledge. To do so, we have taken an uncommon route.

Generally speaking, textbooks provide a more or less exhaustive overview of a research area or topic and discuss the most recent findings and trends therein. The advantage of this strategy is obvious: the reader is provided with maximum information and, thus, with the opportunity to make up her own opinion. But, there is also an often overlooked disadvantage: while experts have enough background to structure new information in the most efficient way, novices can be over-challenged by the sheer amount of information that standard textbooks provide and, often, are not able to appreciate all the subtle implications that the most recent findings might have. This is, at least, the experience that we have had when we were students and that our current students often report when struggling with textbooks.

We therefore opted for another strategy. It consists of focusing on, in our opinion, the most basic principles and theoretical figures of thought in the historical development of the research area. As a result, we have used only a few empirical findings as examples for how theory and data are connected. Also, we translated domain-specific jargon into our own preferred terminology, which makes it easier to relate the concepts we discuss. Furthermore, we offer a general organizational framework of how we think action planning and action control is working, which will help the reader to organize the information we provide. Our approach has obvious disadvantages: it is necessarily much more selective and often refers to classical papers that have introduced particular lines of thought rather than the most recent applications of these thoughts, which explains why the average age of the literature we cite is unusually high. Most importantly, this selectivity implies that this textbook is likely to reflect the theoretical preferences and biases of its authors more than others. It is, thus, important that the reader does not forget that our approach is just one of several possible approaches. “Doubt comes *after* belief,” Wittgenstein says in his last book, *On Uncertainty*, meaning that we can start doubting only once we are done with building up our basis of knowledge. It is in this sense that we encourage readers to use this book as a jumping board to build their first basis and later try questioning it based on what other authors say, wherever appropriate and necessary.

Many of the questions that we discuss in the following chapters, and many of the answers that we suggest, emerged from Nattkemper and Hommel’s numerous collaborations with members of the “Cognition & Action” Unit of the Munich Max Planck Institute for Psychological Research and its spiritual leader, Wolfgang Prinz. We were unable to reconstruct exactly which ideas and speculations have motivated which of the considerations that we will present in the following chapters, but we are 100 % certain that our theoretical preferences and biases, and the style of reasoning about cognition and action, were strongly shaped by “Cognition & Action.”

Leiden, The Netherlands Bernhard Hommel
Leiden, The Netherlands Stephen B.R.E. Brown
Berlin, Germany Dieter Nattkemper

Contents

About the Authors

Dr. Stephen B. R. E. Brown studied developmental and cognitive psychology at the University of Amsterdam. After obtaining his Master's degree there, he moved to Leiden University to obtain a Ph.D. under supervision of Professors Sander Nieuwenhuis and Bernhard Hommel. His dissertation focused on the locus coeruleus, a brainstem nucleus that provides the entire brain with the neuromodulator noradrenaline, which is released whenever we experience arousal. Currently, Dr. Brown works as a postdoctoral researcher at Leiden University, and studies heart rate variability as a predictor of psychosocial stress in collaboration with Professor Jos Brosschot. He is also employed as an instructor at the University of Amsterdam, where he coordinates and teaches a course on Cognition for interdisciplinary students.

Dr. Bernhard Hommel holds the chair of "General Psychology" at Leiden University since 1999, after having worked as senior researcher at the Max-Planck Institute for Psychological Research (Ph.D. at the University of Bielefeld in 1990; Habilitation at the Ludwig-Maximilians University of Munich). He is a co-founder and board member of the Leiden Institute for Brain & Cognition (LIBC), secretary of the International Association for Attention and Performance, and member of the German National Academy of Sciences. His research focuses on cognitive, computational, developmental, neural, and neurochemical mechanisms of human attention and action control, and the role of consciousness therein. Recent work also addresses the role of emotion, creativity, and religion in human cognition. He is chief editor of two journals and has (co-)authored more than 300 articles in international journals and more than 60 chapters in readers and psychological textbooks, (co-) edited three books on action control and the relationship between perception and action, and (co-)edited several special issues on attention and action control.

Dr. Dieter Nattkemper studied psychology at the University of Münster. His professional career started at the University of Osnabrück where he was engaged in investigating mechanisms of eye movement control in reading. In the 1980s, he worked (with W. Prinz) at the University of Bielefeld on issues related to understanding how internal, cognitive models of the environment contribute to controlling human behavior and how these supposed internal models are generated, consolidated, and modified in response to changes in our external world. He achieved his Ph.D. at the University of Bielefeld with studies elucidating the role of eye movements in tasks that require processes of continuous selection (continuous visual search). Thereafter, he joined the research group “Cognition and Action” of the Max-Planck-Institute for Psychological Research (headed by W. Prinz) where his research focused on questions relating to the mechanisms governing performance in serial reaction tasks. At the end of the past century, he moved to Humboldt University Berlin where he studied (with M. Ziessler and P. Frensch) the mechanisms underlying human action planning. Dr. Nattkemper has co-authored several articles in international journals, chapters in readers and psychological textbooks, a German textbook on human action planning and control (with B. Hommel), and co-edited a special issue on human action control. Due to steadily increasing problems caused by multiple sclerosis, he finished his professional career as an experimental psychologist and retired at the end of 2013.

Chapter 1
Introduction and Overview

The theme of the present textbook has accompanied the discipline of psychology for a long time, without actually being considered to be a core area. Textbooks on cognitive psychology generally only discuss perception, attention, memory, and cognition, without taking into account to what purpose (i.e., to control which actions) people use these cognitive processes. Of course, history has seen different approaches that have attempted to complete the cycle from perception to action and back, such as Lotze's (1852) considerations on executive ignorance, James' (1890) treatise on the human will, Woodworth's (1938) psychomotor studies, and many behaviorist programs (e.g., Thorndike 1898). However, in principle, cognitive psychology remained focused on the registration of information from the environment, and the processing thereof through increasingly higher-level cognitive processes. Some authors have even gone so far as to restrict their definition of the entire field of research to this processing (e.g., Neisser 1967). Historically, psychology as an autonomous science has preferentially focused on understanding the cognitive operations that serve to take information from the environment and to establish mental representations of the outside world. By contrast, studies on **action and motor control** had only modest presence in psychological research, to the extent that it has even been called the "Cinderella of psychology" (Rosenbaum 2005).

1.1 Action and Movement as a Theme in Psychological Research

It has only been recently that psychology has succeeded to study action once again. Various trends are responsible for this development. Firstly, the victory march of the **cognitive neurosciences** has made the various subfields of psychology draw closer to each other. From a neuronal point of view, it is indeed less crucial to tease apart where, for example, perception ends and memory starts, or which processes still represent perception, and which ones already constitute attention. This also concerns the apparently clear distinction between perception and action, especially

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_1

because many more sensory and motor brain areas are occupied with the transition from perception to action than with local processing issues.

Secondly—this trend is not completely independent of the first—the various scientific disciplines have been drawn closer together. In psychology, the traditional borders between cognitive psychology and neighboring areas like social psychology or developmental psychology are clearly disappearing slowly, which has led to new descriptive terms like "**cognitive social neuroscience**" or "**the social neurosciences**."

Thirdly, cognitive psychology has rediscovered the concept of the will, although it is now referred to as "**cognitive control**" or "**executive functions**." The increase in our knowledge of the frontal cortex (mainly through studying patients with lesions in this brain area) and the increasing interest in the question of how people are able to perform various tasks with the same environmental stimuli and actions and can construct and implement different strategies has focused the research interest on processes that precede the processing of environmental information. This has led to a clear widening of the identity of cognitive psychology, which had traditionally focused on the processes that occur between the registration of a stimulus and the evaluation thereof through high-level cognitive operations. However, if these operations are not independent of the context and action goals of the person performing them, as is suggested increasingly by research, then the connection between action and the cognitive processes on which these actions depend should also be entered into the theoretical analysis. In other words, pure perceptional, attentional, and memory theories are becoming increasingly more complicated and comprehensive models, that take the action-specific function of cognitive processes into account.

1.2 Deficits in Theory and Research

There are a number of reasons why the psychology of action and movement has been so slow to emerge. We have already mentioned one reason, namely that many disciplines have concerned themselves with this theme, without actually resulting in fruitful interdisciplinary and integrative approaches. However, the various **research traditions** are responsible too, as they have often artificially narrowed the view of the conditions of action control (Hommel et al. 2001).

1.2.1 Sensorimotor Approach

Descartes (1664) instigated a very important and influential research tradition which asserted that **actions** are, so to speak, the **continuation of perception through different means**. As can be seen in Descartes' sketch, perception (in this case visual) leads incoming information to a central cortical coordinating point, where a fitting response is selected and initiated through control of the necessary muscles (Fig. 1.1). Descartes described **three types of processes**, which are important for the control of actions:

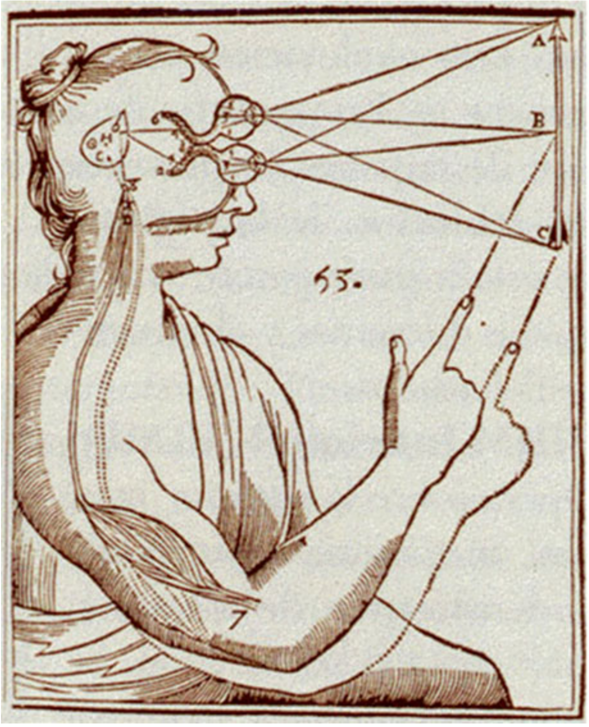


Fig. 1.1 Descartes's view on the relation between perception and action (Descartes 1664)

- **Afferent processes**, through which information that enters through the sensory organs is passed to the central organ.
- **Efferent processes**, through which movement commands from the central organ are passed to the musculature of the body's periphery.
- **Central processes**, which generate efferent commands on the basis of afferent information.

Descartes pictured afferent processes as thin threads, which are put in motion by stimulus information that hits the sensory organs, and that connect the specific sensory organ and the brain. Through the movement of these threads, the information from the sensory organs is carried to the pineal gland in the brain—the assumed perception-action interface. There, central exchange between afferent and efferent processes takes place, according to Descartes, because the pineal gland, which itself has been put in motion through the motion of the threads, secretes nervous fluid from its surface, which causes muscle contractions on the efferent side.

This sensorimotor conceptual model remains the theoretical basis of many approaches in contemporary cognitive psychology. Donders (1868/1969) made a significant step forward by proposing to dissect the general processing pathway into **subprocesses**, ranging from the earliest perceptual processes all the way to the subsequent movement, and to attempt to measure the duration of each of these subprocesses. Donders identified no less than 12 processing steps, from the influence of environmental energy on sensory receptors to the overcoming of corporeal inertia through the activation of muscles. Condensing these 12 steps reveals **four processing categories**, which until today remain the basis for stage models of information processing: sensory (pre-)processing, stimulus identification, response selection, and response initiation.

These types of **stimulus-centered linear stage models** are well suited to illustrate the setup of a typical psychological experiment. In such an experiment, a stimulus is generally used to manipulate the independent variable(s) while a given action is signaled, or the preparation of a given action is required. Processes related to an action are then, in fact, a function of stimulus presentation, and one can ascertain which processing stage is particularly affected by the manipulation of the independent variable. **Behavioristic approaches** have also emphasized the crucial role of the stimulus in the description of action control, although interactions with previous learning experiences are also important in these accounts. However, psychological experiments aside, people hardly ever wait for stimulus signals to make decisions on how to act; to the contrary, they often consider stimulus events to be the consequence of such decisions—just like the letters of this sentence were actively generated by the authors of this book. In fact, many actions are planned in the absence of external stimuli, which is less easy to capture in a conventional stage model.

1.2.2 Ideomotor Approach

The **ideomotor approach** has a long and varied history (Stock and Stock 2004), but is associated mainly with the names Lotze (1852), Carpenter (1852), and James (1890). A seemingly simple question underlies this approach: how can we, on the one hand, carry out arbitrary, goal-directed actions, but on the other hand know very little about how we actually do so?

Ask yourself *how* you actually ride a bicycle, or *how* you tie your shoelaces. Can you really answer that question spontaneously? Or do you imagine these actions first and then describe what you imagine? If the latter applies to you, then you experience a phenomenon that was referred to as **executive ignorance** by Lotze (cf. Turvey 1977). How one is capable of intentional action, notwithstanding this failing insight in one's own motor functioning, is discussed in greater detail later (Chap. 3).

In brief, the ideomotor theory suggests that intentional action presupposes **knowledge** about what one can attain with a given action; that is, which **action effects** can be obtained with a given action. The choice of an action therefore follows on the basis of a comparison between the expected action effect and the desired action goal: when one wants to tie one's shoelaces, one selects those actions that are expected to result in tied shoelaces.

From an ideomotor perspective, the theoretical analysis of action control does not start with a stimulus that precedes an action, but with an interplay between an intention and actions that support that intention, or properties of those actions. It is thus not external stimulus events that cause actions but the anticipation of their outcome: actions serve for the production of events (action effects), which are perceived and evaluated in light of the current intention. Ideomotor approaches study how this works, but tend to neglect the origin of action intentions and how actions are informed by, and adjusted to environmental conditions.

1.2.3 Interplay of Perception and Action

Sensorimotor and ideomotor approaches clearly reflect their origins. **Sensorimotor approaches** originate from neurophysiology and use the reflex arc as a guiding principle: just as a sudden blast of air induces a blink response of the eyelids, the sight of an undone shoe makes us perform shoelace-tying actions. **Ideomotor approaches**, in turn, stem from the era of introspective psychology and therefore deal with the connection between processes of consciousness, such as the experience of an intention and the conscious execution of an action outcome. However, this is not the only difference. Both approaches concern themselves with two separate halves of what is in fact one **perception-action cycle**.

Various authors have pointed out how strongly perception and action are intertwined. Von Uexküll's (1921) **concept of the subjective environment** encompasses both the perceivable properties of environmental events (*die Merkwelt*, or perceivable world) as well as the activities one can perform with them (*die Wirkwelt*, the operational world). In von Uexküll's model, sensory receptors register properties from the environment and pass them to a perceptual organ, which changes the environmental properties with the help of an operational organ. Neisser (1976) uses a very similar description of a circular relationship between three processes: internal knowledge schemata control the goal-directed exploration of the environment. This leads to the perception of the properties of objects, which either confirm the schema or adapt it (to reality). Neisser considers perception to be the result of a **continuous cycle** from the registration of environmental information, the integration of this information into object schemata, the goal-directed exploration which is controlled by those schemata and leads to more incoming information, and so on and so forth. Perception is also active, because generally, perceivable information generates goal-directed actions. Finally, because eye movements are crucial to visual perception, and hand motions are crucial to tactile perception, actions are also **receptive**, to the extent that they allow for new insights into the world. In the light of this relationship, it might be better to stop referring to perception and action, but to refer to the receptive and productive functions or aspects of human behavior.

In any case, it should be clear that sensorimotor and ideomotor approaches to action control do not illustrate the complex interplay between perception and action comprehensively. Sensorimotor approaches emphasize the influence of environmental factors on action production and generally ignore the antecedents or prerequisites of purposive or voluntary actions that necessitate generating and integrating representations that capture the demands of the task at hand and the means to satisfy them (forming and implementing intentions, planning actions, etc.), which makes actions appear to be stimulus-driven responses. In contrast, ideomotor approaches emphasize the intentional aspects of action control and thereby latently underestimate the contributions of the environment. Due to these blind spots in the theories discussed here, one could consider these approaches to be **complementary perspectives**. Recent years have seen an accumulation of attempts to integrate stimulus-oriented and intentional approaches into more complicated models.

1.2.4 Homunculi

A further problem for research into action control comes from the tendency to build **homunculi** ("little men") into theories. In the first, introspection-based ideomotor approaches from Lotze and James, **the will** is considered to be the decisive force behind goal-directed actions. How exactly this organ comes to its decisions and how it transforms decisions into actions is hardly discussed and is not really reflected by theory. James (1890) explains this reluctance through the self-explanatory nature of the phenomenon: "Desire, wish, will, are states of mind which everyone knows, and which no definition can make plainer" (p. 486). However, modern theories, too, have often done little more than replace the old-fashioned concept of the "will" with more familiar, technologically inspired expressions, without actually defining them, or lending them any theoretical support. For example, Baddeley and Hitchs' (1974) central executive or Norman and Shallice's (1986) attentional supervisory system (Chap. 9) are little more than placeholders for organizational processes that we still do not fully comprehend (Baddeley 1986).

Mysterious, homunculoid systems of this kind often emerge through the common tendency to objectify. Take, for example, the observation that humans do not follow every action tendency: we do not do everything that suits our fancy, do not eat everything that tastes good, and do not buy everything we would like to. The thought of an action can bring about foresaid action, but it definitely does not always do so. There are various theoretical options to account for this phenomenon. According to James (1890), thoughts will only lead to action when they are not in competition with other thoughts (e.g., the desire to eat healthier food) and when they are accompanied by an impulse to act ("fiat"). Freud (1923) claimed that undesirable thoughts could be suppressed actively, which would again imply a suppression mechanism. Although this second solution is more complicated from a theoretical perspective and posits a more intelligent mechanism that is harder to explain (how does it know when to suppress? How does it do that?), it appears to be so evident to many authors that they do not seem to look for and test other solutions (MacLeod 2007). This tendency is not limited to inhibitory models, but can be seen in all areas of cognitive psychology (and science in general): a phenomenon or behavior is described and “explained” through postulating a (functional or neuroanatomical) system, which produces the exact same phenomenon or behavior. Willed behavior is produced by the system of the will, behavioral suppression by an inhibitory system, selective attention by an attentional system, and so on and so forth. It ought to be clear that, in these instances, sham explanations are provided, which lead to nothing as long as the proposed systems are not analyzed further.

1.3 An Organizing Working Model

The psychological study of action is at a turning point, in which small, partial theories are increasingly integrated into comprehensive models, and borders between sub-disciplines fade away. Clearly, this complicates understanding the field of study. We would therefore like to propose a comprehensive descriptive working model, which should serve to structure the discussion of research results and theoretical concepts in this book (Fig. 1.2; the numbers in the figure refer to the chapter in which the concerned theme is treated). This working model distinguishes several layers that organize the processes of action control regarding their duration and scope.

On the bottommost layer, we find processes that we have already encountered in the context of the sensorimotor approach. Processes that are associated with stimulus processing are summarized with the term **perception**. We do not limit this term to conscious perception, but use it for every type of action-related stimulus processing.

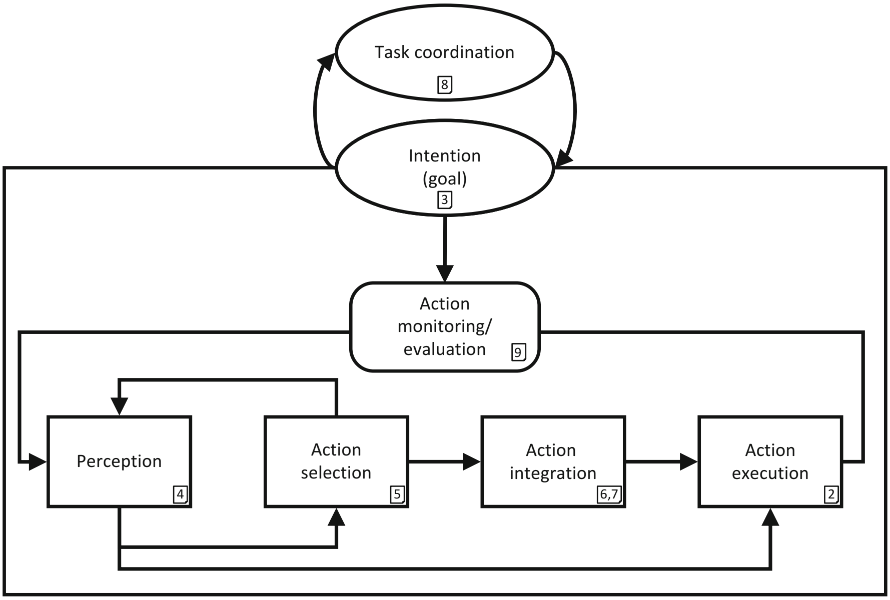


Fig. 1.2 Organizing working model

The outcomes of these processes are used both to select and to adapt actions (Chap. 4). Chapter 5 is devoted to processes that **select actions** and determine the properties of an intended action. As depicted in Fig. 1.2, they work in close coordination with perception. The next step consists of the **planning** of actions through the **integration** of action features (Chap. 6). It shall be made clear that action representations are not autonomous units, but networks of many different perceptual and action-related codes that, at least in some cases, should be integrated so that an action can be performed. Integrative processes are particularly important when complicated **action sequences** are being planned, such as the preparation of a meal (Chap. 7). As soon as a plan of action is finished, it can be executed. The **performance** requires the **translation** from a cognitive representation of a given action into muscle activity (Chap. 2).

The processes on the lowest performance level have a relatively short duration: once a particular action is set, perception and action specification can be devoted to other tasks. Once an action is performed, other actions can be integrated and performed. The scope of these processes is relatively clear too: they will mainly interact with the processes which cause the effects necessary for the first processes, or with processes that need input from these lower processes. The process of **action control** on the next, somewhat higher middle level, generally has a longer duration and takes more information into consideration; therefore, it is more integrative. It **monitors** whether actions that are about to finish are actually in accord with the actual intention, that is, with the action goal (Chap. 9). **Intentions** organize and instrument the processes of perception and action planning in a way that allow, if all is well, the realization of a set goal (Chap. 3). Finally, in Chap. 8 we discuss how **action goals** themselves are controlled and implemented. Everyday actions often require simultaneously pursuing various goals, also known as **multitasking**, and **switching between goals**. This raises the question of how various functions are actually coordinated. However, let us first look into the most important **neurobiological foundations** of human action control.

References

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *The psychology of learning and motivation* (pp. 47–89). New York: Academic.

Carpenter, W. B. (1852). On the influence of suggestion in modifying and directing muscular movement, independently of volition. *Proceedings of the Royal Institution* (pp. 147–154).

Descartes, R. (1664). *Traité de l'homme*. Paris: Girard.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431. (Original work published 1868)

Freud, S. (1923). *Das Ich und das Es*. Leipzig, Germany: Internationaler Psychoanalytischer Verlag.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC). A framework for perception and action. *Behavioral & Brain Sciences, 24*, 849–937.

James, W. (1890). *The principles of psychology*. New York: Holt.

Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig, Germany: Weidmann.

MacLeod, C. M. (2007). The concept of inhibition in cognition. In D. S. Gorfein & C. M. MacLeod (Eds.), *Inhibition in cognition* (pp. 3–23). Washington, DC: American Psychological Association.

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

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (Vol. 4, pp. 1–17). New York: Plenum.

Rosenbaum, D. A. (2005). The Cinderella of psychology: the neglect of motor control in the science of mental life and behavior. *American Psychologist, 60*, 308–317.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research, 68*, 176–188.

Thorndike, E. L. (1898). What is a psychical fact? *Psychological Review, 5*, 645–650.

Turvey, M. T. (1977). Preliminaries to a theory of action with reference to vision. In R. Shaw & J. Bransford (Eds.), *Perceiving, acting, and knowing. Toward an ecological psychology* (pp. 211–265). Hillsdale, MI: Erlbaum.

Von Uexküll, J. (1921). *Umwelt und Innenwelt der Tiere*. Berlin, Germany: Springer.

Woodworth, R. S. (1938). *Experimental psychology*. New York: Holt, Rinehart & Winston.

Chapter 2
Neurobiological Foundations of Action Planning and Execution

In the mid-1800s, Phineas Gage worked on the construction of railroads in the United States, and his job as a foreman was to get rid of annoying masses of stone by blowing them up with explosives. Unfortunately, during one explosion, his **frontal cortex** was pierced by a chisel. In 1868, the physician John M. Harlow describes in great detail how Gage's serious wound was treated successfully, and how he started to work again after a few months. However, colleagues and superiors had to conclude that he was "no longer Gage": he lacked motivation, had difficulty in making plans, and showed strong **personality changes**, which were not in his benefit. Nevertheless, he was still able to work, and so he took a job in a horse stable; however, he found it increasingly difficult to develop action plans and to translate them into appropriate actions. Harlow described Gage as a person who always made plans for future activities, only to abandon them and to replace them with other, apparently better plans.

The analysis of the case, and especially the skull of Phineas Gage, has made substantial contributions to our **understanding** of the interplay between the human brain, cognitive processes, and action control (cf. Sect. 2.6.2). In fact, the actual performance of cognitive functions is usually understood best when they cease to exist for some reason, be it through lack of exercise, natural aging, illness, or accidents. This does not just apply to perception and memory, but also to the planning of actions and action control.

Particularly interesting in this context, are patients who show deficits in the planning or execution of actions, for example, as a result of **brain lesions**. This is interesting because the failure of control of action in patients with specific, accurately described lesions in the brain can give us preliminary insights into which brain areas are involved with action control. Additionally, results from **physiological animal experiments** and **neuroimaging methods** (Box 2.4) have contributed to a better understanding of the neuronal foundation of human action control. Although the mapping and understanding of the neuronal basis of the processes of action planning and action control is currently not as detailed as, say, that of the visual cortex,

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_2

it is becoming clear that successful planning, initiation, and execution of actions require an intact functional loop, which encompasses the **frontal cortex**, the **premotor** and **motor cortices**, the **basal ganglia**, and the **cerebellum**. All of these areas (as well as many others that will not be treated here, from a didactic viewpoint) make specific contributions to action control.

When we attempt to describe the most important contributions of these areas in the text that follows, we should not forget that it is the interplay and integration of these areas that produces effective actions. The performance of a given brain area must always be viewed in conjunction with the **functional loop** to which it contributes. Therefore, we do not have the intention to present a comprehensive overview of neuroscientific research into human action control. Instead, we would merely like to point out some properties of neuronal information processing that have direct consequences for a psychological understanding of action control, and to discuss the actual functions of the neuroanatomical functions that are important for action control.

To aid in orientation and get a grasp of where in the human brain the areas we will discuss are located, we can use a **map of the brain** which was published by the German neurologist Korbinian Brodmann in 1909 (Fig. 2.1). On the basis of his cyto-architectonic studies, Brodmann divided the cortex into 52 areas, known today as **Brodmann's areas (BA)**. A series of these areas is generally considered to have functionally separate roles in cerebral information processing. In the next section, we will first have a look at the question of how these different anatomical areas actually communicate with each other.

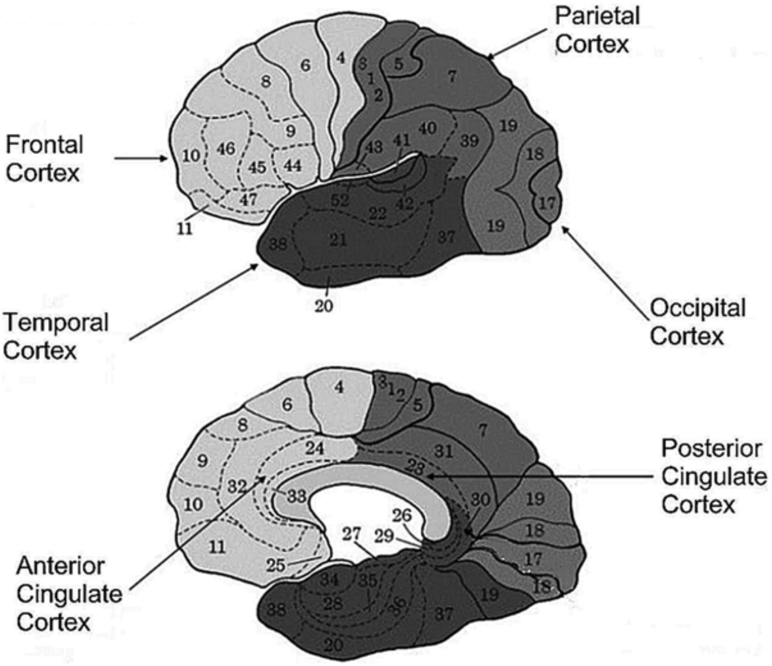


Fig. 2.1 Brodmann's (1909) map of the human brain

2.1 Neuronal Communication

The smallest functional unit of the brain is the **nerve cell** or the **neuron**. There are about 100 billion of them in a human brain. The number of neurons generally remains fairly constant from birth to way beyond the 65th year of age. A neuron has a cell body (soma) with relatively short projections (dendrites), which take up information from other neurons as input, and lead it to the soma. A neuron also has a relatively long projection (the axon), which conducts electrical impulses from the soma to the dendrites of other neurons. The site where the axon of a neuron comes into contact with the dendrite of another neuron is called a **synapse**. When the electrical impulse which is conducted through the axon exceeds a certain threshold, a chemical messenger (called a neurotransmitter) is released at the end of the axon. When the neurotransmitter at the synapse contacts the dendrite of the second and other proximal neurons, the electrical impulse is carried on by the second and a host of other neurons.

Single neurons appear to be highly specialized. This is suggested by studies in which extremely thin microelectrodes are inserted into the brains of animals. When the soma of an active neuron lies close to the electrode’s tip, the minute electrical potentials that are caused by the activity are registered by the electrode. This signal is then amplified acoustically, for example, so that the activity of the neuron can be heard: the stronger the electrical activation of the neuron, the louder the noise. When visual or auditory stimuli are now presented to the animal, it becomes clear that single neurons are tuned to process highly specific information: some cells respond exclusively to the shape or orientation of objects, some exclusively to visible motion in a specific direction. Some cells in the auditory cortex respond to tones of a specific frequency, others to tones of a specific volume, and others yet respond to tones that change their frequency and get higher or lower. Other cells respond to faces, some to specific faces, others to all faces that are oriented in a specific direction. Then there are cells which are active when an animal moves in a particular way, but also when the animal observes that movement in another animal. So, when a neuron is confronted with the specific stimulus information it is sensitive to, it reacts with an **increase in activity** and signals in that way, that at that very moment, highly specific information is available, for example, the movement of an object in a given direction. The only thing this specific neuron “knows” is that, for example, something moves in the specific direction for which it is specialized, that is, it operates completely **feature specific**. It merely codes this one feature, without “knowing” anything about the other features of this object that is moving. It has no information about the object’s shape, color, size, or identity, that is, about features that are coded in other cortical brain areas, which are often located relatively far away and are occasionally organized in a different manner (so-called distributed coding of features).

The **principle of distributed coding** of features is very applicable to the processing of visual information: the various features of visual stimuli are coded into different cortical color, form, orienting, and motion maps (DeYoe and van Essen 1988).

Apparently, it also applies to coding the various features of action. It has been shown in monkeys, for example, that the direction, expenditure of energy, and range of a movement are coded in a distributed manner; in humans, comparable indications for the duration, expenditure of energy, and the effector, with which a movement is executed (see review by Hommel and Elsner 2009). The principle of distributed coding in separate modules offers a range of **evolutionary benefits**: phylogenetically, it allows for a continuous adaptation and the steady expansion of the brain, as individual modules can be modified, added, or eliminated, without the entire brain having to be "rebuilt." Ontogenetically, it gives rise to a comparable measure of tolerance to damage of the brain (occasionally reversible due to plasticity of the brain), which often manifests itself as the loss of subprocesses, which does not necessarily interfere with the functioning of the entire cortex (see Box 2.1).

Box 2.1 Plasticity of the Brain

How exactly do specific cortical areas come to be and how do neuronal networks are connected in the way they are? The fact that our brains resemble both those of other humans and those of primates quite well suggests that genetic wiring plans play a large part in the development of the brain. However, genes do not determine everything because the structure of our brain is largely dependent on experiences. This is demonstrated by classical experiments of the Nobel prize winners David Hubel and Torsten Wiesel, who bandaged up one eye of kittens before they came into contact with direct light (Hubel and Wiesel 1963). These kittens could do everything they liked, but could only use one eye to do so. After multiple months, the scientists took the bandages off, and assessed the neuronal connections between the two eyes and the brain. The surprising result was that the bandaged eye, although **optically intact**, was not connected with the visual areas of the brain. It was **functionally blind**. Clearly, under these circumstances the neurons had formed networks in such a manner that connections had only established between the retinal cells of the seeing eye and the visual cortex.

These early tests with animals have illustrated an important fact: neurons do not form networks based on a fixed blueprint, but based on a **functional, activity-dependent scheme**. Which connections are established, must be at least partially fixed in genetic codes; for example, retinal cells in the eye only connect with cells in the visual cortex in the occipital lobe of the brain, and not with neurons in other parts of the brain, such as the motor cortex. But, apart from that, neuronal networks are generally plastic and flexible, and they adapt continually to the organism and his or her activities through modification, installation, and elimination of connections. Numerous clinical and experimental studies show that this does not just apply to the developing brains of babies and children, but to those of fully developed adults too.

(continued)

Box 2.1 (continued)

It can be shown in animal experiments that loss of specific neuronal populations can be **compensated** in a surprisingly short time, as other neuronal populations take over the functions of lost populations. Sanes et al. (1992) bisected the nerve that innervates the musculature of whiskers in rats. This led to a loss of the neuronal populations in the primary motor cortex that are responsible for control of the whiskers. Within hours, the neuronal network that controls movements of the facial muscles was reorganized in such a manner that neurons in neighboring areas of the motor cortex replaced the lost neuronal populations.

Pascual-Leone et al. (1993) have shown that the size of the finger area in the motor cortex varies depending on activity level: while the finger areas of blind persons who have little expertise in reading Braille are roughly the same size for both hands, in blind Braille experts, the cortical area that represents the "reading" finger is larger than the corresponding area for the finger of the other hand. Complementary to such observations, which indicate that neuronal representations in the motor cortex can **expand** activity-dependently, it has also been demonstrated that motor areas in the brain **decrease** in size when abilities to move are restricted either temporarily or for a longer period of time. Liepert et al. (1995) have studied patients whose ankles were restricted in freedom of movement, without a peripheral nerve lesion being present. They found that the motor areas responsible for control of the damaged ankle were smaller than the corresponding areas responsible for the non-afflicted ankle.

These observations indicate that neuronal representations are plastic and can adapt flexibly to the circumstances and activities of the organism. In which **temporal dimensions** such adaption processes can be completed, can be studied with experiments, in which participants acquire motor skills. Pascual-Leone et al. (1995) had their participants execute movement sequences of five fingers on the keys of a piano over the course of 5 days, and they analyzed the changes in hand representations in the motor cortex. The spatial expansion of the hand area increased as expertise in executing the movement sequences increased. That this growth was actually caused by the acquisition of a skill and not, for example, a random side effect of merely moving the fingers of one hand repetitively was demonstrated by the observation that isolated, non-sequential movements of the fingers were not associated with expansion of the hand area.

The plasticity of the human brain is also demonstrated in the often remarkable successes in the rehabilitation of stroke victims. **Strokes** (**cerebrovascular accidents**) are caused most often by a lack of perfusion in the brain following the obstruction of blood vessels and less often by a hemorrhage (e.g., following an accident). A lack of perfusion leads to a disruption in the oxygen supply to the brain, which results in the death of many nerve cells in the brain. The consequences can be motor disabilities in, for example, the arm, hand, leg or feet on one side of the body as well as loss of speech.

(continued)

Box 2.1 (continued)

These disabilities often restrict a victim's performance of everyday chores in the long term. Almost all of the victim's **motor activities** like, for example, opening doors, getting dressed, reading the newspaper, brushing teeth, and playing cards can often only be performed with the unaffected arm.

Additionally, many patients suffer from stroke-induced **speech disorders** (aphasias) which can be manifested during writing, reading, comprehending, or speaking. Aphasias are caused by damage to the neuronal populations which participate in speech production (Broca's speech area; BA44 and BA45) and/or comprehension of speech (Wernicke's area; BA42 and BA22). Damage to Broca's area mainly leads to problems with speech production accompanied by generally intact speech comprehension, while damage to Wernicke's area is characterized by generally intact speech production while speech comprehension processes are interfered with (review by Kolb and Whishaw 1996). During the rehabilitation of such disturbances, surprising improvements can often be attained. The ability to speak, for example, can be recovered, or paralyses can disappear almost completely. It is the plasticity of the human brain, its ability to adapt and modify neuronal structures continuously, so that neurons in other brain regions can manage to take over the functions of damaged areas (Hallett 2001), that makes these remarkable types of recovery possible.

Problems arise when this system represents multiple different features that are represented **at the same time**—which is always the case in daily life. In these cases, the problem is to distinguish which features belong to which perceptual and action events. To illustrate the problem, imagine a table with two pieces of fruit on it: on your left, quite close to you, is a green, not entirely ripe apple; to the right, a bit farther away, is a red strawberry. Now imagine you want to grasp both fruits at the same time, the strawberry with your right hand, to eat it, and the apple with your left hand, to put it away. How is this scenario represented neuronally? Probably like this: The information that emanates from both fruits activates a great number of neurons, which signal, for example, that the following features are available: *red, green, left, right, large, small, nearby, faraway, sweet, sour*. What we perceive, however, is not a bunch of separate, unconnected features, but a coherent whole, namely a red strawberry, which is on a table in front of us, to the right of a somewhat unripe apple. The preparation of the movements of the two hands also contains a number of feature-based codes, such as *left, right, nearby, faraway*, and many more. Movements, too, are not represented phenomenologically as single features or elements, but as coherent events, namely as the action plan that the right hand will grasp the strawberry, say, while the left hand will grab the apple.

Now, how can a system which is based on the principle of distributed representations distinguish which activated codes belong to which perceptual or action event? Do the features *red, right,* and *small* belong to the same fruit? Do the movement features *right* (for the hand) and *faraway* (for the amplitude of the movement) go together? Should the right hand make the larger and the left hand the smaller movement, or should it be the other way around? The solution to the problem probably requires the integration or binding of related cognitive or cortical feature codes (Singer 1994). How might this binding work?

A simple **solution** would be presented if the brain would have an area where the codes which are represented in a distributed manner would be collected and assembled; in other words, a center like the pineal gland from the Cartesian tradition, to which Descartes ascribed the function of central processing of afferent and efferent processes. However, such a **center** does not exist in the brains of humans and other higher species.

Another solution to the binding problem is based on the idea that spatially distributed populations of neurons, which encode different information, can **communicate** with one another. Individual neurons communicate with a great number of other neurons and form so-called **functional networks**. This proceeds via synapses, through which the axon of a given neuron makes contact with the dendrites of other neurons. At birth, every neuron has about 2500 synapses. In the first 3 years of life, their number increases massively (up to about 15,000 synapses per neuron), only to return to the numbers common in the adult brain (10,000–20,000); this happens somewhere between the tenth year of life and puberty (*synaptic pruning,* Huttenlocher 1994). Therefore, our brain consists of an incredibly complicated network of nerve cells, which are each in direct contact with thousands of other nerve cells through synapses. Most synapses are excitatory in nature (meaning they carry on the stimulation); some are inhibitory and preclude an uncontrolled stimulation in the cluster of neurons.

In recent years, a (however still controversial solution) to the binding problem is being discussed increasingly. Von der Malsburg (1995) proposed that spatially distributed neuronal populations, which encode separate aspects of the same stimulus, **synchronize their discharge patterns in time** and thereby signal which of the activated codes belong together and which codes do not. In fact, single-cell research with cats and monkeys has shown that neuronal populations in separate parts of the cortex, which are sometimes removed quite far from each other, do couple their activities in time. In monkeys, synchronized activity between the premotor and motor cortices and between neurons in the motor and somatosensory areas before initiating a finger movement was found. In cats, temporal synchronization between neurons of the visual and the parietal cortex as well as between neurons in the parietal and the motor cortex was found.

In humans, such temporal synchronizations of neuronal clusters can be measured with electro-encephalogram (EEG). The temporal couplings between distributed neuronal clusters that are reported in animal research are accompanied by **oscillations** in the beta (13–20 Hz) and/or gamma bands (30–80 Hz) and these oscillations can be extracted from the EEG frequency spectrum by using so-called wavelet-analyses. In such experiments, it can be shown that EEG oscillations can occur in relation with both perceptual and action processes. For example, Tallon-Baudry and Bertrand (1999) found an increase of oscillatory activity in the gamma range when their participants observed visual stimulus configurations. EEG oscillations also occur in connection with actions. Pfurtscheller et al. (1994) found gamma oscillations directly prior to the onset of movement of the left or right index finger, the right toe, or the tongue; these oscillations were found in the somatosensory cortex, where these body parts are represented (see Sect. 2.2). In short, ballistic movements (i.e., short, rapid movements which cannot be interrupted), the oscillatory activity starts directly before the execution of a movement and ends with the onset of the movement. In slow, guided movements, the oscillations can last throughout the execution of the movement (Kristeva-Feige et al. 1993).

Now that we have sketched how the brain is built and how it works, we will turn to the question of which cortical and subcortical structures are **involved** in the planning and execution of action and which **part** they play (Fig. 2.2). We will see that the prefrontal cortex is always involved when we act in a goal-directed manner. The neuronal networks of the primary motor and lateral premotor cortices are responsible for the execution of movements. The supplementary motor area (SMA) is concerned with the planning of actions and the sequencing of single action elements. The dorsolateral prefrontal cortex (DLPFC) represents the goal of an action and is responsible for the activation, implementation, and configuration of executive control processes which coordinate our actions and adapt them to changing conditions. The anterior cingulate cortex (ACC) monitors our actions and their success, and signals the DLPFC when updating of action goals would be beneficial. The selection of actions in an interplay with the DLPFC considers expected rewards. These are computed or made available by the orbitofrontal cortex (OFC). We will also see that subcortical structures play a decisive role in the control of actions: Expected

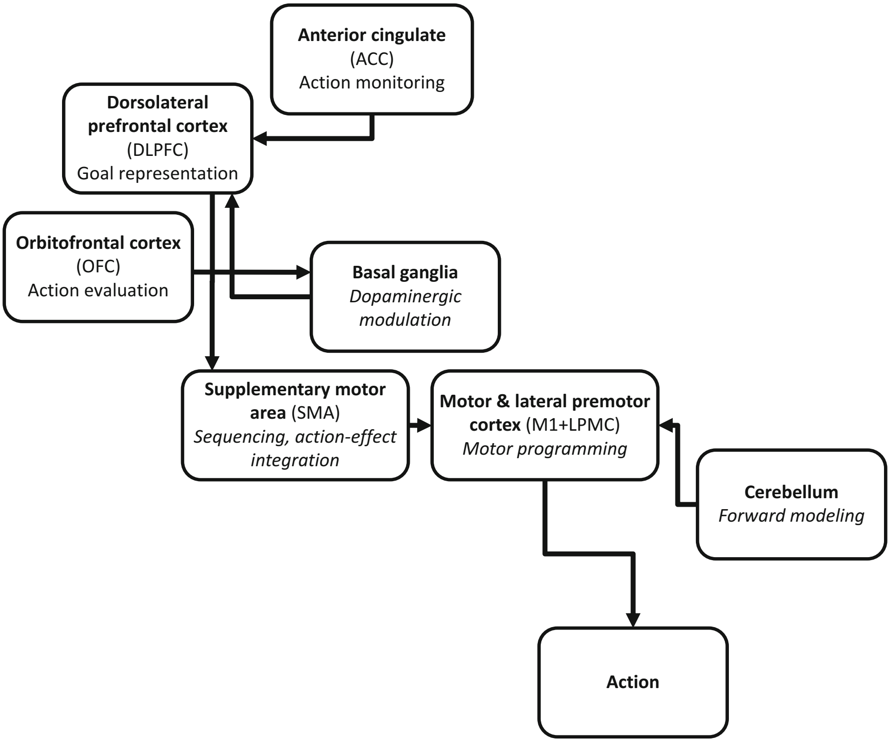


Fig. 2.2 Overview of the major contributions of various cortical and subcortical structures to the control of human actions

rewards influence the dopamine production in the basal ganglia, which modulate the workings of the DLPFC. Finally, accurate and fluid execution of movements depends on intact structures of the cerebellum (small brain), which monitors the success of concrete movement elements on the basis of forward models.

2.2 Primary Motor Cortex and Lateral Premotor Cortex (BA4/6)

Near the end of the nineteenth century, the German neurologists Gustav Fritsch and Eduard Hitzig discovered that electrical stimulation of a dog's cortex caused muscle contractions on the contralateral side: When the right cortex was stimulated, muscles on the left side of the body moved, and stimulation of the left cortex led to muscle contractions on the right side of the body. At approximately the same time, the English neurologist Hughlings Jackson discovered that epileptic insults are caused by lesions in the cortical motor area in the large brains. In the beginning of the twentieth century, the English neuropsychologist Charles Sherrington demonstrated that it is easiest to induce muscle contractions in monkeys when the electrodes are placed in the gyrus praecentralis in either of the hemispheres of the large brain. This area is today known as the **primary motor cortex (M1)** (Fig. 2.3).

M1 is found in the central areas of the two hemispheres (BA4) and borders on the sensory areas (i.e., on the somatosensory cortex). Numerous observations suggest that M1 is an important **coordinating point where cognition and motor activity join** and that its functioning is crucial for the execution of **movements**. For example,

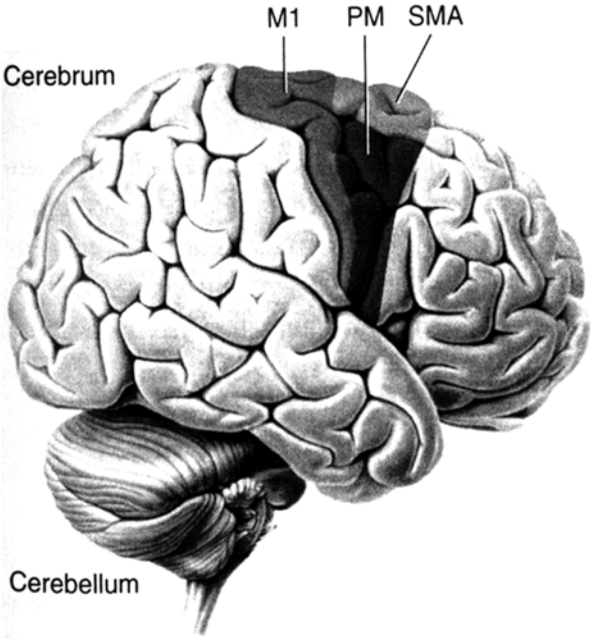


Fig. 2.3 Motor areas of the human cortex (adapted from Konczak 2008; with permission from Spektrum Akademischer Verlag)

patients with damage in just the motor areas have little difficulty with remembering action goals, maintain them, or switch between them. However, they have massive problems with executing bodily movements to realize these goals successfully. Depending on which hemisphere of the brain is afflicted, lesser (paresis) or more severe (plegia) paralyses of the limbs on the **contralateral side of the body** are manifested. If the left hemisphere is damaged, effectors on the right side of the body are paralyzed and vice versa: Effectors on the left side are paralyzed when brain damage is localized to the right hemisphere. Therefore, each half of the motor cortex controls the contralateral effectors. (This applies, at least, to parts of the facial mimic and the hands, but not for control of the movements of the feet.)

2.2.1 Motor Homunculus

In the 1930s, the Canadian neurosurgeon Wilder Penfield started to electrically stimulate various areas of the cortex of patients, whose skulls were opened for surgery. As there are no touch and pain cells in the brain, this process was painless for the patients. Penfield discovered that stimulation of the gyrus postcentralis induced **tactile sensations** in specific parts of the body. Furthermore, he found that these stimulation points in the brain were not scattered haphazardly throughout the brain, but were actually organized according to a **systematic map**. He also concluded that such a map was not present for the surface of the body. In a neighboring area of the brain, in the gyrus praecentralis, a similar map exists, which represents the skeletal musculature of the body. Depending on which part of this map he stimulated, contractions of specific muscle groups occurred. When he stimulated regions near the very top near the central sulcus (furrow) that separates the two brain hemispheres (medially), contractions of muscles in the contralateral leg were induced, while stimulation more to the sides, in the lateral motor cortex, resulted in movements of the hands or facial musculature. The systematic mapping of the primary motor cortex demonstrated that this brain region entails a somatotopic map of the complete skeletal musculature. This map was named the **motor homunculus** (cf. Sect. 1.2.4), which lies opposite to its sensory equivalent (the somatosensory homunculus on the other side of the central sulcus).

As is shown in Fig. 2.4, this representation is strongly distorted. Particularly important parts of the motion apparatus like the hands and mouth are strongly overrepresented while other parts like the trunk are strongly underrepresented. This is probably because the size of the cortical fields is not associated with the size of the innervated muscles, but with the complexity of the motor functions we have at our disposal. This would explain how the hand, with which a large number of different operations can be performed, can be represented by a much larger area than the foot, which is stereotypically used for locomotion only. These kinds of **somatotopic maps** are also located in cortical areas directly anterior to M1. The medially located area near the central sulcus that separates the two hemispheres was designated **supplementary motor area (SMA)** by Penfield (medial BA6, cf. Sect. 2.3); the area lateral to that, **premotor cortex (PM)** (lateral BA6). PM and M1 cooperate intensely, and much information that M1 receives is modulated by the PM.

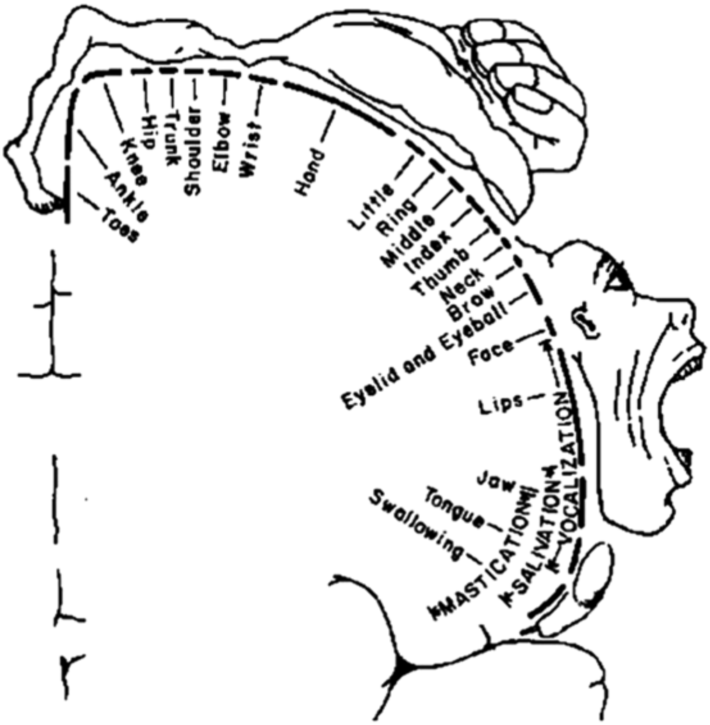


Fig. 2.4 Motor homunculus (adapted from Penfield & Rasmussen 1950, with permission from Gale, a part of Cengage Learning Inc.)

Which function could the motor homunculus have? What do the neurons that constitute it represent? The classical answer to this question is it represents the skeletal musculature of the body, and controls the activity of the various muscles of the bodily periphery. Therefore, the motor homunculus was seen as a type of a marionette's control bar to which the muscles are connected as with threads, and on which they will act like a puppeteer who operates a marionette. For this to work, there should be a 1:1 connection between, for example, M1 neurons to a specific group of muscle fibers. This is clearly not the case, however. Although it is possible to distinguish, for example, leg, hand, and face areas in a somatotopic map, there is no differentiation within these areas: it was impossible to demonstrate a somatotopic representation of the hand muscles when the hand area of primates was stimulated systematically (Schieber 1999). Therefore, repetitive stimulation of the same M1 neurons during various motor tasks activates various muscle fibers (Georgopoulos et al. 1999). Furthermore, various muscles can be represented in the same area in M1 and various areas in M1 can activate the same group of muscle fibers (e.g., Penfield and Boldrey 1937). Therefore, somatotopic maps in M1 do *not* appear to represent the skeletal musculature of the body and do not appear to be the address where control of the various muscular fiber groups takes place.

Box 2.2 Stimulating the Primary Motor and Premotor Cortex

Graziano et al. (2002) electrically stimulated various areas within the primary motor and premotor cortex of two monkeys. Unlike the classical studies by Penfield, in which stimulation of short durations was used (ca. 50 ms), these authors stimulated the neurons of the motor cortex for 500 ms. Instead of contractions of single muscles, they evoked **fluent, spatially and temporally well-coordinated movements that were directed at spatial goals** under these conditions. The stimulation of a specific area, for example, led the monkey to close its finger, move its hand to its head, and then open its mouth. This movement sequence occurred reliably, independent of the hand's position when stimulation was commenced. When neighboring areas were stimulated, the same motion sequence was initiated, but with one significant difference: depending on the site of the stimulation, the hand assumed **different goal positions** after finishing the movement, a little more below or farther away from the midline of the body.

So, the (longer-term) electrical stimulation of neuronal populations in the motor cortex and lateral premotor cortex appears to evoke relatively complex movements towards spatially specific goals. This might mean that the somatotopic maps of the motor cortex do not correspond to specific muscle groups, but to positions in space, that is, to potential targets of movements in the grasp or manipulation space close to the body. In fact, Graziano et al. found a very **close relationship** between the stimulated areas in the brain and the spatial targets of the movements evoked by the stimulation. They stimulated the motor cortex in the right hemisphere in eight different positions within the hand-arm area and found, again independent of the starting position, eight different end positions of the hand (Fig. 2.5; the circle on the sketched brain hemisphere shows the area in which the stimulation took place). The left hand of the monkey moved to a position in the upper, middle, or lower grasping space, either to the right side of the body (ipsilateral to the stimulated hemisphere), to the middle of the body, or to the left side of the body (contralateral to the stimulated hemisphere). Through further measurements and stimulation of the motor and premotor areas (around the areas marked by letters in Fig. 2.5), Graziano et al. were able to demonstrate that a whole series of complex movements actions is represented in **somatotopic maps**. While stimulation of one area evoked hand movements towards the middle of the body at chest height, combined with a precision grip, a fist, an open hand with spread fingers, or a rotation, stimulation of another area led to hand movements to the snout of the monkey, combined with a precision grip and the opening of the mouth.

These observations demonstrate that the neuronal populations of the lateral premotor and the primary motor cortex control complex and coordinated movements. It is particularly interesting that these movements kept occurring consistently and unaltered, even after hundreds of stimulations. Even when an obstruction was placed between the hand and the target position, the movement did not change, so that the hand would hit the obstruction and kept

(continued)

Box 2.2 (continued)

applying pressure to it for as long as the stimulation continued. The induced movements were even independent of what the monkey was doing: they occurred when it sat still, moved spontaneously, grasped a piece of a fruit, or was even anesthetized.

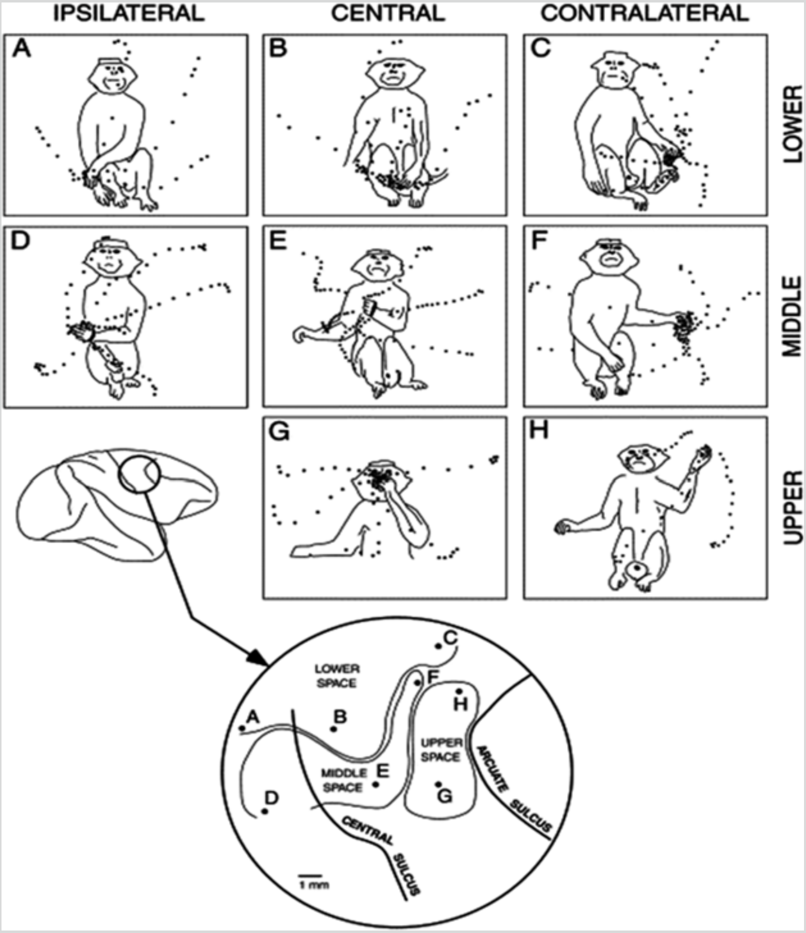


Fig. 2.5 Various hand postures induced by stimulating monkey motor cortex (adapted from Graziano et al. 2002, with permission from Elsevier)

So, if the neurons in M1 and PM do not control single muscles, what *do* they code? Which information do they make available? Research with primates suggests that they encode a sequence of **motor parameters** such as the direction and speed of movements, the position of joint angles, and muscle force (see above), and that they can **directly control** the course of more complex movements. The motor cortex and parts of the premotor cortex (which is also involved in the perception of movements, see Box 2.3) can be seen as the **final station in the action control of humans** and provide the actual muscular activation. These brain areas do not take part in the adaptive planning themselves: they eventually execute what has been planned by other cortical systems. One of the systems that is concerned with the planning of movements is the so-called SMA, to which we turn next.

Box 2.3 Mirror Neurons

The premotor cortex was typically ascribed functions related to action planning and control. This made it all the more surprising when neurons with sensorimotor properties were discovered in the premotor cortex of monkeys (di Pellegrino et al. 1992), which serve both perception and action. These so-called mirror neurons were not just active when a monkey performed a grasping movement itself, but also when it observed the same movement in a member of its species, or even in a human being. The activity of these neurons is highly specific and clearly related to action; these neurons are active only when a specific grasping motion with a specific target object is made (review by Rizzolatti and Craighero 2004). A comparable shared system for the performance and observation of movements appears to exist in humans too (Decety and Grezes 1999). Fadiga et al. (1995) have demonstrated that humans who observe others perform a given action, activate the same muscles they would use to perform that action. The neuronal network which is activated by the observation of actions in humans encompasses, besides the premotor cortex, parietal areas and the superior temporal sulcus (STS) (Grafton et al. 1996; Rizzolatti et al. 1996), but not the primary cortex and the SMA.

These observations have created quite a sensation; partially rightfully so, and partially not rightfully so. Unrightfully so in cases where the existence of mirror neurons is interpreted as an **explanation of a number of hard-to-explain phenomena**, such as imitation learning, empathy, or the understanding of fellow humans. For example, when the perception of another person's action activated one's own mirror neurons, this might explain why that action would be easy to imitate and why people often do so, possibly unconsciously. Of course, it is entirely possible that mirror neurons are involved in these cognitive performances, but their bare existence cannot be considered a satisfactory explanation.

(continued)

Box 2.3 (continued)

In fact, how can one’s own motor neurons actually represent another person’s visual action effects in one’s own motor functions, while the other person often has a completely different body structure, and must therefore use completely different muscles in a completely different way to perform the same action? In a certain sense, mirror neurons do not solve any theoretical problems, nor do they offer a better functional understanding of the mechanisms that lie at their foundation.

However, there is a number of reasons why the discovery of mirror neurons justifies the sensation it caused. One reason is that the existence of mirror neurons indicates the **close relation between action and perception**, a relation that is overlooked in many textbooks, not to mention research. When the pure observation of an action leads to the activation of motor areas, the question arises if or to what extent action-related knowledge influences the observation. That it does do so, is demonstrated by, for example, the fMRI study by Calvo-Merino et al. (2005). They studied professional ballet dancers, Capoeira dancers, and laymen who did not master either of these dancing styles, and presented sequences of ballet and Capoeira movements to all three groups. The neuronal mirror system was only activated during the observation of these movements when the observing person had mastered the dancing style himself. Further evidence for the role of one’s own motor expertise in the perception of movements is provided by the fMRI study of Grèzes et al. (2004). Here, participants saw videos of people (with blurred faces) who lifted a weight, and in some videos, the participants themselves were shown. The mirror systems were activated significantly earlier when participants observed their own movements in the video.

A further ground that makes the discovery of mirror neurons interesting has to do with the **relation between the perception of oneself and others**, which could receive a whole new theoretical meaning in the light of mirror neurons. Let us consider how mirror neurons can produce the relationship between an observed action and an action that is performed by oneself. One possibility is that we first obtain the systematic relationship between our own movement and the sensory consequences thereof (Sects. 2.3 and 2.4). We generally experience our own movements proprioceptively or kinesthetically and, to a certain extent, also visually. In general, perception is multimodal, however, and as children we quickly learn to generalize between modalities (Spelke 1976): we see what things feel like, and we feel what they look like. These two learning processes—the association of motor commands and sensory effects on the one hand, and generalization over modalities on the other hand—are sufficient to establish a mirror system that responds equally strongly to perceived and one’s own movements and that depends on one’s own experiences (see also Keysers and Perrett 2004).

2.3 Supplementary Motor Area (BA6 Medial)

The SMA is the medial part of the premotor cortex and is important for the **selection, planning,** and **sequencing** of goal-directed actions (Chaps. 5–7). The SMA also appears to be important for the perception of the intentionality of an action. Together with the lateral part of the premotor cortex, the SMA is one of the most important information sources for the primary motor cortex.

2.3.1 Role of the SMA in the Sequencing of Action Elements

Evidence for a decisive role of the SMA in the sequencing of movements is derived from **patient studies, fMRI and TMS studies in healthy participants, and animal research**. For example, patients with unilateral lesions in the SMA have deficits in the execution of sequential movements with the contralateral arm, or difficulties in reproducing rhythms from memory (Dick et al. 1986; Halsband et al. 1993). In healthy persons, the SMA is much more active during self-initiated—as compared to stimulus-induced—movements (Deiber et al. 1999), and disruption of the SMA through targeted transcranial magnetic stimulation (TMS)-pulses produces errors in the generation of complicated movement sequences (Pascual-Leone et al. 2000).

Tanji and Shima (1994) were able to demonstrate the existence of different **types of neurons** in the SMA of monkeys, neuron types that apparently encode different aspects of the sequencing of movements. The monkeys learned to execute various movements, and then had to string them together from memory in various successions. Three types of SMA neurons could be identified:

- Neurons that fired during the preparation of a movement sequence
- Neurons that were only active in the interval between two movements
- Neurons that appeared to represent the succession of single movements

In a follow-up study, Shima and Tanji (1998) temporarily disabled the neuronal population of the SMA pharmacologically, and found that the monkeys made more errors in the execution of movement sequences from memory under these circumstances, but not when the to-be-executed movements were signaled by visual cues.

2.3.2 Role of the SMA in the Integration of Actions

Besides its key role in the sequencing of movements, the SMA is also involved significantly with the **encoding** of intentions and action goals and the **selection** of intentional actions. Goal directed actions are movements which are executed for the purpose of producing very specific, intended effects: one operates the light switch to turn on the light, rides a bicycle to reach another area, talks to convey a message to others. Intentions and goals are therefore focused on the relationship between movement patterns and desired outcomes (Chap. 3).

The SMA also appears to play an important role in the **integration** of movements and effects. For example, in the study by Elsner et al. (2002) participants first acquired new auditory action effects when they pressed buttons that produced specific tones. Later, they were instructed to wait for the presentation of another tone, while they were in a **Positron emission tomography** (PET) scanner that measured their brain activity (cf. Box 2.4). During this waiting period, the action effects which were acquired previously, that is, the tones that were previously elicited by button presses, were also presented. The auditory action effects activated not just the auditory cortex, but the SMA and the hippocampus (a structure that is important for episodic memory) too. This observation was replicated recently by Melcher et al. (2008) in a fMRI experiment. This showed that acquired action effects are integrated with the associated movement patterns, and that this integration produced a connection between the SMA and the sensory representations in the episodic memory. As we will see in later chapters, this connection constitutes an important **condition** for the **selection of movement patterns** based on the number of effects that can be attained with them. In other words: it is this connection that allows us to perform goal-directed actions.

Box 2.4 Methods to Study Brain Processes

The classical method to study the active brain is to record the fluctuations in potentials on the surface of the skull, which occur before, during, or after a sensory, motor, or psychological stimulus or reaction in the EEG. The largest contribution to potential fluctuations is made by spontaneous activity of cortical neurons. This changes immediately when the brain is occupied with the processing of a stimulus or the preparation of a motor response. This causes systematic activation patterns that can be averaged across many trials and made visible in an **evoked potential** (also known as event related potential, ERP).

The evoked potentials are extracted for the spontaneous EEG through averaging across a series of responses (generally, a couple dozen) to visual or acoustic signals. They are classified according to their polarity, positive or negative, and their time of occurrence. For example, the first positive part of the potential (generally in the range from 90 to 140 ms) is designated P1, a component that is associated with early stimulus processing. The specification of polarity and latency alone do not define potentials sufficiently: the area(s) on the skull from which the different evoked potentials are recorded should be added. This provides information on which cortical areas are involved in the processing of stimuli or the performing of specific tasks (albeit with fairly poor spatial resolution). In sum, evoked potentials have an excellent temporal resolution (in the millisecond range) and provide valuable

(continued)

Box 2.4 (continued)
information about the **electrical activity** that accompanies perceptual and action-related processes.

The traditional shortcoming of evoked potentials, their poor spatial resolution, can be countered by **imaging techniques** that have been developed in recent years. **Positron emission tomography (PET)** and **functional magnetic resonance imaging (fMRI**) are relatively new techniques for the imaging of brain areas that are activated function-dependently. PET is based on the measuring of radioactive marker substances that are injected into the bloodstream beforehand. These radioactive markers are used more strongly in actively metabolizing cells, in other words, cells that are engaged in performing a given task. A positron detector, which is placed around the head, counts the emitted particles so that a computer can identify areas of stronger or weaker radiation. PET attains a spatial resolution which allows localization in the millimeter range, but has a very poor temporal resolution (in the range of multiple seconds, up to 10 s).

The particular **benefit** of functional magnetic resonance imaging (fMRI, also known as functional nuclear spin tomography) is that radioactive marker substances are not needed. The only downside is that the processes whose brain activation is studied (e.g., reading, arithmetic, finger movements) must be performed in the small space that is offered by an MRI scanner.

The MRI technique uses the fact that our brain, like other body tissues, consists for a significant percentage of water. This fact is used to **image** the structures of the brain: the hydrogen molecules in our brains possess magnetic properties; each of their atoms functions as a magnetic dipole. When these dipoles are put in a magnetic field, they align with the surrounding magnetic field, just like a compass needle. For this to happen, an extremely strong **magnetic field** is necessary. Typically, MRI scanners use magnetic fields that are 50,000 times as strong as the magnetic field of the earth. When the alignment of magnetic dipoles is disrupted through high-frequency energy impulses, after which they return to their previous aligned position, impulses that are recorded and then amplified, arise. These signals allow the identification of hydrogen molecules and the assessment of their relative proportion in various brain areas. As other body tissues, the brain consists of 70 % of water, and various areas in the brain have different proportions of water. Nerve cells, for example, are relatively rich in water, while the myelin sheath that covers the axon is relatively poor in water. This generates **intensity differences** between signals from different tissue types, which are used to identify various structures in the brain in relatively high levels of detail.

Up to this point in our description, MRI presents us with an image of the **architecture of the brain**. With their high spatial resolution of less than 0.5 mm, MRI images show that the surface of the brain is not particularly spectacular, but that it is richly structured internally. However, no matter how detailed these images are, they provide no clues as to what extent brain areas are involved in performing various tasks. To draw conclusions like that, one would have to observe the brain as it is working, so to speak, and to analyze the

(continued)

Box 2.4 (continued)
task- and function-dependent activity of various brain areas. This is possible through functional magnetic resonance imaging (fMRI), which indirectly measures the **metabolism** of the brain. To understand how the brain's metabolic activity can serve as an indicator of the activation of neuronal populations, we must first have a look at the functional units that constitute the brain.

The basic functional unit of the brain is the **neuron** (Sect. 2.1). A neuron has a cell body, with relatively short projections, dendrites, which take up information from other neurons as input, and lead it to the soma. A neuron also has a relatively long projection, the axon, which conducts electrical impulses from the soma to the dendrites of other neurons. The site where the axon of a neuron comes into contact with the dendrite of another neuron is called a **synapse**. When the electrical impulse which is conducted through the axon exceeds a certain threshold, a chemical **neurotransmitter** is released at the end of the axon. When the neurotransmitter at the synapse contacts the dendrite of the second and other proximal neurons, the electrical impulse is carried on by the second and a host of other neurons, which form a network together. It is important to note that after they have been released, neurotransmitters are "recycled" and are transported back into the neuron. This process requires energy, which has a consequence: the brain's **perfusion** increases locally (the so-called hemodynamic response), to ensure that sufficient amounts of oxygen are available.

This effect is utilized to identify the areas of the brain that are particularly **active** during the performance of a given task: the properties of the dipoles of the hydrogen atoms depend, amongst other factors, on the amount of oxygen in the surrounding blood. Local changes in the amount of oxygen follow increased neuronal activity with a short delay, so that differences in the intensity of signals from the dipoles can be indirectly used to identify those regions of the brain that are responsible for the increased oxygen requirements.

Take, for example, a classical experiment in which participants had to move their fingers ("tapping"). Such movements are prepared in the premotor and primary motor cortices and are led to the motor neurons in the spinal cord after modification by the extrapyramidal system and the cerebellum. To display the neuronal activations in fMRI, participants had to tap his fingers for 30 s while in the MRI scanner, then remain still for another 30 s. During this time, the scanner recorded a few hundred images of both movement phases. These were then averaged and showed, after complicated post-processing, the distribution of neuronal activity during finger movements of the right hand.

While interpreting fMRI data, one should understand that it is not actually the neuronal activation that is being displayed, but a **surrogate**, the hemodynamic response to increased energy requirements *as a result of* neuronal activity. This surrogate is much **slower** (it is developed during multiple seconds after setting of a task) than the neuronal activation, which lasts less than 100 ms. Therefore, there is always a delay of up to 6 s between the time point of neuronal activation and the hemodynamic response. This would not be a problem if the hemodynamic

(continued)

Box 2.4 (continued)

response would always start after a fixed interval following the neuronal activation. However, this is not the case at all. There are indications that the time point at which the hemodynamic response sets in varies between persons, brain regions, and tasks. Therefore, one cannot be certain that the measured hemodynamic reaction actually represents the neuronal activity that is elicited by a given task, or whether it reflects activity that started later.

These problems of interpretation can be neutralized through different experimental designs. For starters, one can use a **block design**: intermix blocks in which the participant performs the task of interest and in which he performs no task or, even better, a control task. The control task should be constructed in such a manner that it is completely similar to the experimental task, but does not evoke the cognitive processes of interest that are associated with the experimental task. This allows researchers to subtract the neural activations from the control task from those of the experimental task, and thus isolate the activity of those brain areas which are associated with the cognitive processes of interest.

Secondly, one can choose a so-called **parametric design** and vary the manipulation of interest systematically in intensity. When a given brain area is systematically affected by this manipulation, the extent of its activation should vary systematically with the intensity of the manipulation. To illustrate: in a study on motivation, one could vary the incentive for solving a problem in equidistant steps and test which brain area is more activated during stronger incentives.

Thirdly, one can present stimulus configurations as isolated stimuli with sufficiently long time intervals, so that individual reactions to single stimuli can be identified. Such **event-related designs** evade potentially confounding factors such as fatigue and habituation, which are common during repetitive stimulation.

The methods to study brain processes that we have sketched up to this point, all attempt to open a window onto the brain and to observe it during its “everyday” activities without attempting to intervene in these activities from the outside. But this is not the only way to study the workings of the brain. The other possibility is to artificially simulate the different structures and to observe the sensory and motor consequences of external stimulation. This path is followed by **stimulation studies** that are generally performed with animals, although occasionally also with humans, whose skull caps are opened, for example, because of neural surgery.

The electrical simulation of neurons in the cortex goes back to Fritsch and Hitzig (1870) who used electrodes to stimulate the surface of the cortex and demonstrated that the motor cortex of dogs shows a somatotopic organization. These observations were confirmed in monkeys and humans during the next decades. Asanuma et al. (1976) developed the method when they simulated

(continued)

Box 2.4 (continued)

neurons through microelectrodes with low intensity. This technique was used in many experiments, in which it was shown that short (often shorter than 50 ms) stimulation of neurons in the motor cortex evokes muscle activity, which is in fact the contraction of single or multiple groups of muscle fibers.

Graziano et al. (2002) used **longer stimulation** (500 ms) in addition to short simulation and this longer stimulation was close to the time monkeys need to execute arm or hand movements. They found that under these conditions, complicated, well-coordinated movements were sold the monkeys. The contractions of single muscle fibers may well have been the beginning of these longer movement sequences that were evoked by longer stimulation.

In such studies, narrowly defined areas of the cortex were stimulated locally, under the assumption that stimulating a relatively small number of neurons locally would spread to widely branching networks of neurons, that control the action. However, the electrical stimulation of neuronal populations does not resemble actual biological or physiological processes. Therefore, any and all effects should be interpreted with caution. For example, it is possible that external stimulation generates an artificial, unnatural activation pattern in the affected neuronal population. This method can be convincing when the evoked movements resemble biological movements that can be associated with known functions of the brain area that is being studied.

A less invasive method to stimulate the cortex externally is **transcranial magnetic stimulation (TMS)**. With this technique, the electrical activity of cortical neurons is influenced through a magnetic field, which is created by a coil that can be placed in various positions on the surface of the skull. Depending on such stimulation parameters as duration and intensity, local neuronal populations of the brain can be inhibited in a temporally precise manner (this would lead to a, so to speak, transient, functional lesions of the neuronal population), but they can also be activated. For example, when the magnetic field is located over the visual areas of the occipital cortex near V5, an area that is assumed to be important for the perception of movement, then it can be demonstrated that TMS selectively disrupts the perception of the motion direction of an object, without influencing its identification at the same time. When neurons in the motor cortex are stimulated, it can be shown in reaction time experiments that TMS shortly before the execution of a movement delays that movement, without influencing its form. Finally, it can be demonstrated that stimulation of the SMA selectively disrupts the execution of complex motion sequences, but not the execution of simpler sequences.

Further evidence for the role of the SMA in the integration of movements and their effects is presented by Haggard et al. (2002). They could first demonstrate that participants systematically underestimate the timing of the effects of actions that they produced themselves: even when an effect occurred after a delay, participants still experienced it as occurring temporally close to the action. This observation also suggests an integration of action and action effects, which led to a temporal illusion here. Interestingly, this illusion breaks down when the activity of the SMA is interrupted through targeted TMS pulses (Haggard et al. 2002). Therefore, the SMA appears to be important both for the integration of action effects and for specific aspects of the experiencing of intentionality. This conclusion is supported by work of Lau et al. (2004). Here, participants made spontaneous finger movements and they were instructed to pay attention either to external stimuli, or to their own intentions. Activation of the SMA was greater in the latter condition, so the greater intentionality of the action corresponded to a stronger contribution of the SMA.

Patient studies suggest a strong relation between intentionality and SMA too. Patients with damage to the SMA often respond to objects in their surroundings with actions that are not accompanied by feelings of intentionality. For example, they may grasp a pen that lies in front of them and start to write, apparently unintentionally. Highly comparable is the so-called **alien hand syndrome**, in which patients know that they are executing particular movements, but cannot report on plans to perform this movement (Marcel 2003).

The observations of Fried et al. (1991) are very interesting in this context: they stimulated the SMA of epileptic patients directly during preoperative neurosurgical measures. During low-intensity stimulation, patients indicated that they occasionally felt the desire to move a specific limb. When stimulation of a greater intensity was applied to the very same area, actual contractions of the limb were observed. This observation, like the alien hand syndrome, indicates a close connection between the subjective experience of intentions and the neuronal processes in the SMA.

2.4 Cerebellum

The SMA is involved in the integration of actions and their consequences. This brain area is important because it allows the selection of alternative actions depending on the intended effect, as well as the evaluation of the consequence of an action by comparing the intended and actually obtained effects (Chap. 9). These effects which we have spoken of up to this point, and which are generally related to intentions, are relatively abstract when compared to the concrete muscular activities and motor parameters of the movements that are necessary to attain these effects. Where does the **information** to control these activities and to specify their parameters come from? It seems plausible that the **cerebellum** makes this information available.

The cerebellum is located below the large brain, in the occipital cortex (the lobe in the back of the head). Damage to the cerebellum does not result in complete failure of motor functions, but is manifested by problems in **coordinating movement processes**. Luciani (1891), for example, studied swimming movements of dogs in which one of the cerebellar hemispheres was removed. He observed that the basic movement pattern was retained in all four paws. However, the coordination of the two paws located ipsilaterally to the lesion was disrupted: their movements were irregular, uncoordinated, and lost their fluidity.

2.4.1 Consequences of Damage to the Cerebellum

Clinical studies by the British neurologist Gordon Holmes of soldiers who suffered gunshot wounds to the cerebellum during the First World War showed that the cerebellum is involved with the regulation of **muscle tone**, with the control of **support and gait motor function**, and with the **coordination** of movement segments (Holmes 1917, 1939). Small lesions of the cerebellum can be compensated fairly well, while larger lesions can cause so-called **ataxic motion disorders**. **Ataxia** describes a lack of coordination, which can be manifested in the motor skills of the eyes, speech, torso, and extremities. Dissymmetric movements would be an example: healthy persons are generally able to close their eyes and move their arms from the side of their torso in such a manner that the tips of their index fingers touch in the center of their torso without many problems. Patients with cerebellar lesions are unable to generate motion impulses that make the two hands move in synchrony spatially and temporally during such pointing movements. This often leads to over- and undershooting movements of the fingers.

These patients often have problems with speech, which is manifested as slowed, halting speech with poor articulation and uneven stressing of syllables. Beyond that, they have problems with the execution of fast, alternating movements, which require a rapid switching between agonists and antagonists, often have an insecure, wide-legged gait, and occasionally show a(n intentional) tremor, which, contrary to the resting tremor in Parkinson's disease patients, occurs during the execution of a movement, especially during its final phase, where the accuracy requirements are often greatest.

2.4.2 Cognitive Functions of the Cerebellum

The cerebellum has many recurrent (that is, reciprocal, interactive) connections with almost all areas of the cerebral cortex (Middleton and Strick 2000). It receives input from the motor cortex and almost all sensory areas in the cerebral cortex. Through ascending projections in the spinal cord, it receives proprioceptive information about the conditions of the skeletal musculature and about the current

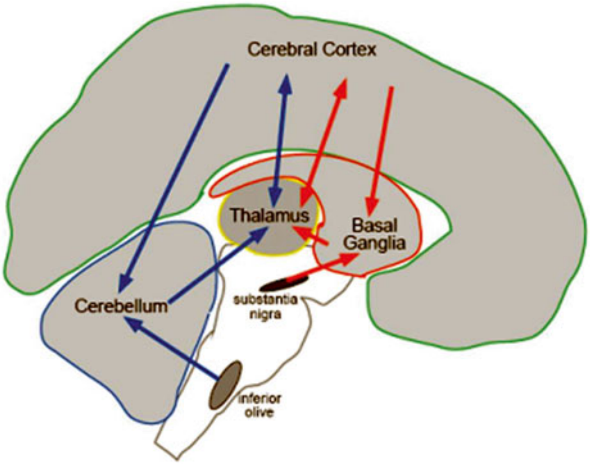


Fig. 2.6 Connections between the cerebellum and various cortical and subcortical structures of the human brain (from Doya 1999; adapted with permission from Elsevier)

positions of the various joints. It projects through the thalamus to the motor areas of the large brain and to further areas of the prefrontal parietal and temporal cortices (Fig. 2.6). This strong networking suggests that the cerebellum is involved with other cognitive processes besides its motor tasks.

In addition to the control of action, **two** partially related **functions** of the cerebellum are of particular importance: predicting the sensory consequences of concrete movements (so-called forward modeling; Box 2.5), and the control of motor learning (Sect. 9.4). Evidence for the involvement of the cerebellum in motor learning is provided, for example, by Imamizu et al. (2000). These authors had participants in a fMRI study follow a moving stimulus with a mouse, and to attempt to cover it with the mouse cursor. The requirement of motor learning was induced through a manipulation of the relationship between the movements of the mouse and the mouse cursor so that, for example, moving the mouse upwards led the mouse cursor to move 120° to the right. To actually move the cursor upwards, participants had to move the mouse 120° to the left. In this condition, activity in the cerebellum was considerably greater than in a control condition. Given increasing practice with the oddly behaving mouse, activity declined proportionally to the reduction of motion errors of the participants (i.e., the difference between the position of the moving stimulus and the mouse cursor). This suggests that the cerebellum is involved with the **acquisition of new motor models**. However, even after the participants had learned to use the new mouse, activity of some areas in the cerebellum was increased as compared with the control condition, possibly because the newly acquired models in this task should be maintained actively.

Box 2.5 Forward and Inverse Models of Action Control

The concept of forward modeling stems from systems theory, as it is applied, for example, in engineering in the construction of robots. For example, assume that you are attempting to grasp an object, such as a coffee cup that is placed in front of you, for the first time in your life. In such a case, no motor experience is available to you, and so, you have no other choice but to try out various movements on a trial and error basis. As your experience increases, you will, however, acquire the successful movement, so **motor learning** takes place. In the future you will be able to reliably execute the desired grasping motion.

Expressed in terms of systems theory or cybernetics, you can thereby transfer a **desired** state into an **actual** state (Fig. 2.7). At the beginning of the movement, the desired state is compared to the actual state and the difference (estimated state error) is computed. This activates a motor control structure that, in turn, passes on instructions to the motor system (see outer loop in Fig. 2.7). The activity of the motor system leads to perceivable sensory changes: you can see and feel how your hand moves towards the cup. This changed state is compared to the desired state; when they resemble each other (so when the estimated error is close to zero), then the action is finished, otherwise the whole loop is reiterated until the desired state is attained. You may be familiar with this principle from central heating: the heating is kept active until the desired temperature is reached.

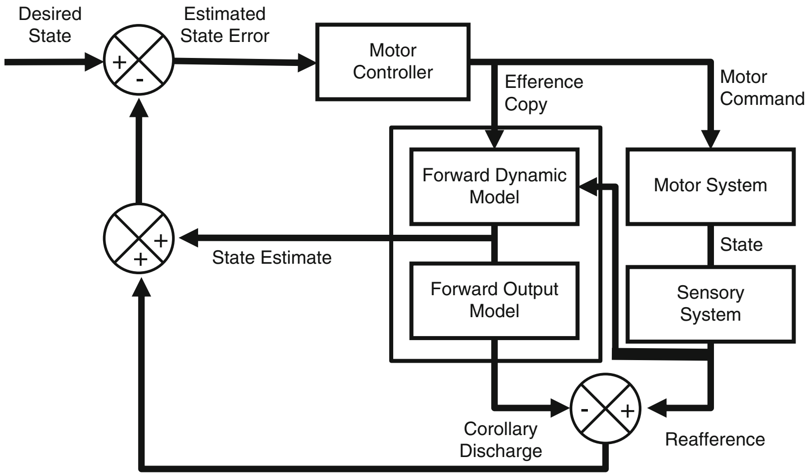


Fig. 2.7 The principles underlying forward modeling (after Wolpert et al. 1998. © Cell Publishing)

(continued)

Box 2.5 (continued)

Going through such loops again and again in the course of motor learning allows the learner to discover and acquire knowledge about systematic relationships between the motor commands (or rather the motor activities that result from them) and their sensory consequences. The acquisition of these relationships allows two things: the forward modeling and the inverse modeling of the relevant action. **Inverse models** provide information about which motor commands are required to achieve specific movement effects—like how the hand can be controlled to grasp a coffee cup in a specific location. The relationships between motor commands and sensory consequences are used here to determine the former on the basis of the latter. **Forward models** serve, conversely, to predict sensory consequences on the basis of motor commands (Wolpert et al. 1998).

What purpose could such forward models serve? Why would one even want to predict sensory consequences? There are a number of reasons for that. Most importantly, a prediction allows us to check internally, before a movement even starts, whether the selected motor commands will actually realize the intended movement (i.e., the desired sensory effects) and, if necessary, make corrections to the motor commands. You have probably experienced breaking off or not even having initiated a movement or the uttering of a word, simply because it somehow seemed wrong. It is very likely that the prediction of (in this case) apparently incorrect sensory consequences played a part in that. In the depicted model, this possibility has been taken into account: a copy of the motor command (so-called efference copy, a term introduced by Von Holst and Mittelstaedt 1950; cf. Box 9.2 on the **Reafference Principle**) is used to predict the execution of a movement in parallel with motor control and execution of the likely sensory consequences. This prediction can be compared directly to the desired goal state and, when the outcome is negative (that is, when the estimated error is large), passed on directly to the motor control structures and be used to change the motor commands there.

The possibility to monitor the performance of a movement and to correct it if necessary, independently of its actual consequences, facilitates action control in many situations (cf. Chap. 9). Sensory feedback about movements (e.g., visual feedback about the position of a moved hand) is not available until a considerable amount of time (usually multiple hundreds of milliseconds) has passed after the movement has ended; that depends on neuronal transfer and processing times. If the next movement elements would have to wait for this feedback first (in which it is ensured that the previous movement has been terminated correctly), then more complex movement terminations would last very long, would not look particularly smooth, and would be very difficult to control. In reality, the elements of an action are superimposed (cf. Chap. 7). This is clearly visible in grasping, where the hand already opens

(continued)

Box 2.5 (continued)
itself and adjusts its position to the size of the object as it moves towards the object that is to be grasped (Chap. 6). Now, when the putative outcomes of a movement are available through forward modeling even before it is terminated, then subsequent movements can already be planned and partially initiated. Both the planning of various movement elements as well as their execution can therefore overlap temporally, and the complete movement termination will become very efficient and smooth.

A further advantage of forward modeling is the possibility to control motor learning (Sect. 9.4). Hitherto, we have assumed that forward modeling will lead to the same outcome as the actual movement, so that the predicted and the actual states are identical at the end of a movement. At the start of the acquisition of a movement, that is not always the case, as the predictions are often fragmentary and unreliable due to lack of practice and experience. In such a case, if the outcome does not correspond with the outcome of a movement, so when, to use system theoretical jargon, an error has been detected, then this is an indication that the movement has not yet been sufficiently learned motorically. In other words, the identity of forward modeled predictions and actual movement outcomes signals successful motor learning, while differences, in contrast, signal further learning needs (Doya 2000; Wolpert et al. 1998). In this sense, a comparison of forward modeling and movement can take on the role of a trainer, without requiring the judgement of other persons.

2.5 Basal Ganglia

The basal ganglia are a collection of **subcortical nuclei** (nucleus caudatus and putamen, together known as the striatum, globus pallidus, substantia nigra, and nucleus subthalamicus), which receive strong afferent projections from not just the motor areas of the large brain, but also from the frontal eye fields, the limbic system, and the orbitofrontal and prefrontal cortices. In the basal ganglia, signals from the various brain areas remain topographically separated, are processed in parallel, and are sent back to the original cortical areas (Alexander 1995). That the basal ganglia are connected with not just the motor areas of the large brain, but also with the limbic and associative systems, indicates that they are not involved exclusively with **voluntary motor skills**. A particularly important function of the basal ganglia is the production of dopamine, a very influential neurotransmitter that modulates numerous cognitive and motor processes.

2.5.1 Consequences of Damage to the Basal Ganglia

Damage to the basal ganglia leads to a number of disruptions in **voluntary motor skills**, which are manifested in different manners, depending on which structures are damaged: a strong retardation of movements (bradykinesia), which is reflected by increased reaction times to visual and acoustic stimuli and altered speed profiles in goal-directed movements; a rest tremor (involuntary, rhythmic oscillations of the limbs); and a heightened stiffness of the musculature (rigor) can be the result of damage to the substantia nigra. Disruptions of this kind often occur in **Parkinson's disease** (morbus Parkinson), one of the most common neurological ailments (famous Parkinson's patients are the late pope John Paul II and the ex-boxing champion Muhammad Ali). The neurophysiological cause of the disease is the massive loss of dopamine-producing neurons in the substantia nigra. Damage in the area of the striatum is causes of **Huntington's disease**, which is manifested by massive gait and postural instabilities. Finally, lesions of the nucleus subthalamicus lead to involuntary heavy, large-amplitude movements.

Both Parkinson's and Huntington's diseases are explained by damage to the basal ganglia (and the accompanying dopaminergic failure regulation), which leads to a **disturbed balance** between excitatory and inhibitory influences on the frontal cortex (Alexander et al. 1990). According this conception, the execution of voluntary action is blocked by Parkinson's patients, while the inhibitory influences are increased drastically, which suppresses the initiation of voluntary movements, or at least makes their initiation harder. In Huntington's patients, the reverse is true: here the inhibitory influence of the basal ganglia is reduced drastically, which lowers the threshold to initiate motor activities dramatically. This would account for the common execution of involuntary, fast, and choppy movements.

2.5.2 Cognitive Functions of the Basal Ganglia

Dopamine that is produced in the basal ganglia does not only influence action regulation directly, it also modulates the acquisition of cognitive and motor skills. The unexpected success of an action leads to a temporary increase of the dopamine level, while an unexpected failure leads to reduction of the level (Schultz 1998). This makes information available to learning processes, about whether a given action was beneficial or detrimental, and whether it should be learned or avoided in the future. Groundbreaking research by Schultz et al. (1993) supports these hypotheses. These researchers have, for example, trained monkeys in operant conditioning paradigm, to grasp for objects following the presentation of visual stimuli; the monkeys were rewarded with fruit juice. Dopamine-producing neurons were very active during the starting phase of the training, when the monkeys received their reward following a successful performance. Following sufficient training, these neurons started to fire when the visual stimulus was presented, and they adjusted their activity when a reward was received. Therefore, the activity of the dopaminergic neurons first signaled the rewards that were unexpected in the beginning of the training, then, later, increasingly signaled the expectation of a future reward (for comparable observations in humans, see Haruno et al. 2004). According to Doya (2000), based on experience, the basal ganglia can maintain information about which rewards are to be expected under the current circumstances and which action alternatives are associated with which rewards. Such a system can play a decisive role in the evaluation and selection of actions, and bias the action alternative that offers the greatest reward.

2.6 Prefrontal Cortex

The frontal and in particular the prefrontal cortex are relatively voluminous in humans as compared to other species. For example, in dogs, the frontal cortex makes up about 7 % of the cortex, in monkeys about 17 %, and in humans, more than 30 %. Until fairly recently, this part of the brain was considered to be quite mysterious. By now, however, it has become clear that the neuronal populations of the frontal cortex are not just involved with the primary processing of light and sound stimuli, olfaction, taste, and tactile sensations, but mainly with so-called **executive functions**. When a human develops plans, makes judgments, forms intentions and transforms them into actions, then the frontal cortex is active. Besides this, the frontal cortex is crucial for **working memory**. This is a type of short-term storage that is required, for example, when people converse; to understand a spoken sentence, one must recall at the end what was said at the beginning. Such short-term memory is also required at a party, where one should remember who is already greeted, and who is yet to be greeted. More importantly, the frontal cortex checks the **processes** that occur on the way from intention to action, and recognizes **possible errors**. Even before you put sugar (instead of grated Parmesan cheese) on your spaghetti, your frontal cortex starts corrective measures and steers your hand away from the sugar and towards the Parmesan cheese.

That the frontal region of the human brain is involved with intentions, and their conversion into actions, is corroborated by people with **damage** to this region. They often act thoughtlessly and prematurely, often change plans, ignore important information, and often commit errors without noticing it. Unless motor areas are affected, these patients are not noticeably or visibly disabled. However, their behavior is often curiously **inflexible and environmentally dependent**: they have problems with planning actions, remembering action goals, and switching between action goals (Burgess 2000). The control of their actions appears to have shifted to the outside world, so to speak, so that the confrontation with objects leads to the execution of actions that are often associated with those objects: many patients smoke when they encounter cigarettes, drink when they see a glass of liquid, and grasp or manipulate objects without an apparent goal (see Sect. 3.1.2). Altogether, their long-term actions appear to be **less intentional**; they appear to lose interest in attaining specific goals, and to merely respond to external stimuli. We will now discuss three particularly important areas of the (pre)frontal cortex and outline their significant contributions to human action planning.

2.6.1 Dorsolateral Prefrontal Cortex (BA9/46)

The **dorsolateral prefrontal cortex (DLPFC)** is connected with almost all areas of the human brain, especially with the basal ganglia, the hippocampus, and the temporal, parietal, and occipital cortices. Its mode of operation depends strongly on **dopamine**, which is delivered to the DLPFC through the so-called mesocortical pathway (that originates in the ventral tegmentum). The DLPFC is the human brain area that requires the longest time to mature (often up to young adulthood) and that degenerates particularly quickly through aging processes and is therefore responsible for many side effects of aging.

Two important, strongly overlapping **functions** of action control are ascribed to the DLPFC. According to an older interpretation (Goldman-Rakic 1987), this structure serves as **working memory** in the meaning of Baddeley (1986). This was based on the repeated observations from cell recordings in animal research, that neurons in the DLFPC are activated through use of the working memory. In a typical task, monkeys were shown target positions for eye or hand movements, but there had to be a delay between these cues and the actual movements—these target positions had to be kept in working memory. It was demonstrated that neurons in the DLPFC fired more strongly during this **retention** interval. Therefore, it would seem plausible that these neurons represented either the target positions themselves, or actively maintained other representations during the retention interval.

More recent interpretations suggest that the DLPFC is mainly involved in **cognitive control** (Miller and Cohen 2001). It represents the goals of cognitive and motor actions and actively supports all processes that are relevant for the execution of these actions. In other words, the DLPFC could be the **neural correlate of the human will**. In fact, a number of neuroimaging studies have shown that the DLFPC is activated especially during the preparation of a new task and also that the more difficult the task, the stronger the participant's will must be to complete the task. MacDonald et al. (2000) presented Stroop stimuli, that is, color words that are presented in incongruent colors (for example, the word "red" in a green color, see also Sect. 9.3). In numerous trials, participants either had to read the word, which ought to be a natural reaction, or name the color, which is a complex assignment, as in these cases, reading of the word has to be inhibited. The activation of the DLPFC clearly increased during the preparation for the upcoming trial, especially during the preparation of the difficult color-naming trials. When we assume that the DLPFC is necessary to actively maintain action goals and to guarantee them the necessary influence on the relevant cognitive processes, then it ought to be clear why frontal lesions lead to large deficits in prioritizing, organizing, and coordinating various actions (Burgess et al. 2000).

In a liberal interpretation, the two assumed functions of the DLPFC are not mutually exclusive, by the way. For example, the concept of working memory does not suggest that it is a type of container into which the contents of our thoughts are transferred. Much more plausible is the assumption that activation in the DLPFC leads to **maintenance of relevant representations** in the sensory or thought areas. The codes of the working memory are therefore not necessarily copies of representations from other areas of the brain, but merely **pointers** to these representations. Action goals can function in the same way, that is, they can indicate or refer to the processes that are vital for attaining a specific goal and thereby actively maintain them.

2.6.2 Orbitofrontal Cortex (BA10-14/47)

The neuroanatomical localization of the **orbitofrontal cortex (OFC)** is still under discussion (Kringelbach and Rolls 2004), and its borders and their relation to neighboring areas are viewed different by various authors; some authors consider the OFC and the ACC (Sect. 2.6.3) to be one area, and they refer to it as **ventromedial cortex**. The OFC receives information from all sensory systems, such as the hippocampus, the amygdala, and the cingular cortex, but also from other areas of the prefrontal cortex. It sends information to, amongst other areas, the striatum (where it might possibly influence dopamine production), to the amygdala, entorhinal cortex, the hippocampus, and the inferior temporal cortex.

On the one hand, the OFC is clearly involved in **affective processes**, and it is assumed that it plays a crucial role in the **association of stimulus features** and the rewards (or punishments) that are associated with them (Rolls 1999). On the other hand, the OFC is important for **action planning**. The first clues for this were provided by the extensive analysis of the brain lesion of Phineas Gage, which we discussed at the start of this chapter. You will remember that the frontal cortex of Mr. Gage was pierced by a chisel during an explosion. Fortunately, both Gage's skull and the chisel have been retained, which allowed Hanna Damasio and colleagues (1994) to reconstruct the injury exactly in a computer simulation.

According to the reconstruction, the OFC was particularly damaged, which suggests that the difficulties Gage had in planning may well have been caused by the lesion of this brain area. This raises the question of to what extent the OFC is involved in action planning, and how this involvement can be united with the role the OFC plays in affective processing. The explanation of this association rests on the assumption that action planning involves **distinguishing** between alternative stimuli (Rolls 1999) or actions (Damasio 1994) and that these distinctions are made on the basis of the reward someone is trying to obtain. Damasio (1994) assumes that every action is associated with a representation of its **affective consequences**. This way, we learn what it "feels like" (or would feel like) to perform a specific action. These so-called **somatic markers**, that is, the representations of expected affective bodily reactions, allow for relatively fast, often also intuitive decisions; one simply chooses the action that "feels best." More recent studies of patients with OFC damage support this notion. These patients especially have problems with making risky decisions, when compared with healthy persons or people with lesions of other brain areas (Bechara et al. 1998); also, these patients do not demonstrate the typical sweating response prior to making a very risky decision that is common in healthy persons.

2.6.3 Anterior Cingulate Cortex (BA24)

The name of the cingular cortex (*cingulum* is Latin for belt) is derived from the fact that it wraps around the corpus callosum (the connection between the two cortical hemispheres) like a belt. It is part of the **limbic system** that plays an important role in the emergence of emotions and the regulation of memory and behavior. The ACC receives afferent signals mainly from thalamic nuclei and sends efferent signals to other areas of the prefrontal cortex, the anterior nucleus and other limbic areas. The ACC is thought to be important for **monitoring action control**, for which it collaborates closely with the DLPFC. Evidence from Botvinick et al. (2001) suggests that the ACC registers conflicts between stimulus and response alternatives, and strengthens the representations of the action goal in the DLFPC (Sect. 9.3). This **strengthening of a target**, in turn, leads to **increased attention** to the information relevant to the action (Egner and Hirsch 2005).

References

Alexander, G. E. (1995). Basal ganglia. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (pp. 139–143). Cambridge, MA: MIT Press.

Alexander, G. E., Crutcher, M. D., & De Long, M. R. (1990). Basal ganglia-thalamocortical circuits: Parallel substrates for motor, oculomotor, “prefrontal” and “limbic” functions. *Progress in Brain Research, 85*, 119–146.

Asanuma, H., Arnold, A., & Zarzecki, P. (1976). Further study on the excitation of pyramidal tract cells by intracortical microstimulation. *Experimental Brain Research, 26*, 443–461.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Bechara, A., Damasio, H., Tranel, D., & Anderson, S. W. (1998). Dissociation of working memory from decision making within the human prefrontal cortex. *Journal of Neuroscience, 18*, 428–437.

Botvinick, M. M., Braver, T. S., Carter, C. S., Barch, D. M., & Cohen, J. D. (2001). Evaluating the demand for control: Anterior cingulate cortex and crosstalk monitoring. *Psychological Review, 108*, 624–652.

Brodmann, K. (1909). *Vergleichende Lokalisationslehre der Grosshirnrinde in ihren Prinzipien, dargestellt auf Grund des Zellbaues*. Leipzig, Germany: JA Barth.

Burgess, P. W. (2000). Strategy application disorder: The role of the frontal lobes in human multitasking. *Psychological Research, 63*, 279–288.

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologica, 38*, 848–863.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex, 15*, 1243–1249.

Damasio, A. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.

Damasio, H., Grabowski, T., Frank, R., Galaburda, A. M., & Damasio, A. R. (1994). The return of Phineas Gage: Clues about the brain from the skull of a famous patient. *Science, 264*, 1102–1105.

Decety, J., & Grezes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences, 3*, 172–178.

Deiber, M. P., Honda, M., Ibanes, V., Sadato, N., & Hallett, M. (1999). Mesial motor areas in self-initiated versus externally triggered movements examined with fMRI: Effect of movement type and rate. *Journal of Neurophysiology, 81*, 3065–3077.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences, 11*, 219–226.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research, 91*, 176–180.

Dick, J. P. R., Benecke, R., Rothwell, J. C., Day, B. L., & Marsden, C. D. (1986). Simple and complex movements in a patient with infarction of the right supplementary motor area. *Movement Disorders, 1*, 255–266.

Doya, K. (1999). What are the computations of the cerebellum, the basal ganglia, and the cerebral cortex. *Neural Networks, 12*, 961–974.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10*, 732–739.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8*, 1784–1790.

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, et al. (2002). Linking actions and their perceivable consequences in the human brain. *Neuroimage, 17*, 364–372.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*, 2608–2611.

Fried, I., Katz, A., McCarthy, G., Sass, K. J., Williamson, P., et al. (1991). Functional organization of human supplementary motor cortex studied by electrical stimulation. *Journal of Neuroscience, 11*, 3656–3666.

Fritsch, G., & Hitzig, E. (1870). Über die elektrische Erregbarkeit des Grosshirns. *Archiv der Anatomischen und Physiologischen Wissenschaften der Medizin, 37*, 300–332.

Georgopoulos, A. P., Pellizzer, G., Poliakov, A. V., & Schieber, M. H. (1999). Neural coding of finger and wrist movements. *Journal of Computational Neuroscience, 6*, 279–288.

Goldman-Rakic, P. S. (1987). Circuitry of the prefrontal cortex and the regulation of behavior by representational knowledge. In F. Plum & V. Mountcastle (Eds.), *Handbook of physiology* (pp. 373–417). Bethesda, MD: American Physiological Society.

Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research, 112*, 103–111.

Graziano, M. S. A., Taylor, C. S. R., & Moore, T. (2002). Complex movements evoked by microstimulation of precentral cortex. *Neuron, 34*, 841–851.

Grèzes, J., Frith, C. D., & Passingham, R. E. (2004). Inferring false beliefs from the actions of oneself and others: An fMRI study. *NeuroImage, 21*, 744–750.

Haggard, P., Aschersleben, G., Gehrke, J., & Prinz, W. (2002). Action, binding and awareness. In W. Prinz & B. Hommel (Eds.), *Common mechanisms in perception and action: Attention and Performance* (Vol. XIX, pp. 266–285). Oxford: Oxford University Press.

Hallett, M. (2001). Plasticity of the human motor cortex and recovery from stroke. *Brain Research Reviews, 36*, 169–174.

Halsband, U., Ito, N., Tanji, J., & Freund, H. J. (1993). The role of premotor cortex and the supplementary motor area in the temporal control of movement in man. *Brain, 116*, 243–266.

Haruno, M., Kuroda, T., Doya, K., Toyama, K., Kimura, M., et al. (2004). A neural correlate of reward-based behavioral learning in Caudate nucleus: A functional magnetic resonance imaging study of a stochastic decision task. *Journal of Neuroscience, 24*, 1660–1665.

Holmes, G. (1917). The symptoms of acute cerebellar injuries due to gunshot injuries. *Brain, 40*, 461–535.

Holmes, G. (1939). The cerebellum of man. *Brain, 62*, 1–30.

Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 371–398). New York: Oxford University Press.

Hubel, D. H., & Wiesel, T. N. (1963). Single-cell responses in striate cortex of kittens deprived of vision in one eye. *Journal of Neurophysiology, 26*, 1003–1017.

Huttenlocher, P. R. (1994). Synaptogenesis in human cerebral cortex. In G. Dawson & K. W. Fischer (Eds.), *Human behavior and the developing brain* (pp. 35–54). New York: Guilford.

Imamizu, H., Miyauchi, S., Tamada, T., Sasaki, Y., Takino, R., et al. (2000). Human cerebellar activity reflecting an acquired internal model of a new tool. *Nature, 403*, 153–154.

Keysers, C., & Perrett, D. I. (2004). Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences, 8*, 501–507.

Kolb, B., & Whishaw, I. Q. (1996). *Neuropsychologie*. Heidelberg, Germany: Spektrum Akademischer Verlag.

Konczak, J. (2008). Motorische Kontrolle. In J. Müsseler (Ed.), *Allgemeine Psychologie* (pp. 739–766). Berlin, Germany: Springer.

Kringelbach, M. L., & Rolls, E. T. (2004). The functional neuroanatomy of the human orbitofrontal cortex: Evidence from neuroimaging and neuropsychology. *Progress in Neurobiology, 72*, 341–372.

Kristeva-Feige, R., Feige, B., Makeig, S., Ross, B., & Elbert, T. (1993). Oscillatory brain activity during a motor task. *NeuroReport, 4*, 1291–1294.

Lau, H. C., Rogers, R. D., Haggard, P., & Passingham, R. E. (2004). Attention to intention. *Science, 303*, 1208–1210.

Liepert, J., Tegenthoff, M., & Malin, J. P. (1995). Changes of cortical motor area size during immobilization. *Electroencephalography and Clinical Neurophysiology, 97*, 382–386.

Luciani, L. (1891). *Il cervelletto. Nuovi studi di fisiologia normale e patologica*. Firenze: Le Monnier.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulated cortex in cognitive control. *Science, 288*, 1835–1838.

Marcel, A. J. (2003). Introspective report: Trust, self-knowledge and science. *Journal of Consciousness Studies, 10*, 167–186.

Melcher, T., Weidema, M., Eenshuistra, R. M., Hommel, B., & Gruber, O. (2008). The neural substrate of the ideomotor principle: An event-related fMRI analysis. *NeuroImage, 39*, 1274–1288.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: Motor and cognitive circuits. *Brain Research Reviews, 31*, 236–250.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Pascual-Leone, A., Cammarota, A., Wassermann, E. M., Brasil-Neto, J. P., Cohen, L. G., et al. (1993). Modulation of motor cortical outputs to the reading hand of Braille readers. *Annals of Neurology, 34*, 33–37.

Pascual-Leone, A., Dang, N., Cohen, L. G., Brasil-Neto, J. P., Cammarota, A., et al. (1995). Modulation of muscle responses evoked by transcranial magnetic stimulation during the acquisition of new fine motor skills. *Journal of Neurophysiology, 74*, 1037–1045.

Pascual-Leone, A., Walsh, V., & Rothwell, J. (2000). Transcranial magnetic stimulation in cognitive neuroscience—Virtual lesion, chronometry, and functional connectivity. *Current Opinion in Neurobiology, 10*, 232–237.

Penfield, W., & Boldrey, E. (1937). Somatic motor and sensory representations in the cerebral cortex of man as studied by electrical stimulation. *Brain, 60*, 389–443.

Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localization and function*. New York: Macmillan.

Pfurtscheller, G., Flotzinger, D., & Neuper, C. (1994). Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG. *Electroencephalography and Clinical Neuropsychology, 90*, 456–460.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience, 27*, 169–192.

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Experimental Brain Research, 111*, 246–252.

Rolls, E. T. (1999). *The brain and emotion*. Oxford: Oxford University Press.

Sanes, J. N., Wang, J., & Donoghue, J. P. (1992). Immediate and delayed changes of rat motor cortical output representation with new forelimb configurations. *Cerebral Cortex, 2*, 141–152.

Schieber, M. H. (1999). Somatotopic gradients in the distributed organization of the human primary motor cortex hand area: Evidence from small infarcts. *Experimental Brain Research, 128*, 139–148.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *Journal of Neurophysiology, 80*, 1–27.

Schultz, W., Apicella, P., & Ljungberg, T. (1993). Responses of monkey dopamine neurons to reward and conditioned stimuli during successive steps of learning a delayed response task. *Journal of Neuroscience, 13*, 900–913.

Shima, K., & Tanji, J. (1998). Both supplementary and presupplementary motor areas are crucial for the temporal organization of multiple movements. *Journal of Neurophysiology, 80*, 3247–3260.

Singer, W. (1994). The organization of sensory motor representations in the Neocortex: A hypothesis based on temporal coding. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance XV: Conscious and nonconscious information processing* (pp. 77–107). Cambridge, MA: MIT Press.

Spelke, E. S. (1976). Infant's intermodal perception of events. *Cognitive Psychology, 5*, 553–560.

Tallon-Baudry, C., & Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends in Cognitive Sciences, 3*, 151–162.

Tanji, J., & Shima, K. (1994). Role for supplementary motor area cells in planning several movements ahead. *Nature, 371*, 413–416.

Von der Malsburg, C. (1995). Binding in models of perception and brain function. *Current Opinion in Neurobiology, 5*, 520–526.

Von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften, 37*, 464–476.

Wolpert, D. M., Miall, R. C., & Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*, 338–347.

Chapter 3
Intentions and Action Goals

An **action** is an intended, goal-directed movement; the execution of an action therefore prerequisites the existence of action intentions and action goals. But what is an intention? How is an action goal represented cognitively? How can such a representation control an action? Let us start with the relationship between intention and goal. Many **definitions** of these terms have been suggested and debated strongly, especially in action philosophy (Meggle 1993). When referring to a goal, we generally refer to the desired product of an action, to the final state that should be attained through the action; therefore, the concept of intention sometimes also includes the (conscious) striving to attain a goal (e.g., Heckhausen and Gollwitzer 1987). Other authors assume that goals are always represented consciously, while intentions are unconscious structures and processes that convert these goal representations into action (Baars 1988).

In this book, we would first leave it open whether intentions and/or goals are necessarily experienced consciously (Sect. 3.1), and we will therefore use the terms *intention* and *goal* interchangeably. In other words, we mainly understand intentions and goals as **functionally characterized states and processes of action control** and ignore for the most part how these states are experienced personally.

3.1 Function of Action Goals

Ach (1910) already assumed that the establishment of an action goal, be it out of one's spontaneous consideration or by being required to do so, calls upon a "determining tendency." The idea is that a goal restructures the cognitive processing system, and the processes associated with the processing, in such a manner that the attainment of the goal is made possible. How should one conceptualize that?

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_3

In his groundbreaking work on will psychology, Ach (1910, 1935) assumed that the human will is continuously in conflict with **habits**. When we follow our habits, so it is assumed, willpower is not required. Ach conceptualized habits as strongly overlearned stimulus-response associations, which evoke the associated response more or less automatically whenever a stimulus is encountered. If, however, a habit must be broken, this needs a strong exertion of will, as, for example, in a smoker who would like to quit. To study this association, Ach developed his **combined method**. In the first phase, his participants learned new stimulus-response associations; for example, they learned to respond to meaningless syllables (e.g., "zup") with rhyme words ("tup"). After an extensive practice period, participants had to respond to these same syllables with different, new reactions, for example, with reversed syllables (e.g., "puz"). As hypothesized, Ach found longer reaction times and more errors to syllables that were previously paired with other reactions than in reaction to new syllables. He also observed strong inter- and intraindividual differences in performance, which he considered to reflect differences in willpower and temperament.

The studies by Ach have led to a true boom of experiments in the psychology of the will (Ach 1935), but as the results thereof were only published in German, they did not have a profound influence on international research. However, there were also arguments that spoke against the concept of the will battling against habits. Lewin (1922) could only replicate Ach's findings under very specific conditions, namely by encouraging participants to focus very strongly on their newly acquired habits. By changing these conditions, Lewin attempted to bring his participants to a more intentionally controlled processing strategy, and the performance differences between practice-congruent and practice-incongruent conditions actually disappeared. According to Lewin, these findings indicate the great importance of the **task set** that apparently directly influences the extent to which previously acquired associations influence information processing.

3.1.1 Action Goal as Context

After a long dormant period, the relationship between will and habit—or, as people tend to say now, between intentional and automatic processes—has seen a renaissance. One of the first detailed models that occupied itself with this question is from Cohen and Huston (1994, see Fig. 3.1) and is called a **parallel distributed processing (PDP) model** (Rumelhart and McClelland 1986). It seeks to explain, through use of the Stroop effect, how one can respond in various ways to the same stimulus. The Stroop effect occurs in tasks where participants have to name the color of color words (MacLeod 1991). This is easier when the color word is congruent (for example, when the word *green* is printed in green and should be named as *green*), than when it is incongruent (when the word *red* is printed in green and should be named as green). Apparently, an incongruent word induces two response tendencies (to say "red" and to say "green") that compete with each other and thereby delay the choice

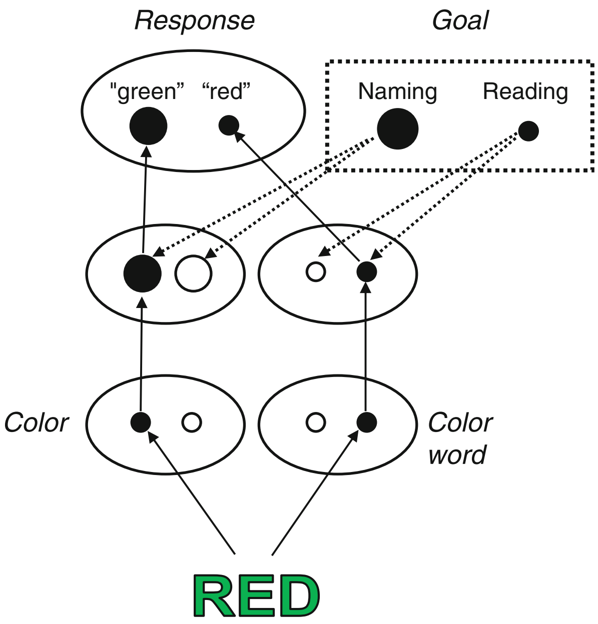


Fig. 3.1 PDP-model of the Stroop effect (modified from Cohen & Huston 1994) by permission of MIT Press

of the correct response, or even elicit the wrong reaction—we immediately recognize the logic of Ach's (1910) combined method here. The problem with the Stroop effect (as with Ach) lies therein, that we can perform various actions that are incompatible with each other on the basis of the same stimulus.

The PDP-model explains how one can choose between actions and in which way incongruent stimuli can delay the choice of reactions, or even lead it astray. Figure 3.1 displays a simplified version of this model. The stimulus is the word *red* printed in green. The two task-relevant properties of the stimulus (the word and its color) are processed simultaneously in the respective property domains, where they activate their corresponding representations. This activation is led to the choice of a reaction via a relay station (hidden units), where the appropriate (in this case verbal) reaction is activated.

If the model had no further properties, it would be impossible to decide whether one should say "green" or "red" upon seeing this stimulus. Of course, one can assume that this model is associated with some variability, so that sometimes one and sometimes the other reaction wins, for example, because either the color or color name can be processed faster at times, or the respective representation is activated more strongly. However, this does not solve our problem. The proposition of Cohen and Huston is merely to let the processing stream on the mid-level (i.e., between stimulus identification and selection of a reaction) be influenced by the **action goal**. If one's goal is to name the color of the stimulus, then activation of this goal supports the stream of information from color coding response selection (see the dotted arrow in Fig. 3.1). In other words, the **color** of a stimulus has a stronger influence on the selection of a reaction than its **name**. Accordingly, in our case, the reaction “green” is activated more strongly than “red.” When one assumes that the influence of the color name cannot be ignored completely (for example, because the reactions also consist of color names), it can be understood why incongruent stimuli can impair the choice of a reaction: the correct response may be activated more strongly, and will win in the end, but the competition with a weaker competitor takes time. The model can also explain why one can name the color without reading the color name: one activates the representation of the goal “to read,” which supports the information stream of the name encoding to the selection of a reaction more strongly than the stream from the color encoding to the selection of a reaction.

3.1.2 Role of the Prefrontal Cortex

Other, partially better models of the Stroop effect have been developed (MacLeod 1991) but the general idea of Cohen and Huston, that goal representations support goal-relevant stimulus reactions, is generally maintained (e.g., Gilbert and Shallice 2002). Note that this idea is very compatible with the evidence provided by Ach (1910, 1935) and Lewin (1922); in agreement with Ach, the processes that are supported by goal-representations (the will) compete with other, non-supported processes (the habits). In agreement with Lewin, it is possible to come up with circumstances under which the support is sufficiently strong to restrict the conflict that is induced through automatic processes to a minimum.

Neuroscientific findings are consistent with this idea too (cf. Sect. 2.6.1). As Miller and Cohen (2001) indicate, the prefrontal cortex (PFC) possesses a range of attributes that are useful for the representation of goals. For example, the PFC is able to autonomously keep neural representations active, and to protect them from disruptions; it has the relevant connections to sensory and motor centers, and it has sufficient plasticity to quickly establish short-term connections and to adapt them to changing circumstances.

Additionally, disruptions of functions of the PFC have the exact **effects** that one would expect on the basis of the model by Cohen and Huston (cf. Sect. 2.6.1); patients with prefrontal lesions tend towards habitual, stimulus-induced actions and to perseveration (persistent repetitions of movements; Milner 1963), and have great difficulties with naming colors in the Stroop task (Perret 1974). Especially serious consequences of prefrontal lesions are reported by Lhermitte (1983). In this study, he confronted patients with frontal lobe damage with everyday objects and observed that the patients used these objects immediately and without an apparent reason: paper and pencil induced uncontrollable writing, a knife and an apple led to spontaneous peeling and eating, and a cigarette and lighter to the smoking of the cigarette.

The lack of apparent motives was obvious, because patients ate and drank even immediately after their lunch, and they even put on glasses that were placed before them if they did not need any prescription glasses themselves. These observations suggest that the PFC controls the extent to which stimuli evoke automatic reactions, that is, the extent to which overlearned stimulus-response associations determine our behavior.

3.1.3 Control as a Specification of Control Parameters

So, prefrontal goal representations influence information processing through the support of goal-relevant connections between stimuli and actions that are associated with them (also see Duncan 2001). But how exactly do they do this? According to Cohen and Huston goal representations control neither the identification of stimuli nor the selection of actions directly, but the exchange of information between these two functions. Numerous empirical observations are in accordance with this assumption, but one could wonder whether goals cannot cause more. In fact, in their **Executive Control of TVA model** (see Box 3.1) Logan and Gordon (2001) postulate four different ways in which information processing can be influenced by goals.

The implementation of a goal in working memory, so it is assumed, leads to the specification of four variable control parameters: c, β, π, and K (Fig. 3.2). One of these parameters sets which categorization one intends to make (i.e., to what extent

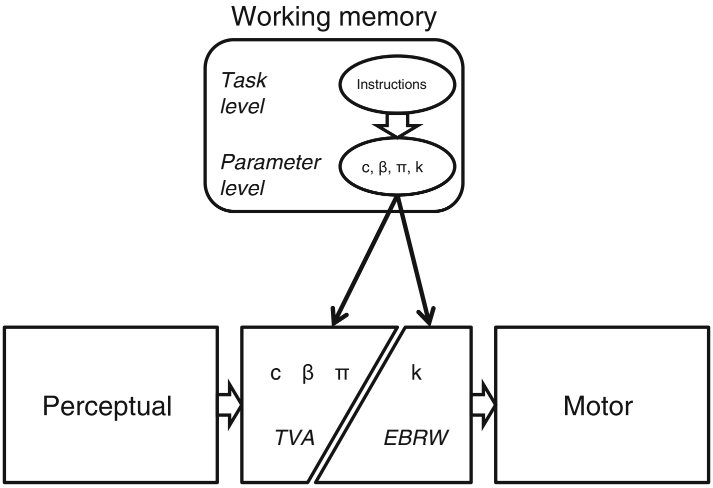


Fig. 3.2 ECTVA-model (adapted from Logan and Gordon 2001; with permission of APA)

Box 3.1: ECTVA Model

The **Executive Control of TVA (ECTVA) model**, an extension of Bundesen's (1990) **Theory of Visual Attention** (TVA) model, was presented by Logan and Gordon (2001). It is based on the idea that task instructions, as given to a participant in an experiment, are stored in working memory (Fig. 3.2). In working memory, the instruction is analyzed in terms of particular control parameters, which serve to inform and control (i.e., parameterize) lower-level cognitive processes. The original model was developed to account for action control in dual task situations, which according to the authors requires the distinction between four parameters.

Parameter β ("beta") stems from Bundesen's TVA and serves to determine the degree to which perceived stimuli are categorized in particular ways. Assume, for instance, you are instructed to look out of a window and overtly categorize each bypassing person as "female" or "male." Performing this categorization task would require the establishment of high values of the β parameters of your representations of women and men, while the β parameter of alternative representations (i.e., of gender-unrelated features of these people) should receive low values.

Parameter π ("pi") also stems from TVA. It serves to select task-relevant objects. If you look out of the window and keep watching for men and women, chances are that you will not only see humans but animals, cars, and bicycles as well. Before categorizing a stimulus as female or male, you will thus first need to determine whether the present stimulus is relevant for your task at all. In particular, you would be willing to focus on humans only, and only in the case that they are actually passing your window, and only as long as you are busy with this particular task. This will be achieved by establishing a high value for the π parameters of the defining features, i.e., "human" and "passing window"

Parameter c serves to control the perceptual organization of complex stimuli. The exact function is difficult to capture in a few words (see Logan 1996), but the basic idea is intuitive. Complex stimuli allow for different grain sizes or foci of attention. Assume, for instance, you are sitting in your living room. You might perceive the room holistically, taking into account many different objects and the overall arrangement, but you may also focus on a particular board, on the books therein, on a specific book, on the title of this book, on a letter of one particular word, or on a stroke of this letter. Depending on how you focus your attention, it will be easy or difficult to register a particular feature. Parameter c determines this difficulty, and it therefore depends on the chosen perceptual organization of the stimulus situation, which in turn depends on the action goal, and on the organized stimulus information.

(continued)

Box 3.1 (continued)

Parameter *K* stems from the decision-making model of Nosofsky and Palmeri (1997). Modern decision-making models assume that people solve decision-making under uncertainty by collecting evidence for the available alternatives. If you for instance are again busy with categorizing the gender of people passing your window, it will often take some time before you are able to determine the correct gender. Unconsciously, you are processing all sorts of features that might be diagnostic for gender, and it often may take some time before you are able to make the correct decision with sufficient certainty. There are indeed not many cues that are so unique that they allow for just one conclusion, just think of the way male and female hairdo changes over the years and fashion waves. If we assume that each piece of information increases the degree to which the corresponding gender representation is activated, this means that decision-making will often involve the activation of more than one alternative (e.g., some cues may suggest that the person is male while others may suggest she is female). It would thus not be particularly reasonable to follow the first tendency, as this may not favor the most informative cues but, rather, those that are processed the fastest. Decision-making will often be better if one first accumulates some amount of information and then picks the most strongly activated (i.e., supported) alternative. Ideally, this representation is much more active (i.e., much better supported) than the competing representations, as this allows one to make the decision with some certainty.

But what does “much more” actually mean? This is where parameter *K* comes into play. Nosofsky and Palmeri (1997) have used *K* to quantify “more,” that is, to express the degree to which the activation of one representation needs to exceed the activation of the others to be chosen. If the value for this parameter is low, this would allow following the first slight tendency, which means that the fastest-processed information has the strongest impact on the decision. If it is high, a lot of evidence has to be collected before the decision is made. Logan and Gordon (2001) assume that acting individuals adjust the value of parameter *K* to the current action goal, so that they can optimize either accuracy (implying a high value of *K*) or speed (implying a low value of *K*).

and how one wants to evaluate stimulus events); a second parameter ascertains the properties that characterize the action-relevant stimulus events (i.e., the properties that can be used to recognize action-relevant events); the third influences how complex stimulus events are structured perceptually; the fourth controls whether one responds quickly and riskily, or slowly and accurately.

Executive Control of TVA (ECTVA) is by no means a general model of human action, but applies to **action control in dual tasks**. A more general model should probably contain considerably more parameters. However, ECTVA is a step in the right direction, to the extent that it makes the generally accepted assumption that action goals influence information processing more concrete, so that it can be tested in detail empirically.

It is interesting to ask which relation exists between the PDP and the ECTVA models. Although none of the original model versions were related to neuronal structures, more recent work suggests that these models are compatible with one another, and that the ECTVA model can be considered to be a more detailed version of the PDP model. As mentioned before, Miller and Cohen (2001) assumed that action goals are represented in the PFC and regulate the information flow between the sensory and motor areas in the brain from there. This concept is corroborated by the most recent attempt by Bundesen et al. (2005) to capture the functional Theory of Visual Attention (TVA) from Bundesen (1990), from which Gordon and Logan took two of their four parameters, in a neuronal model. Bundesen et al. assume that the attention-related parameters β and π modulate processes in the visual system but are themselves located in the PFC.

3.1.4 Action Goal as a Global Mediator

So far, we have only discussed situations in which the action goals are clearly defined, in which the acting person knows exactly what is to be expected, with what stimuli he or she will be confronted, and how to react to them. These situations exist in real life, but they are not very common. More often, we experience situations in which an action goal is **defined vaguely** at first, and then is made more precise through processing of the situational conditions and further thought. For example, when we are looking for a friend we have not seen for many years in a bar, at first we may not have a clear idea of how we are going to distinguish him from the other patrons of the bar (because we do not know that his hair has turned grey and he now wears glasses), how we can search the bar in the most convenient manner, and how we are going to react when we locate him. In this case, we cannot specify the ECTVA-parameters c, β, π, and K adequately. These types of situations are the starting point of Baars' (1988) **global workspace theory (GWT)**.

GWT follows the assumption of James (1890) that the concrete action goal is always **conscious**. Conscious content is located in the so-called **global workspace** (Fig. 3.3a), which has two functions. Firstly, its contents influence the way "specialized processors" that work in parallel operate, which applies to all innate or learned cognitive skills, that is, to the executive processes on the lower level of our working model (Fig. 1.2). In this regard, Baars' model resembles the other models we have discussed so far: goals structure information processing.

The second function of the global workspace is new, however. As the specialized processors are organized in a modular manner, they can process and pass on

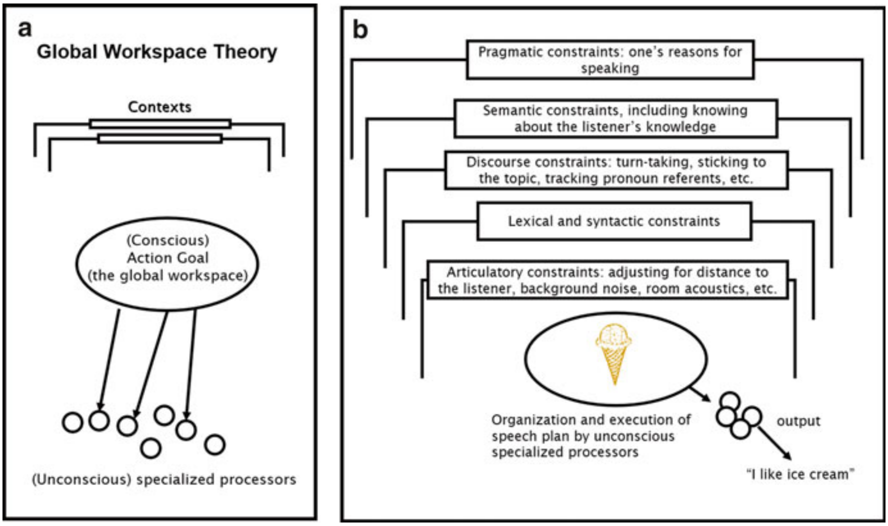


Fig. 3.3 Global Workspace Theory (partly redrawn from Baars 1988, modifications and adjusted terminology ours)

information, but they cannot exchange it directly with other processors. This exchange works indirectly through the global workspace, to which the specialized processors pass on their outcomes and make them available for the entire cognitive apparatus and all involved modules. The specialized processors are, so to speak, the **factories** that transform information into cognitive products, without revealing their production secrets to prospective customers, and the global workspace is the **market** in which the finished products are displayed and sold openly. As the global workspace is global, and serves to exchange information between modules, its capacity is restricted to specific content, an outcome, or a fact—otherwise a module could not "be certain" about which outcome is actually being communicated. According to Baars this is in accordance with the fact that our consciousness cannot have more than one content at a given time.

What is interesting and innovative about the GWT is the assumption that the action goals that are represented in the global workspace can be **adapted continuously**. They are determined by an increasingly abstract, nested hierarchy of goals, which ranges from relatively concrete goals (e.g., "do this without making errors!") to very vague life goals (e.g., "Enjoy life!"). Figure 3.3b gives an example of a verbal action ("I want some ice cream!"). Here, the goal hierarchy ranges from a general expression of an intention (one can express a specific desire) through situational and language-related action conditions, which are all represented unconsciously in the GWT by the way, to concrete, conscious goals images, which finally organize the work of specialized processors. Beyond that, goals are also influenced by the outcomes of the specialized processors (Box 3.2).

Box 3.2: Decision-Making According to Global Workspace Theory

In the **global workspace theory** of Baars (1988), goals are represented in a global workspace, where they can be influenced by top-down processes and by the outcomes of specialized processors. Let's take the example of choosing between two **alternatives**, for example, one of two reactions in a psychological experiment. We make such **decisions** within 100 ms, or even faster, which excludes extensive considerations. Although we are generally aware of the alternatives, we can say little about why we choose either of the alternatives. In the description of mechanisms of response selection in the decision-making model of Nosofsky and Palmeri (1997) and in ECTVA (see above) we already discussed how such decisions come into being: registered evidence for a given response alternative that is derived from perceptual or other processes increases the level of activation of its representation, until one representation reaches a certain threshold and the associated response is executed.

In GWT, this **process** operates as shown in Fig. 3.4. The two alternatives A and B are located in the global workspace, and are therefore capable of becoming conscious in principle. However, because consciousness is limited in capacity with regard to the number of representations it can simultaneously present (to only one event at a time), the two alternatives compete for selection and, thus, for conscious representation. The outcome of this conflict depends on how much **support** the alternatives receive through the processing results of the specialized processors (which continuously accumulate evidence, according to Nosofsky and Palmeri). In the illustrated example, alternative A receives the strongest support, wins the competition for selection and representation in the global workspace (and therefore in consciousness), and thereby causes the execution of reaction A.

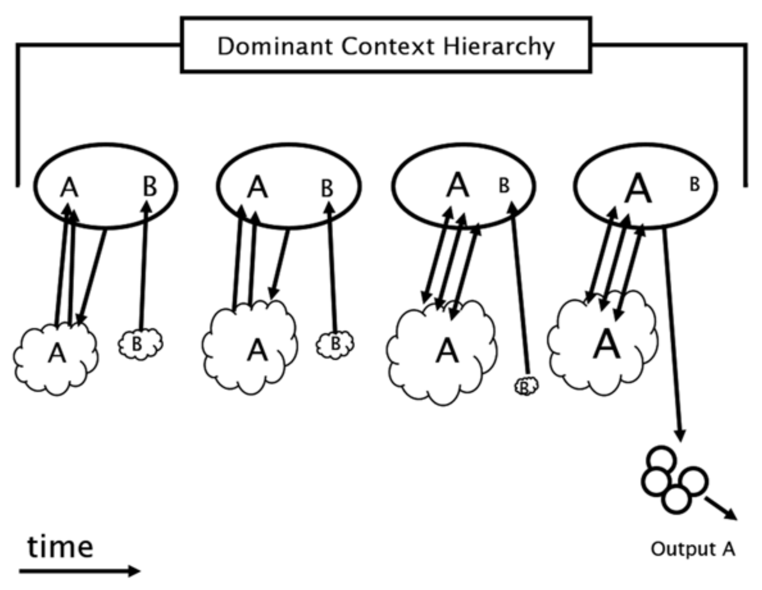


Fig. 3.4 Decision-making according to Global Workspace Theory (partly redrawn from Baars 1988, modifications ours)

3.2 Representation of Action Goals

So far, in our discussion of the function of action goals, we have not truly defined an **action goal**. In the PDP model, an action goal consists, for example, of little more than a representational code (e.g., "name colors" in the Stroop task), and its function within the network (i.e., to support the information stream from color coding to response selection more strongly than the information stream from name encoding to response selection) is actually much more straightforward than its contents. Other attempts and models of action control are not more explicit, no matter how strongly the integrative meaning of action goals is emphasized. With respect to the **contents** of action goals, two important questions present themselves:

- In which format are action goals represented? Must they have a sensory nature, as the discussion of the GWT suggests, or can they be abstract?
- Must action goals truly be represented consciously, to influence information processing?

3.2.1 The Role of Consciousness

When discussing action control, most people (including the authors of psychological theories) assume implicitly or explicitly that effective action goals must necessarily be conscious (e.g., Atkinson and Shiffrin 1968; Baars 1988; Norman and Shallice 1986; Umiltà 1988). Often, the connection between consciousness and action control appears to be so evident that authors always speak of "conscious control," as if the very idea of unconscious control is absurd. For example, Norman and Shallice (1986) contrast automatic, stimulus-driven actions and actions that are under deliberate conscious control, and the only control-related glossary entries in Johnson and Proctor's (2004) textbook on attention are "controlled and automatic processing" and "conscious control," which appears to imply that action goals must be represented consciously to be able to influence information processing. It is definitely plausible to assume that our action goals are always conscious, but is that really the case?

Wegner (2002) has proposed many arguments that cast doubt upon this possibility. When goal-directed actions and the consciousness to act in a goal-directed manner are necessarily related, so he argues, then it should be impossible to:

- Act in a goal-directed manner without being conscious of it
- Believe to act in a goal-directed manner without actually doing so

Actually, many examples show that this is, in fact, possible.

Illustrations of the ability to act without having a consciously represented goal are the so-called **automatisms**. Various authors have reported on people who have a tendency to write automatically (reviews by Hilgard 1986; Koutstaal 1992). These persons produce many short or long handwritten texts, without having conscious knowledge of the contents or reason thereof. Often, they are convinced that they are abused by an “agent” or “higher power” to transmit a message. For example, Mattison (1855, p. 63) cites an automatic writer as follows: “My hand was frequently used, by some power and intelligence entirely foreign to my own, to write upon subjects of which I was uninformed, and in which I felt little or no interest.” Various observations have been made of patients who suffer from the alien hand syndrome (cf. Sect. 2.3). This neuropsychological disorder leads a patient to have normal sensations in a hand, but to be convinced that it does not belong to his or her own body and cannot be controlled by him- or herself. Actually, the hand “behaves” unproductively and independent from the rest of the body of the patient, and occasionally, the hand may even attack the patient. These and many further examples (Wegner 2002) suggest that goal-directed actions like the writing of a text do not require the conscious representation of an action goal.

A comparable conclusion is suggested by a famous experiment conducted by Libet et al. (1983). In this experiment, participants had to move a finger at will. The point in time when they expressed their intention to move the finger was measured. To this end, they were presented with a small hand moving around in a circle (a kind of super-fast clock) and indicated the hand position (“clock time”) at which their intention to make a movement had formed. Additionally, the so-called readiness potential (Bereitschaftspotenzial) was measured through EEG electrodes; this is a cortical signal that occurs a few hundred milliseconds before the commencement of a voluntary movement (Kornhuber and Deecke 1965). Figure 3.5 shows the average time points to which the readiness potential and the conscious experience occurred in relation to the finger movement. It is no great surprise that the readiness potential precedes the movement with about half a second. Much more surprising is the observation that the physiologically measured commencement of an action **temporally clearly precedes** the intention; so, the participants prepared their action long before they wanted to execute it. This finding is hard to reconcile with the assumption that action goals must be conscious.

Other examples suggest that it appears to be possible to **merely believe** that one is acting in a goal-directed manner, but without actually doing so. A particularly dramatic—albeit not particularly representative—example is the reports of persons with amputated limbs. In a study with 300 patients (Henderson and Smyth 1948),

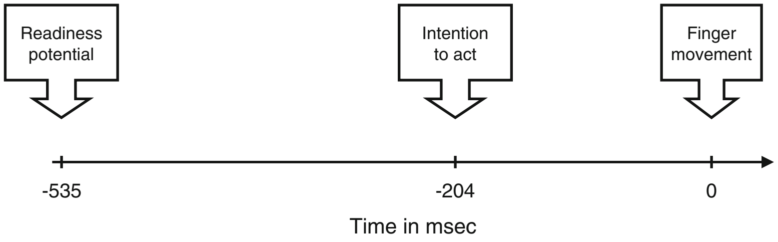


Fig. 3.5 Temporal relations between electrophysiological, volitional, and motor aspects of actions in the experiment of Libet et al. (1983)

98 % reported the experience of movement of the amputated limb ("phantom limbs"). The experienced movements can be voluntary or involuntary and they can be experienced for any limb (fingers, elbows, arms, toes, or legs) and any type of movement.

Comparable illusions of control are also observed in intact individuals. Nielsen (1963) instructed his participants to put on gloves and then to copy figures with a pencil. Their movements were performed inside of a box, and the participants could only see their hands through a hole on top of the box. At least, that is what they believed. In fact, they were looking at a mirror, which actually reflected the gloved hand of the researcher that was next to them. The researcher performed the same movements as the participants but deviated from the correct drawing line in some cases, which led the participants to attempt to correct "themselves" by making compensatory movements. In other words, the participants actually believed the researcher's movements were their own.

A further illustration is derived from Delgado (1969), who was able to stimulate the motor cortex of patients whose skullcaps were lifted for surgery. This stimulation induced head and body movements to either the left or the right side. Delgado asked the patients why they have made a movement, immediately after they had done so. Interestingly, the patients always gave plausible reasons, for example, that they were looking for their shoes, or tried to find the source of a sound, or wanted to look under their beds.

Wegner (2002) indicates that these examples do not really suggest that there is a relationship between the conscious experience of an action goal or a voluntary impulse and the performance of a goal-directed action. The experience of a goal and action control are without doubt highly correlated but, according to Wegner, they do not have a causal relationship. His counter-design is sketched in Fig. 3.6. The critical assumption is that the experience of an intention ("thought") and the execution of an action ("action") are two processes that are caused **independently** of one another. Having an intention is therefore not the cause of an action. However, Wegner believes that it is possible that the actual causes of these two processes ("unconscious cause of thought" and "unconscious cause of action") communicate with one another, so that the execution of an action is generally accompanied by the experience of an intention. Now, if one considers that the performance of an action is normally completed after experiencing an intention, and that the outcome of the action can only be observed at that point, then most actions are accompanied by the outcome sequence "intention precedes action effect." From studies of causal perception (e.g., Heider and Simmel 1944), we know that these types of event sequences are spontaneously interpreted as **causal relationships** when the outcomes are merely correlated (i.e., can also occur independently of one another). Therefore, it is not very surprising that humans have the tendency to consider their intentions to be the causes of their actions. If Wegner is correct, then this consideration rests on an illusion.

It will not be surprising that this radical theoretical approach has evoked skeptical reactions (e.g., Haggard 2005). And not just because the reasoning that is its foundation is completely new: Münsterberg (1888) and Ziehen (1927) had already

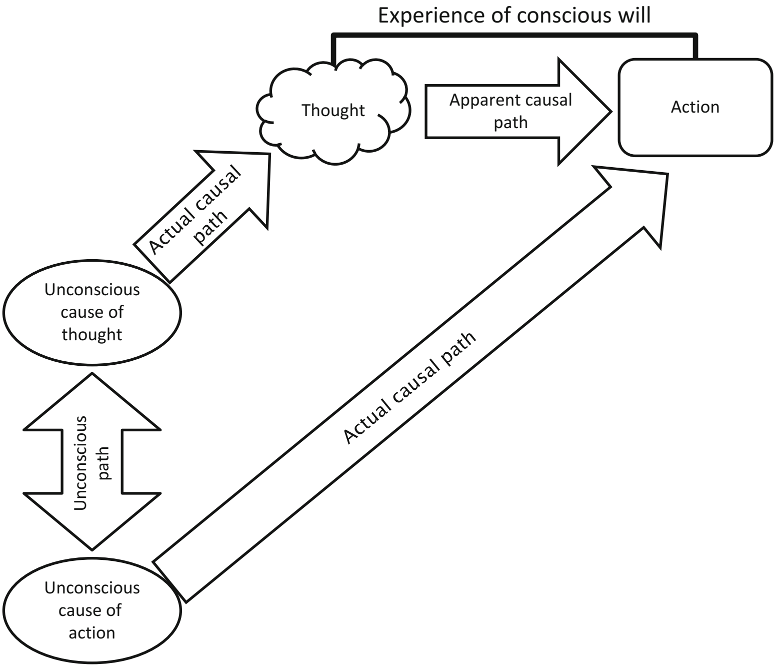


Fig. 3.6 On the parallelism between apparent mental causation and actual causal processes in action production (adapted from Wegner 2002 by permission of MIT Press)

considered the possibility that the actual factors of our action control are of **unconscious nature**. However, our intuition does not correspond with this, and there are better arguments to remain skeptical. For example, Wegner does not discuss how, at what events, and through which factors actions and intention outcomes are resolved. As an illustration, let us assume that the causal processes are completely intertwined with one another, or are even identical, that is: actions and intentions have the same causes. On the one hand, then Wegner's assumption that the experiencing of an intention does not cause an action, would still be correct. On the other hand, it would then be accurate to designate the identified cause as an intention, voluntary impulse, or action-goal, and to assume that "having" or activating an intention is more relevant from a causal point of view than experiencing it. Intentions would then be the cause of human action, whether they are conscious or unconscious.

Wegner's ideas are not incompatible with Baars' (1988) GWT either. Baars postulates a close relationship between the conscious representation of an action-goal and its global availability. From Wegner's perspective, the critical question would be **how close** this relationship would actually be. It would be possible, for example, that goals only (or mainly) exert control over action control when they are represented in the global workspace (as Baars assumes) and that with this type of global representation, the probability of being represented consciously is increased, but that the actual conscious experience does not play a causal role. In other words, the global representation may be more relevant than the question of whether it is connected with conscious experience. To test this possibility empirically is obviously complicated, but not impossible. For example, when it would be possible to find a reliable neuronal indicator for a global processing mode, then it could be measured whether this processing mode only occurs when it is accompanied by conscious experiences (as Baars assumes), or whether exceptions are possible (which would be assumed by Wegner). Actually, Gross et al. (2004) have managed to demonstrate global information processing by using magneto-encephalographic techniques. The challenge is now to combine those methods systematically with reliable measurements of conscious experience.

3.2.2 Format of Goal Representations

We will now turn to the question of how, that is, in what format, action goals are actually represented. In the context of our discussion of GWT, we have already encountered the assumption that concrete goals are defined within a system of abstract goals that can be structured in a complicated, nested manner (Fig. 3.3b). We therefore pursue many general goals in parallel and usually ensure that they are consistent with one another. The most concrete goal, that is, the one that ends up in the global workspace, is always defined in a **sensory** manner in GWT. It also refers to the perceivable effects of a movement, like the experienced amount of pressure, the audible click, and the visible movement of a light switch. That goals are always defined in a sensory manner in the GWT is because the GWT is founded in the ideomotor tradition of Herbart (1825), Lotze (1852), Carpenter (1852), and James (1890) (cf. Sect. 1.2, see Stock and Stock 2004, for a broad historical overview).

Ideomotor theories intend to explain how one can perform a voluntary action, given that one knows little about one's own motor apparatus. At least, that is how Lotze (1852) viewed the theoretical problem that was raised by the work of Herbart (1825). Logically, he assumed that the acquisition of action control does not consist of making the body obey the will, or of achieving more insight in one's own motor functioning. Instead, one is much more likely to be an **observer** of one's own behavior and studies the relationship between one's own mental predisposing conditions (perceptions, conclusions, feelings, etc.) and the primarily reflexive motor responses that flow forth from them (Fig. 3.7). In this way we learn, mostly completely unconsciously, that we, for example, respond to the perception of a given sensory event (s_1) with a specific motor response (m_a), that is, we register specific event-movement sequences ($s_1 \rightarrow m_a$, s2$\rightarrow m_b$, etc.). Mind that the behavior we observe in ourselves in the beginning (e.g., as a child or as a beginner in a given sport) is not voluntary, but **reflexive**, because we do not yet know how we can evoke the observed reactions intentionally. The acquisition of stimulus-response rules offers an opportunity to do so: once we have learned that the movement m_a is

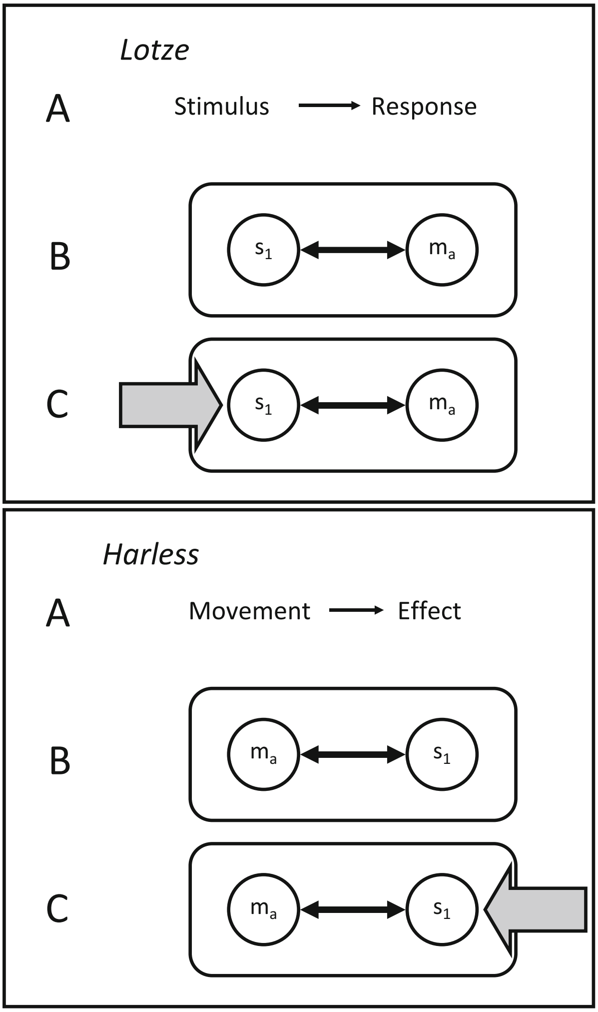


Fig. 3.7 Acquisition of willed actions

often evoked by the perception of a specific stimulus s_1, we only need to simulate this stimulus mentally (i.e., to imagine it) to generate m_a intentionally. This way, one persuades the motor system, so to speak, to generate specific, intended acts, given that one creates the initial conditions that generally evoke this specific action. In the words of Lotze (1852, p. 55; translation is ours):

> Here, as everywhere, the will can only generate those inner psychological states that the course of nature has intended as starting points for the external action; the execution of these actions, however, must be left to their own involuntary force with which those states bring about their consequences.

Harless (1861) focused more on the actual learning process, that is, on the **acquisition of intentional action**. He agreed with Lotze that the motor system cannot be controlled directly at first, and already generates coincidental or reflexive movements prenatally. He too believed that the acquisition of regularity of these movements and their accompanying sensory accompaniments is crucial for the intentional control of actions. Although Lotze mainly discussed the sensory predisposing conditions, that is, the events that evoke a movement, Harless was more interested in perceived movement effects, that is in the events caused by movements. This way, we learn that the execution of a specific movement (m_a) regularly evokes specific perceptions (s_1), for example, that the movement of the fingers evokes a specific visual and kinesthetic, and often also tactile and auditory, perception. The repeated registration of such movement-outcome sequences ($m_1 \rightarrow s_a$, $m_2 \rightarrow s_b$, and so on and so forth) leads to the creation of bidirectional movement-effect associations according to Harless ($m_1 \leftrightarrow s_a$, $m_2 \leftrightarrow s_b$, etc.). This allows us to turn the spear around, so to speak: we can imagine a desired action effect (i.e., the desired perception) and thereby intentionally evoke the movement that is associated with it. Using modern terminology, we could say that the acquisition of associations between motor patterns and their sensory effects makes the representations of these effects to be retrieval cues or primes of the motor patterns. Therefore, we do not need to know anything about our motor system, nor do we need to activate motor sequences directly; we merely activate the representation of a desired effect (i.e., that which we seek to attain with the desired action). This activation is then "relayed" automatically to the associated motor sequence.

The approach of Harless is the basis for the **ideomotor principle** popularized by James (1890), which was forgotten for a long time and has been rediscovered only recently, and has formed the basis for models of intentional action control (Baars 1988; Greenwald 1970; Hommel et al. 2001). The background of this principle clarifies why the concrete, action controlling action goal in Baars' GWT is defined in a sensory manner (ideomotor theories generally speak of a "goal image," but Harless referred to this as "Effektbild," lit. "effect image"). According to the ideomotor principle, only representations of sensory action effects are immediately linked with motor sequences. Abstract action goals are therefore principally inadequate to directly trigger actions; at best, they can be helpful when looking for a concrete, sensory action goal (Fig. 3.3b). Only this concrete action goal is able to interfere with action control by itself. Baars (1988) calls that the "impulsivity" of sensory-defined action goals. The imagination of a movement does, in fact, not activate merely the sensory brain areas that are involved with processing movement feedback, but also the premotor cortex and the supplementary motor area (SMA) (Decety et al. 1994). Even the observation of actions (i.e., their sensory effects) leads to motor activity and to the activation of the muscles that one would use for the execution of the specific movement (Fadiga et al. 1995).

Not all authors assume that actual effective action goals must necessarily be defined sensorially. Above all, Vygotsky (1934) has pleaded for a central role of speech and especially "**inner speech**." He assumed that there is a dramatic change in the relationship between inner speech and action control in the first years of life of a child. According to Vygotsky, shortly after acquiring the ability to speak, speech typically follows an action: young children act first and then describe what they have done. After some time, speech evolves into a more accompanying, commenting role. Children then describe what they are doing at that moment. In the final developmental stage, speech is considered to have a more preparatory, announcing function: older children first say what they want to do, before actually doing it. According to Vygotsky, they do not do this merely for communicative purposes, but because the linguistic designation of an action serves to **control action**. By verbalizing actions and action elements, agents acquire the advantage of being able to extensively plan independently of situational conditions, and verbalized action plans are, when combined with inner, covert or overt speech (i.e., talking aloud), relatively easy to implement and maintain in working memory. A further benefit of speech is that one can use it to more easily encode relationships between action elements (e.g., "first put on shoes and then get out into the garden") (Zelazo 1999).

Accumulating evidence suggests that speech is also involved in the control of non-speech actions (Box 3.3). However, how compatible is the apparently large

Box 3.3: Evidence for the Verbal Control of Action

The first empirical support for the verbal control of action was brought forth by the developmental psychological studies by Luria (1961). For example, in one study, Luria asked 1.5–2-year-old children to squeeze a rubber ball with their hand (measured pressure was the dependent variable) whenever a light was switched on in front of them. As soon as this instruction was given, the children started to squeeze the ball, even before the light signal was presented. According to Luria, the mere processing of the verbal description of action immediately evoked the associated motor action. Therefore, it would appear that in this case, a non-sensory, symbolic representation of an action goal can have "impulsive" qualities, in the terminology of Baars (1988). In a next step, Luria et al. attempted to motivate the children to **encode their action verbally**, by instructing the children to say "go!" whenever they squeezed the ball. Children in the age range of 1.5–2 years were hardly able to perform this task; most of them had already stopped to squeeze when they verbalized their action. Children of 3–4 years old, however, demonstrated perfect performance with the aid of this encoding trick. Luria assumed that in this age range, verbal control functions already function well and the children could therefore use verbal codes for action control.

More recent studies of children and adults support both the assumption that there is a direct link between verbal representations and actions, as well as the presumption that (internal or external) speech plays a central role in the control of actions. Gentilucci et al. (2000) had their participants grasp objects which varied in size, position, and distance to the participant for the various persons participating. There were **verbal labels** on the objects, which were irrelevant to the task. However, the labels were reflected in the execution of

(continued)

Box 3.3: (continued)

the movements: for example, when the label was "far," movements were performed faster than when the label was "near," and the hand was opened wider when the label was "large" than when the label was "small." This suggests that the mere reading of an action-related word suffices to influence corresponding action parameters.

Hauk et al. (2004) presented their participants with action-related verbs that were related to specific body parts, in this case the face (e.g., to lick), arm (to grasp), or foot (to step). Brain activity was measured during the presentations by using fMRI and compared with the activity of brain areas that were activated during the corresponding actions. As Penfield and Rasmussen (1950) had already concluded, the movements of one's own limbs are represented in somatotopic maps in the human brain, in the form of the so-called Penfieldian homunculus (cf. Sect. 2.2). Action-related verbs activate the motor cortex in the exact same manner: verbs like "to step" activated areas that overlap with foot representations, verbs like "to grasp" activate hand-related areas, and verbs like "to lick," face-related areas.

As studies with task switching (cf. Chap. 8) have suggested, **verbal naming** of each action goal is also helpful. In such studies, participants switch to and from between two or more tasks, for example, between naming the color and the form of visual stimuli. Usually, **switch costs** occur in such tasks, that is, participants are slower when they must switch to another task. However, these switch costs are clearly reduced when participants verbally name the new action goal while preparing the task switch, so, for example, say either "color" or "shape" (Goschke 2000). Manipulations that prevent the verbal naming, for example, by having to articulate irrelevant words, have the opposite effect by increasing switch costs (Emerson and Miyake 2003). Interestingly, this is only the case when a task is not or not clearly signaled by external signals; when the external signals are clear (e.g., the words "color" and "shape"), then the disruption of the articulation has no effect on the switch costs (Miyake et al. 2004). This suggests that **internal** or **external speech** during the definition or implementation of an action goal is helpful.

benefit of verbal encoding with the assumption that effective (impulsive) action goals must be **defined sensorially**? On the one hand, it is possible that the advantage of sensory encoding is restricted to the first years of life, and that verbalizing action control gains the upper hand with increasing development. However, studies of the acquisition of action effects (cf. Sect. 3.3) do not suggest that the importance of sensorially defined action goals decreases during development. On the other hand, it is also possible that verbal action control merely enhances sensory action control and thereby expands the human ability to plan. This option would be compatible with Pavlov's (1927) differentiation between a primary signaling system (which is based on sensory representations and associations between them) and a secondary signaling system (which is based on verbal representations). The secondary, verbal signaling system could dock with the primary when associations between sensorically represented action effects on the one hand and the associated verbal descriptions on the other hand are qualified.

3.3 Acquisition of Action Goals

In our previous discussion of the function and representation of action goals, we assumed that humans have these kinds of goals and follow them actively. But where do action goals actually come from? From an ideomotor perspective, concrete action goals are nothing more than representations of initially randomly and **involuntarily** generated movement effects which one uses to generate the appropriate actions **voluntarily**. This implies that the execution of movements is connected to the more or less automatic acquisition of new effects and that the representations of these effects are actually associated with the movements they have produced. These associations between motor movement (m) and sensory effect (s) must be **bilateral** (m ↔ s), so that activations can spread out from m to s (m → s), as well as from s to m (s → m) (Fig. 3.7, lower panel). The basis is that during the acquisition of associations, the relevant representations are activated in reversed order (m → s) with respect to later voluntary activations of the movement representations through the active anticipation of the desired effect (s → m). If this is true, then one should be able to investigate the acquisition of action effects and of movement-effect associations relatively simply: one first combines specific movements with new, initially random effects (hoping to induce movement → effect associations) and then tests whether the presentation of one of these new effects also activates the associated movement (effect → movement), so whether the effects have become effective **movement primes**.

Studies of infants and children suggest that we actually perceive and acquire new movement and action effects from the first years of life (possibly even earlier) and attempt to actively control them (see Box 3.4). Although the acquisition of movement-effect associations already starts quickly after birth (and maybe even earlier), an infant's behavior is very **reactive** and **stimulus-driven** at first. At about 7–10 months a further important developmental stage sets in, which allows the growing child to, for example, **suppress counterproductive grasping and avoidance reflexes** in manual actions (Diamond 1990) and temporarily suppress **competing action goals** (Diamond and Gilbert 1989). This suggests that a relevant action goal can be maintained increasingly better in the working memory (or the Global Workspace in GWT terminology) and can be shielded increasingly well from disturbance by alternative action goals or action tendencies. This developmental leap can probably be traced back to the maturation of the (dorsolateral) PFC, which is important for the representation of action goals; its metabolic circuits take on adult

Box 3.4: Studying the Acquisition of Action Effects in Infants and Children

The experimental work with young participants is often very costly and arduous. Studying children who cannot speak yet is particularly difficult, because they cannot be instructed in conventional manners. Resourceful developmental psychologists have, however, developed possibilities to measure learning processes practically from birth (and occasionally even earlier!). Excellent examples from the area of development of the memory of action effects stem from the group of Carolyn Rovee-Collier.

These scientists have, for example, given infants the opportunity to move a mobile that was hung over their crib with one of their feet. To this end, one of the feet of an infant was wrapped with a band; the other end of the band was connected either to a mobile that was hung over their crib or to an elastic bar (e.g., Rovee and Rovee 1969). In the mobile group, infants were therefore enabled to create new effects with their foot, while in the control group (in which their foot was connected to an elastic bar), the mobile was moved independent of the infant's foot. In further studies of this nature, infants were able to move a mobile by exerting pressure onto their pillow, or to change the pitch of a tone (Rochat and Striano 1999) or the presentation quality of a movie (Kalnins and Bruner 1973) by varying the amount of pressure applied to a pacifier. In somewhat older children action effects are often offered during play, for example, in the controlling of a toy train with a joystick. Here, the joystick movement serves as an action and the movement of the train as an action effect.

Research of this nature has demonstrated two things: first of all, infants are generally more active under conditions of action-dependent feedback. This could suggest that the possibility to control a new action effect works in a motivating manner. However, since the frequency of new events (e.g., the number of mobile-movements) was generally not strictly controlled, one cannot exclude the possibility that the event merely occurred more often in the motion-dependent condition and was therefore more stimulating.

Secondly, and this is theoretically more interesting, the findings suggest the existence of bilateral motion-effect associations: when infants were again confronted with the mobile, they again demonstrated the same foot movements. The experiencing of a contingency between motor movements and sensory events clearly induced a bilateral association between their cognitive representations (bilateral because the associations went in both directions). Accordingly, the further reactivation of the representation of sensory events to an activation ("priming") of the movement that was previously associated with it. Associations between movement- and effect-representations are already acquired in the age of 2 months, and persist for 2 or more days from a very early age onwards (Butler and Rovee-Collier 1989; Fig. 3.8).

(continued)

Box 3.4: (continued)

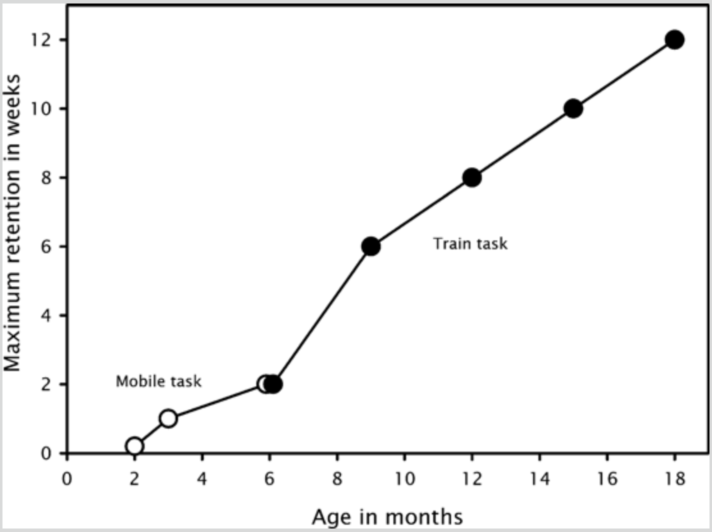


Fig. 3.8 Retention intervals for action-effect relations as a function of age (adapted from Rovee-Collier & Cuevas 2008, by permission of the APA)

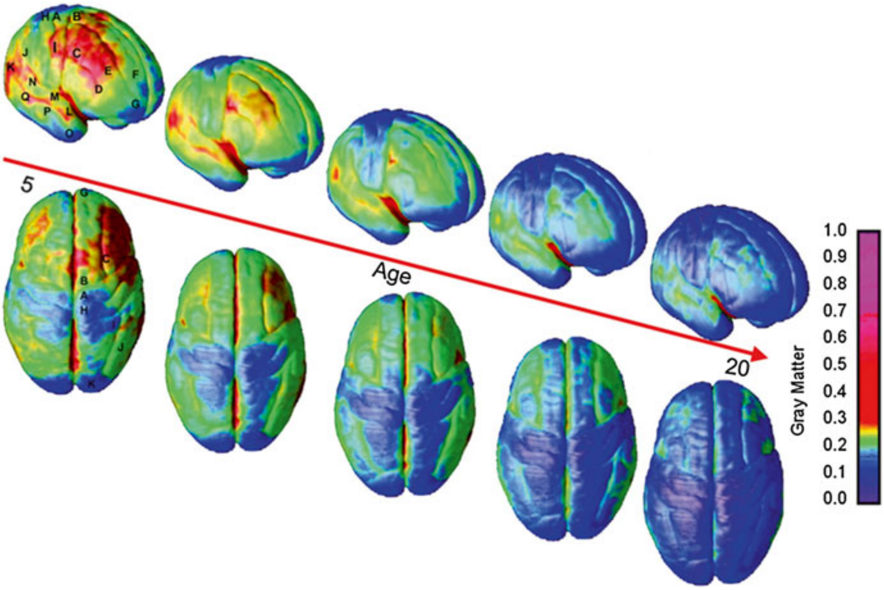


Fig. 3.9 Maturation of the human brain between the fifth and 20th year of life (from Chugani 1994 by permission of Guilford Press)

dimensions with about 8–12 months (Chugani 1994). In comparison to other brain areas, the frontal cortex matures very slowly. As Fig. 3.9 displays, the maturation (myelination) of most other areas is about complete in puberty, while frontal areas maturate well into early adulthood.

With 5–6 years a further important step in the development of action control can be observed: relevant action goals can then be **shielded** even better from competing action tendencies. In this age, that succeeds even when competing tendencies are induced by currently observed stimuli. In other words, children are increasingly able to pursue important interests, even in the face of environmental temptations (Hommel and Elsner 2009). For example, younger children have considerably more difficulties with waiting for a larger reward when smaller rewards are immediately available; however, children of about 5–6 years of age can master this (Mischel and Mischel 1983). In this age, they also have increasingly less difficulties with performing stimulus-incompatible reactions and with suppressing non-desired actions (Gerstadt et al. 1994). An example of that would be the day and night task used by Gerstadt et al. (1994), a child-friendly version of the Stroop task. In this task, children are confronted with picture cards, to which they must respond with a semantically wrong associate. For example, they should respond to a picture of the sun with the word "night" and to a picture of the moon with the word "day." As in the Stroop task, children must suppress their strongly overlearned habits (e.g., to react with the word "day" to a picture of the sun), by strongly activating the representation of the current, incompatible action goal and maintaining it actively. As already mentioned, this activation depends on the dorsolateral PFC, and therefore, it is plausible that the reliability of being activated along with the action goals increases with the maturation of the frontal cortex.

The aforementioned findings suggest that children and infants acquire new action goals very early on and connect them with the associated motor patterns. The **utilization** of these associations for intentional action control begins a bit later and takes many years to reach perfection. The **acquisition** of new action goals is, naturally, not restricted to childhood. **Novices** of a given action or sport experience similar problems to children: they must learn which movements lead to which effects and then produce these effects intentionally.

Numerous studies show that adults acquire new action effects more or less **automatically** (see review by Hommel 2009). In the study by Elsner and Hommel (2001), participants underwent two phases. In the first phase, they pressed left and right buttons randomly. Every button press produced a specific tone, for example, a high tone when the left button was pressed and a low tone when the right button was pressed. The expectation was that this experience would lead to appropriate movement-tone associations, which ought to turn the tones into effective movement-primes. To test this, some participants were instructed in the second phase of the experiment to press the left and right buttons to high and low tones, respectively. Other participants received the reversed instruction, which was incompatible with the practice phase: they should respond to high and low tones with the right and left buttons, respectively. As predicted, the performance in the practice-compatible group was better than in the practice-incompatible group. Although the relationship between buttons and tones was completely irrelevant to the task in the first study phase, participants nevertheless associated the two, which led to an advantage in the practice-compatible condition. Similar studies have demonstrated that adults **integrate** action effects of the most various natures with the actions that cause them (see reviews by Hommel 2009; Hommel and Elsner 2009), such as the position of auditory and visual stimuli, visually presented letters and words, or the affective valence of stimuli. As already discussed (cf. Sect. 2.3), the SMA plays a decisive part in the integration of actions and their effects. During this integration, new connections between action plans in the SMA and representations of the accompanying effects in the sensory cortical areas are made, under mediation of the hippocampus; these connections are reactivated during the perception of these effects, for example (Elsner et al. 2002; Melcher et al. 2008).

3.4 Intraindividual Dynamics and Interindividual Differences

Action goals must be maintained actively to be effective. James (1890) already assumed that alternative action goals steadily compete for action control. An action goal can only win this competition when it receives support from "higher levels," that is, through long-term and stabile **overall objectives** (Fig. 3.3b; Bargh and Barndollar 1996). However, these are not the only factors that determine how strongly an action goal is maintained in working memory. In addition to interindividual differences, we must also assume the presence of **intraindividual fluctuations**. Some of these fluctuations can depend on the outcomes of action monitoring. For example, the registering of a reaction conflict or an error in a response can lead to an adaptation of the degree of activation of an action goal, that is, the action goal could undergo additional support when problems in execution are identified.

The **strength of activation** of action goals can also fluctuate spontaneously. How such fluctuations can be captured statistically and be used for the prediction of behavior was demonstrated by de Jong (2000). De Jong assumes that people in repetitive tasks, such as in psychological experiments with many trials, sometimes are prepared, while at other times they are not. Therefore, in an experiment, there should be trials in which the relevant action goal is activated beforehand, as well as trials in which this is not the case. In other words, the total performance in an experiment, such as the average reaction time within a given condition, should represent a mixture of two statistical distributions: the distribution of the prepared and the distribution of the non-prepared trials. Using the multinomial maximum-likelihood method (MMLM) of Yantis et al. (1991), De Jong could determine the probability with which a data point was drawn from one or the other distribution. This technique allowed de Jong to compute the relationship between prepared and unprepared trials for every individual participant. The preparatory indices that were obtained in this manner allowed a reliable prediction of individual performance in the Stroop task (de Jong et al 1999) and in task switching studies (de Jong 2000).

Altmann and Gray (2002) reported another, more systematic trend in the activation strength of action goals. They performed a task-switching experiment in which every (rare) switch to a new task was followed by 16 task repetitions. Performance declined continuously with repetition, that is, reaction times increased with the number of repetitions. This suggests that an action goal is activated particularly strongly shortly after a task switch (and support the task specific processes especially efficiently), but that this activation spontaneously decreases as time progresses.

Besides intraindividual fluctuations, **interindividual differences** have been studied increasingly in recent years. Duncan et al. (1996) have observed that participants with low scores in tests of general intelligence (Spearman's factor *g* (Spearman 1927)) "lose sight" of action goals more often than others, that is, they forget the criteria of task instructions more often. Moreover, as Duncan et al. also demonstrated, the action patterns of these participants strongly resemble those of patients with lesions in the frontal cortex. This suggests that, to a significant extent, indices of general intelligence measure cognitive functions that are related to the maintenance of action goals and depend upon an intact frontal cortex.

If it is true that interindividual (and possibly also intraindividual) differences are somehow involved in the maintenance of action goals, then one would expect to find a difference between these differences and the individual capacity of working memory. Indeed, participants with good performances in working memory tests also show better performance in tasks in which maintaining action goals is particularly important, such as the Stroop task or task switching tasks (see review by Kane and Engle 2002).

References

Ach, N. (1910). *Über den Willensakt und das Temperament*. Leipzig, Germany: Quelle & Meyer.

Ach, N. (1935). Analyse des Willens. In E. Abderhalden (Ed.), *Handbuch der biologischen Arbeitsmethoden* (Vol. VI). Berlin, Germany: Urban & Schwarzenberg.

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*, 27–33.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *Psychology of learning and motivation: Advances in research and theory* (pp. 89–195). New York: Academic.

Baars, B. J. (1988). *A cognitive theory of consciousness*. Cambridge, MA: Cambridge University Press.

Bargh, J. A., & Barndollar, K. (1996). Automaticity in action: The unconscious as repository of chronic goals and motives. In P. M. Gollwitzer & J. A. Bargh (Eds.), *The psychology of action: Linking cognition and motivation to behavior* (pp. 457–471). New York: Guilford.

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

Bundesen, C., Habekost, T., & Kyllingsbaek, S. (2005). A neural theory of visual attention. Bridging cognition and neurophysiology. *Psychological Review, 112*, 291–328.

Butler, J., & Rovee-Collier, C. (1989). Contextual gating of memory retrieval. *Developmental Psychobiology, 22*, 533–552.

Carpenter, W. B. (1852). On the influence of suggestion in modifying and directing muscular movement, independently of volition. *Proceedings Royal Institution of Great Britain, 54*(1), 147–154.

Chugani, H. T. (1994). Development of regional brain glucose metabolism in relation to behavior and plasticity. In G. Dawson & K. Fischer (Eds.), *Human behavior and the developing brain* (pp. 153–175). New York: Guilford.

Cohen, J. D., & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umilta & M. Moscovitch (Eds.), *Attention and performance* (Vol. XV, pp. 453–456). Cambridge, MA: MIT Press.

De Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance* (Vol. XVIII, pp. 357–376). Cambridge, MA: MIT Press.

De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica, 101*, 379–394.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., et al. (1994). Mapping motor representations with positron emission tomography. *Nature, 371*, 600–602.

Delgado, J. M. R. (1969). *Physical control of the mind: Toward a psychocivilized society*. New York: Harper & Row.

Diamond, A. (1990). Developmental time course in human infants and infant monkeys, and the neural bases of inhibitory control of reaching. *Annals of the New York Academy of Sciences, 608*, 637–676.

Diamond, A., & Gilbert, J. (1989). Development as progressive inhibitory control of action: Retrieval of a contiguous object. *Cognitive Development, 4*, 223–249.

Duncan, J. (2001). An adaptive coding model of neural function in prefrontal cortex. *Nature Reviews Neuroscience, 2*, 820–829.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30*, 257–303.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 229–240.

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., et al. (2002). Linking actions and their perceivable consequences in the human brain. *NeuroImage, 17*, 364–372.

Emerson, M. J., & Miyake, A. (2003). The role of inner speech in task switching: A dual-task investigation. *Journal of Memory and Language, 48*, 148–168.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology, 73*, 2608–2611.

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., & Gangitano, M. (2000). Language and motor control. *Experimental Brain Research, 133*, 468–490.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of 3.5 to 7 year olds on Stroop-like day–night test. *Cognition, 53*, 129–153.

Gilbert, S. J., & Shallice, T. (2002). Task switching: A PDP model. *Cognitive Psychology, 44*, 297–337.

Goschke, T. (2000). Intentional reconfiguration and involuntary persistence in task set switching. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance* (Vol. XVIII, pp. 331–355). Cambridge, MA: MIT Press.

Greenwald, A. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review, 77*, 73–99.

Gross, J., Schmitz, F., Schnitzler, I., Kessler, K., Shapiro, K., et al. (2004). Long-range neural synchrony predicts temporal limitations of visual attention in humans. *Proceedings of the National Academy of Sciences of the United States of America, 101*, 13050–13055.

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences, 9*, 290–295.

Harless, E. (1861). Der Apparat des Willens. *Zeitschrift für Philosophie und philosophische Kritik, 38*, 50–73.

Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in human motor and premotor cortex. *Neuron, 41*, 301–307.

Heckhausen, H., & Gollwitzer, P. M. (1987). Thought contents and cognitive functioning in motivational vs. volitional states of mind. *Motivation and Emotion, 11*, 101–120.

Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology, 57*, 243–249.

Henderson, W. R., & Smyth, G. E. (1948). Phantom limbs. *Journal of Neurology, Neurosurgery and Psychiatry, 11*, 88–112.

Herbart, J. F. (1825). *Psychologie als Wissenschaft neu gegründet auf Erfahrung, Metaphysik und Mathematik. Zweiter, analytischer Teil*. Königsberg: Unzer.

Hilgard, E. R. (1986). *Divided consciousness: Multiple controls in human thought and action*. New York: Wiley.

Hommel, B. (2009). Action control according to TEC (theory of event coding). *Psychological Research, 73*, 512–526.

Hommel, B., & Elsner, B. (2009). Acquisition, representation, and control of action. In E. Morsella, J. A. Bargh, & P. M. Gollwitzer (Eds.), *Oxford handbook of human action* (pp. 371–398). New York: Oxford University Press.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849–878.

James, W. (1890). *The principles of psychology*. New York: Holt.

Johnson, A., & Proctor, R. W. (2004). *Attention: Theory and practice*. Thousand Oaks, CA: Sage.

Kalnins, I., & Bruner, J. (1973). The coordination of visual observation and instrumental behavior in early infancy. *Perception, 2*, 307–314.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin & Review, 9*, 637–671.

Kornhuber, H. H., & Deecke, L. (1965). Hirnpotentialänderungen bei Willkürbewegungen und passiven Bewegungen des Menschen: Bereitschaftspotential und reafferente Potentiale. *Pflügers Archiv, 284*, 1–17.

Koutstaal, W. (1992). Skirting the Abyss: A history of experimental explorations of automatic writing in psychology. *Journal of the History of the Behavioral Sciences, 28*, 5–27.

Lewin, K. (1922). Das Problem der Willensmessung und das Grundgesetz der Assoziation. *Psychologische Forschung, 1*, 191–302.

Lhermitte, F. (1983). "Utilization behaviour" and its relation to lesions of the frontal lobes. *Brain, 106*, 237–255.

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). *Brain, 106*, 623–642.

Logan, G. D. (1996). The CODE theory of visual attention: An integration of space-based and object-based attention. *Psychological Review, 103*, 603–649.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*, 393–434.

Lotze, R. H. (1852). *Medicinische Psychologie oder Physiologie der Seele*. Leipzig, Germany: Weidmann.

Luria, A. R. (1961). The development of the regulatory role of speech. In J. Tizard (Ed.), *The role of speech in the regulation of normal and abnormal behavior* (pp. 50–96). New York: Liverlight.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin, 109*, 163–203.

Mattison, H. (1855). *Spirit-rapping unveiled!* New York: Derby.

Meggle, G. (1993). *Handlung, Kommunikation, Bedeutung*. Frankfurt am Main: Suhrkamp.

Melcher, T., Weidema, M., Eenshuistra, R. M., Hommel, B., & Gruber, O. (2008). The neural substrate of the ideomotor principle: An event-related fMRI analysis. *NeuroImage, 39*, 1274–1288.

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience, 24*, 167–202.

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of frontal lobes. *Archives of Neurology, 9*, 90–100.

Mischel, H. N., & Mischel, W. (1983). The development of children's knowledge of self-control strategies. *Child Development, 54*, 603–619.

Miyake, A., Emerson, M. J., Padilla, F., & Ahn, J. (2004). Inner speech as a retrieval aid for task goals: The effects of cue type and articulatory suppression in the random task cuing paradigm. *Acta Psychologica, 115*, 123–142.

Münsterberg, H. (1888). *Die Willenshandlung. Ein Beitrag zur physiologischen Psychologie*. Freiburg, Germany: Mohr.

Nielsen, T. I. (1963). Volition: A new experimental approach. *Scandinavian Journal of Psychology, 4*, 225–230.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation* (pp. 1–18). New York: Plenum.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological Review, 104*(2), 266–300.

Pavlov, I. P. (1927). *Conditioned reflexes*. London: Oxford University Press.

Penfield, W., & Rasmussen, T. (1950). *The cerebral cortex of man: A clinical study of localization and function*. New York: Macmillan.

Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behavior. *Neuropsychologia, 12*, 323–330.

Rochat, P., & Striano, T. (1999). Emerging selfexploration by 2-month-old infants. *Developmental Science, 2*, 206–218.

Rovee, C. K., & Rovee, D. T. (1969). Conjugate reinforcement of infants' exploratory behavior. *Journal of Experimental Child Psychology, 8*, 33–39.

Rovee-Collier, C., & Cuevas, K. (2008). Infant memory. In H. L. Roediger, III (Ed.), Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference (pp. 687–714). Oxford: Elsevier.

Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge, MA: MIT Press.

Spearman, C. (1927). *The abilities of man*. London: Macmillan.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research, 68*, 176–188.

Umiltà, C. (1988). The control operations of consciousness. In A. J. Marcel & E. Bisiach (Eds.), *Consciousness in contemporary science* (pp. 334–356). Oxford: Oxford University Press.

Vygotsky, L. (1986). Thought and language. Cambridge, MA: MIT Press. (Originally published 1934)

Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge, MA: MIT Press.

Yantis, S., Meyer, D. E., & Smith, J. E. K. (1991). Analyses of multinomial mixture distributions: New tests for stochastic models of cognition and action. *Psychological Bulletin, 110*, 350–374.

Zelazo, P. D. (1999). Self-reflection and the development of consciously controlled processing. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind*. Hove, England: Psychology Press.

Ziehen, T. (1927). *Sechs Vorträge zur Willenspsychologie*. Jena, Germany: Fischer.

Chapter 4
Perception and Action

"In the beginning was the act," Faust concluded in Goethe's (1808/1963) same-named play. About six generations later, Neisser (1976, p. 93) reasoned: "The act of motion, that *requires* more information if it is to be performed successfully, also *produces* more information for the moving perceiver. Were that not to be the case, evolution would hardly have produced moving animals." Both quotes point out, in differing ways, that we do not, or not only act in response to our perceptual experience, but that we actively create the perceptual experiences themselves. Not in the sense that we perceive things that are not there, but in the sense that we actively expose us to some perceptual events but not others, and that we actively use our body to generate these events.

That applies on both small and large scales. For example, whatever you see, you only see because you have first focused your eyes on the stimulus event in question and have moved your body towards it correspondingly. The TV program about which you complain you have most likely switched on yourself. Therefore, to perceive something does not mean to being passively exposed to an external stimulus stream, but to have actively sought out and actively brought your sensory receptors in contact with that stream. In fact, almost all sensory processing presupposes systematic action to provide the sensory channels with the information they need. This implies that perception and action are in a sense two sides of the same sensorimotor coin, which we refer to as perception when emphasizing the fact that it leads to informational uptake and as action to emphasize the fact that it actively changes the relationship between our body and our environment.

Insights into the tight relationship between perception and action had very little impact on psychological theorizing and research until the late 1980s. Instead, psychological thinking about the connection between perception and action was dominated by the idea that actions represent the final component of the sensorimotor arc, which runs through several processing stages from signals processed in the sensory organs to the final contraction of muscles (cf. Sect. 1.2.1). According to this scheme,

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_4

actions are little more than the **consequence** of perceptual processes; they were seen as reactions to the registration and perception of external stimuli and, thus, as being caused by stimulus events. This viewpoint has shown to be **misleading** in various respects. Although sequential stage models of information processing still serve as useful first approaches to understanding phenomena that are new or not yet well understood, severe doubts have been cast on at least two of its implications.

For one, the processing of information is not as unidimensional as the sensorimotor approach assumes. Stimulus information is not first subjected to perceptual processes and then passed on to initiate an action; rather, various perceptual and action-related processes occur **simultaneously**. Information about the environment is concurrently processed by numerous specialized modules and is sent, at different time points, to various action-related modules that construct or specify action plans, in spatially and functionally distributed ways. This is obvious from numerous **dissociative phenomena**, that is, observations suggesting that perception and action are affected differently by demands and situational conditions. This suggests that there is no such thing as *the* perception or *the* action planning, at least not in the sense that perception always precedes action planning and determines its contents (cf. Sect. 4.1).

For another, the processing of information is not as unidirectional and as stage-like as the sensorimotor approach suggests. Not only do different representations of the same stimulus trigger actions at different points in time, but action planning moderates and to some degree controls perception (cf. Sect. 4.2). Hence, perception and action are interwoven more tightly than traditionally assumed. This has spawned new theoretical considerations and approaches that consider **perception and action as two sides of the same coin** (see Sect. 4.3).

4.1 Dissociations of Perception and Action

4.1.1 Sensorimotor Adaptation and Sensorimotor Learning

An obvious way to study the relationship between perception and action lies in changing the nature of this relationship. In the nineteenth century, so-called **prism adaptation experiments** were conducted that did exactly this (Helmholtz 1866; Stratton 1896). One of the first systematic reports comes from the American psychologist George Stratton, who covered his own eyes with a mask in a self-study. The mask fully covered his left eye and inverted the input to his right eye, so that left and right was inverted and the entire visual world turned upside down. Stratton would wear the mask for 7 days in a row and replaced it by eye patches during the night. This experiment was intended to solve an age-old conundrum: how can we see the world upright when the image projected on our retina shows the world upside-down? Is it necessary, so asked Stratton, for the image on the retina to be upside-down for us to be able to see the world upright? Or can the neuronal networks in our brain also adapt to another orientation of the image of the world?

Stratton found that he was able to adjust rather quickly to this “distorted” perceptual world. At the start of the experiment, it was impossible for him to take notes and look at the paper while writing. If he wanted to grasp something, he would often move the wrong hand. When he did not see his arms or legs, he felt them in the usual location, but when they entered his field of vision and he bumped into something, he felt the impact on the side where he saw the leg. Within a few days, these bizarre effects disappeared and on the fifth day, he could move through his house without having to grope everything with his hands. When he took off the glasses, the perceptual and movement disturbances encountered during the starting phase of his self-experiment returned briefly. After 87 h with the reversing glasses, Stratton (1896) concluded that the inverted image of the world on the retina is not required for perceiving the world upright. The brain can also produce coherence between what one sees, hears, and feels from a “wrong” image of the world.

About 50 years later, the Austrian researchers Theodor Erismann and Ivo Kohler (Kohler 1951) have largely confirmed and expanded the findings of Stratton. Their participants wore reversing glasses that inverted up and down and left and right through a mirror system. In the beginning of the experiment, massive disturbances in the control of actions and movements occurred. After just a few days, however, participants could already undertake bicycle and ski tours, and they again appeared to perceive the world as upright. When the reversing glasses were taken off, the aftereffects as described by Stratton occurred. When the participants grasped for an object without wearing the inverse glasses, they would reach to a location in space where an object would have been perceived with the glasses on.

These **sensorimotor adaptation** studies of relatively long duration demonstrate that movement behavior normalizes first and that perceptual impressions follow only later. It thus seems that the control of movements does not depend on the conscious representation of the contents of perception. Not any less interesting is a second observation in these types of experiments: successful adaptations to an altered perceptual world appear to require that humans (and animals) actively explore and interact with their surroundings (Kohler 1951). Hence, perception and action depend on each other but not in the sense that the latter is a consequence of the former. Successful adaptation is not only required in exotic self-studies of the sort we have described but in everyday life as well, albeit in less dramatic proportions. Every wearer of glasses is probably familiar with the phenomenon of seeing the world in a distorted fashion when one wears glasses for the first time or has one’s prescription strength changed. Depending on the strength of the glasses the distortion can be so strong that the control of one’s own movements creates problems, for example, when grasping for objects, walking, climbing stairs, or typing on a keyboard. However, typically, these distortions disappear within a short amount of time. Apparently, our cognitive system can adapt relatively rapidly to the new situation, so that the environment quickly appears to be normal again.

Further scientific studies have investigated why prism adaptation affects perception and action differently. These studies typically consist of three phases: a pretest, an adaptation phase, and a posttest. In the pretest, the participant does not wear a prism, and is asked to point at a perceived object with a hand that is hidden from view.

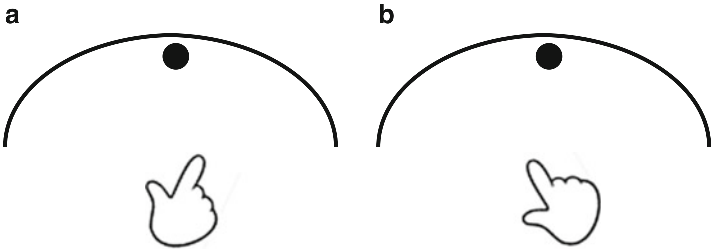


Fig. 4.1 Pointing behavior (**a**) in the adaptation phase and (**b**) after successful adaptation (adapted from Bedford 1993, by permission of Elsevier)

During the adaptation phase, the participant would wear a prism that, for example, displaces objects in the horizontal and/or vertical dimension, inverts the dimensions, or simply distorts the view in irregular ways. If the prism displaces objects to the right, say, pointing at an object with an unseen hand would result in a movement pointing to the right of the actual object (cf. Fig. 4.1a). However, participants are informed about pointing accuracy after each movement, so that the spatial error decreases with every trial and eventually disappears completely. As time passes, the participant will have adapted manually to the prismatic distortion and can point flawlessly at the object, even when still seeing it in the wrong location. The procedure of the posttest is identical to that of the pretest and the difference in performance between the pretest and the posttest serves to determine the adaptation effect. In our example, were the prism made the object appear to the right of its actual location, the participant will point to the left of the object when not seeing it through the prism (cf. Fig. 4.1b). Hence, the participant will automatically overcorrect the pointing motion away from the perceived object even though that is no longer necessary. This demonstrates that the pointing movement depends on spatial information other than that provided by conscious perception.

Observations of this kind raise two questions. The first is based on the finding that perception and action behave independently in these types of adaptation experiments. They are affected differently by the experimental manipulations, adapt on different time scales, and yield different **aftereffects**. Why is that the case and what does this say about the processing of information that is necessary for perception and action planning? (cf. Sect. 4.1.2). The second question is based on the observation of aftereffects. That one can learn, given adequate feedback in the adaptation phase, to adapt one's actions to the prismatic distortions, is not very surprising. But that the acquired knowledge continues to influence the control of actions even when it is no longer useful is more difficult to understand. What knowledge is acquired here and why does it produce these aftereffects?

One possibility is that participants learn new **stimulus-response rules** in the adaptation phase (cf. Chap. 5). For example, when one perceives a stimulus shifted

10° to the right, then one merely needs to learn to point a location 10° to the left of the stimulus in order to give a correct response. This type of learning is not necessarily related to the prism, because clearly one can also learn to point 10° to the right of a presented stimulus without the involvement of any prism. Precisely this condition was studied experimentally by Bedford (1993, 1999), but the pattern of results was considerably different from the observations in adaptation experiments. Apparently, stimulus-response rules do not appear to explain what people find in adaptation experiments—but then what is the explanation?

First of all, it is important to understand which types of representations are affected by the adaptation to a prism. These are the visual representation of the goal location, the proprioceptive representation of the index finger, and the motor command that moves the finger into a specific spatial location. All three representations refer to **spatial locations**. When we see an object in everyday life and then point at it, then these locations typically correspond, which however is not determined genetically but the result of deliberate sensorimotor learning (cf. Sect. 2.4). For example, an infant impossibly knows which perceived location corresponds to which felt location, and which motor command is required to bring the hand in this location. The correlations between visual, proprioceptive, and motor location codes must not just be **learned** but also be continuously **updated** throughout life. All changes in the perceptual stream (such as aging of the eyes or the wearing of glasses) or in the body (e.g., through maturation and becoming an adult or through aging) change these correlations and therefore require a continuous revision of the sensorimotor knowledge. These types of **adaptations** appear to be touched upon in adaptation research. For example, visual perception that appears to be shifted to the right is paired with a new motor command and therefore also with a new proprioceptive perception of the location of the hand. Such adaptations do not necessarily change perception itself (so that the visual stimulus can still be seen as shifted even after the adaptation), but they modify the stored associations between visual, proprioceptive, and motor codes. When the prism is removed, the sensorimotor system assumes that the successful adaptations are still valid (so that the hand movement is "adapted"). As this is no longer accurate, practice is required to update the knowledge.

4.1.2 Pathways of Visual Information Processing

The initially described sensorimotor approach rests, among other things, on the assumption that human information processing can be described as a sequence of processing stages. However, in recent years, this assumption has been doubted increasingly, especially in the light of neurophysiological and neuropsychological evidence that suggests the parallel processing of information from the retina through two separate cortical pathways: a so-called **ventral pathway** that projects from the primary visual cortex (BA17 or V1) into the inferotemporal cortex and a so-called **dorsal pathway** that projects from the primary visual cortex into the posterior parietal cortex (Fig. 4.2). In groundbreaking experiments with monkeys,

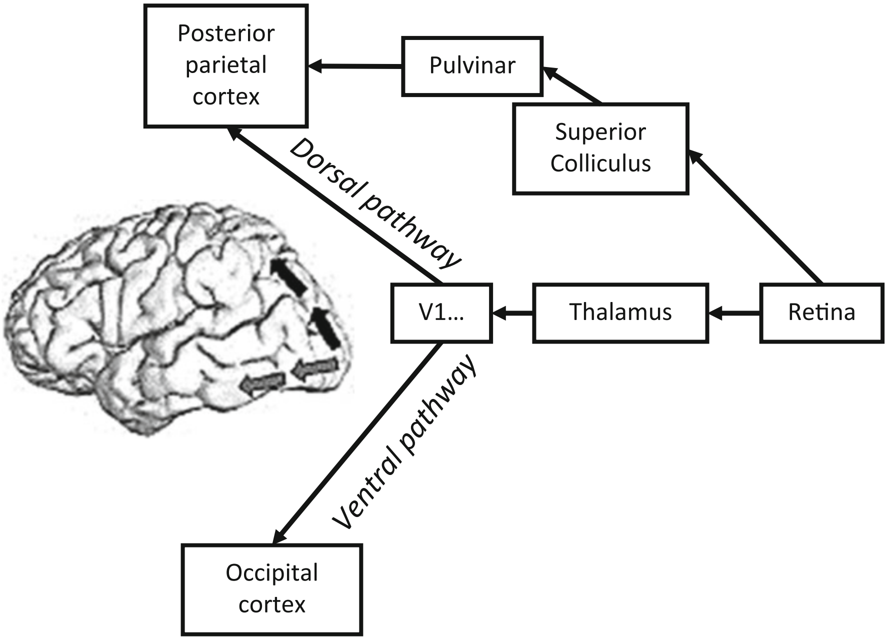


Fig. 4.2 Anatomy of the dorsal and ventral pathway of visual information processing in the human brain

Ungerleider and Mishkin (1982) have shown that the ventral pathway plays a central role in the recognition of objects and provides information about the identity of objects (hence it is also known as the “what” pathway). The dorsal pathway, in contrast, provides information about the spatial location of objects and allows for the localization of objects in space (“where” pathway).

4.1.2.1 Consequences of Selective Damage to Visual Pathways

Neural systems in the ventral processing pathway are specialized in, among other things, the registration and identification of visual stimulus properties like color, form, and texture, but also more complex stimulus configurations, such as faces. Therefore, damage to the ventral processing pathway can lead to massive losses in recognition performance and is often associated with **visual agnosia**. Patients with visual agnosia can often recognize simple visual stimuli, like points of light or lines, but fail to recognize more complicated configurations. They also fail to identify the outlines of objects, nor can they associate colors and objects with each other (cf. Box 4.1). So, visual agnosia does not influence the identification of isolated features of objects but the integration of individual features into a unified and coherent perceptual impression (“percept”).

Box 4.1: Types of Visual Agnosia

Coherent perceptual experiences rely on integrative processes that can be impaired in various, very specific ways. **Apperceptive agnosia** leads to deficits in recognizing and distinguishing simple shapes, such as when telling a square from a rectangle. The inevitable consequence of such a basic impairment is that all subsequent processes involved in object recognition will be affected too. It is suspected that the disturbance is due to a damage of either the neuronal populations that represent the shapes of objects or the connections to information-providing neurons in the primary visual cortex.

Patients with **associative agnosia**, in contrast, have no problems with recognizing and discriminating between shapes, but they find it difficult to identify objects. This problem probably reflects difficulties in assigning the correct meaning to the perceived stimulus configuration, which in turn might be caused by a disturbance in the communication between the neuronal populations that are crucial for object representation and semantic representation.

Patients with **color agnosia** are unable to integrate the colors and shapes of objects, although both the perception of colors and shapes are usually intact. The patients are able to discriminate and identify colors, but they cannot associate them with the shape of objects (e.g., “the red square”) or recall the typical color of an object (e.g., the color of a ripe strawberry). The cause of this deficit is probably a disturbance in the communication between the brain areas that process shape and color representations.

A fourth type of agnosia is related to the recognition of faces. People who suffer from **prosopagnosia** can identify other people by their voice or other typical characteristics, but not by their faces. This deficit could be due to a selective damage of neural populations in the so-called “face area” in the fusiform gyrus within the ventral processing pathway.

Despite strong deterioration in the identification of objects and their properties, agnosia patients display almost flawless performance in the control of actions that rely on visual information. An example is provided by patient D. F., whose inferotemporal cortex was affected by a severe carbon monoxide poisoning (see below). Interestingly, damage to the dorsal pathway often has consequences that, to a certain extent, represent the opposite of those caused by damage to the ventral pathway (review by Milner and Goodale 1995). For example, despite bilateral damage to the posterior parietal cortex, patient R. V. did not have any difficulties in identifying objects that were difficult to identify for D. F. However, R. V. did have massive difficulties with grasping these objects and he was often unable to place his fingers in the required opposite location when he was asked to pass an object to someone else.

Box 4.2: Damage to the Ventral Visual Processing Pathway: The Case of D. F.

Patient D. F. suffered a serious carbon monoxide poisoning, which led to considerable damage in her inferotemporal visual cortex. As she woke up out of her coma, she was completely blind at first, but after a few days, she could recognize colors and shapes again. However, she was not yet able to recognize the shapes of objects or faces, not even her own face or that of her husband. She did not have any problems with recognizing people by their voices or to name objects when they were placed in her hands. It was surprising that she could not indicate the orientation of objects, but she could execute movements perfectly, even if they presupposed knowledge about object orientation (Goodale et al. 1991). For example, she was unable to indicate either verbally or with her hands whether the slit in a letterbox was oriented horizontally or vertically. When she was given a letter and asked to put it in the letterbox, however, she was able to do so without any problems: she brought the letter to the box and turned her hand in such a manner that it aligned with the orientation of the slit in the letterbox. Therefore, her hand "knew" more about the orientation of the slit in the letterbox than her conscious perception did.

The patient did not have any conscious representation of the size or the shape of perceived objects either. If one placed two wooden blocks of differing sizes before her, she could not indicate which was the bigger one. When she was asked to indicate the size of wooden blocks with her thumb and index fingers, there was no systematic relationship between her indication and the actual size of the wooden blocks. However, when she was asked to grasp a wooden block, she moved her hand in a goal-directed manner and spread her thumb and index finger in a way that corresponded with the size of the wooden block. So, she scaled the distance between thumb and index finger with the dimension of the object that she intended to grasp, even though she appeared unable to perceive these object dimensions consciously.

To successfully lift an object, it does not suffice to merely rotate the hand to match the orientation of the object and to adjust the distance between the fingers to the dimensions of the object. One should also place the thumb and index fingers on appropriate, opposite locations of the object, and therefore process information about the appearance and the circumference of the to-be-grasped object. D. F. also mastered this requirement without any name worthy problems; although she could not verbally describe the appearance and contours of such non-symmetrical, rounded objects (Fig. 4.3), she did position her fingers in the appropriate positions around the object and raised them just like a healthy participant would. We can therefore conclude that damage of the ventral path is associated with an impairment of conscious perceptual performance, while visuomotor coordination of manual object handling is unaffected. In other words, conscious perception appears to avail itself of other visual information than sensorimotor action control.

(continued)

Box 4.2: (continued)

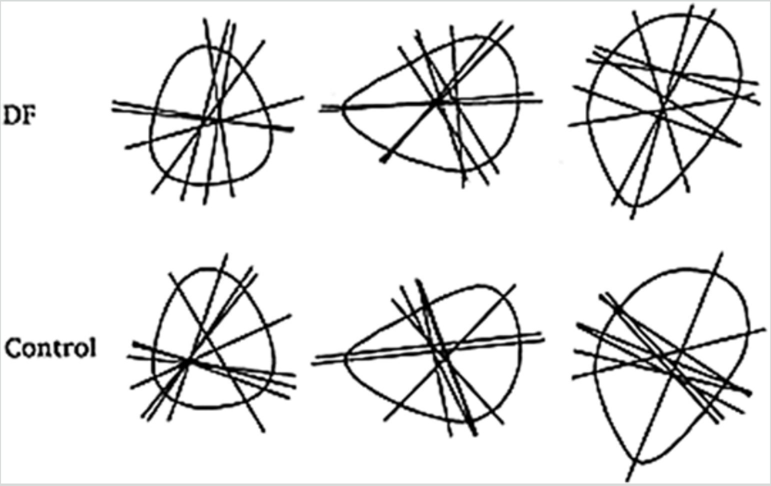


Fig. 4.3 Patient D.F.'s hand orientation when grasping objects (adapted from Goodale & Humphrey 1998, by permission of Elsevier)

These types of problems are typical for patients with dorsal damage, who are often unable to coordinate hand movements in the visual field contralateral to the lesion site (optic ataxia; Balint 1909). These patients generally have difficulties to move their hand in the correct direction and to adjust the orientation of their hand to the orientation of objects, although they can verbally describe the relative location and orientation of objects. They can estimate the size of objects correctly, but when they grasp for these objects, they fail to adjust their hand to its size. We can therefore conclude that damage to the dorsal pathway influences the visuomotor coordination of the manual manipulation of objects without influencing conscious perception. This idea suggests that conscious perception makes use of **other kinds of visual information** than sensorimotor action control does.

4.1.2.2 Milner and Goodale's Perception and Action Model

The findings about the effects of ventral and dorsal lesions reveal a so-called **double dissociation**: ventral lesions influence conscious perception, but not sensorimotor action, while the opposite applies to dorsal lesions. These observations support the suspicion of Ungerleider and Mishkin (1982) that both processing pathways serve distinct functions, although these functions do not seem to fit the distinction between a What and a Where pathway. While it makes sense to assume that the conscious identification of objects relies on What-information, sensorimotor action control requires more than mere Where-information. Accordingly, Goodale and Milner (1992; also see Milner and Goodale 1995) proposed to replace the what/where distinction by one between a **ventral perception pathway** and a **dorsal action pathway**. According to this suggestion, the ventral pathway is mainly responsible for the identification of objects and their properties. Only ventrally processed information is accessible to consciousness (although not all ventral information needs to become conscious), and only the ventral pathway has access to semantic memory—which is necessary for identification. The dorsal pathway, in turn, provides information for the nonconscious visual control of actions. It has no access to memory but provides continuous online information that feeds into action control: information about the size of objects, their orientation and location in space, as well as information about the direction and speed of moving objects.

Milner and Goodale's suggestion cannot just explain the double dissociations observed in patients, but it also allows making a number of predictions about the behavior of healthy participants. It should be possible that one and the same sensory input influences conscious perception and the control of manual actions in different ways. To test this prediction, Aglioti et al. (1995) used the so-called **Ebbinghaus-Titchener illusion**, which arises when one is confronted with two circles as shown in Fig. 4.4. One of the circles is surrounded by a ring of smaller circles, the other by a ring of larger circles. In the standard version of the illusion (a), the two central circles are physically identical, but most people perceive them as having different sizes. In another version of the illusion (b), the two central circles appear to be of equal size to most people, even though the right one is physically larger than the left.

Instead of using drawings, Aglioti et al. (1995) constructed analogue arrangements with discs that the participants had to hold between thumb and index fingers. When both inner discs appeared to be the same size to the participants, they had to lift the left disc, but when they appeared to differ in size, the right disc. In this manner, Aglioti et al. collected (indirect) perceptual judgements, and found that participants were sensitive to the illusion: which disc was chosen as the goal of the grasping action depended on the size contrast between the inner discs and the surrounding discs. Physically different discs were judged as being of similar size, and physically identical discs as differing in size (**Ebbinghaus-Titchener illusion**). The

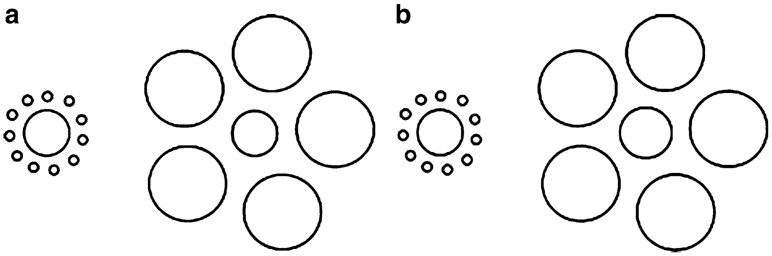


Fig. 4.4 Ebbinghaus-Titchener illusion (adapted from Aglioti, DeSouza, & Goodale 1995, by permission of Elsevier)

question was whether the movement with which a disc was grasped was influenced by its perceived size. This was not the case: the distance between the thumb and index fingers in grasping posture depended on the physical size of the discs and was not modulated by the perceived size. In other words, the visual system supplied action control with more objective information than conscious perception. Comparable dissociations between the perceptual judgment and the control of spatial movements (e.g., during the specification of the target of a movement or the control of the distance between the fingers when grasping an object) were also found with other visual illusions. For example, Bridgeman et al. (1981) induced apparent motion of a static target stimulus by presenting it on a moving background (e.g., a moving point pattern that moved horizontally across the screen). Although participants perceived the stimulus as moving, they could nevertheless point at it with their fingers without any problems.

Goodale and Milner assumed that both processing pathways operate in parallel and independently of each other. This assumption is controversial, though, and the proposed functional division has been criticized too. For example, neurophysiological findings suggest early interactions between the dorsal and the ventral processing pathway (Van Essen and DeYoe 1995). Moreover, experimental observations show that in some cases, control of movements is not nearly as immune against perceptual illusions as was originally thought (review by Glover 2004). For example, when a motor response is not executed immediately, but following a temporal delay (in the range of seconds), then perceptual illusions also influence action control (Gentilucci et al. 1996). Conditions of this nature require the employment of (probably ventral) memory representations for action control, which, in turn, increases the influence of the illusion. This need not speak against the division between a ventral and a dorsal pathway, but it does not seem to fit with the characterization of the dorsal pathway as an "action pathway"—either the dorsal pathway is sensitive to illusions or it is not the only pathway subserving action.

How strongly action control is influenced by visual illusions also depends on the availability of **visual feedback**. Gentilucci et al. (1996) instructed participants to indicate the beginning and end of visual lines arranged in a way that produces the **Müller-Lyer illusion** (Fig. 4.5) by means of their fingers. When fingers and lines were both visible, the effects of the illusion on movement control was mod-

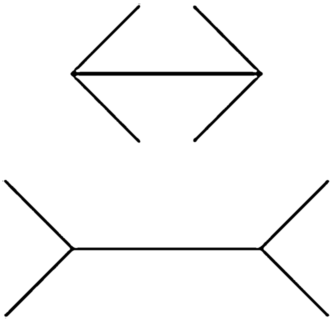


Fig. 4.5 Müller-Lyer illusion

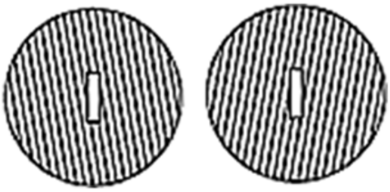


Fig. 4.6 Illusory orientation of a rectangle before a structured background (adapted from Glover & Dixon 2002, by permission of the Psychonomic Society)

est, but these effects increased as visual feedback was reduced. Further studies have demonstrated that some movement parameters are influenced by visual illusions, while others are not. Van Donkelaar (1999) reported, for example, that the accuracy of pointing movements is not influenced by the Ebbinghaus-Titchener illusion, while the movement times are. When participants pointed at circles that looked smaller, then their movement times were shorter than when they pointed at circles that appeared larger. Glover and Dixon (2002) studied changes in the hand opening during the course of movement in the Ebbinghaus-Titchener illusion and the **orienting illusion** (Fig. 4.6). Participants had to use their thumb and index finger to grasp objects of differing orientations that were placed on a background of differently oriented line configurations. The orientation of the hand as it approached the object was analyzed: the effects of the visual illusion on the orientation of the hand were large at the beginning of the movement but became smaller as the movement progressed.

Such observations are difficult to combine with the idea of strongly separated, independently operating processing pathways as propagated by Goodale and Milner. Rather, the two pathways appear to interact in multiple and complicated ways during the planning and execution of actions. Glover (2004) and Hommel et al. (2001; Hommel 2006) have therefore proposed modifications of the original idea of Milner and Goodale, according to which both processing streams can be involved in action control, but with different functions. The ventral pathway would be responsible for the selection and preparation of actions, and for determining their goal-related, relevant properties. The dorsal pathway, in contrast, would continuously feed online information into action control, which would allow specifying the action's "open parameters," that is, the situation-dependent, not-yet-determined of the action (cf. Sect. 6.2.1).

Even though the framework suggested by Milner and Goodale seems to require some adjustments, it has still contributed significantly to understanding the relationship between perception and action control. Most importantly, the framework has provided a useful basis for interpreting dissociations between perception and action, and for understanding the neural underpinnings of parallel perception-action pathways. As we have pointed out, the available evidence does not support the long-cherished assumption of linear models of human information processing that perception and action represent separable and sequentially operating stages of a linear processing stream.

4.2 Interactions Between Perception and Action

There is another basic assumption of linear models of human information processing that has been criticized. Following the logic of linear models, the transition from perception to action requires some translation process that selects the currently relevant information and transforms it into an appropriate response. Doubts concerning this idea were raised when it was demonstrated that not just relevant information is translated into action, but completely irrelevant information as well. In fact, there is considerable evidence suggesting that irrelevant stimulus information can directly activate action tendencies and sometimes even trigger the execution of actions.

4.2.1 The Impact of Stimulus Processing on Action Control

Many observations that show how stimuli can directly affect action selection come from studies on phenomena of **stimulus-response compatibility**. This term was proposed by Fitts and Seeger (1953) to refer to the frequent observation in studies on hardware ergonomics that some stimulus arrangements facilitate action control more than others. For example, in a comprehensive study, Loveless (1962) compared numerous combinations of complex visual displays and manual control arrangements, to see which combinations allow for the best performance under realistic working conditions. It turned out that introducing structural similarities between display and control arrangements produces the best performance. For instance, tasks requiring the control of temperature benefit from combining a display in which temperature runs from left (low temperature) to right (high temperature) with a control arrangement that requires movements to the left to reduce, and movements to the right to increase the temperature. In other words, action control seems to be supported by **spatial correspondence** (i.e., compatibility) between stimuli and responses.

Fitts and colleagues were the first to systematically study such stimulus-response compatibility effects in a series of experiments. Fitts and Seeger (1953) found, for example, that spatially defined joystick responses are particularly fast and accurate if the location of the stimulus spatially corresponds to the action's end location. Fitts and Deiniger (1954) made similar observations with a more symbolic "spatial" relationship between stimuli and responses: for example, participants responded considerably faster when the actions were signaled by symbolically compatible stimuli (3:00, 6:00, 9:00, and 12:00 to indicate left, down, right, and up movements, respectively) than if the stimulus-response combinations were incompatible (e.g., 3:00 and 9:00 indicating right and left movements, respectively) or arbitrary.

Later, Simon and Rudell (1967) showed that spatial correspondence or compatibility affects action control even if stimulus location is **not task-relevant**. For example, if participants are to press a left response key in response to a square and a right key in response to a circle, the location of a stimulus is not relevant for identifying the correct response. And yet, responses are faster and more accurate if the stimulus is presented in a location that corresponds to the required response—a phenomenon known as the **Simon effect**. Note that this effect can be considered a spatial variant of the **Stroop effect**, which can also be interpreted as an effect of stimulus-response compatibility: even though participants are to name the color of words instead of reading them, responses are faster and more accurate if they match the task-irrelevant word.

Various effects of stimulus-response compatibility have been reported over the years (review by Kornblum et al. 1990; Proctor and Vu 2006; Hommel and Prinz 1997). Many are spatial in nature but there are also non-spatial effects other than the Stroop effect. Greenwald (1970) observed, for example, that spoken responses can be produced faster when they are evoked by auditory rather than visual signals, while the opposite holds for written responses. Even though the manifold of compatibility effects may appear confusing, they do follow a rather systematic pattern (Kornblum et al. 1990; Prinz 1990); in general, one can perform actions better and faster:

- If **stimulus and response sets** overlap dimensionally, so that stimulus and response share features on the same (e.g., spatial) dimension; this has been called **set level compatibility** by Kornblum et al.
- If the feature values of **individual stimuli and responses** correspond to each other; this has been called **element level compatibility** by Kornblum et al.

Note that these subtypes of compatibility refer to different units and that the relationship between them is asymmetric. Set-level compatibility refers to **structural similarities** between the ways a stimulus display and a set of responses are organized. If stimulus and response sets are organized in the same dimension, such as if they are arranged from left to right (thus sharing the horizontal spatial dimension), they are set-level compatible. This does not necessarily mean that an individual pairing of stimulus and response is element-level compatible. For instance, pairing a stimulus on the left with a response on the right would be element-level incompatible, even though the arrangement is set-level compatible. However, *set-level compatibility is a requirement for element-level compatibility to emerge*: if stimulus and response sets would not share any dimension (e.g., if responses are signaled by colors appearing at the center of a screen), no stimulus-response pairing is more or less element-level compatible than any other.

Effects of stimulus-response compatibility show that the selection of an action consists of more than the selective translation of the relevant stimulus properties into instructed reactions. For example, if a stimulus appears on the left side of a screen in an experiment, why would one initiate a left response faster than a right response?

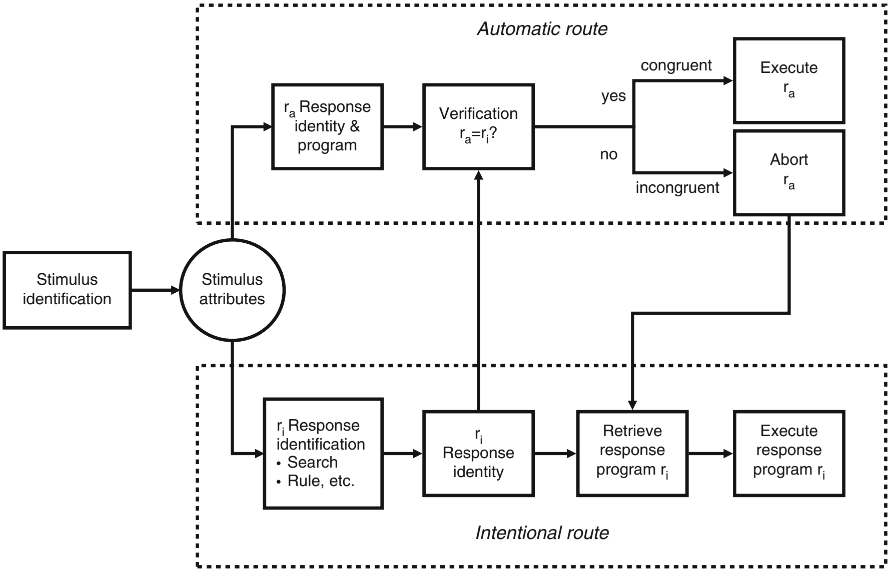


Fig. 4.7 Model of dimensional overlap (adapted from Kornblum, Hasbroucq, & Osman 1990, by permission of the APA)

Questions of this nature have motivated the development of numerous **dual-route models of response selection**. In principle, these models follow Ach's (1910) logic of pitting will against habit. Kornblum et al. (1990) assume, for example, that stimuli in compatibility tasks are processed along two routes: an intentional and an automatic route (Fig. 4.7).

The **intentional route** more or less represents the processing logic suggested by the classical stage model: the relevant stimulus property is used to identify the correct response. This is eventually selected and the corresponding motor program is retrieved and executed. However, along the **automatic route**, both relevant and irrelevant stimulus properties can activate feature-overlapping reactions without the intervention of control processes. If a reaction is activated automatically in this manner, then it must be compared to the intentionally activated reaction. If they match, as is to be expected with stimulus-response compatibility, then the activated reaction is executed. However, if they do not match, then the automatically activated reaction must first be aborted before the intentionally activated reaction can be executed. This takes time, which accounts for a corresponding delay in reaction time—the compatibility effect that is.

The compatibility model of Kornblum et al. (1990) can account for numerous compatibility phenomena, like the Simon effect. If, for instance, a square signaling a left button press appears on the left side of a screen, then the stimulus is processed in two ways: on the one hand, the stimulus shape is identified along the intentional route and translated into the correct (i.e., left) response; on the other hand, processing the stimulus along the automatic route leads to the involuntary activation of the spatially corresponding reaction, i.e., a left button press. The **verification process** confirms that the automatically activated reaction is congruent with the intended response, so that it can be executed immediately. It is a different matter if the square appears on the right side. The processing along the intentional route does not change, but the processing of the stimulus location along the automatic route now results in the activation of the right (and, thus, incorrect) response. The verification process signals that the automatically activated response is incongruent with the intended reaction, so that it needs to be aborted before the intended reaction can be performed. This takes time and these temporal costs are manifested in the Simon effect.

This scenario suggests that irrelevant stimulus information can actually directly activate reactions. How realistic is this assumption? To test this, some authors have made use of the so-called **lateralized readiness potential** (**LRP**; de Jong et al. 1988)—an electrophysiological measure. It occurs shortly before executing responses with the left or right hand and is commonly interpreted as a measure of the activation of a response (Eimer 1997). If it were true that left and right stimuli can activate spatially corresponding actions directly, then they should evoke corresponding LRPs, irrespective of the actually correct response. This can indeed be tested (Sommer et al. 1993). Especially interesting is the result pattern in incompatible trials, if thus stimulus and response do not correspond spatially. As predicted, a stimulus indeed evokes an LRP for the incorrect (i.e., spatially corresponding) response, that is later replaced by the LRP of the correct reaction (cf. Sect. 5.3, Fig. 5.4). Further support for the assumption of **automatic processing** is found in the observation that even nonconsciously perceived stimuli evoke an LRP (Eimer and Schlaghecken 1998).

4.2.2 The Impact of Action Control on Stimulus Processing

The available evidence for the existence of parallel processing streams in human information processing does not quite fit with Descartes' and Donders' (Sect. 1.2.2) characterization of the relationship between perception and action as the conscious translation of encountered information unto subsequent reactions. Another reason to doubt this characterization is the consideration of action as the necessary **consequence** of perception, and the underlying assumption that the perception-action relationship is a **one-way street**. As already explained (Sect. 1.2.3), this assumption overlooks the fact that most actions are **driven by inner causes** (i.e., goals), so that actions are at least as much aiming to create particular stimuli (i.e., intended action effects) as they are triggered by external stimulus events. This suggests that the relationship between perception and action is bidirectional: perception can influence action control but action control can also influence perception. There is in fact ample evidence that the processing and perception of external events is systematically affected by the current action goals and action-control processes. Just like the effects of perception on action control, the effects of action control on perception and attention can refer to the degree of dimensional overlap between stimulus and response sets (as referred to by Kornblum's concept of **set level compatibility**, Sect. 4.2.1) and to the degree of feature overlap between perceptual and action events (as implied by Kornblum's concept of **element level compatibility**). Let us first consider the effects of dimensional overlap.

4.2.2.1 Effects of Dimensional Overlap

In an fMRI study on the ability to detect regularities in visual stimulus sequences, Schubotz and von Cramon (2003) presented their participants with sequences of stimuli that followed particular rules, so that each stimulus was for instance bigger than the one before or that stimulus colors alternated in a particular rhythm. Participants were to discover each rule by themselves and the task was simply to indicate at the end of each sequence whether the final three stimuli followed the rule. The task was thus clearly perceptual in nature and yet, the authors found that each task was systematically activating particular areas in **premotor cortex**. More specifically, the perceptual dimension to which the rule referred was predicting the particular premotor area. For instance, paying attention to shape activated a premotor area that is known to be responsible for the control of grasping movements (BA 6), while paying attention to temporal or auditory stimulus dimensions activated areas associated with the control of rhythmic manual or vocal movements (BA 44). Schubotz and von Cramon suspected the existence of specific bidirectional connections between particular perceptual dimensions and particular movement-control areas, so that activating one component would automatically activate the other.

Support for this assumption comes from the behavioral study of Fagioli et al. (2007), who reversed the perspective: if attending to a perceptual dimension automatically primes an associated action-control area, preparing an action that is controlled by this area should automatically draw attention to the associated stimulus dimension. This prediction was motivated by the assumed share of labor between ventral and dorsal pathways discussed in Sect. 4.1.2. As described, the ventral pathway plays an important role in action preparation but leaves the provision of online information to the dorsal pathway. But how does the dorsal pathway know which information is important for the current action? One way to solve this problem would be to have the ventral pathway biasing information processing towards those perceptual dimensions that provide action-relevant online information, so that stimulus information from these dimensions could be preferred by the dorsal pathway (Fig. 4.8) (Hommel 2010). To test this, Fagioli et al. instructed their participants to prepare a pointing or grasping action, but to not execute it yet. Between preparing and executing, participants had to identify a rule-deviating stimulus from a sequence of visual stimuli, similar to the study of Schubotz and von Cramon. Importantly, the stimulus deviance could be defined in terms of size or in terms of location.

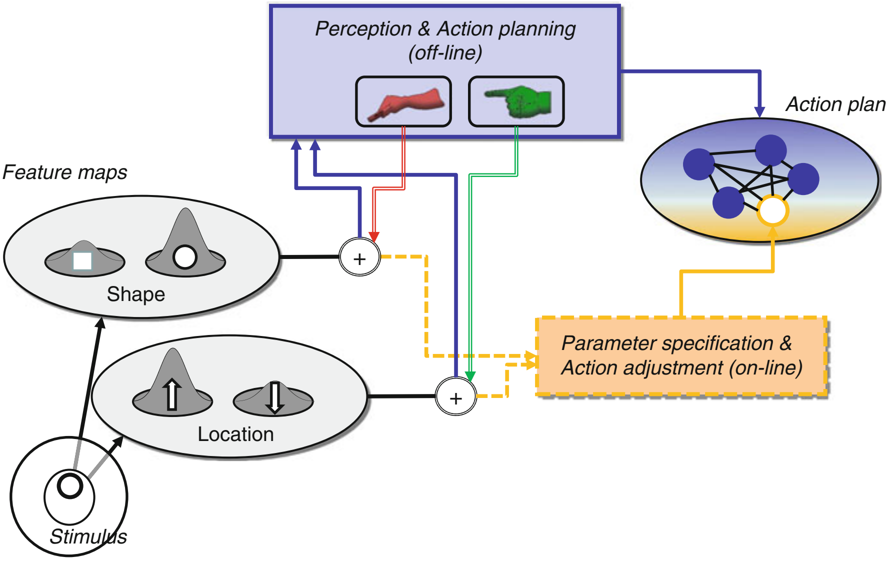


Fig. 4.8 Effects of action control on attention (modified after Hommel 2010, by permission of MIT Press)

As expected, stimulus identification depended on the type of **action preparation**: size-related deviants were recognized faster when having prepared a grasping action while location-related deviants were recognized better when having prepared a pointing action. The preparation of an action therefore appears to facilitate the processing of those features that are particularly important for the fine-tuning of this action.

This mechanism can also account for a number of other observations that imply that action affects perception, such as with the perception of apparent motion. Apparent motion is perceived under conditions that provide cues for motion even though the stimulus is actually stationary. An everyday example can be experienced when watching a train out of the window of another train: if the watched train is starting to move in one direction, one briefly experiences the apparent motion of one's own train into the other direction. A similar, more artificial but better-studied example is the **Barber-Pole illusion**, which is generated by the motion of a stripe pattern behind an opening. If this opening has the shape of a horizontally extended rectangle, the stripes appear to move horizontally (either from left to right or from right to left). If the opening has the shape of a vertically extended rectangle, the stripes appear to move vertically (either from top to bottom or from bottom to top). However, if the opening is quadratic, then the direction of the movement is ambiguous: the stripes can move either horizontally or vertically. Interestingly, the perceived direction is affected by concurrent actions: Moving one's hand(s) in the vertical plane while observing the ambiguous arrangement induces the perception of vertically moving stripes, while moving one's hand(s) in the horizontal plane induces the perception of horizontally moving stripes (Ishimura and Shimojo 1994). Hence, the spatial dimension to which action control refers seems to be used by perceptual processes to interpret the features of ambiguous stimuli.

Actions do not just influence our perception in such rather artificial situations, but also when we observe actual dynamic events. In a large number of experiments, Viviani and colleagues have studied the relationship between the speed of movements with which geometric figures like ellipses or circles were manually drawn, and the movement trajectory that was produced. The authors were able to demonstrate the existence of a lawful relationship between speed and movement trajectory. In particular, speed depends on the radius of the trajectory: the smaller the curve, the slower the movement. Viviani and colleagues have developed a mathematical formula describing this relationship between angular speed and curve, called the **two-thirds power law** (Viviani and Terzuolo 1982; Lacquaniti et al. 1983). A similarly lawful relationship between curve and speed appears to apply to perception: the speed of a point that moves, say, along an elliptical trajectory, is perceived as uniform only if the speed follows the same laws that apply to the production of such a movement. Conversely, the same motion will not look uniform when it is presented with constant speed. The same applies to linear motion, which is perceived as constant only if there is an acceleratory phase in the beginning of the motion (Mashhour 1964; Rachlin 1966; Runeson 1974; Viviani and Stucchi 1989). It thus seems that motion perception is affected by (not necessarily conscious) knowledge about how to produce the particular motion oneself, suggesting that action informs the perception of dynamic events. In other words, the perception of dynamic events appears to be based not just on the physical properties of perceived motion, but more so on the interplay between sensory information and the principles of the motor systems.

4.2.2.2 Effects of Feature Overlap

In addition to such demonstrations that perception is affected by dimensional overlap between action and stimulus sets, there are also observations suggesting the direct interaction between action- and perception-related features (set level compatibility). Depending on the task and the circumstances, such interactions can result in facilitation or inhibition (see Sect. 4.3.2). An example of **facilitation** was reported by Craighero et al. (1999). These authors instructed participants to grasp an object that was oriented to the left or right, but only after being presented with a go signal. The specific properties of this signal were actually irrelevant to this task and could safely be ignored. However, in some trials it showed an object that was oriented the same way as the to-be-grasped target and in other trials it showed an object with a different orientation. In other words, the orientation of the go signal was either compatible or incompatible with the grasping action. It turned out that the grasping action could be executed more quickly when go signal and action were compatible. To test whether this effect was due to more efficient preparation of the action or more efficient perception of the go stimulus, the authors had participants occasionally respond with their foot—i.e., manual and pedal actions were mixed. Foot responses were also executed more quickly, suggesting that preparing the grasping movement facilitated the perception of compatible go signals.

Perceiving a particular feature is typically facilitated by action preparation in tasks in which the perceived stimulus is functionally related to the action, such as in the case of the Craighero et al. study. If it is not, **interference** is more typical. For example, Müsseler and Hommel (1997) had participants prepare manual actions (pressing a left or right button) but to withhold it until a go signal was presented. In between, participants were presented with an arrow pointing to the left or right, and they were to identify the arrow's direction. The arrows were presented only very briefly and were then masked with a random pattern, so that they were very difficult to see and could only be identified with considerable effort. The question in this experiment was whether arrow identification would be affected by the compatibility between the prepared action (left or right button press) and the direction of the to-be-identified arrow. Interestingly, arrow identification was worse for compatible than for incompatible trials: An arrow pointing to the left (right) was almost impossible to see after having prepared a left-hand (right-hand) button press. These effects are relatively specific to the relevant features. For example, preparing a left-hand movement hinders the perception of a left-pointing arrow but not of the word "left," while the opposite is true for preparing to say the word "left" (Hommel and Müsseler 2006). Such negative compatibility effects can also be found under less artificial circumstances. For example, Hamilton et al. (2004) presented their participants with short videos showing a hand that grasped an object and then put it on a tray (Fig. 4.9). Participants were to estimate the weight of the object while grasping objects of various weights putting those on trays themselves. Interestingly, participants tended to overestimate the weight of the perceived object when lifting a lightweight object but to underestimate the weight of the perceived object when lifting a heavy object. Again, we see that planning and executing an action has a systematic impact on perception—an impact that tends to be facilitatory in nature if the perceived event is functionally related to the action and inhibitory in nature if it is not.

4.3 Integration of Perception and Action

In his book *Cognitive Psychology* that was published in 1967 and that became a manifesto of cognitive psychology, Ulric Neisser suggested that cognitive processes should be analyzed by tracing the stimulus through the cognitive system. According to this approach, human cognition is understood as the consequence of external stimulation, which triggers increasingly complex information-processing procedures. While this conceptualization of cognition as a linear processing chain that is initiated by the stimulus fits with the ideas of Descartes and Donders (Sect. 1.2.1), it is not truly consistent with the empirical observations reviewed in the

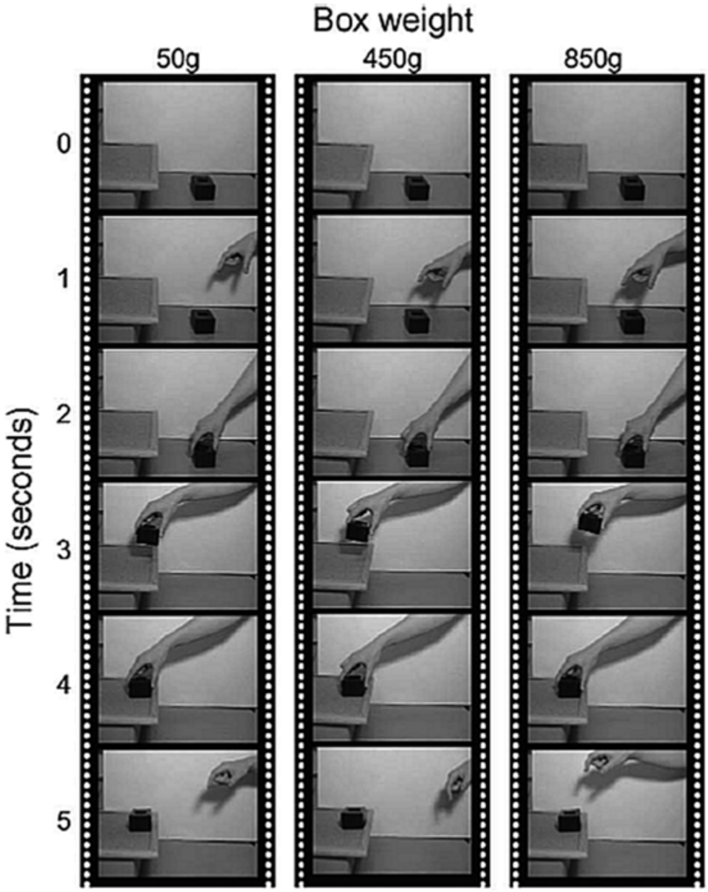


Fig. 4.9 Estimating the weight of objects while lifting boxes of different weights (from Hamilton et al. 2004, by permission of Elsevier)

present chapter. In fact, the failure of the linear, stimulus-centered approach was anticipated not much later by Neisser himself in his however less influential book *Cognition and Reality*. As we have pointed out, stimulus information is processed through multiple pathways that operate in parallel and that have different functions in perception, action planning, and movement control. We also have seen that multiple actions can be activated by external information in parallel, which is inconsistent with the idea of selective stimulus-response translation, and that action preparation and execution can impact perception and attention. All this does not seem to be captured well by the traditional unidirectional stimulus-response approach. But how else can the available findings be theoretically integrated in a meaningful way?

4.3.1 Theory of Event Coding

A framework that allows this integration was developed by Prinz (1990, 1992) and his colleagues in what was then the Max Planck Institute in Munich, and it became known as the **Theory of Event Coding**, abbreviated **TEC** (Hommel et al. 2001). At present, TEC is the most comprehensive attempt to replace the classical stimulus-response model by a model of human cognition and action control that does not make the stimulus the starting point of analysis. Instead, it conceptualizes humans as active agents that have particular goals, which they transform into overt actions whenever situational circumstances allow. The key question is thus no longer how people respond to stimuli but how they transform action intentions into actions, given suitable stimulus conditions. TEC makes three basic assumptions, which we will explain in more detail below:

- Perceived events and produced events (one's own actions) are coded in a shared representational medium ("common coding"; cf. Prinz 1990).
- Events are represented through codes of their features; events are therefore represented in a distributed fashion and not, for example, through abstract symbols.
- These feature codes refer to the distal properties of the represented events, but not to the proximal properties of the system registering or producing them.

4.3.1.1 Common Coding of Perception and Action

Let us first turn towards the first of these assumptions, which deviates particularly strongly from previous models. Why and in what sense should an action be represented like a perceived event, from a cognitive viewpoint? TEC can be placed in the tradition of ideomotor approaches (Lotze 1852; Münsterberg 1888; James 1890; review by Stock and Stock 2004) (Sects. 3.2.2 and 3.3). They assume that we can experience our motor skills only **indirectly**, namely, by perceiving the sensory effects they produce. We thus actually have no direct access to our motor system, so we first have to **learn** which sensory effects we can produce by activating particular motor patterns (Elsner and Hommel 2001). Let us play through this process, which is captured in Fig. 4.10.

Every body movement is produced by the activation of particular patterns of muscle activities, that arise more or less by chance in newborns or newcomers to a certain sport, but that become more systematic and purposeful with increasing development or practice. In Fig. 4.10, the network labeled MP (for motor pattern) signifies all the representations that are activated for realizing a particular movement. TEC and ideomotor theory assume that this network is necessarily activated **randomly** for the first time. After all, one cannot normally know what consequences an action will have before having executed it at least once. Without knowing the consequences, one cannot logically execute it *to attain* specific consequences, which again represents a logically necessary condition of goal-directed action

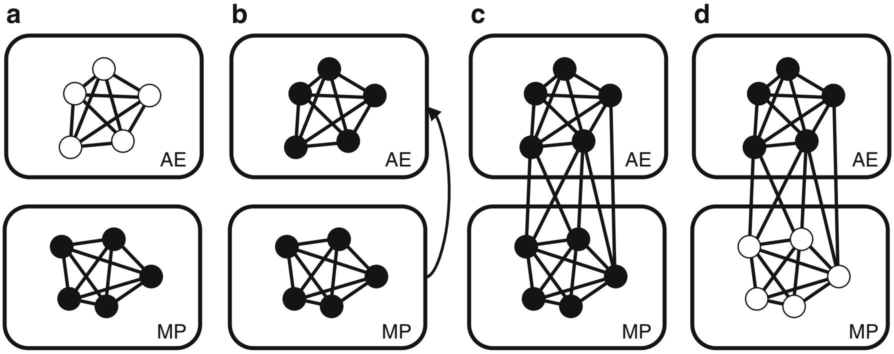


Fig. 4.10 The development of sensory-motor structures underlying voluntary actions (from left to right)

(i.e., an action is goal-directed only to the degree that it is performed to reach that goal). The first time, MP is therefore activated, and the respective movement is executed, without knowing the consequences (by coincidence or through a reflex) (Fig. 4.10a).

Movements often produce a large number of perceivable effects that are manifested more or less reliably during or following the execution of a movement. For example, when we operate a light switch, we feel and see the movement of our fingers and the light switch, and we perceive the visual changes that follow from switching the light on or off. According to TEC, we represent all of these effects of our actions through cognitive **feature codes**, that is, through a number of codes that represent properties of the sensory effects that follow from the executed actions (action effects or AE). When you execute a movement, two types of representations are activated in close temporal proximity: the motor pattern that produces the relevant movement and the codes representing the thereby created and perceived action effects (Fig. 4.10b). From learning studies we know that neural structures that are activated close in time, so that their activation overlaps, become associated by means of the Hebbian principle: "what fires together wires together" (Hebb 1949). This association leads to the integration of motor representations and the codes of the action effects that are realized by the corresponding action; so, both networks (MP and AE) become one functional unit (Fig. 4.10c).

Through the integration of motor representations and sensory effects, **sensorimotor units** are created. Such a unit represents the action and provides the individual with indirect but intentional cognitive access to the action. From now on, the individual can activate the action voluntarily by simply imagining the desired consequences (i.e., by activating the codes of the desired action effects endogenously; Fig. 4.10d). This activation spreads to the associated motor pattern, so that motor activity is from now on under voluntary control. Moreover, the created sensorimotor unit in a sense represents both the perceptual event one can from now on actively produce and the action necessary to produce it, and it is this co-representation of perception and action that the common-coding principle of TEC refers to.

4.3.1.2 Distal Representation of Perception and Action

These considerations already provide some insight into why TEC assumes that events are represented in a distributed manner (that is, through networks of feature codes as opposed to single symbols) and that these codes refer to the distal properties of the represented event. The concept of **distal representation** was introduced in perceptual psychology by Heider (1926), who contrasted it with the concept of **proximal representation**. We commonly assume that objects in our perceptual world have a specific location, form, color, and other properties that can be measured physically, independent of our perception. These properties can be perceived visually when they reflect the ambient light in a specific manner and when these reflections reach our retinas. The light waves are registered by receptors and transformed into electrical signals that eventually reach our brain and somehow trigger the conscious perception of an object.

Heider indicated the complexity of this process and pointed out an interesting phenomenon: without exception, our conscious perception refers to the distal properties (i.e., those properties that can be measured in the physical world), even though there is no obvious geometric-physical relationship between the three-dimensional distal properties and the two-dimensional proximal properties (i.e., the representation of an object on the human retina). An example of this **ambiguity** is the fact that large objects in the distance are represented similarly on the retina as smaller objects that are nearer. Not any less ambiguous is the relationship between the internal representation of proximal properties (that result from the retinal activation pattern) and the internal representation of distal properties.

As Brunswik (1944) pointed out, a very similar theoretical problem exists for goal-directed action. The cognitive representations of action goals invariably refer to their distal properties (i.e., the properties of action consequences that can be measured objectively such as the new location of a hand after moving it), but not to the proximal properties (the muscle commands that are required for movement execution), even though the proximal properties cause the distal properties. We cannot resolve this rather complicated theoretical problem here (cf. the more detailed discussion in Prinz 1992), but it seems clear that cognitive representations refer to the distal, but not the proximal properties of perceived and produced events. To explain how one can produce proximally defined motor patterns from distal action representations, TEC assumes that codes of distally defined action effects are integrated with proximally defined codes of motor patterns that occur in close temporal proximity.

4.3.1.3 Distributed Representation of Perception and Action

TECs assumption that events are represented in a distributed manner is motivated by neuroscientific evidence. The human brain does not represent external events through single neurons or local neuronal networks, but through widely distributed neural activity (Sect. 2.1). It is known, particularly for the visual system, that the features of perceptual events are coded in parallel in various **feature maps**

(DeYoe and Van Essen 1988) and the same is true for features processed in other sensory modalities. TEC generalizes this principle and assumes that all perceived and produced events are represented through networks of codes of their features. An important implication of this assumption is that it allows for graded similarities between representations (as representations can share none, some, or many feature codes), which allows for various degrees of feature overlap between event representations. This holds for different perceptual representations and for different action plans, but also for perceptual events and action plans—just as assumed (but not explained) by Kornblum et al.'s (1990) dimensional-overlap model. The assumption of representational overlap is particularly relevant for explaining interactions between perception and action, as we will see later on.

As we have discussed in Chap. 3, the current **action goal** plays an important part in controlling and regulating cognitive processing. In TEC, action goals have a key function in moderating the relative contribution of particular feature codes to representing a given event. The **intentional weighting** principle underlying this moderation was already touched upon in Sect. 3.1.1. To explain the Stroop effect, Cohen and Huston (1994, Fig. 3.1) have assumed that the action goal "name color" increases the influence of those codes that represent the color of a stimulus, while the action goal "read word" increases the influence of word codes. Very similarly, TEC assumes that action control does not just select and activate action codes describing the relevant features of the appropriate action, but it also enhances the processing of feature dimensions that provide goal-relevant information. For example, intending to grasp a cup of coffee with your preferred (dominant) hand will lead to the facilitation of spatial information regarding the cup's location and shape, and of the orientation of the cup's handle. After all, the hand must "know" where it should go and the fingers what they are supposed to grasp.

So far, these considerations might sound rather abstract and perhaps even counterintuitive. We will therefore conclude by applying TEC to a few theoretically interesting empirical findings that either guided the development of the theory or were discovered when testing novel predictions derived from it.

4.3.2 Empirical Implications of the Theory of Event Coding

4.3.2.1 Simon Effect

One of the many examples demonstrating a close connection between perception and action is the previously described **Simon effect** (Sect. 2.1). This effect manifests itself when, for instance, participants are to press a left or right button in response to a visual square or circle, respectively. As we have described earlier, presenting the stimuli randomly to the left or right of some reference point produces a spatial compatibility effect: participants are faster if the square appears on the left and the circle appears on the right. Given that TEC emphasizes individual goals, a TEC-motivated analysis of the Simon effect begins before the first stimulus is

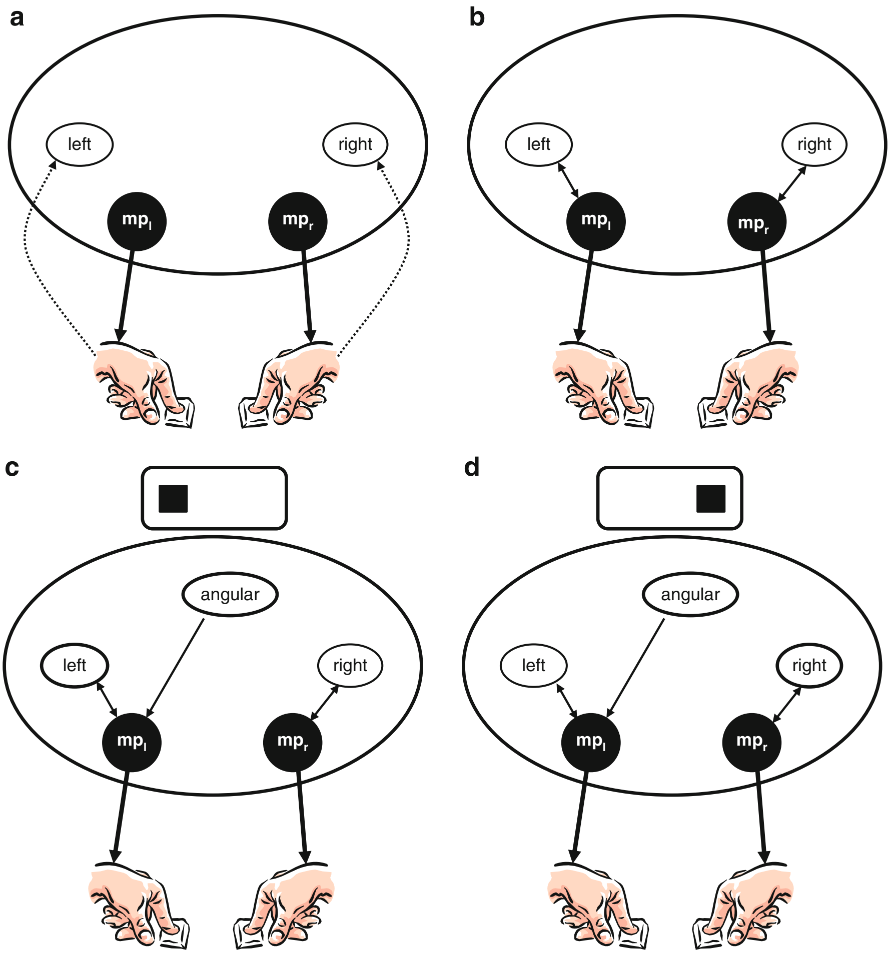


Fig. 4.11 Analysis of the Simon-effect according to TEC

perceived and the first response is performed. To respond effectively in a Simon task, participants need to prepare their cognitive system according to the instruction. Even though that might sound trivial, participants need to know how a left or right button can be pressed and they need to make use of this knowledge when preparing for the task. According to TEC, the participant must have previously learned that moving his or her left or right index finger by activating the corresponding motor patterns (mp_l and mp_r) causes left or right sensory effects, respectively (as indicated in Fig. 4.11a). This learning process should have established an associated structure as sketched in Fig. 4.11b. (Note that, for the sake of simplicity, we ignore other codes, for example, those representing the identity of the fingers and the hand.)

Now, when, for example, the square is presented on the left of the screen, this will lead, among other things, to the activation of codes representing shape ("angular") and the spatial stimulus location ("left"). To be able to perform the task correctly, the participant must have successfully associated the shape code with the appropriate motor pattern (Fig. 4.11c, d). Being presented with a square would thus activate the shape code "angular," which in turn will activate the associated motor pattern mp_l. But this is not all that happens: the stimulus will also activate the code representing its spatial location, which in turn will activate the appropriate motor pattern mp_l if it appears on the left (Fig. 4.11c) but the wrong motor pattern mp_r if it appears on the right (Fig. 4.11d). If we assume that motor patterns are activated faster if being concurrently activated by more than one feature but more slowly if a competing motor pattern is also activated, we can explain why the Simon effect exists. What is more, this reasoning can be easily generalized to other effects and observations. The general principle of code overlap is no way restricted to spatial tasks but can be generalized to any task in which stimulus representations and response representations overlap with respect to their feature codes (Sect. 4.2.1).

4.3.2.2 Reversal of the Simon Effect

The way in which TEC operationalizes the influence of action goals on the interplay between perception and action can be demonstrated particularly well by analyzing an experiment conducted by Hommel (1993). In this experiment, participants responded with left or right button presses to low and high tones that were presented randomly through left or right loudspeakers—a classical Simon task. However, there was something special about this task: each response button was connected with a light-emitting diode (LED) on the opposite side, just as indicated in Fig. 4.12, so that pressing the left button would switch on a light on the right side and pressing the right button a light on the left side.

The task was performed by two groups of participants. Both would perform the exact same task but they were instructed in different ways. The first group was asked to respond to low tones by "pressing the left button" and to high tones by "pressing the right button," and the two lights were not mentioned in the instruction. In contrast to this "button instruction," the second group was asked to respond to low tones by "switching on the right light" and to high tones by "switching on the left light"—a "light instruction." As the right light was switched on by pressing the left key and the left light by pressing the right key, an outside observer of the following task performance would not be able to judge which group received which instruction: both groups were physically carrying out the same task. And yet, the outcome shows that the different instructions must have induced different representations of the task. With button instruction, responses were faster if the stimulus for the left response was appearing on the left side, or the stimulus for the right response appeared on the right—a standard Simon effect that is. Light instruction inverted this pattern entirely, however, responses were faster if the stimulus for the left response appeared on the right and the stimulus for the right response appeared on the left. How is that possible?

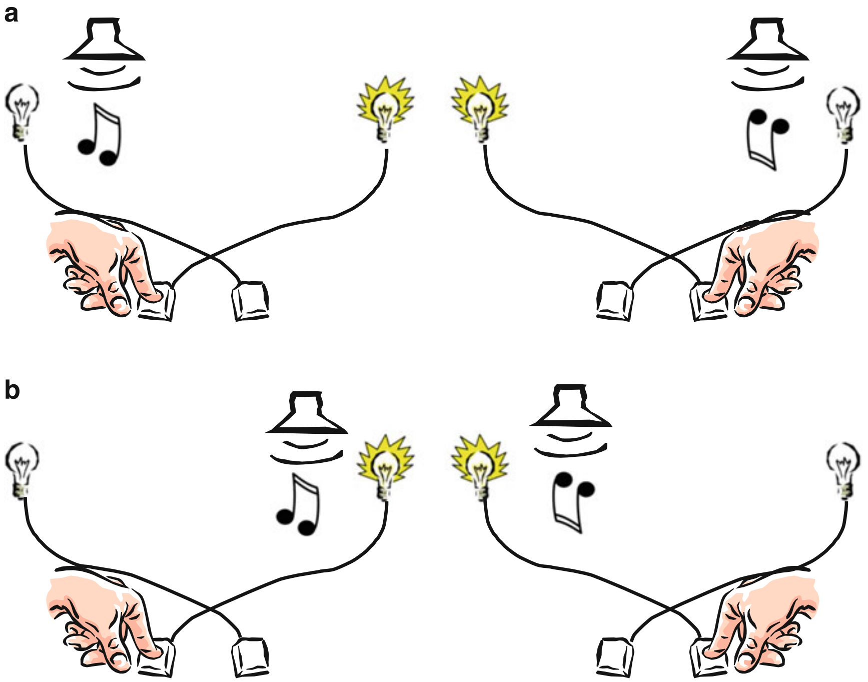


Fig. 4.12 Compatible and incompatible conditions in the study of Hommel (1993). (**a**) Tone-key spatially compatible/tone-light incompatible. (**b**) Tone-key spatially incompatible/tone-light compatible

We have already discussed how TEC accounts for the Simon effect proper, so let us now consider how the light instruction could change cognitive processing in such a manner that the effect was reversed. According to TEC, actions are represented through the codes of their effects, so that we can assume that Hommel's (1993) participants had already acquired stable associations between the motor patterns for moving the left and right index fingers (mp_l and mp_r) and their respective spatial codes ("left" and "right") before they even volunteered for the experiment (Fig. 4.13a). As soon as the experiment started, participants also experienced that moving the left index finger caused an extra action effect on the right side (namely the lighting up of an LED), while moving the right index finger caused an extra action effect on the left side (the lighting up of the left LED). Now, the fact that both motor programs evoked effects on the left and the right (Fig. 4.13a) provides each action with both left and right action effects (Fig. 4.13b), so that instructions could emphasize one or the other.

The two types of instruction indeed referred to just one of the two types of effects. Describing the action in terms of button pressing rendered this aspect the actual goal of the action, while describing it in terms of light switching rendered that the actual

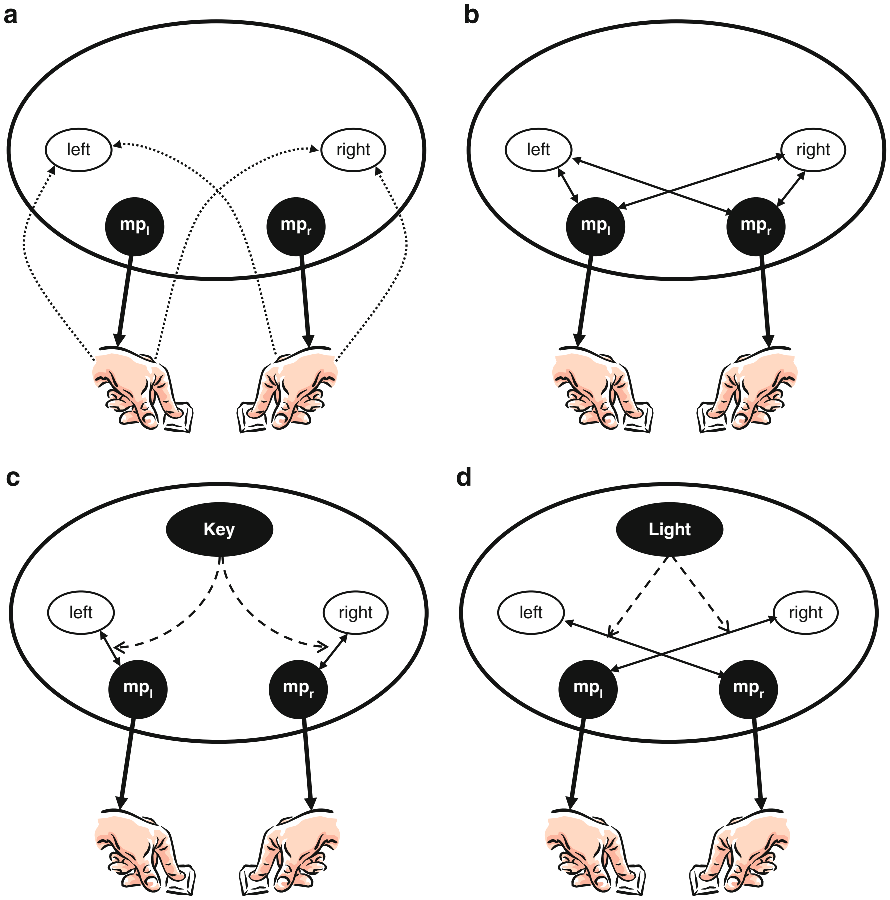


Fig. 4.13 Analysis of instruction effects on the Simon-effect according to TEC

goal. According to TEC, the current goal affects the weighing of codes by which a given action is represented. The button instruction should therefore lead to a stronger weighting of the action effects that are related to button-pressing, and therefore highlight the associations between mp_l and "left" and between mp_r and "right." In contrast, the light instruction should lead to a stronger weighting of light-related action effects and therefore emphasize the associations between mp_l and "right" and between mp_r and "left" (Fig. 4.13d). Under this instruction, left and right button presses would therefore be represented as "right" and "left" actions, respectively. The appearance of a stimulus on the right side would therefore activate the right light code that is now mainly associated with a motor pattern that carries out a left button press, which renders stimuli on the right compatible, rather than incompatible, with a left response, and the same logic applies to the right response. More generally speaking, stimuli tend to activate not all actions they share features with, but mainly those actions that share features on a goal-related feature dimension.

4.3.2.3 Action Effect-Blindness

Another somewhat peculiar phenomena was discovered by Müsseler and Hommel (1997; see Sect. 4.2.2). As we had mentioned above, participants in this study prepared, but did not yet execute a left or right button press before being presented with a masked and therefore barely visible visual arrow pointing to the left or right. It was almost impossible for participants to recognize action-compatible arrows, that is, left-pointing arrows after having prepared a left button press and right-pointing arrows after having prepared a right button press. How can that be the case and why does the compatibility between stimulus and reaction have a negative effect here?

To make sense of this relationship, we will now consider the processing assumptions of TEC. TEC assumes that actions, just like other events, are represented in a distributed fashion, that is, through a network of feature codes. Now how does one prepare such a network for the execution of an action? Firstly, the actual feature codes need to be **activated**, that is, to be put in a condition of heightened preparedness. That alone does not suffice, however. After all, a large number of other codes are active at about the same time, such as those required for the identification of the stimulus. To ensure that the relevant codes of concurrently represented events do not get mixed up, TEC assumes the existence of an **integration** process that ties together the codes that belong to the same event (Sect. 6.4). So, the phase of activation of the codes is followed by a phase of integration (Stoet and Hommel 1999).

For example, let us assume that you are preparing a left button press. Among other things, this leads to the activation of the relevant motor pattern (mp_l) and the associated spatial code (“left”). The integration of these activated components into an action plan temporarily binds a spatial code to the relevant, planned event (i.e., to the representation of perceived properties of the event). In a sense, the spatial code is therefore already occupied when a left stimulus must be identified as “left”; it is therefore not available for representing the left stimulus. As long as the spatial code remains bound to the plan, this can make you blind with regard to the property “left” as an aspect of other events—at least if conditions are so demanding that you cannot switch back and forth between the representations (which was prevented by the masking manipulation). This blindness should disappear as soon as you have executed the action, which is indeed what has been found (Wühr and Müsseler 2001).

4.3.2.4 Action-Effect Compatibility

TEC does not just perform well in explaining observations that seem peculiar at first sight, but it has also motivated new research questions and experimental approaches. An example would be studies into the compatibility between actions and their effects (review by Kunde 2006). Conventional linear models of information processing do not suggest that the consequences of an action have an effect on the reaction time that is required to initiate this action. After all, these consequences are only manifested through, and thus after the initiation, so that they should therefore not influence any processes that precede them temporally. According to TEC, however, actions serve to achieve specific effects and are thus represented through codes of these effects.

The selection of an action therefore occurs on the basis of these codes (see Chap. 5), so that the nature of the expected consequences is not unlikely to affect the speed of this selection. In fact, participants in choice reaction experiments respond better and faster if the required response matches the action effect that it produces, so when action and acquired action effect are compatible (Kunde 2001, 2003; Kunde et al. 2004). For example, left and right responses are selected faster when the left response produces a visual effect on the left and the right response an effect on the right. Also, a forceful or weaker button press is initiated faster when this button press is followed by a loud or soft auditory effect, respectively. So, participants appear to already anticipate the consequences of their actions during the selection of an action, and these anticipations influence the selection process (see Chap. 5).

References

Ach, N. (1910). *Über den Willensakt und das Temperament*. Leipzig, Germany: Quelle & Meyer.

Aglioti, S., DeSouza, J. F., & Goodale, M. A. (1995). Size-contrast illusions deceive the eye but not the hand. *Current Biology, 5*, 679–685.

Balint, R. (1909). Seelenlähmung des ‘Schauens’, optische Ataxie, räumliche Störung der Aufmerksamkeit. *Monatschrift für Psychiatrie und Neurologie, 25*, 51–81.

Bedford, F. L. (1993). Perceptual and cognitive spatial learning. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 517–530.

Bedford, F. L. (1999). Keeping perception accurate. *Trends in Cognitive Sciences, 3*, 4–11.

Bridgeman, B., Kirch, M., & Sperling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics, 29*, 336–342.

Brunswik, E. (1944). Distal focussing of perception. Size constancy in a representative sample of situations. *Psychological Monographs, 56*(254), 1–49.

Cohen, J. D., & Huston, T. A. (1994). Progress in the use of interactive models for understanding attention and performance. In C. Umiltà & M. Moscovitch (Eds.), *Attention and performance* (Vol. XV, pp. 453–456). Cambridge, MA: MIT Press.

Craighero, L., Fadiga, L., Rizzolatti, G., & Umiltà, C. A. (1999). Action for perception: A motor-visual attentional effect. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1673–1692.

De Jong, R., Wierda, M., Mulder, B., & Mulder, L. J. M. (1988). The use of partial information in response processing. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 682–692.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences, 11*, 219–226.

Eimer, M. (1997). The lateralized readiness potential as an on-line measure of automatic response activation in S-R compatibility situations. In B. Hommel & W. Prinz (Eds.), *Theoretical issues in stimulus-response compatibility* (pp. 41–73). Amsterdam: Elsevier.

Eimer, M., & Schlaghecken, F. (1998). Effects of masked stimuli on motor activations: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1737–1747.

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 229–240.

Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. *Psychological Research, 71*, 22–29.

Fitts, P. M., & Deiniger, R. L. (1954). S-R compatibility: Correspondence among paired elements within stimulus and response codes. *Journal of Experimental Psychology, 48*, 483–492.

Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46*, 199–210.

Gentilucci, M., Chieffi, S., Daprati, E., Saetti, M. C., & Toni, I. (1996). Visual illusion and action. *Neuropsychologia, 34*, 369–376.

Glover, S. (2004). Separate visual representations in the planning and control of action. *Behavioral and Brain Sciences, 27*, 3–78.

Glover, S., & Dixon, P. (2002). Dynamic effects of the Ebbinghaus illusion in grasping: Support for a planning/control model of action. *Perception & Psychophysics, 64*, 266–278.

Goodale, M. A., & Humphrey, G. K. (1998). The objects of action and perception. *Cognition, 67*, 181–207.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences, 15*, 20–25.

Goodale, M. A., Milner, A. D., Jakobson, L. S., & Carey, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature, 349*, 154–156.

Greenwald, A. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review, 77*, 73–99.

Hamilton, A., Wolpert, D. M., & Frith, U. (2004). Your own action influences how you perceive another person's action. *Current Biology, 14*, 493–498.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Heider, F. (1926). Ding und Medium. *Symposion, 1*, 109–157.

Helmholtz, H. (1866). *Handbuch der Physiologischen Optik* (Vol. 3). Hamburg, Germany: Voss.

Hommel, B. (1993). Inverting the Simon effect by intention: Determinants of direction and extent of effects of irrelevant spatial information. *Psychological Research/Psychologische Forschung, 55*, 270–279.

Hommel, B. (2006). How we do what we want: A neuro-cognitive perspective on human action planning. In R. J. Jorna, W. van Wezel, & A. Meystel (Eds.), *Planning in intelligent systems: Aspects, motivations and methods* (pp. 27–56). New York: Wiley.

Hommel, B. (2010). Grounding attention in action control: The intentional control of selection. In B. J. Bruya (Ed.), *Effortless attention: A new perspective in the cognitive science of attention and action* (pp. 121–140). Cambridge, MA: MIT Press.

Hommel, B., & Müsseler, J. (2006). Action-feature integration blinds to feature overlapping perceptual events: Evidence from manual and vocal actions. *Quarterly Journal of Experimental Psychology, 59*, 509–523.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC). A framework for perception and action. *Behavioral & Brain Sciences, 24*, 849–937.

Hommel, B., & Prinz, W. (Eds.). (1997). *Theoretical issues in stimulus-response compatibility*. Amsterdam: Elsevier.

Ishimura, G., & Shimojo, S. (1994). Voluntary action captures visual motion. *Investigative Ophthalmology and Visual Sciences, 35*, 1275.

James, W. (1890). *The principles of psychology*. New York: Holt.

Kohler, I. (1951). *Über Aufbau und Wandlungen der Wahrnehmungswelt, insbesondere über "bedingte Empfindungen"*. Wien, Austria: Rohrer.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus response compatibility: A model and taxonomy. *Psychological Review, 97*, 253–270.

Kunde, W. (2001). Response-effect compatibility in manual choice reaction tasks. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 387–394.

Kunde, W. (2003). Temporal response-effect compatibility. *Psychological Research, 67*, 153–159.

Kunde, W. (2006). Antezedente Effektpräsentationen in der Verhaltenssteuerung. *Psychologische Rundschau, 57*, 34–42.

Kunde, W., Koch, I., & Hoffmann, J. (2004). Anticipated action effects affect the selection, initiation and execution of actions. *The Quarterly Journal of Experimental Psychology. Section A: Human Experimental Psychology, 57A*, 87–106.

Lacquaniti, F., Terzuolo, C., & Viviani, P. (1983). The law relating the kinematic and figural aspects of drawing movements. *Acta Psychologica, 54*, 115–130.

Lotze, R. H. (1852). *Medicinische Psychologie oder die Physiologie der Seele*. Leipzig, Germany: Weidmannsche Buchhandlung.

Loveless, N. F. (1962). Direction-or-motion stereotypes: A review. *Ergonomics, 5*, 357–383.

Mashhour, M. (1964). *Psychophysical relations in the perception of velocity*. Stockholm: Almquist.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: Oxford University Press.

Münsterberg, H. (1888). *Die Willenshandlung. Ein Beitrag zur physiologischen Psychologie*. Freiburg, Germany: Mohr.

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 861–872.

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

Neisser, U. (1976). *Cognition and reality*. San Francisco: Freeman.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action: Current approaches* (pp. 167–201). Berlin, Germany: Springer.

Prinz, W. (1992). Why don't we perceive our brain states? *European Journal of Cognitive Psychology, 4*, 1–20.

Proctor, R. W., & Vu, K.-P. L. (2006). *Stimulus-response compatibility principle: Data, theory, and application*. Boca Raton, FL: Taylor & Francis.

Rachlin, H. C. (1966). Scaling subjective velocity, distance, and duration. *Perception & Psychophysics, 1*, 77–82.

Runeson, S. (1974). Constant velocity—not perceived as such. *Psychological Research, 38*, 3–23.

Schubotz, R. I., & von Cramon, D. Y. (2003). Functional-anatomical concepts on human premotor cortex: Evidence from fMRI and PET studies. *NeuroImage, 20*, 120–131.

Simon, J. R., & Rudell, A. P. (1967). Auditory S-R compatibility: The effect of an irrelevant cue on information processing. *Journal of Applied Psychology, 51*, 300–304.

Sommer, W., Leuthold, H., & Hermanutz, M. (1993). Covert effects of alcohol revealed by event-related potentials. *Perception & Psychophysics, 54*, 127–135.

Stock, A., & Stock, C. (2004). A short history of ideo-motor action. *Psychological Research, 68*, 176–188.

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1625–1640.

Stratton, G. M. (1896). Some preliminary experiments on vision without inversion of the retinal image. *Psychological Review, 3*, 611–617.

Ungerleider, L., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549–586). Cambridge, MA: MIT Press.

Van Donkelaar, P. (1999). Pointing movements are effected by size-contrast illusions. *Experimental Brain Research, 125*, 517–520.

Van Essen, D. C., & DeYoe, E. A. (1995). Concurrent processing in the primate visual cortex. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 383–400). Cambridge, MA: MIT Press.

Viviani, P., & Stucchi, N. (1989). The effect of movement velocity on form perception: Geometric illusions in dynamic displays. *Perception & Psychophysics, 46*, 266–274.

Viviani, P., & Terzuolo, C. A. (1982). Trajectory determines movement dynamics. *Neuroscience, 7*, 431–437.

von Goethe, J. W. (1963). *Faust* (Walter Kaufman, Trans.). Garden City, NY: Anchor. (Originally published 1808)

Wühr, P., & Müsseler, J. (2001). Time course of the blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 1260–1270.

Chapter 5
Action Selection

The expression "the agony of choice" refers to the fact that having to choose between several alternatives is often experienced as pain rather than pleasure. The main character of Luke Rhinehart's book (1971) "The Dice Man" associates choosing with so much pain that he decides to delegate all further decision-making to throwing a die. Prior to making a decision, he would number the conceivable alternatives from one to six, and then has the die make the decision for him. For the readers' enjoyment, this strategy renders his further lifestyle rather adventurous in ways that one may or may not enjoy oneself, but the story makes one thing clear: decisions are difficult and they often torment us. Indeed, every day, hour, and minute, we have so many choices that it is difficult for a lot of people to choose one alternative. That goes for both abstract action alternatives, such as opting for one of several possible lifestyles or partners, and concrete alternatives, such as grasping a cup of tea—an action one could perform in a thousand different manners. How do we choose actions? How do we ensure that they serve our goals?

Disregarding nuances, two perspectives can be distinguished: **sensorimotor** approaches (in this context often also referred to as chronometric or information processing approaches) that mainly occupy themselves with the stimulus-induced selection of reactive actions, and **ideomotor** approaches that are more concerned with the question of how the action goal makes possible and controls the selection of an action. We will see that these two perspectives are not mutually exclusive, but that they merely emphasize different facets in action control and occupy themselves with different phases of action selection. In our description, we will mainly follow the suggestions of Fitts and Posner (1967) and Rasmussen (1980) (Sect. 9.1), who proposed three different **modes of action control**:

- The **cognitive** (Fitts), respectively **knowledge-based** (Rasmussen) **mode** in which the acting person transforms an action goal into a suitable action, that is, makes a goal-induced action selection (Sect. 5.1).

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_5

- The **associative**, respectively **rule-based mode**, in which one transforms specific, predefined stimuli into assigned responses on the basis of stipulated rules, that is, makes a rule-based action selection (Sect. 5.2).
- The **autonomous**, respectively **skill-based mode**, in which stimuli trigger more or less automatically associated responses, so in which a type of automatic action selection takes place (Sect. 5.3).

Even though these three selection modes cover a large part of our everyday action selections, a further modus is lacking: one often makes a decision based on "gut feeling," so without following a clear goal, utilizing an acquired rule, or pursuing a habit, for example, when purchasing a new but not urgently needed article of clothing or some technical gadget. This type of **intuitive decision-making** has attracted a lot of research interests in recent years (Sect. 5.4).

5.1 Goal-Induced Action Selection

Actions are motoric means to achieve particular goals. An action goal is an anticipation, that is, a consciously or unconsciously represented sensory or verbal description of an intended, yet to be created event or state (Chap. 3). Now how does one get from the representation of this state to the motor means that are required to create it?

Theories on problem solving suggest two possible **solution strategies** for problems of this nature:

- **Hill climbing technique**: this consists of randomly trying out several possible alternatives (e.g., Robertson 1999). The best alternative is found by trial and error and simulated mentally to see to which outcome it would lead. When the outcome can be considered an improvement over the current situation, so when it decreases the perceived distance to attaining the goal, then this alternative is realized, and otherwise, the next alternative is tried out.
- **Means-end analysis**: this strategy is considerably more efficient. The analysis starts with the goal that is to be achieved and then works its way back all the way through the current starting point, so as to move along the path of solving the problem and to achieve the goal in manageable steps. Both strategies assume that possible alternatives are evaluated to check to what extent the states that one can achieve with their help match the goal that is to be attained. When we apply this logic to the selection of concrete actions, then we must assume that the representations of actions also contain information about the states that can be achieved by these actions. Therein lies the core assumption of the ideomotor approach, namely that actions are represented through codes of their sensory effects.

For example, let us assume that you are driving a car and that you end up in a dangerous situation. You have learned that in such circumstances, it is a good idea to brake, and this is your action goal. The action goal is represented in a sensory

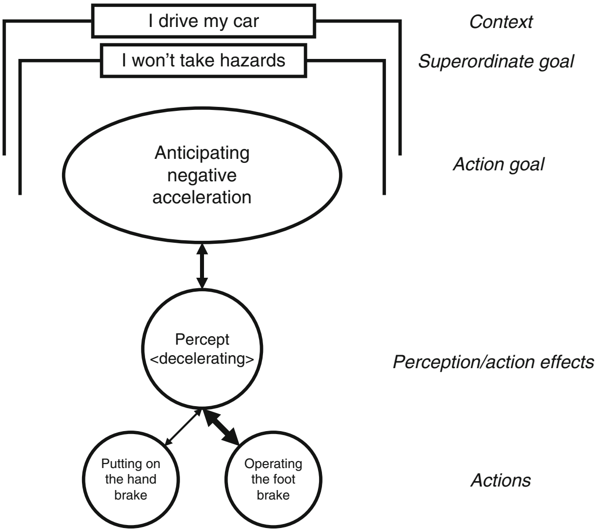


Fig. 5.1 Role of action goals in action selection

and/or verbal manner and refers to the desired state. Concretely, this means that you activate representations of the sensory impressions that you have previously had when braking in a car (Fig. 5.1; Percept "braking"), possibly accompanied by the activation of representations of descriptive words like "to brake."

Activations of cognitive representations spread, so they tend to activate other representations that are associated with them. If you are an experienced driver, then you will have learned to associate decreased acceleration mainly with braking maneuvers. You will have typically performed braking maneuvers with the brake pedal, so that the action "step on brake pedal" has become associated with the action goal "decrease acceleration." In some cases, for example, when the brake pedal did not work, you may have caused the car to stop with the handbrake; in that case, the action "engage handbrake" has also become associated with the action goal "decrease acceleration," but more weakly so than the action "step on brake pedal" (Fig. 5.1).

The activation of the action goal therefore results in the activation of at least two actions. This has various consequences. Firstly, a mild decision conflict between the two action alternatives can arise, but that should be settled quickly in favor of putting one's foot to the brake pedal due to the different weights between the actions and the different action effects. Secondly, it will become clear that in every case a fitting solution will be found, that is, that a goal-appropriate action will be selected. Whether this will be the case for a given goal depends, of course, on whether **action-effect associations** have been acquired whose effect components are similar to the currently represented goal. Otherwise, trial and error remain.

The sensory representation of an action goal is of benefit to the finding of an appropriate action alternative, but not always necessary. Verbally represented action goals also allow for the selection of appropriate actions, provided that the used verbal concepts are associated with the relevant actions. That this functions well, will become immediately clear when you realize that you are able to perform a given task on the basis of oral or written instructions, and that you are able to cook a dish by following a recipe. Importantly, these examples also suggest that the selection of actions on the basis of verbal representations is more complicated and less unambiguous than selection on the basis of sensory representations, for example, when a TV cook shows you how to prepare a dish.

The ideomotor approach suggests, amongst other things, how problems can be solved creatively. Let us assume that in the given example, you have performed the two obvious actions (step on brake pedal and engage handbrake) in sequence, but both do not have an effect due to a technical defect. What will you do next? Let us assume that the action that is associated most strongly with the goal-directed perceptual impression (percept), namely, to brake, is currently without success, and therefore it is inhibited. This would suggest that weaker associations of the percept now also come into play, as well as associations that actually occur in other contexts. Another context in which the action goal and the percept "braking" plays a part is bicycling. While riding a bike, depending on the model, you can either use the handbrake (which would require a different movement than it would in a car) and/or use the coaster brakes. Both actions will be of little use in a car. So what else could you do to make your bike stop? You could brake with your foot, that is, put your foot on the street to induce so much friction that your bike will eventually come to halt.

5.2 Rule-Based Action Selection

Fortunately, unforeseen emergencies are not very common in our everyday behavior. Often, we have experienced situations many times before, and often, situational stimuli are available, that point us in the direction of the best action alternative. These indicators can be of a semantic nature: for example, when we know that we must get off of the subway on Piccadilly Circus, then the sign labelled "Piccadilly Circus" will inform us about the location to perform that action. However, indicators can also be of an episodic nature: although the color green does not have a direct association with the movements of pedestrians and vehicles, over the years we have learned that when a traffic light is green, we can safely cross the street. In other words, whenever a systematic association between stimuli and actions exists or can be expected, we are able to develop rules that can steer the selection of the relevant action.

This **rule-based action selection** is typically modelled in two ways. From the perspective of connectionist (also known as PDP or Parallel-Distributed-Processing) models, rules are considered to be contextually adapted stimulus-response connections.

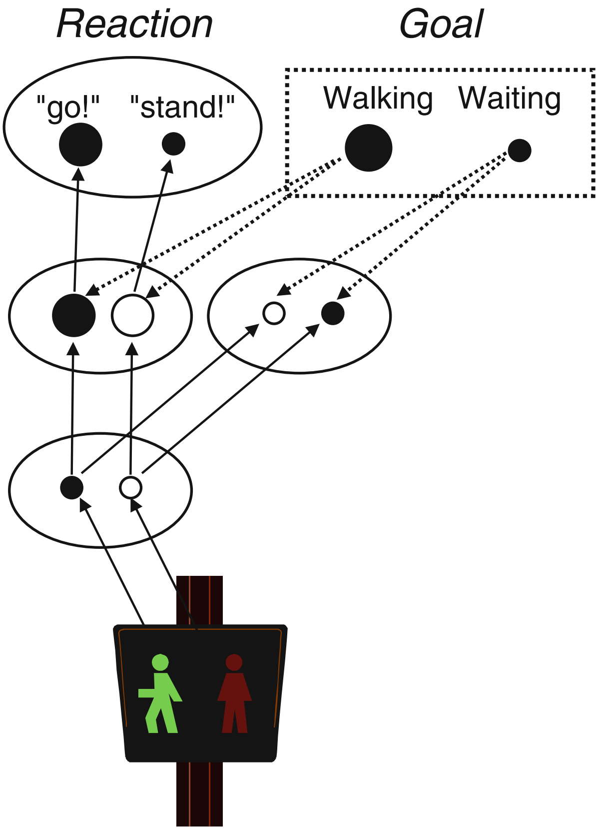


Fig. 5.2 Rule-based action selection

How such a thing might look is represented by the example of the pedestrian traffic lights (Fig. 5.2, cf. Fig. 3.1). In a first feature domain, the properties of stimulus alternatives are coded, in our case the green or red color of the little guy in the traffic lights or his two shapes, for example. Activation of this domain is fed forward, along a middle layer (a type of relay station) that is modulated by action goals, to action selection. In our example there are two such relay stations. One leads to the action domains that are relevant in this situation, in which the two action alternatives "to go" and "to stay" are distinguished, the other station can be connected with any other, currently irrelevant action alternative (which we have not depicted here, for simplicity's sake). We have distinguished two goals in this example: the relevant goal "to walk" and the currently irrelevant goal "to wait." On the basis of the system configuration, activation of the goal "to walk" leads to a strengthening of the activations that lead from the input domain to the alternatives "to go" and "to stay." So, the lighting up of the green guy on a traffic light activates the alternative "to go," while the lighting up of the red guy will lead to activation of the alternative "to wait." This rule (*if* green, *then* go; *if* red, *then* wait) is contextually adapted, for if another goal than "to go" would have been activated (for example, "to wait"), then the action alternatives "to go" and "to stay" would only have been activated marginally by the traffic light. Note that the requirements of a Stroop task can also be conceptualized as a system of contextually adapted rules (Fig. 3.1). The following variants flow forth from that:

- *If* the word "red" is presented in green *and* the goal is "to name color," *then* say "green"
- *If* the word "red" is presented in green *and* the goal is "to read word," *then* say "red"
- *If* the word "red" is presented in red *and* the goal is "to name color," *then* say "red"
- *If* the word "red" is presented in red *and* the goal is "to read word," *then* say "red," etc.

An alternative strategy to model rule-based action has become known under the name **Adaptive Character of Thought** (**ACT**); the most current version is ACT-R (Anderson 1993). This is a framework theory that encompasses the acquisition, representation, and use of declarative and procedural memory. The actual units of this theory are called **productions**. Productions consist of a conditional part, which represents the "if"-aspect of a rule, and of an action part, which represents the "then"-aspect. From the perspective of ACT-R, a minimum of two productions are required for a successful processing of the little guy in the traffic light: one in which "green" is represented in the conditional part and "to go" is represented in the action part, and a second one in which "red" is represented in the conditional part and "to wait" is represented in the action part.

The selection of actions depends on a number of factors. Some of these factors are general, and although they influence the efficiency of action selection, their effect is not specific to selection processes. This includes, for example, strategies that allow us to respond either very quickly or very accurately (Woodworth 1899), or the time to prepare a transition (Niemi and Näätänen 1981). In what follows, we will focus on factors that influence the selection of actions in a specific, theoretically interesting manner.

5.2.1 Number of Action Alternatives

Donders (1868) already pointed out that the selection of an action takes time. In his study, he compared, amongst other things, the reaction times (RTs) in three different tasks:

- **A-response tasks**, in which a previously defined response had to be performed when a previously defined stimulus (a light switching on) appeared.
- **B-response tasks**, in which one of two stimuli was presented, each requiring a separate response.
- **C-response tasks**, in which the response was fixed, but in which only one of two stimuli should be responded to (a so-called go-nogo task).

Donders' consideration was that these tasks require, at least partially, different processes and that by comparing the corresponding reaction times, one can measure the duration of these processes. In this context, the a-task represents a measure of the required basic sensory and motor processes. In this task, the reaction time corresponds to the time it takes for a stimulus to be detected and for a movement to be executed. The difference with a c-response is that two stimuli need to be distinguished in a c-response task. The costs of this **stimulus discrimination** are determined by subtracting the mean reaction time in the a-response task from the mean reaction time in the c-response task (stimulus discrimination time = RTc – RTa). The time that is required to select an action can be determined in a similar manner: the only difference between b-responses and c-responses is that an action needs to be selected; this means that the costs of action selection are represented by RTb – RTc.

The Dondersian subtraction logic contains some deficiencies that have often been criticized and that have been corrected by Sternberg (1969). The observation that the selection of an action takes measurable time has nevertheless been confirmed repeatedly. On top of that, Merkel (1885) discovered that reaction time increases as the number of response alternatives increases. As the upper panel of Fig. 5.3 shows, this increase is not linear: the larger the number of alternatives

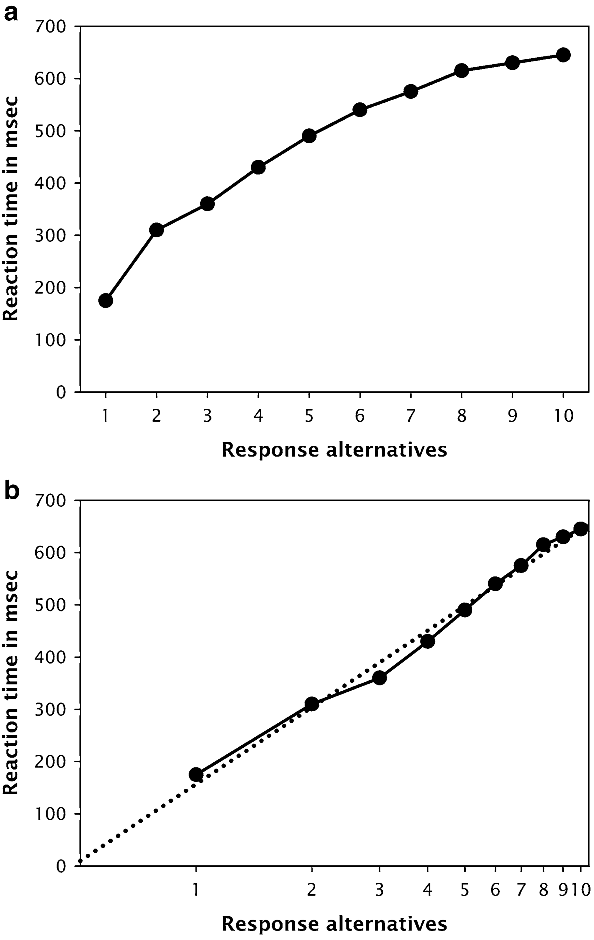


Fig. 5.3 Relationship between reaction time and, respectively, the number of alternatives and the logarithm of the number of alternatives (representative data)

already is, the fewer costs the addition of yet another alternative adds. This observation was taken up again and formalized by Hick (1952). As Hick demonstrated, the relationship between reaction time and number of alternatives can be characterized by a simple formula:

$$\text{reaction time} = a + b\left(\log 2N\right),$$

where the constant a represents sensory and motor basic principles (that Donders intended to measure through his a-response task), constant b represents the increase in reaction time per unit (that is, the slope of the function), and *N* the number of response alternatives. In other words, reaction time increases linearly with the logarithm of the number of alternatives (Fig. 5.3b). As Hyman (1953) was able to show, not much later, Hick's logarithmic definition corresponds to the definition of the *X*-axis of the information-theoretical measure Ht (standing for the amount of transferred information). As information theory was very popular in those years, Hyman suggested an equivalent variant of Hick's formula:

$$\text{reaction time} = a + b\mathrm{Ht}.$$

This postulation was named after its inventors, the **Hick-Hyman Law**, and it has been confirmed often, but exceptions to it are known too: for example, reaction times in tasks with highly compatible stimulus-response assignments are independent of the number of alternatives (Leonard 1959), and in tasks with a very large number of alternatives, the formulae of Hick and Hyman do not allow for adequate descriptions of the results (Longstreth et al. 1985). But why is there a relationship between reaction time and number of alternatives? Two explanations have been discussed in particular.

According to one explanation, this relationship can be an expression of a searching process. For example, authors like Hick (1952) or Falmagne et al. (1975) assume that the possible response alternatives in a given task are held in a response buffer. Following identification of a stimulus, this buffer needs to be searched through step by step to check whether the current response is the correct one. In the terminology of the ACT, one could imagine, for example, that response alternatives are stored in the form of productions and that for every production needs to be checked whether the current stimulus corresponds to its conditional ("if") part. The more response alternatives the response buffer holds, the longer the average searching time. This model can definitely depict some of the aspects of the relationship between reaction time and alternatives, but it is not very realistic. For example, it cannot explain why, under specific circumstances, multiple reactions can be activated simultaneously (Hommel 1998a; Logan and Schulkind 2000).

In contrast, other authors assume that multiple stimulus response-rules can be applied and checked at simultaneously. For example, Meyer and Kieras (1997) assume that rules that are held in short-term memory in the form of ACT productions are compared in parallel to the stimulus. Accordingly, a stimulus can simultaneously activate all the productions whose conditional parts it satisfies (for example, activate all actions that are associated with the condition “I am located in a church”). Connectionist/PDP models, which are designed to process in parallel by definition, work in a similar way. Such models therefore do not explain the **number-of-alternatives effect** by assuming serial response selection. Instead, they assume that response alternatives that are held available simultaneously interfere with each other. For example, the representations of such alternatives, that are mutually exclusive to each other, are connected in an inhibitory fashion in network models, so that the activation of a response leads to the inhibition of other responses. Now, if one considers that in a task, most response alternatives are activated to a certain degree (for example, from the previous trial, on the basis of false expectations or misleading interim results of the stimulus identification), which in turn leads to the inhibition of all other alternatives, then ultimately every correct response alternative needs to overcome more inhibition as more alternatives are available. In other terms, reaction time increases with the number of alternatives.

5.2.2 Stimulus-Response Compatibility

Stimuli and actions are not represented completely independently from each other, so that action representations can influence stimulus processing and stimulus representations can influence action selection (Chap. 4). Stimuli that are compatible with responses facilitate response selection (Sect. 4.2.1), probably because stimuli automatically activate responses that have overlapping properties (Sect. 4.3.2). **Compatibility effects** are interesting for a number of reasons: they demonstrate, for example, that perception and action are connected more intimately than process models often assume (Chap. 4). They show, therefore, that the selection of an action cannot merely consist of the activation of relevant rules or productions. For example, when a stimulus appears on the left in an experimental setup, then why would one be able to find a production with the feature “left” in the action part (that is, in the *then*-part of the production rule) more quickly than a production with the feature “right”?

Questions of this nature have spawned a number of **dual-process models** of response selection. An example of one of those is the **model of dimensional overlap** by Kornblum et al. (1990), which states that stimuli in a compatibility task are processed along intentional and automatic routes (Fig. 4.7). The **intentional route** follows the logic of rule-guided action selection, while the **automatic route** assumes direct associations between stimuli and responses (Sect. 5.3). In a certain way, the theory of event coding (TEC, Sect. 4.3.1) also employs a dual-route solution, while relevant stimulus properties are connected by short lasting, intentionally created associations, the feature overlap between stimulus and response leads to the partial identity of the corresponding cognitive representations. This could be interpreted as an extreme version of the automatic route. The existence of such automatic routes suggests that action selection does not just depend on the application of stimulus-response rules, but that the perceived stimuli can also directly influence action selection (Sect. 5.3).

5.2.3 Repetition

It has been known since the work of Hyman (1953) and Bertelson (1961) that the **repetition of responses**, or of stimulus-response combinations, influences performance (see the reviews by Kirby 1980; Kornblum 1973). When the temporal distance between reactions is short (<500 ms), one mainly observes facilitation, that is, repetition of responses facilitate performance.

This effect has been interpreted in a number of ways. Bertelson (1963) assumed that, before a trial starts, persons check whether the current stimulus matches that of the previous trial. If that is the case, then the stage of response selection is skipped, so to say, and the previous response is repeated. At first glance, this approach appears intuitive, but it is more of a description than an explanation. Besides, this approach cannot explain why the advantage of response repetition remains even when the stimulus alternates (Bertelson 1965). Falmagne et al. (1975) assumed that response selection occurs through the sequential checking of stimulus-response rules. They assumed that the respectively last rule is more available, that is, is checked first. Other authors have postulated that the activation of a response representation slowly subsides following the execution of the response, so that for a short while, further repetitions can still benefit from this residual activation (Vervaeck and Boer 1980).

As the temporal distance between the responses increases, the advantage of repetition decreases to the point where, for longer temporal distances, it may actually become a disadvantage, so that **response alternations** become an advantage (Bertelson 1961; Soetens et al. 1985). This effect has been interpreted more strategically. It has been known from roulette and other games of chance that people systematically overestimate the probability of alternations, that is, following red they expect black, and vice versa (Keren and Wagenaar 1985). The more time passes since the last event, and the stronger the immediate repetition tendencies have subsided, the stronger such strategic considerations dominate.

A further effect that has only been studied recently is episodic in nature. When stimulus and response properties independently differ from one another, then the partial repetition of a stimulus-response combination leads to a decrease in performance. For example, when you have to make left and right button presses to red and green circles, your performance will be better when both the stimulus and the response repeat themselves. However, if the stimulus alternates, the performance is better when the response alternates as well (Hommel 1998b). In other words, performance will be optimal when both stimulus and response are **repeated**, or when they both **alternate**. This observation suggests that the common occurrence of stimulus properties and responses leads to an **integration** of both (*stimulus-response binding*). In the repetition of one element of the episodic memory trace that has thus been created, the entire trace is reactivated and thereby, the related other element is activated too. If this other element does not fit to the current trial (as is the case in partial repetitions), the result is **interference**.

All three types of effect demonstrate that the selection of an action also depends on short-term events and situational circumstances. Above all, they show that the selection of action is a dynamic process that sensibly reflects the current state of the cognitive system and the various response tendencies.

5.2.4 Practice

When the selection of actions occurs on the basis of stimulus-response rules, and when those are represented in the form of associations between stimulus and response codes, then one should assume that the **overlearning** of such associations considerably facilitates action selection. Indeed, long-lasting practice leads to sometimes dramatic accelerations of processes that are involved in the selection of actions. For example, Mowbray and Rhoades (1959) asked participants to practice a reaction time task with two stimulus-response alternatives 1500 times and a task with four stimulus-response alternatives 3000 times, so that every single stimulus-response association was learned very well. As the practice progressed, responses did not just become considerably faster, but the difference in performance between the tasks also got increasingly smaller. In other words, the Hick-Hyman Law did no longer apply. Comparable observations have been recorded by Seibel (1963), who found only a marginal difference in performance between tasks with 1023 alternatives and a task with 31 alternatives after 75,000 practice trials.

When performance is improved that much through practice, then the question arises for the cause of this improvement. Which processes are affected? From a connectionist perspective, this should be caused by a strengthening of the associations between the representations of the relevant stimulus and the relevant response features, a bit like that between the green guy in the traffic light and the response "to go!" in Fig. 5.2. Through this strengthening, not only will the activation spread faster from stimulus to response representations, but interference from competing stimulus-response associations can be avoided or dealt with more easily. In other words, practice does not just lead to a stronger connection between stimulus and response, but also to a greater dominance of the relevant stimulus-response associations relative to other associations. From an ACT perspective, a comparable consideration can be developed: practice strengthens the appropriate productions and gives them dominance vis-a-vis possible competitors. Anderson (1983) also discusses the possibility that through practice, productions can be adapted better to **situational demands**, for example, through specification of the conditional part.

Nevertheless, all of these arguments only apply to the relevant, intentional aspect of stimulus-response translation, so to the **intentional route** in the model of Kornblum et al. (1990; Fig. 4.7). As the other, **automatic route** completely depends on completely automatized stimulus-response associations (an assumption that we will restrict further on in this chapter), its processing abilities shouldn't change notably through practice. So if it is true that the effects of stimulus-response compatibility can be attributed to the automatic activation of responses along the automatic route, then it should not be possible to eliminate compatibility effects through practice. In fact, this does indeed appear to be impossible: neither the effect of spatial stimulus-response compatibility (Brebner 1973), nor the Simon effect (Simon et al. 1973), nor the Stroop effect (Stroop 1935) disappears with practice.

5.3 Automatic Selection of Actions

Theories about the acquisition of skills often assume that a large amount of practice leads to automatization of the practiced action. According to Fitts and Posner (1967), extensive training leads to a so-called **autonomous learning phase**, in which the corresponding action no longer requires conscious control and can be performed together with other actions without cognitive costs. In a very similar manner, Anderson (1983) assumes that long practice transforms the initially declaratively available (i.e., verbally describable) knowledge into procedural knowledge that is no longer consciously available. In the face of these reflections, it could be argued that one might be able to skip the process of action selection when specific stimulus-response combinations are practiced extensively and one leaves action selection to the stimulus. Indeed, this idea forms the basis of the dual process model of Ach (1910), which combines the automatic **habit** (the automatic route) with the **will** (the intentional route). Phenomena like the Stroop effect are also compatible with this idea, as the effect demonstrates that one can hardly suppress reading a word. So, through extensive practice, can one leave action selection to the stimulus?

The answer is yes and no. There are indications that the selection of actions gets extraordinarily efficient through practice, and the disappearance of the number-of-alternatives effect is a particularly strong indicator that the demands of action selection are reduced drastically through practice. Furthermore, training allows for the practically simultaneous execution of various actions without cognitive costs (cf. Chap. 8). The observation that highly compatible stimulus-response assignments can both directly activate actions and make the alternative effect disappear (Sect. 5.2.1) suggests that stimuli are able to take over the selection of actions.

Insights into the **automaticity** of stimulus-induced action activation are provided by, for example, electrophysiological studies of the temporal properties of the lateralized readiness potential (Sect. 4.2.1). Valle-Inclán and Redondo (1998) have performed such a study using the Simon effect. The participants made spatially defined responses (press a top or bottom button, executed with the left or right hand, respectively) to colored stimuli. The stimuli appeared randomly at the top or bottom of the computer screen and were therefore sometimes compatible and sometimes incompatible with the response. As is shown in Fig. 5.4a, in a compatible trial (so when the stimulus position is equivalent to the manual response, for example, top stimulus → top hand), the activation (that is, the LRP) of the correct response starts about 200 ms after the appearance of the stimulus (S). The activation then increases steadily, until the response is finally executed (in this case about 560 ms following the appearance of the stimulus). In contrast, the development of the activation in

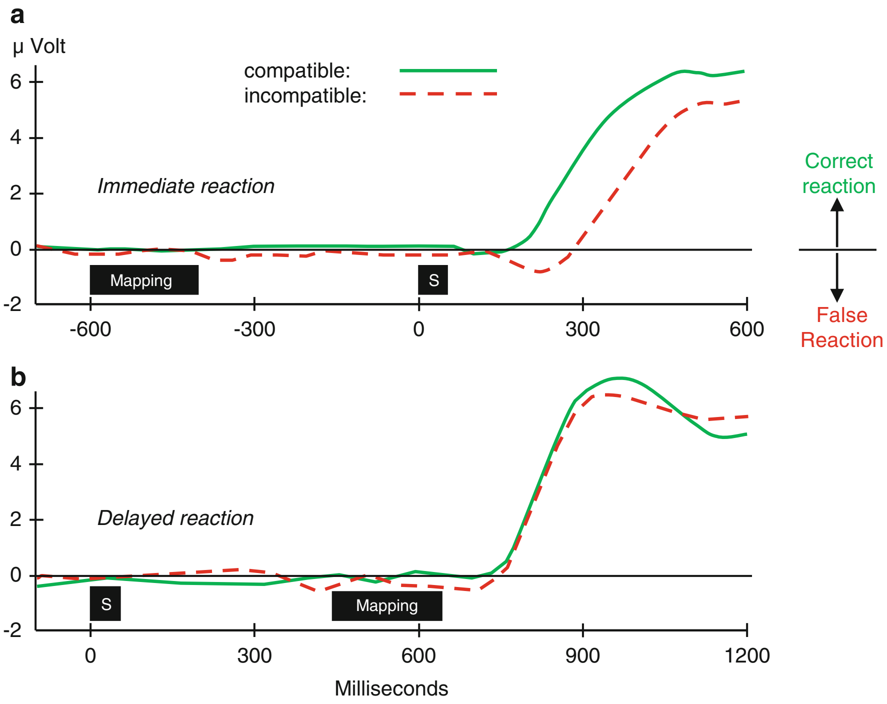


Fig. 5.4 The development of activating the correct or wrong response in incompatible (when the stimulus position evokes the incorrect response) and compatible trials (when the stimulus position evokes the correct response) (adapted from Valle-Inclan & Redondo 1998, by permission of Wiley Blackwell Publishing; coloring ours)

incompatible trials, so when the stimulus position evokes the incorrect response, is completely different. Here, activation starts in the wrong direction, that is, the incorrect response is activated firstly, before the activation in the correct direction becomes perceivable. Accordingly, the peak of the activation of the correct response is reached later. In this case, the incorrect response has been activated by the irrelevant, spatial stimulus property, which was facilitated by the fact that spatial information can be processed extraordinarily quickly (Hommel 1993).

On the one hand, such observations suggest that we cannot completely control the amount of stimulus information that we want to process and the transformation of this information into action tendencies. So, in this manner, action selection can be automatic. On the other hand, there is simply no clear evidence that stimuli activate responses completely independently from the current action intention—except perhaps in patients with frontal lobe damage (Lhermitte 1983) (Sect. 3.1).

The processing of responses is strongly intertwined with the processing of stimuli, and the way in which response alternatives have been defined has immediate consequences for the attention to stimulus properties (Sect. 4.2.2). For example, the intention to execute a grasping movement leads to focusing attention to shape properties and the intention to execute a pointing movement, to spatial properties

(Bekkering and Neggers 2002; Fagioli et al. 2007). That means that the **intentional** preparation of a specific type of action leads more or less **automatically** to the processing of properties that fit to and inform the action (Fig. 4.8), as can be observed in the Simon task. This logic also applies to the Stroop effect, which is interpreted as evidence for the automatic processing of words. It is manifested under circumstances in which words play a considerable role; after all, participants respond verbally.

Considerations of this nature have led authors like Bargh (1989) to question the existence of the principle of pure automaticity. Therefore, on the one hand, automatic processes represent the actual basis of our everyday lives, on the other hand, automaticity is always at the service of intention and is only made possible through intention even when the outcomes of automatic processing are not always useful to intention. In other words, what looks like automaticity is actually **conditional automaticity**. In a certain way, this approach represents a rebirth of an idea by Exner (1879). In his discussion on the role of the will in action control, that was mainly based on introspection, Exner failed to observe any willful experience between the appearance of a stimulus and a response. Following this theory, the intention to perform a given action appears to transform the cognitive apparatus into a type of reflex machine. The actual action therefore resembles a mere reflex, but it is an **intentionally prepared reflex** (Hommel 2000).

A particularly convincing argument for the existence of a prepared reflex of this nature can be gleaned from the previously mentioned experiment by Valle-Inclán and Redondo (1998). In that experiment, the assignment of responses to the color stimuli was not fixed, but it varied from trial to trial (Fig. 5.4). Sometimes, a red stimulus had to be responded to with a top button, and a green stimulus with a lower button, and sometimes the assignment was reversed. If the assignment appeared before the stimulus (Fig. 5.4a), so that the participants could completely prepare themselves for the trial, then stimulus presentation had the described "automatic" effect: the stimuli activated the spatially compatible response even when it was actually incorrect. Valle-Inclan and Redondo also studied a second condition, in which the stimulus was presented before the assignment (Fig. 5.4b). In this condition, participants were unable to prepare themselves for the trial, as they could not accomplish much with just the stimulus until they had learned the stimulus-response association. Please note what influence the stimulus presentation had on response activation in this condition: none! So when one interprets the stimulus-induced activation of the incorrect response as an automatic process in the top panel of the figure, then the automaticity of this process is apparently dependent on the intentional preparation of the task, so in this case, dependent on the implementation of the relevant stimulus-response associations—in other words, a prepared reflex.

5.4 Intuitive Action Selection

In the late 1970s, Dörner started an encompassing study of human decision-making in complex tasks, the **Lohhausen project** (Dörner et al. 1983). The city of Lohhausen actually only existed in the framework of a computer simulation.

Student participants received the assignment to control the city for 10 years (that were temporally compressed) as a mayor. This required dozens of decisions, for example, with respect to collecting taxes, social politics, and measurements to reduce unemployment, which altogether could influence about 2000 variables in the simulation.

These decisions were very difficult, and the participants met with various levels of success, measured by the socioeconomic status of the city at the end of the virtual 10 years. An important research question was related to identifying personal characteristics that were correlated with successful decisions. Some surprises surfaced there: whereas a long list of obvious predictors like intelligence, motivation, creativity, gender, age, or occupation revealed no relationship with the quality of the decisions, some unexpected variables turned out to be reliable predictors, namely self-confidence, extraversion, and the tendency to intellectual exploration. Interestingly, the predictors that were unsuccessful in this context are much more suitable to predict correct decisions in more simply structured logical problems; but the more complex the problems got, the less this was the case. Actually, successful decision-makers are often unable to name the criteria of their decisions in more complex contexts; they appear to make their decisions more **intuitively**. But how could that work?

One variable that may play a central role in intuitive decisions are the **affective consequences** of the associated actions. Classical learning theories have consistently demonstrated that the affective consequences of actions (their valence) influence their selection: actions that have repeatedly led to positive affective consequences are selected with a higher probability in the future, while the selection probability of actions with negative affective consequences diminishes (Thorndike 1927). In other words, the valence of actions supplies criteria to prefer specific actions in favor of others. Even though these observations are supported by an enormous body of animal and human studies, the underlying selection mechanisms are still not well understood.

A possible approach has been suggested by Rolls (1999) (Sect. 2.6.2). He assumes that we acquire associations between stimuli from the environment and the rewards that are associated with them. When we are then faced with various stimuli, we consider the rewards that are expected by the actions that can be selected. So, we select the response that is related to the stimulus with the highest expected reward. Another suggestion was made by Damasio (1994). He assumes that representations of actions are provided with so-called **somatic markers**, that is, representations of body-related feelings that are associated with an action (Sect. 2.6.2). When various actions are available and there is either no time or no adequate information available for rational deliberation, then we will choose the action that has the most positive feeling associated with it. In other words, we choose the action that “feels best” (see Box 5.1).

A number of empirical observations support the assumption that the expected (or previously experienced) affective consequences of actions play an important role in action selection (Sect. 2.6.2). For example, Bechara et al. (1997) have demonstrated that people show stronger **autonomous affective responses** and sweat more immediately before taking risky decisions (see Box 5.1). So, they appear to anticipate the possible negative consequences of their decision during the decision-making process. Conversely, patients with damage in the ventromedial prefrontal cortex, which

Box 5.1: Intuitive Decision-Making

The now classical study of Bechara et al. (1997) aimed to unravel the **neural mechanisms underlying intuitive decision-making**. Participants played a kind of card game, the so-called **Iowa Gambling Task**. In each move of this game, one card is drawn randomly from one of several (commonly 4) decks, and some cards lead to the winning of a particular amount of money while others lead to a loss. Importantly, the probability of receiving a winning or losing card is not random, so that, after a short introductory phase in which all decks lead to gains, some decks lead to overall gains ("good decks")—resulting from a mixture of small gains and slightly smaller losses—while others lead to overall losses ("bad decks")—resulting from high gains and even higher losses. As the goal of the game is to win as much money as possible, it is not surprising that participants learn to choose from good and less risky decks more often as the game progresses.

Figure 5.5 shows how the choices of healthy participants change during the game (see upper left panel). In the first, loss-free introductory phase, cards are drawn from the two types of decks with equal probability. In the next phase, where participants are confronted with losses, choice behavior is slowly changing and increasingly favoring the good and less risky decks. In the beginning of this phase, participants do not yet know why they are changing their behavior, which is why this part is often referred to as the "clueless" period. With increasing experience, however, the preference becomes more and more pronounced, even though participants cannot yet fully explain why they are having this preference (hunch period). The last, conceptual period is commonly reached by no more than 70 % of the participants: they now have acquired a basic concept about the rules according to which the game is working, why some decks are better than others, and why drawing cards from them is beneficial.

In the original study, the galvanic skin conductance response (GSR) of participants was continuously measured. The GSR reflects how strongly people are sweating, which in turn is a good indicator of the affective arousal and fear of the particular individual. The panel shown at the bottom left indicates the level of GSR during decision-making in the different periods of the study. It is easy to see that the affective response increases with the introduction of losses after the first period. However, while the affective response is gravitating towards baseline when choosing from the good decks, it stays at a high level when choosing from the bad decks. In other words, people learn to become afraid of losses and tend to show fearful reactions when choosing from a bad and risky deck that makes high losses particularly likely.

Compare the behavioral and physiological data on the left with the corresponding data on the right, which stem from patients with lesions in ventromedial prefrontal cortex. The behavior in this group of participants is comparable with the behavior of healthy participants only in the first period. The introduction of losses leads to an entirely different behavioral pattern however,

(continued)

Box 5.1 (continued)

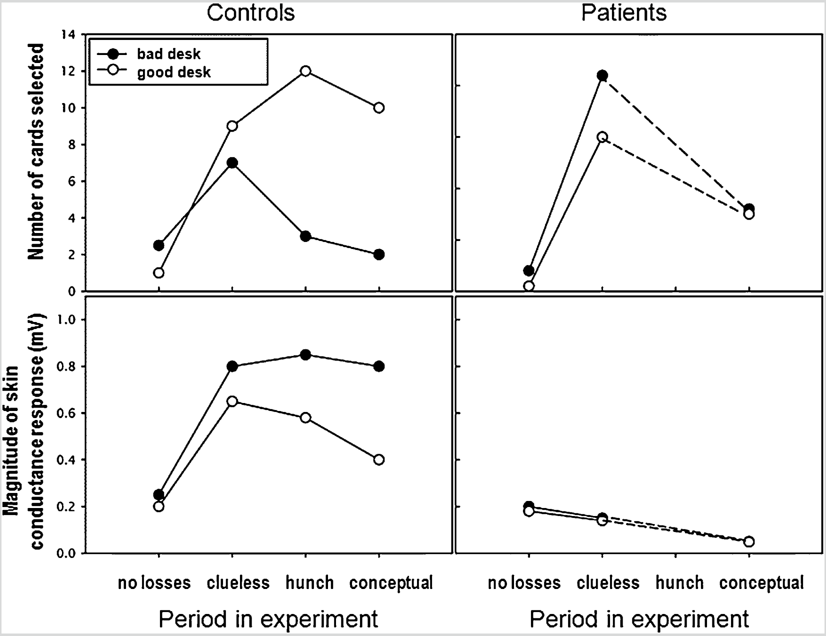


Fig. 5.5 Behavioral and psychophysiological signatures (Galvanic skin response) of decision-making in healthy persons and patients with damage in the ventromedial prefrontal cortex (redrawn after Bechara et al. 1997, by permission of the American Association for the Advancement of Science)

with a clearly visible preference for the riskier and less beneficial decks. The last, conceptual period is reached by 50 % of the patients only, and even the insight that some decks are more beneficial than others does not lead to a corresponding preference for the good decks. The patients are thus not able to make beneficial choices. It is interesting to see that this inability is associated with a complete lack of affective responsivity. In other words, overly risky decision-making seems to be paired with fearlessness.

While the findings from this study show only a correlation between risky decision-making and a lack of fear, the authors have speculated that the connection may be causal in nature, in the sense that people are risky decision-makers if, and because their decisions are not accompanied by fear. The increasingly beneficial decision-making strategies in healthy participants might thus in turn be a consequence of learning to be afraid of losses and loss-related decisions, so that fear in a sense makes people risk-aversive. According to the **theory of somatic markers** (Damasio 1994; see Sect. 2.6.2), decision-making may be based on the simulation of to-be-expected affective consequences, so that people would tend to prefer choice alternatives that are expected to produce positive affective effects. The observation that this mechanism apparently no longer works with lesions in prefrontal cortex fits with considerations of Damasio (1994) and Rolls (1999) that parts of the prefrontal cortex are responsible for computing the **affective valence of action alternatives**.

is important for affective processing, tend to make extremely risky decisions, which suggests that the consideration of affective consequences is imperative for at least some decisions.

From a theoretical perspective, one can imagine that the affective concomitants and consequences of actions are coded and represented cognitively in a similar manner to the previously discussed external, physical effects of actions (Sect. 3.2.2). When we put ourselves in the perspective of our own cognitive system, this makes sense: the cognitive system does not experience the environment directly, but merely through the activation of sensory receptors. What we experience as affect (be it pleasure or pain) is principally nothing but the decisive receptors responding to events inside of our bodies. However, for the cognitive system, our body is just as much outside world as is the environment outside of our bodies. In other words, whether the consequences of a specific action stimulate receptors in our eyes or in our stomachs does not make a principal difference, as it always concerns experienced, sensory consequences of our action (James 1884).

In this sense, the expression of making a decision by "gut feeling" may not just be literally true, but it may also not necessarily indicate an odd or inferior decision-making process. Admittedly, the criteria of these kinds of decision-making processes are logically less accessible and therefore less comprehensible for other people: when you receive either $1000 or 2 kg of chocolate by performing action A (two external events that can be observed by other people), and receive an emotional shiver by performing action B (an event that cannot be directly perceived by other persons), then your choice of action A will be easier to understand for others than a choice of action B. Perhaps that is what causes the "bad press" for intuitive decisions.

References

Ach, N. (1910). *Über den Willensakt und das Temperament*. Leipzig, Germany: Quelle & Meyer.

Anderson, J. R. (1983). *The architecture of cognition*. Cambridge, MA: Harvard University Press.

Anderson, J. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 3–51). London: Guilford.

Bechara, A., Damasio, H., Tranel, D., & Damasio, A. R. (1997). Deciding advantageously before knowing the advantageous strategy. *Science, 275*, 1293–1295.

Bekkering, H., & Neggers, S. F. W. (2002). Visual search is modulated by action intentions. *Psychological Science, 13*, 370–374.

Bertelson, P. (1961). Sequential redundancy and speed in a serial two-choice responding task. *Quarterly Journal of Experimental Psychology, 13*, 90–102.

Bertelson, P. (1963). S-R relationships and reaction times to new versus repeated signals in a serial task. *Journal of Experimental Psychology, 65*, 478–484.

Bertelson, P. (1965). Serial choice reaction-time as a function of response versus signal-and-response repetition. *Nature, 206*, 217–218.

Brebner, J. (1973). S-R compatibility and changes in RT with practice. *Acta Psychologica, 37*, 93–106.

Damasio, A. (1994). *Descartes' error: emotion, reason and the human brain*. New York: G.P. Putnam's Sons.

Donders, F. C. (1868/1969). On the speed of mental processes. *Acta Psychologica, 30,* 412–431. (Original work published 1868)

Dörner, D., Kreuzig, H. W., Reither, F., & Stäudel, T. (Eds.). (1983). *Lohhausen. Vom Umgang mit Unbestimmtheit und Komplexität*. Bern, Switzerland: Huber.

Exner, S. (1879). Physiologie der Grosshirnrinde. In L. Hermann (Ed.), *Handbuch der Physiologie* (Vol. 2, pp. 189–350). Leipzig, Germany: Vogel.

Fagioli, S., Hommel, B., & Schubotz, R. I. (2007). Intentional control of attention: Action planning primes action-related stimulus dimensions. *Psychological Research, 71*, 22–29.

Falmagne, J. C., Cohen, S. P., & Dwivedi, A. (1975). Two-choice reactions as an ordered memory scanning process. In P. M. A. Rabbitt & S. Dornic (Eds.), *Attention and performance* (Vol. V, pp. 296–344). San Diego, CA: Academic.

Fitts, P.M., & Posner, M.I. (1967). Human Performance. Brooks/Cole Pub. Co. Belmont, CA.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11–26.

Hommel, B. (1993). The relationship between stimulus processing and response selection in the Simon task: Evidence for a temporal overlap. *Psychological Research, 55*, 280–290.

Hommel, B. (1998a). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1368–1384.

Hommel, B. (1998b). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183–216.

Hommel, B. (2000). The prepared reflex: Automaticity and control in stimulus-response translation. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance* (Vol. XVIII, pp. 247–273). Cambridge, MA: MIT Press.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*, 188–196.

James, W. (1884). What is an emotion? *Mind, 9*, 188–205.

Keren, G., & Wagenaar, W. A. (1985). On the psychology of playing Blackjack: Normative and descriptive considerations with implications for decision theory. *Journal of Experimental Psychology: General, 114*, 133–158.

Kirby, N. H. (1980). Sequential effects in choice reaction time. In A. T. Welford (Ed.), *Reaction times* (pp. 129–172). London: Academic.

Kornblum, S. (1973). Sequential effects in choice reaction time: A tutorial review. In S. Kornblum (Ed.), *Attention and performance* (Vol. IV, pp. 259–288). New York: Academic.

Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus response compatibility: A model and taxonomy. *Psychological Review, 97*, 253–270.

Leonard, J. A. (1959). Tactual choice reaction times. *Quarterly Journal of Experimental Psychology, 11*, 76–83.

Lhermitte, F. (1983). "Utilization behaviour" and its relation to lesions of the frontal lobes. *Brain, 106*, 237–255.

Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. *Journal of Experimental Psychology: Human Perception & Performance, 26*, 1072–1090.

Longstreth, L. E., El-Zahhar, N., & Alcorn, M. B. (1985). Exceptions to Hick's law: Explorations with a response duration measure. *Journal of Experimental Psychology: General, 114*, 417–434.

Merkel, J. (1885). Die Zeitlichen Verhältnisse der Wiellensthätigkeit. *Philosophische Studien, 2*, 73–127.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: Part 1. Basic mechanisms. *Psychological Review, 104*, 3–65.

Mowbray, G. H., & Rhoades, M. U. (1959). On the reduction of choice reaction times with practice. *Quarterly Journal of Experimental Psychology, 1*, 16–23.

Niemi, P., & Näätänen, R. (1981). Foreperiod and simple reaction time. *Psychological Bulletin, 89*, 133–162.

Rasmussen, J. (1980). What can be learned from human error reports? In K. D. Duncan, M. M. Gruenberg, & D. Wallis (Eds.), *Changes in working life* (pp. 97–113). Chichester, England: Wiley.

Rhinehart, L. (1971). *The dice man*. St Albans, England: Frogmore.

Robertson, S. I. (1999). *Types of thinking*. London: Routledge.

Rolls, E. T. (1999). *The brain and emotion*. Oxford: Oxford University Press.

Seibel, R. (1963). Discrimination reaction time for a 1,023-alternative task. *Journal of Experimental Psychology, 66*, 215–226.

Simon, J. R., Craft, J. L., & Webster, J. B. (1973). Reactions toward the stimulus source: Analysis of correct responses and errors over a five-day period. *Journal of Experimental Psychology, 101*, 175–178.

Soetens, E., Boer, L. C., & Hueting, J. E. (1985). Expectancy or automatic facilitation? Separating sequential effects in two-choice reaction time. *Journal of Experimental Psychology: Human Perception and Performance, 11*, 598–616.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.

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

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39*, 212–222.

Valle-Inclán, F., & Redondo, M. (1998). On the automaticity of ipsilateral response activation in the Simon effect. *Psychophysiology, 35*, 366–371.

Vervaeck, K. R., & Boer, L. C. (1980). Sequential effects in two-choice reaction time: Subjective expectancy and automatic aftereffect at short response-stimulus intervals. *Acta Psychologica, 44*, 175–190.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review; Monograph Supplements, 3*, 1–114.

Chapter 6
Action Planning

Cognitive process models of human information processing are traditionally rather vague if it comes to the planning and execution of actions. The actually interesting cognitive work, so they often imply, ends with the selection of an action, while the rest is considered "physiology." We have been equally vague so far. We have pointed out that actions are represented by codes of their sensory effects, and that these codes become associated with motor patterns realizing these effects. But how do these motor patterns look like, how are they organized to control an action? The present chapter is devoted to this question.

6.1 Characteristics and Functions of Action Plans

Motor patterns involved in performing intentional actions are often referred to as "**motor programs**." This term has been introduced by Keele (1968, p. 387), who defined a motor program as "a set of muscle commands that are structured before a movement sequence begins, and that allows the entire sequence to be carried out uninfluenced by peripheral feedback." This terminology obviously derives from the 1960s, in which cognitive psychology was enriched by numerous terms and metaphors from the computer sciences. Keele thought that motor programs may be created just like computer programs, by defining each single step towards the end goal in terms of muscle instructions (corresponding to machine language in programming) and then compiling the instructions into a program. This program could be stored, retrieved, and run until completion whenever needed. Before we discuss whether this metaphor is sufficient to capture the flexibility of human action, let us first consider the reasons and observations that led Keele and other researchers to believe that motor programs may exist.

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_6

6.1.1 Independence from Sensory Feedback

Keele (1968) explicitly claimed that programmed actions run through completion **without any impact of sensory feedback**. This presumes that actions can be performed without feedback in the first place. Indeed, Lashley (1917) reported about a wounded soldier who, despite the loss of all kinesthetic feedback, was able to move his leg towards spatial goals at different speeds even with his eyes closed. The control of these actions was apparently independent from the availability of external information, suggesting the existence of an internal control structure. Similar observations come from other deafferented patients. For instance, even people who lost their ability to process proprioceptive information about their exerted effort and the relative position of effectors and other body parts through accidents or diseases are able to perform rather complex activities, like drawing figures (Rothwell et al. 1982) or synchronizing their finger movements with an external sequence (Bard et al. 1992). Likewise, experimentally deafferented monkeys are able to walk, jump, and climb without any kinesthetic or visual feedback (Taub and Berman 1968). Once acquired, actions can thus be performed in the absence of any sensory feedback. However, this does not mean that acting individuals do not benefit from using sensory feedback if it is available (see Box 6.1). It is indeed easy to imagine various activities that would be much more difficult, and that would be performed much more poorly without sensory feedback—just think of downhill skiing with your eyes closed. Moreover, the acquisition of novel actions is almost impossible without information about relative success (Thorndike 1927).

6.1.2 Anticipation of Future Action Elements

Motor programs in the sense of Keele (1968) contain entire actions and, thus, all elements a given action comprises. The underlying assumption that action planning spans multiple components receives strong support from the observation of so-called **anticipation effects**. Indeed, the way a given element of an action sequence is performed often reflects characteristics of forthcoming elements of the same sequence. An everyday example for such anticipation effects is **goal-directed grasping**. If you record the kinematics of a grasping action and analyze it in detail (e.g., by watching a recorded movie in Slo-Mo), you can see that the posture of the hand is adjusted to the size and shape of the to-be-grasped object long before reaching it: the larger the object, the more the hand is opening even several centimeters and seconds before the object is actually touched (Jeannerod 1981). The movement thus anticipates the object to be grasped, which suggests that object-relevant movement parameters have been programmed beforehand.

Anticipation effects can also be observed in speech, where they are called "**coarticulation effects**." Just pronounce the word "screen" and watch what you're doing with your lips. Now pronounce the word "screws." What you will see is that, in contrast to the first example, you will round your lips from the first sound on, even

Box 6.1: The Role of Visual Information in Controlling Manual Movements

The probably first empirical study on the role of visual information in the control of manual movements was reported in 1899 by Woodworth. His participants were to copy lines of particular length by means of a pen, with eyes open or closed. The dependent measure was accuracy, that is, the difference in length between the model line and the copied line. Apart from the length of the lines, Woodworth also manipulated the speed of the movement: each movement was to be carried out within one stroke of a metronome, the speed of which was varied.

Figure 6.1 shows the results for one of the four investigated participants, whose performance was comparable. It is easy to see that performance was better with eyes open than with eyes closed, but even with eyes closed performance was not random. Participants were thus able to perform the required movement without visual feedback, but accuracy benefited from the availability of such feedback. However, this was the case only when the movements were slow. With increasing speed, the benefit from visual feedback tended to disappear and there was not any benefit left with the fastest movements.

Why that is becomes clear if we consider the temporal demands of the processes involved. Carrying out 200 movements per minute leaves no more than 300 ms for each given movement. Transmitting information from the human retina to the relevant cortical circuits requires about 100–140 ms, and it takes another 80 ms to have the muscles realizing a movement. With very fast movements, it is thus almost impossible to process the relevant information and adjust the ongoing action plan. The slower the movements are, the more time and opportunity is available to carry out corrections to increase the accuracy of the movement.

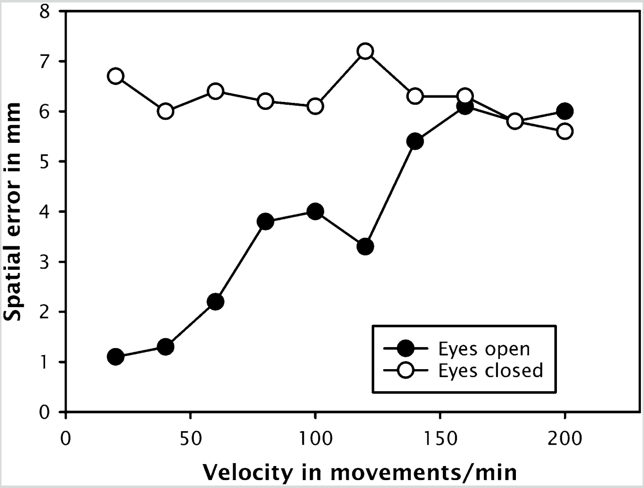


Fig. 6.1 The role of visual feedback in controlling the accuracy of manual movements

though the first four letters of the two words are identical. You apparently anticipate the fact that you will need to round them for the last ("u") part of the world, which demonstrates that the ending of the sound sequence is already known and considered from the start. There are many more examples showing that each given sound is produced and pronounced differently in the context of (i.e., before or after) other sounds. Your speech production is thus anticipatory and considers more than just the next element of a sound sequence.

Very similar effects can be observed in the **planning of manual actions**. In the study of Rosenbaum et al. (1990), participants were to move cylinders from one location and position to another. Participants spontaneously adopted awkward postures (i.e., joint angles at or near the extremes of the joint angle range) if those postures permitted comfortable postures (i.e., joint angles at or near the middle of the joint angle range) at the end of the transport task—the so-called *end-state comfort effect*. Maximizing end-state comfort makes perfect functional sense, but it requires that the final element of the action is considered right from the start of planning.

Evidence for the anticipation of later elements in action planning can also be obtained from the analysis of **action errors**. Particular funny examples have been generated by the English philosophy professor William Archibald Spooner—known as **Spoonerisms**. History has it that he tended to have the strong tendency to exchange sounds in successive words, which created remarkable products such as "the queer old Dean" (instead of "the dear old queen"). Errors of this sort can be experimentally induced by means of priming techniques, which has opened new avenues to studying speech errors systematically (e.g., Motley et al. 1982). A simple example you can try at home: have someone say "ork, ork, ork" (the prime) and then ask "with what do you eat your soup?" Similar errors can be observed in typewriting (Rumelhart and Norman 1982), where one particularly frequent error is the "correct" doubling of the incorrect "leeter." All these examples suggest that **action planning goes beyond the next upcoming action element**, suggesting that action plans consider many, often all elements an action comprises.

6.1.3 Complexity Effects

The notion of a motor program implies that many or all aspects and elements of an action are **planned in advance**, that is, before the action is initiated. As the planning of each aspect or element should take some time, this implies that it should take more time to initiate complex as compared to simple actions. Hence, initiation time should increase with action complexity. Henry and Rogers (1960) tested this prediction by having participants execute actions that varied in length and number of elements, and by measuring reaction time from the presentation of a start signal until the first movement. As expected, reaction time increased with a number of elements, even though all elements were known and could be prepared before the presentation of the start signal. Henry and Rogers explained this observation by assuming that action planning consists in transferring instructions for each single element into a motor buffer (a “memory drum”), so that planning duration increases with the number of elements (i.e., the corresponding transfer steps). Numerous studies could replicate and extend the findings of Henry and Rogers in better controlled experiments, showing for instance that initiation time increases with the duration and distance of hand movements (e.g., Klapp 1975), the number of syllables in vocal actions and the number of strokes in word production on a typewriter (e.g., Sternberg et al. 1978).

6.2 The Structure of Action Plans

Numerous observations and experimental findings have provided ample support for Keele’s (1968) assumption that actions are planned in advance and that their execution is guided by the prepared action plan. And yet, the general concept of a motor program has been frequently criticized, in particular the assumption of **muscle-specific representations**. Consider a simple action, such as grasping a cup in front of you. Performing such an action relies on the contribution of numerous muscles and each tiny change in the path or speed requires the reprogramming of many of them. Each (re)programming step requires the parameterization of the respective muscle, which in turn needs to consider all the parameters used to program the other muscles. Where does the motor system get all this information from? Does it really store each single program and maintain programs for each single version of a given action? Even for a single reaching action, this would imply hundreds and hundreds of different programs, and here we are not yet considering the fact that assuming a different posture would require yet another set of programs. And how about new movements that we never performed before? Are we really unable to plan them ahead?

6.2.1 Online and Offline Control

A possible solution to these problems has emerged from observations suggesting that action plans do not specify all aspects of an action. Particularly impressive demonstrations of that principle were obtained in so-called **double-step experiments**. For instance, Prablanc and Pélisson (1990) had participants carry out manual reaching movements from a starting position in front of them to a goal position indicated by a light spot in 30, 40, or 50 cm distance. The participant’s hand was visible during the preparation of the action but was made invisible as soon as it started to move. In some trials, the goal position was changed by moving the respective light spot by 2 or 4 cm. Importantly, this move took place when the participant made an eye movement (during which people are functionally blind), so that participants were unable to consciously perceive the change of the goal.

If a goal-directed hand movement would be controlled by a motor program that specifies all muscle movements in advance, one would expect that in trials with a goal change the hand should first move to the original goal, slow down or stop (reflecting the time needed to set up a new program), and then move to the new goal. This is not what the findings show, however. Instead, the hand was always **moving straight to the new goal** without participants being aware of any change. How is that possible?

It seems that actions are indeed planned ahead but that the planning is restricted to the really important, predictable aspects of an action (such as moving to the light spot in our example) while leaving the specification of less goal-relevant aspects (like the exact distance) to lower-level sensorimotor loops (Heuer 1981). The actual action plan would thus be **incomplete and not include muscle-specific information** but focus on only those aspects and parameters that are crucial for achieving the intended action goal. The further specification of less crucial aspects would be left to the *environment*.

This **sharing of labor between anticipatory action planning and sensorimotor loops** has been associated with the anatomical distinction between a ventral and a dorsal pathway in the human brain (see Sect. 4.1.2). Milner and Goodale (1995) assume that visual stimuli processed along the *ventral pathway* are available for conscious experience while stimuli processed along the *dorsal pathway* serve for unconscious online control of manual actions. According to these authors, the quick and unconscious adjustment of hand movements in double-step experiments is achieved by the dorsal pathway that translates movement-relevant stimulus features into corresponding muscle activities. Glover (2004) has suggested a somewhat different interpretation of the respective roles of ventral and dorsal pathways. According to his suggestion, the ventral pathway serves for anticipatory offline action planning while the dorsal pathway is adding not yet specified action parameters (Hommel et al. 2001) (see Sect. 4.2.1).

The distinction between a **preparatory offline mechanism and a filling-in online mechanism** can explain numerous empirical observations. For instance, people perform saccadic eye movements towards new visual goals in two steps: first they move their fovea centralis (the part of our retina with the best spatial resolution) close to the goal location without actually reaching it, and only then they execute a correction saccade that homes in onto the goal. If the goal object is made to disappear at the beginning of the first movement, no correction saccade is performed (Prablanc and Jeannerod 1975), suggesting that the latter is not preprogrammed. This means that the first movement serves to bring the retina close enough to the goal object that the sensory information it provides allow online sensorimotor loops to take over the fine-tuning second movement. The same logic can be seen in fast pointing and reaching movements, which consist of a preprogrammed ballistic **transport phase**, which is more or less independent from the availability of visual and proprioceptive information, and an approach or **manipulation phase**, which strongly depends on such information (Jeannerod 1984).

6.2.2 Parameters and Features

Action programs are thus less complete than Keele's (1968) notion of motor programs has suggested, which has a number of advantages. Part of the control process can be left to the environment, which makes actions more flexible and easier to plan. Moreover, the assumption that action programs are incomplete helps to address two problems that have been raised by Schmidt (1975): the **storage problem** and the **novelty problem**. If people would indeed create complete, fully specified in muscle-specific action programs, storing all of them arguably exhausts the capacity of a human brain. One may be tempted to argue that we only create, but do not store action programs, which however would raise the question of how we are able to learn motor skills. Related to this question is the novelty problem: if action programs are complete, where do we get all the information from needed to program new movements, such as when learning a new sport?

Schmidt has suggested that both problems could be successfully tackled by dropping the idea of muscle-specific motor programs and replacing this concept by assuming cognitive "schemata." A **schema** is assumed to contain only the invariant features of a particular class of actions (such as manual throwing), while the remaining open parameters (i.e., the not yet specified features) can be used to adjust the action to the current situation. The first step in programming an action would thus consist in activating the relevant schema that includes the invariant features, while the second step would consist in specifying the variable features. For instance, a schema for writing the lowercase letter "A" would contain information about the basic strokes the letter is made up of while slots for variable parameters are left for adjusting the action to produce the letter in several sizes, at several speeds, and on different surfaces.

Figure 6.2 shows a typical empirical observation that can be explained by the schema notion. The figure shows speed profiles that emerge when writing the Latin letter "a." On the one hand, the profiles differ considerably with respect to the total duration of the movement, which is due to the different sizes in which the letter was written. On the other hand, however, the ratio relating the duration of one movement part to another is rather constant. Along the lines of Schmidt (1975), this could suggest that the (invariant) relations between the movement parts involved in writing the letter are controlled by a schema while the specification of the (variable) size of the letter is adjusted according to the current action goal.

In the next chapter, we will get back to the question of how sequences of movements are programmed. However, the considerations of Schmidt already indicate what units programs or action plans are made of: they consist of **representations of movement features**—whether these features are specified in advance or on the flight. This assumption has not only neuroscientific plausibility (Sect. 2.1) but is also shared by theoretical models in other areas, such as Kornblum's model of feature overlap (Sect. 4.2.1) or the theory of event coding (Sect. 4.3.1).

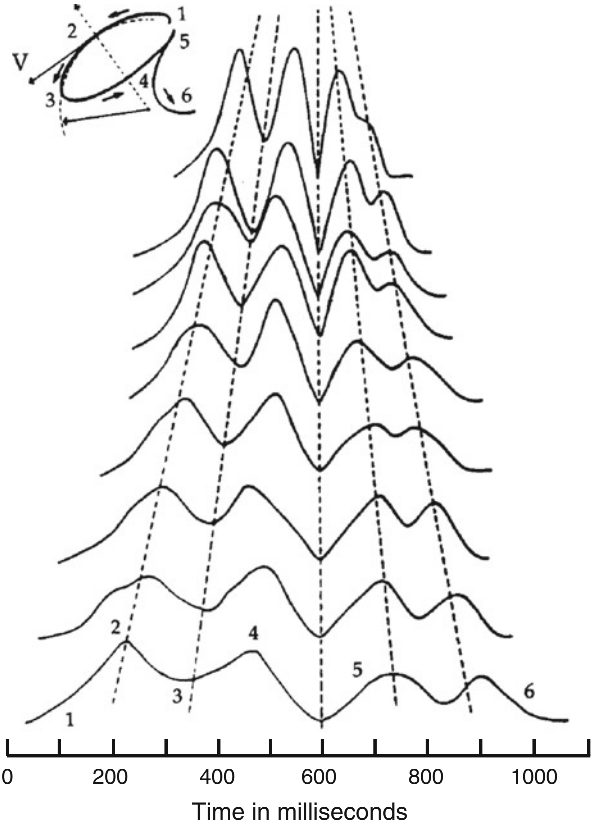


Fig. 6.2 Speed profiles that emerge when writing the Latin letter "a" in different sizes (taken from Viviani & Terzuolo 1980, by permission of Elsevier, Amsterdam, North-Holland)

6.3 Programming an Action

People thus represent actions in terms of their perceived features, so that programming an action consists in specifying its relevant features and activating the codes representing them. First insights into the **process of programming** were obtained in a study by Rosenbaum (1980), who had participants carry out reaching movements from a central home key to spatially defined goal positions. In each trial, a stimulus would signal the required action, which varied with respect to the arm being used, the direction of the movement, and its distance. However, some time before the stimulus, a "precue" was presented which informed participants about some of the features of the upcoming action (e.g., by specifying that the left arm would be required).

If actions are represented in terms of their features and if action programming consists in specifying these features, participants should be able to plan the features specified by the precue. This should reduce their reaction time, which would then reflect the time needed to specify the remaining features. Indeed, Rosenbaum observed that the reaction time decreased with the number of features specified by the precue. Moreover, the feature-specific reductions in reaction time were not dependent on which other features were still to be specified. This suggests that participants were able to **plan the features of their actions separately and in advance**, which in turn supports the notion of action planning as the specification of relevant action features.

Unfortunately, the support is not unequivocal, as **alternative interpretations** of Rosenbaum's findings are possible. For instance, it is possible that participants have represented all eight possible actions as separate programs. The precue could then have served to reduce the number of possible programs, which would have made the final choice easier. According to the Hick-Hyman law, reaction time increases with the number of possible choice alternatives (Hick 1952; Hyman 1953), which opens the possibility to explain the reaction-time benefit in terms of response-choice processes rather than action planning (Goodman and Kelso 1980). To avoid interpretational problems of this kind, **two variants** of Rosenbaum's technique have been suggested and successfully applied (Sects. 6.3.1 and 6.3.2). As we will see, they differ from the original procedure with respect to the validity of the precue.

6.3.1 Action Preparation with a Constant Number of Action Alternatives

One variant of Rosenbaum's technique also uses precues that reduce the number of possible actions. The key difference, however, is that the degree of this reduction is **the same** for all precues while the precued feature varies (Heuer 1982). For instance, Zelaznik and Hahn (1985) used a task with eight possible actions, which resulted from combining the two hands, two fingers, and two durations. Each precue reduced these possibilities to two but different precues allowed for different degrees of preparation. For instance, facing a choice between a

- <brief> keypress with the <index finger> of the <left hand> and a
- <long> keypress with the <index finger> of the <left hand>

provides the opportunity to prepare both the hand and the finger. In contrast, choosing between a

- <brief> keypress with the <index finger> of the <left hand> and a
- <long> keypress with the <thumb> of the <right hand>

does not allow for the preparation of any action feature.

With this version of the task, Zelaznik and Hahn (1985) found hardly any effect on the number of precued action parameters. A possible explanation of this failure to replicate Rosenbaum's (1980) observations could be that Zelaznik and Hahn used a very artificial task and it is for instance questionable whether fingers are represented separately from the hands they are a part of. However, Lépine et al.

(1989) also failed to find systematic effects of the number of prepared features with arguably more natural movements. Another possibility is that choosing between very few action alternatives may allow people to prepare integrated action plans (see Sect. 6.4) that temporarily associate particular parameter values with each other, so that the number of values no longer matters (Lépine et al. 1989). In other words, when facing very few alternatives, people may be able to prepare more than one action at a time.

Taken together, the available studies on the partial preparation of actions provide a rather unsystematic picture and did not yet generate a coherent programming model. What seems to be clear, however, is that informing people about the features of an upcoming action provides planning benefits at least under some conditions, and that these benefits do not depend on the sequence in which the information is provided. This is consistent with the idea that action **planning consists in specifying relevant action features**.

6.3.2 Motor Priming

A second variant of Rosenbaum's (1980) technique is the **method of motor priming** developed by Rosenbaum and Kornblum (1982). Lépine et al. (1989) have combined this method with Rosenbaum's original task and asked participants to carry out pointing movements with the left or right hand, in a forward or backward direction, and with two different distances. However, this time the precue was not always valid but only in 65 % of the trials. Valid precues again reduced the reaction time, that is, participants initiated their actions faster if the precues validly specified action features in advance. More interesting were the findings for the invalid trials however. In these trials, the reaction times were longer the more incorrect parameters were specified by the precue. This suggests that incorrectly specified parameters needed to be reprogrammed, so that the reaction time costs reflected the time needed to reprogram the respective feature. Particularly interesting was the observation that the **reprogramming** of distance produced by far the least costs.

6.3.3 Invariants and Variables

What might be responsible for the time needed to program an action feature? It is possible that programming consists in the activation of particular neural ensembles and that action features differ with respect to the size of the respective ensemble (i.e., the number of neurons making it up). A similar explanation can be derived from Schmidt's (1975) schema theory (see Sect. 6.2). According to this theory, programming an action consists in the activation of a schema and the subsequent parameterization. The schema consists of information about invariant, frequently reoccurring features of the respective action, while the open parameters referred to variable, situation-specific features. If we assume that the activation of a schema takes more time than the subsequent parameterization, we could predict that the **(re) programming of invariant action features is more time-demanding than the (re)programming of variable action features**.

Based on this consideration, Quinn and Sherwood (1983) had participants carry out lever movements of 400 ms duration. In some trials, a signal appearing at or soon after movement onset requested either an increase of movement speed or a reversal of the movement direction. Participants were able to adjust their actions to both of these requests, but the adjustment of the direction took considerably more time than the adjustment of the speed. According to the authors, this suggests that the reversal of the direction requires the (more time-demanding) activation of a new schema while the increase of speed merely requires the adjustment of a parameter. This would fit with Lépine et al.'s (1989) observation that the reprogramming of movement distance was by far the least costly.

Explaining the time demands of programming and reprogramming particular action features by referring to the schema-theoretical distinction between schema-inherent **invariants** and situation-specific **variable** parameters provides a coherent interpretation but unfortunately comes with considerable conceptual and empirical problems. The conceptual problem consists in the lack of agreement regarding the classification of features into invariant and variable features. For instance, in a study of Roth (1988) participants benefited more from valid precues indicating the category of a handball throw than from precues indicating the required force or direction (similarly to handwriting: Stelmach and Teulings 1983). Roth explains this finding by assuming that the category defines the program (or schema) while force and direction represent variable parameters. This interpretation fits with Bairstow's (1987) observation that the direction of manual movements can still be adjusted even after their onset and with the already discussed finding of Prablanc and Pélisson (1990) that the width of manual reaches can be adjusted online. Unfortunately, however, the interpretation of movement direction as a variable contrasts with Quinn and Sherwood's (1983) assumption that movement direction represents an invariant that is a defining part of a schema. Similar problems are raised by the study of Lee et al. (1987). The findings of this study suggest that precues informing about the action category (grasping a ball vs. rotating a potentiometer) do not provide any greater benefit than precues informing about the distance of the goal object, even though according to Schmidt (1975) the category should be an invariant of a schema while distance should be a variable. It is certainly possible that the same feature plays the role of an invariant in some schemata and of a variable in others, but as long as there is no theoretical rationale that allows one to decide that beforehand, clear-cut empirical predictions are impossible.

It is possible that some of the existing inconsistencies have to do with the already discussed distinction between **online control** and **offline control** (Sect. 6.2.1). Assuming that the ventral pathway is responsible for the offline preparation of actions while the dorsal pathway provides online information for the specification of open parameters (Glover 2004), it makes sense to assume that the adjustment of action features handled by the ventral pathway produces considerable processing costs while the adjustment of action features via the dorsal pathway does not.

6.4 Integration of Action Features

The idea that actions are represented through codes of their features suggests that action plans or programs are no unitary entities but **networks of functionally and anatomically distributed representations**. Distributed representations have advantages (they are for instance less vulnerable to the damage of their elements) and disadvantages. Assume, for instance, you are planning to perform a short forward movement with your left hand in a Rosenbaum task. The planning process comprises the activation of the codes of the required feature—the codes representing the features <short>, <forward>, and <left hand>, as indicated in Fig. 6.3a. Let us further assume that you intend to carry out another action before executing that action, a long backward movement with your right hand, say. This would require activating the feature codes <long>, <backward>, and <right hand>, which is

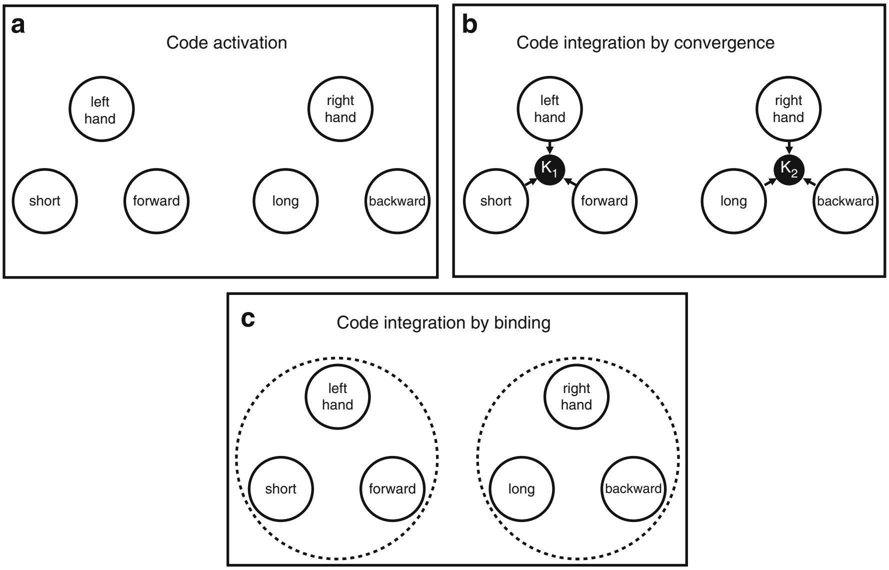


Fig. 6.3 Three possibilities to resolve the binding problem

indicated in the figure as well. The figure also shows the emerging problem: how does the planning process know that the left hand should carry out a forward movement rather than a backward movement and whether it should be short or long? As you can see, having representations of more than one action activated at the same time (which will often be the case when multitasking) creates a mapping problem, which in the literature is often referred to as the "**binding problem**" (Treisman 1996; Hommel 2004). Note that the binding problem is by no means specific to action planning but represents an unavoidable consequence of distributed representations in general.

6.4.1 Integration Mechanisms of Action Planning

One possibility to resolve the binding problem consists in the use of **converging representations**. For instance, there may be codes that do not represent single features of actions but conjunctions of such features. As indicated in Fig. 6.3b, conjunction code C_1 may represent the combination of all features belonging to one action while conjunction code C_2 represents the combination of all features belonging to the other action. The problem of this theoretical option is that it brings us back to Schmidt's (1975) storage problem: if we would create and store conjunction codes for each possible combination of feature codes, the number of required conjunctions codes would be almost infinite.

Another possible solution relies on the **temporary coupling of codes** belonging to the same action, as indicated in Fig. 6.3c. Abeles (1991), von der Malsburg (1981) and others have suggested that elements of distributed neural ensembles might be functionally linked through the temporal synchronization of the firing rates of the respective neurons. Sanes and Donoghue (1993) showed that neural ensembles in different locations of the primary motor and premotor cortex of the monkey show temporal synchrony at time intervals that correspond to the preparation of actions. Also in monkeys, Murthy and Fetz (1992) demonstrated action-related synchronization of the activities of neural ensembles in motor and somatosensory areas, thus covering distances of up to 20 mm.

Studies in humans suggest an equally central role of neural synchronization in the planning of actions. Pfurtscheller et al. (1994) reported rhythmical activity patterns in the EEG preceding the onset of movements with the right or left index finger, the right toe, or the tongue. Of particular interest, the cortical site of the activation corresponded to the anatomical representation of the corresponding movement in Penfield's Homunculus (see Sect. 2.2). Even the temporal dynamics of these neural activities correspond to the specific requirements of action planning: the rhythmical activities end before movement onset with fast, ballistic movements (i.e., with movements that are fully preprogrammed) but continue after movement onset with slow, feedback-sensitive movements (Kristeva-Feige et al. 1993).

6.4.2 *Effects of Action Integration*

If action planning involves the temporary binding of feature codes into integrated action plans, the codes that are integrated into one plan should be difficult or impossible to integrate into another plan. In other words, temporary integration should occupy the integrated feature codes. Evidence supporting this possibility comes from studies on **speech planning** conducted by Meyer and colleagues (Yaniv et al. 1990). In these studies, participants were presented with two possible syllables (e.g., “up” and “ut”) and were asked to prepare the pronunciation of one of them (e.g., “up”). Then a start signal would appear, which requested either the production of the prepared syllable or the production of the alternative syllable. The production of the alternative syllable thus required the reprogramming of the verbal action, which unsurprisingly led to a reaction-time delay. Interestingly, however, this delay was consistently longer if the alternative syllable shared features with the prepared syllable, such as if the ending of both syllables was voiced (“ub”-“ud”) or unvoiced (“up”-“ut”), as compared to combinations without feature sharing (“ub”-“ut”). This might suggest that having integrated a feature, such as voicing, into the first-prepared action plan makes it temporarily less available for another action plan.

Further evidence for this possibility comes from the study of Stoet and Hommel (1999). In this study, participants carried out two actions (A_1 and A_2), which however were planned in the reversed order. That is, participants would first plan A_2, which was a movement with the left or right hand, but not yet carry it out. They would then be presented with a stimulus that signaled A_1, also a movement performed with the left or right hand, which was immediately carried out. Only then the prepared A_2 was performed. If we assume that preparing a left-hand action, say, as A_2 leads to the integration of the corresponding feature <left> into the A_2 plan, this feature should be difficult to integrate into A_1. Hence, the reaction time for performing A_1 should be slower if A_1 and A_2 would be carried out with the same hand. This is indeed what was found. The same pattern was obtained when A_1 was carried out with the left or right foot: foot reactions were slower if their location was shared by the prepared hand action. This suggests that the integration of action features into action plans is indeed blocking the respective codes from becoming parts of other, temporally overlapping action plans (see Sect. 4.3.2).

Feature integration might also be an explanation for an interesting phenomenon reported by Rosenbaum et al. (1986). In this study, participants repeatedly recited sequences of letters of the alphabet with alternating stress (e.g., AbCdAbCdAbCd…, with uppercase indicating stressed letters). Performance (production of letters per minute) was systematically better with even numbers of letters (e.g., AbCdAbCd…) than with uneven numbers of letters (e.g., AbCaBc…). Try it yourself, you will immediately notice what the problem is! Rosenbaum and colleagues interpreted this observation as a “**parameter remapping effect**” and assumed that the parameters of an action plan remain integrated with a plan even after use. Modifying a parameter is time-demanding, which would explain why changing the stress parameter for a particular letter would slow down production. From a feature-binding perspective, this makes sense because even numbers of letters would keep the bindings between letters and <stressed>/<unstressed> codes intact while uneven numbers would require frequent rebinding.

References

Abeles, M. (1991). *Corticonics: Neural circuits of the cerebral cortex*. New York: Cambridge University Press.

Bairstow, P. J. (1987). Analysis of hand movement to moving targets. *Human Movement Science, 6*, 205–231.

Bard, C., Paillard, J., Lajoie, Y., Fleury, M., Teasdale, N., et al. (1992). Role of the afferent information in the timing of motor commands: A comparative study with a deafferent patient. *Neuropsychologia, 30*, 201–206.

Glover, S. (2004). Separate visual representations in the planning and control of action. *Behavioral and Brain Sciences, 27*, 3–78.

Goodman, D., & Kelso, J. A. S. (1980). Are movements prepared in parts? Not under compatible (naturalized) conditions. *Journal of Experimental Psychology: General, 109*, 475–495.

Henry, F. M., & Rogers, D. E. (1960). Increased response latency for complicated movements and a “memory drum” theory of neuromotor reaction. *Research Quarterly, 31*, 448–458.

Heuer, H. (1981). Fast aiming movements with the left and right hand: Evidence for two-process theories of motor control. *Psychological Research, 43*, 81–96.

Heuer, H. (1982). Binary choice reaction time as a criterion of motor equivalence. *Acta Psychologica, 50*, 35–47.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11–26.

Hommel, B. (2004). Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences, 8*, 494–500.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849–878.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 45*, 188–196.

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural objects. In J. Long & A. Baddeley (Eds.), *Attention and performance* (Vol. IX, pp. 153–169). Hillsdale, NJ: Erlbaum.

Jeannerod, M. (1984). The contribution of open-loop and closed-loop control modes in prehension movements. In S. Kornblum & J. Requin (Eds.), *Preparatory states and processes: Proceedings of the Franco-American conference* (pp. 323–337). Hillsdale, NJ: Erlbaum.

Keele, S. W. (1968). Movement control in skilled motor performance. *Psychological Bulletin, 70*, 387–403.

Klapp, S. T. (1975). Feedback versus motor programming in the control of aimed movements. *Journal of Experimental Psychology: Human Perception and Performance, 104*, 147–153.

Kristeva-Feige, R., Feige, B., Makeig, S., Ross, B., & Elbert, T. (1993). Oscillatory brain activity during a motor task. *NeuroReport, 4*, 1291–1294.

Lashley, K. S. (1917). The accuracy of movement in the absence of excitation from the moving organ. *American Journal of Physiology, 43*, 169–194.

Lee, T. D., Elliott, D., & Carnahan, H. (1987). The preparation of actions and parameters of action: A fixed or variable process? *Acta Psychologica, 66*, 83–102.

Lépine, D., Glencross, D., & Requin, J. (1989). Some experimental evidence for and against a parametric conception of movement programming. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 347–362.

Milner, A. D., & Goodale, M. A. (1995). *The visual brain in action*. Oxford: University Press.

Motley, M. T., Camden, C. T., & Baars, B. J. (1982). Covert formulation and editing of anomalies in speech production: Evidence from experimentally elicited slips of the tongue. *Journal of Verbal Learning and Verbal Behavior, 21*, 578–594.

Murthy, V. N., & Fetz, E. E. (1992). Coherent 25- to 35-Hz oscillations in the sensorimotor cortex of awake behaving monkeys. *Proceedings of the National Academy of Sciences of the United States of America, 89*, 5670–5674.

Pfurtscheller, G., Flotzinger, D., & Neuper, C. (1994). Differentiation between finger, toe and tongue movement in man based on 40 Hz EEG. *Electroencephalography and Clinical Neurophysiology, 90*, 456–460.

Prablanc, C., & Jeannerod, M. (1975). Corrective saccades: Dependence on retinal reafferent signals. *Vision Research, 15*, 465–469.

Prablanc, C., & Pélisson, D. (1990). Gaze saccade orienting and hand pointing are locked to their goal by quick internal loops. In M. Jeannerod (Ed.), *Attention and performance* (Vol. XIII, pp. 653–676). Hillsdale, NJ: Erlbaum.

Quinn, J. T., & Sherwood, D. E. (1983). Time requirements of changes in program and parameter variables in rapid ongoing movements. *Journal of Motor Behavior, 15*, 163–178.

Rosenbaum, D. A. (1980). Human movement initiation: Specification of arm, direction and extent. *Journal of Experimental Psychology: General, 109*, 444–474.

Rosenbaum, D. A., & Kornblum, S. (1982). A priming method for investigating the selection of motor responses. *Acta Psychologica, 51*, 223–243.

Rosenbaum, D. A., Marchak, F., Barnes, H. J., Vaughan, J., Slotta, J., et al. (1990). Constraints for action selection: Overhand versus underhand grips. In M. Jeannerod (Ed.), *Attention and performance* (Vol. XIII, pp. 321–342). Hillsdale, NJ: Erlbaum.

Rosenbaum, D. A., Weber, R. J., Hazelett, W. M., & Hindorff, V. (1986). The parameter remapping effect in human performance: Evidence from tongue twisters and finger fumblers. *Journal of Memory and Language, 25*, 710–725.

Roth, K. (1988). Investigations on the basis of the generalized motor programme hypothesis. In O. G. Meijer & K. Roth (Eds.), *The motor-action controversy* (pp. 261–288). Amsterdam: North-Holland.

Rothwell, J. C., Traub, M. M., Day, B. L., Obeso, J. A., Thomas, P. K., et al. (1982). Manual motor performance in a deafferented man. *Brain, 105*, 515–542.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science, 6*, 1–36.

Sanes, J. N., & Donoghue, J. P. (1993). Oscillations in local field potentials of the primate motor cortex during voluntary movement. *Proceedings of the National Academy of Sciences of the United States of America, 90*, 4470–4474.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225–260.

Stelmach, G. E., & Teulings, H. (1983). Response characteristics of prepared and restructured handwriting. *Acta Psychologica, 54*, 51–67.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic.

Stoet, G., & Hommel, B. (1999). Action planning and the temporal binding of response codes. *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1625–1640.

Taub, E., & Berman, A. J. (1968). Movement and learning in the absence of sensory feedback. In S. J. Freedman (Ed.), *The neuropsychology of spatially oriented behavior* (pp. 173–192). Homewood, IL: Dorsey.

Thorndike, E. L. (1927). The law of effect. *American Journal of Psychology, 39*, 212–222.

Treisman, A. M. (1996). The binding problem. *Current Opinion in Neurobiology, 6*, 171–178.

Viviani & Terzuolo (1980), in G.E. Stelmach & J. Requin (eds.), Tutorials in motor behavior (pp. 525–533). © Amsterdam, North-Holland.

Von der Malsburg, C. (1981). *The correlation theory of brain function. Internal report*. Göttingen: Max-Planck Institute for Biophysical Chemistry.

Woodworth, R. S. (1899). The accuracy of voluntary movement. *Psychological Review; Monograph Supplements, 3*, 1–114.

Yaniv, I., Meyer, D. E., Gordon, P. C., Huff, C. A., & Sevald, C. A. (1990). Vowel similarity, connectionist models, and syllable structure in motor programming of speech. *Journal of Memory and Language, 29*, 1–26.

Zelaznik, H. N., & Hahn, R. (1985). Reaction time methods in the study of motor programming: The precuing of hand, digit and duration. *Journal of Motor Behavior, 17*, 190–218.

Chapter 7
Sequential-Action Planning

Experimental investigations of action control often use very simple movements, such as pressing a key, uttering a syllable, looking into a particular direction, or grasping an object. The reasons for that are obvious: simple movements can be carried out quickly and be measured without much cost or effort, which allows even less well-equipped laboratories a high degree of experimental control. From a theoretical perspective, such simplifications should not matter, as even a single keypress on a computer keyboard represents an intentional, goal-directed action. And yet, it is important to consider whether the insights obtained through investigating such simple actions really generalize to the often **more complex actions** we carry out every day. We do not use computer keyboards to press single, arbitrarily determined keys but produce entire words, sentences, and comprehensive texts; and if we grasp objects we commonly do so as part of a more complex, meaningful action sequence, such as preparing a cup of tea. This means that the simple movements favored in psychological experiments represent little more than **elements of everyday action**. This raises a question that we did not yet consider: how do people integrate sequences of simple movements into coherent actions?

In this chapter, we will address three aspects of this question. First, we will discuss under which circumstances and to which degree action elements are integrated into comprehensive action plans (Sect. 7.1). Action plans that comprise multiple components provide an important challenge to the acting person: the components have to be ordered in such a way that the actual goal is eventually achieved, and we will discuss how this might be done (Sect. 7.2). Finally, we will discuss whether and how sequencing and integration changes through practice (Sect. 7.3).

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_7

7.1 Programming Action Sequences

If we observe someone while preparing a cup of tea, we have the impression that we are witnessing a **goal-directed process**. Each single step in the preparation process seems to have a well-defined function within a larger context and each action element brings the eventual goal one step closer. This strongly suggests that the entire action complex has been **planned in advance** and that the plan has defined the specific function of each single contribution to it. But is that really true? It is possible that the bigger picture we envision exists only in our perception, which relies on the fact that we have seen similar actions before, so that we can anticipate the eventual goal from early on. The agent herself, however, might think only one step ahead: seeing the water boiling might trigger the action of filling the water into the cup, which might trigger the next step and so forth. Hence, what we perceive as fully planned may actually consist of an associative chain of movements, in which the next movement is triggered by the completion of the previous one. In fact, we acquire many complex actions in a stepwise, associative way, just think of preparing a new dish or learning a poem from heart.

How can we determine whether people engaging in complex actions are planning ahead more than one step? The previous chapter was posing similar questions such as whether actions are planned ahead at all. Our answer to that question was positive, mainly because of empirical evidence showing that

- Actions can be carried out independently from feedback at least in principle.
- The way action elements are executed sometimes reflects the identity of following elements.
- The time it takes to start the execution of an action increases with the action's complexity.

Similar phenomena have been demonstrated in the context of more extended action sequences. As we will see, such sequences can be **planned ahead** and their execution seems to be **controlled by an action plan**. However, whether and to which degree action sequences are actually planned beforehand depends on a number of conditions, such as the length of the sequence or the dependency of an element on the successful completion of the previous element.

7.1.1 Independence from Sensory Feedback

Let us assume, for the sake of the argument, that we would indeed execute action sequences in a stepwise fashion, that is, without having prepared a comprehensive action plan. We would first perform the starting element, such as grasping for the kettle when preparing a cup of coffee (in the old-fashioned way). The completion of this element would then activate the next: once we have grasped the kettle we would put it on the stove. To organize action sequences in this way, one would thus first have to program and execute the first element, so that the perception of the achieved result (e.g., the feeling of the kettle in one's hand) would **trigger the programming and execution of the next element**. The production of each element and the perception of its results would cost time (of about the order of magnitude of an average reaction time) so that the next action element can be performed no earlier than about one reaction time after the previous one.

You may spontaneously think about a whole bunch of activities for which such an associative-chaining form of action control does not seem particularly plausible. Just consider how fast you can utter entire words or sentences and how little time this high speed leaves for the programming of each single phoneme. Or consider a drum solo. Indeed, humans can produce keypressing sequences in speeds that leave less than 100 ms for each keypress (Sternberg et al. 1978). If we consider Keele and Posner's (1968) estimate that humans need about 190–260 ms for the processing of visual feedback, it is virtually impossible that the individual elements of such sequences are triggered by the feedback about the completion of the previous element.

Empirical evidence **against the associative organization** of action sequences was reported by Gentner et al. (1980). In their study, skilled typists were filmed while writing the phrase "an epic" on a typewriter. Figure 7.1 shows the time points of initiating and the duration of each keystroke in this sequence. The figure shows two examples: The second, shown at the bottom reveals that the movement of pressing the "i" key begins before the keys producing the preceding three elements

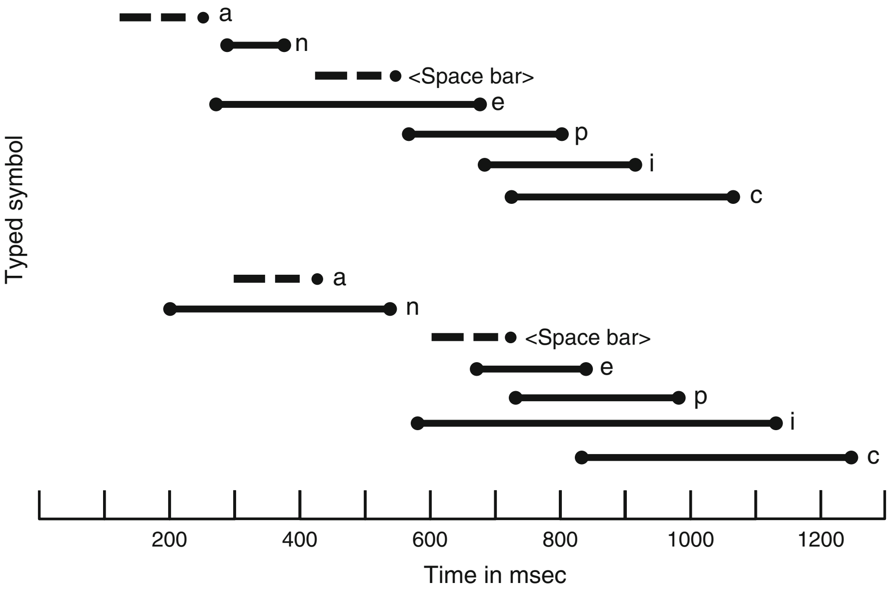


Fig. 7.1 The time points of initiating and the duration of each keystroke when writing the phrase "an epic" on a typewriter (adapted from Gentner, Gruding, & Conway 1980; taken from Rosenbaum 1980)

("_ep") have been pressed. This rules out the possibility that the element "i" could have been triggered by processing the feedback about the programming or execution of the preceding element "p." Similarly, the first example shown at the top reveals that the execution of the element "e" begins before the two preceding elements ("n_") were initiated. We can thus conclude that at least skilled typists are planning manual typing actions ahead, so that their actions are controlled by action plans that consider three or more action elements.

However, while it is easy to find more examples of that sort, there are reasons not to overestimate the argument that elements of sequential actions are **independent from feedback**. For instance, visual feedback might be processed more quickly than Keele and Posner (1968) and others have assumed, as it is possible that other, faster-operating sensory modalities (such as kinesthesis and proprioception) are involved. It is also possible that action elements are not triggered by feedback from the immediately preceding element but from earlier ones, which would leave more time. Furthermore, it is possible that elements are triggered not by the perception of the actual feedback but by the anticipation of expected feedback (see Sect. 7.2.1). In the context of action sequences, the demonstration that action can proceed without immediate feedback is thus less problematic for feedback-based approaches than it is for the control of individual, ballistic actions as discussed in Chap. 6.

7.1.2 Anticipation Effects and Sequence Errors

In Chap. 6 we have already mentioned some examples for **anticipation effects**, such as when the pronunciation of a sound reflects the characteristics of a following sound. Similar examples can be found in the context of sequential action. For instance, in typewriting the time it takes to produce a sequence of letters depends on the word in which this sequence occurs (e.g., Terzuolo and Viviani 1980). Particularly well investigated are **sequence errors** as observed in vocal and manual word and sentence production (Fromkin 1980). Within individual words one can observe reversals of letters, such as when saying "bootfall" instead of "football" (Garrett 1982) or typing "ignroe" instead of "ignore" (Shaffer 1975), and even letters of the following word can induce errors such as when saying "spear bill" instead of "spill beer" (Dell 1986). Anticipation-induced errors have been reported even across multiple words as obvious from sentences like "you hissed all my mystery lectures" (Potter 1980) or "writing a mother to my letter" (Dell 1986). These last two examples are particularly interesting in showing that the syntactical structure of the sentences is correct and thus not affected by the semantic error. It would thus not be correct to say that the verbal action planning failed entirely; the error rather points to a specific problem in assigning sequence elements to syntactically defined sequence positions (see Sect. 7.2). In any case, all these examples show that people are **planning vocal and manual actions ahead** and that the action plans underlying our actions are **more comprehensive than just chaining** one element to the next.

7.1.3 Complexity Effects

In Chap. 6 we have seen that the time it takes to initiate an action can be informative with respect to how comprehensive the underlying action representation is. If we would plan actions in an entirely associative fashion, the time taken to initiate the first element should not depend on the complexity of the entire action and the number of elements that follow. If, however, sequential actions are controlled by complex plans that take many or all action elements needed to reach the eventual goal into account, the time to initiate the first element should increase with the complexity of the action. Just like a computer needs more time to load a more comprehensive program, humans should take more time to initiate a complex action.

Evidence supporting this prediction has been reported by Sternberg et al. (1978). These authors had humans produce verbal sequences consisting of one to four mono- or bisyllabic elements and they made two important observations:

- Reaction time for the first syllable of a sequence increased with the number of words that would follow.
- Reaction time was longer for bisyllabic than for monosyllabic words.

Both observations show that longer and more complex action sequences require **more preparation,** which in turn suggests that the production of these sequences relies on **comprehensive action plans**. Similar outcomes have been reported from other tasks: the reaction time for the first sequence element does not only increase with the number of to be uttered syllables (Eriksen et al. 1970) and words (Monsell 1986), but also with a number of to be pressed keys (Sternberg et al. 1978), to be written symbols (Hulstijn and van Galen 1983), to be performed hand movements (Harrington and Haaland 1987), and to be executed eye movements (Inhoff 1986).

Another empirical approach to investigate the planning of action sequences consists in studying the time it takes to decide between response alternatives that differ in complexity. Rosenbaum et al. (1984) had participants carry out sequences of two or three keypresses. In each trial, a stimulus would indicate which of the two prepared sequences was to be carried out. The two prepared sequences differed only with respect to one element and the position of this element in the sequence was manipulated. The reaction time for initiating the sequence was higher for earlier than for later positions of the differing element (e.g., longer for deciding between AXC and AYC than for deciding between ABX and ABY). This suggests that participants were able to **plan ahead** up to the not yet known element, so that the reaction time reflects only the time needed to complete the plan from that element on.

The available evidence shows that people are able to plan ahead even actions that are more complex than single movements, which in turn suggests that our conclusions from the investigation of simple actions can also be applied to more complex actions. Planning ahead has its **limitations** however. Just think of a skilled pianist: it is not particularly plausible that it would take her twice the time to initiate the first note of a 20-min piece than of a 10-min piece. It is equally implausible that it would take a distance runner twice the time to start in a 10,000 m race than in a 5000 m race. Hence, we are able to plan ahead our actions in principle but we do so only if planning ahead is really necessary or more efficient.

7.2 Sequencing Action Elements

Planning an action that consists of multiple elements presents the planner with a particular challenge: not only does she have to specify the required elements of the action but she also needs to determine the order in which these elements are to be executed. How specifying and planning an individual element might work we have already discussed in Chap. 6. But how might the **sequencing problem** be tackled? Let us for instance assume a skilled English speaker intends to utter the word "word." A skilled speaker will be able to represent the utterance at least at three levels (Fig. 7.2):

- A **conceptual level**, at which the utterance is represented by the concept <word> (concepts are indicated by angle brackets)
- A **stimulus level**, at which the individual stimuli (letters in this example, but one may also think of configurations of letter parts) corresponding to the concept are represented (indicated by uppercase letters)
- A **motor level**, at which the corresponding phonological patterns and the motor patterns producing them are represented (indicated by lowercase letters in quotes)

We can assume that the representations at the different levels are associated with each other, so that the activation of the concept leads to the activation of the stimulus representations, and that the activation of the stimulus representations leads to the activation of the corresponding motor representations. So far so good: being presented with the visual stimulus "word" will thus more or less directly prime the

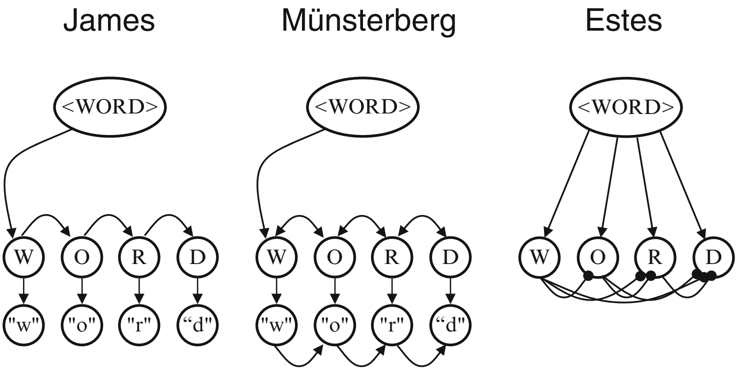


Fig. 7.2 Chaining models for the planning of producing the word "word"

corresponding motor representations, so that all required elements of the action are specified. But if all four motor representations are active at the same time, what prevents us from saying “drow” or “ordw” instead of “word”? How do we make sure that the activated elements of the action are carried out in the right order?

Over the years, various theoretical ideas and models have been suggested to account for this ability, and ongoing discussions favor one over the other. We do not share the underlying opinion that the available models represent mutually exclusive alternatives however. As we have seen and as we will see, action sequences can be planned in very flexible ways. It is thus possible that the exact same sequence is planned in different ways, depending on the situation, the particular conditions, and the degree of practice. It is also possible that different kinds of sequences are planned in systematically different ways—e.g., depending on their length. If so, it is possible that **different models of action sequencing apply to different situations and conditions** (see Verwey et al. 2015).

7.2.1 Chaining Action Elements

7.2.1.1 James’ Model

Early models of action sequencing assumed that the execution of action sequences is controlled through associative chains between representations of action elements. The particularly influential **chaining model** of James (1890) assumed that this chaining process occurs at the level of stimulus representations (Fig. 7.2, left panel). To utter the word “word,” one would thus activate the corresponding stimulus representations W, O, R, and D in a sequence, which would produce a sequential utterance by priming the corresponding motor patterns (W→“w,” O→“o,” R→“r,” D→“d”). Repeating the sequence and the corresponding phonological patterns over and over again would then create direct forward associations between individual stimulus representations and the subsequent stimulus element (W→O, O→R, R→D). Executing the motor pattern corresponding to the first element (i.e., saying “w”) would then produce an acoustic stimulus (a “w”) that would activate the corresponding stimulus representation W, which then would tend to activate the associated stimulus representation O. This would lead to the activation of the corresponding motor pattern and so forth and so on, until the entire word is uttered. Control would thus operate in a stepwise manner at the beginning but would generate a **sequential action plan** through practice.

7.2.1.2 Münsterberg’s Model

Münsterberg (1889) strongly doubted that action sequences can be controlled through direct associations between stimulus representations. Associations between stimulus representations, so he assumed, are necessarily **bidirectional**: if stimulus

X is associated with stimulus Y, then stimulus Y is associated with stimulus X as strongly (see Fig. 7.2, center). If so, proper sequencing in time cannot be guaranteed by associating stimulus representations, as for instance activating O would activate W just as strongly as it would activate the subsequent R. As a solution, Münsterberg suggests that sequences are not controlled by associations between stimulus representations but by **associations between motor representations**, that is, at the motor level. It is true that this theoretical move would solve the possible problem of bidirectional stimulus associations, but it is not quite clear why Münsterberg believes that this problem is restricted to the level of stimulus representation. If associations between motor patterns would be as bidirectional as between stimulus representations, and there is no reason why they shouldn't, the order problem would be no less pressing at the motor level.

Notwithstanding this weakness, the idea that chaining may occur at the motor level is theoretically interesting however. A key argument against chaining models along the lines of James (1890) refers to the role of feedback. The logic of James' model requires that feedback from a preceding element is responsible for activating the subsequent element, and we have seen that some sequential actions seem to be so fast that they do not seem to leave enough time for processing such feedback (see Sect. 7.1.1). But if it would be not stimulus representations that would be associated but motor representations, there would be no need to wait for sensory feedback—activating one motor component could directly trigger the next. Modern models of the acquisition of action sequences have indeed fruitfully adopted Münsterberg's motor-chaining approach (e.g., Hoffmann and Koch 1997; Nattkemper and Prinz 1997).

7.2.1.3 Greenwald's Model

Another possibility to make chaining models more consistent with the available data is to consider that subsequent elements of a sequential action may not be triggered by the actual feedback from the preceding element but by the mere **anticipation of that feedback** (Greenwald 1970). Let us for instance assume that the stimulus representations in the Jamesian model would not refer to the triggering stimuli but to the expected action effects. The sound pattern "w" would thus be generated by activating the representation of the expected feedback W, and this representation would then activate the associated representation of the expected feedback O, and so forth. Given that the representation of an expectation does not need to await the actual stimulus to which this expectation refers, possible temporal delays resulting from feedback processing would no longer be a theoretical argument against chaining. In other words, it is possible that sequential action plans consist of chains of **associations between representations of expected action effects** (Hoffmann et al., 2001; Tubau et al. 2007).

7.2.1.4 Hull's Model

A less satisfying aspect of chaining models consists in the fact that they make no assumptions about the actual **action goal**. Once started, the associative chain runs through to the end without being controlled or guided by some overarching representation of what the action is supposed to achieve. This may be fully sufficient for numerous routine actions or for more extended sequences, such as singing a song or playing a concert. But a control chain would be rather inflexible, so that we would be unable to continue if for instance no kettle is available for heating up the water to prepare our tea. To replace a missing tool or improvise requires information about the ultimate goal of the entire sequence, which raises the question of how sequence production is informed by goal representations.

First measures to address this question were taken by Hull (1931). He proposed that **goals are connected to action sequences in two ways**. First, each sequence will be associated with the motivational state (drive) it is accompanied by. The representation of the sequence is thereby in the sense "colored" by this state, and the state represents the purpose of the sequence. Second, the representation of the ultimate consequence of the action (e.g., the complete utterance "word" in our example) is partially activated throughout the execution of the entire sequence (the so-called "**fractional anticipatory goal response**"). The intended sensory consequences of the sequence (i.e., the representation of the action goal) thus accompany the performance of the entire sequence and will thus be associated with it, so that the way one works through the associative structure representing a given action is controlled through the goal.

Hull's (1931) model assumes that **two representations are active** throughout the entire production of an action sequence: that of the currently performed action element and that of the ultimate action goal. Given the available evidence for anticipation effects (Sects. 6.1.2 and 7.1.2), we also have to assume that **representations of other elements are also active** to some degree, as we have seen in Fig. 7.1, two or more action elements can overlap in time and the execution of the next to next element can begin before the current element is completed. We have also seen that the reaction time for initiating a sequence can reflect the number of elements it comprises, suggesting that multiple, if not all elements of a sequence can be active at the same time. These kinds of observations are very difficult, if not impossible to explain with chaining models.

7.2.1.5 Estes' Model

A model that is trying to tackle that problem up front has been suggested by Estes (1972; Fig. 7.2, right panel). He assumes that preparing an action sequence leads to the pre-activation of all of its elements, which makes it easy to account for anticipation effects and reversal errors. At the same time, however, the planning process also involves the application of a **forward-inhibition structure**, which means that each element inhibits all elements that follow (i.e., that are to be carried out later).

This implies that, when an action sequence is initiated, the very last element of the sequence is the most inhibited, while the first element is not inhibited at all. As a consequence, the first element is the most strongly activated one, so that it in a sense can execute itself. Each executed element will inhibit itself, so that all inhibition produced by that element will be released. As a consequence, the second element will be the most activated, so that it executes itself, inhibits itself, and so forth and so on, until the last element is produced. Based on this model, Rumelhart and Norman (1982) have developed a comprehensive model of typewriting that accounts for typical typewriting errors, letter reversals, and more. However, it is not yet clear how the assumed forward-inhibition structures can be developed and implemented (Rosenbaum 1991) (see Chap. 8).

Taken altogether, it seems clear that pure chaining models are not sufficiently flexible to capture all aspects of sequential action planning. This means that there are styles and modes of sequential action planning that are more complex and complicated than chaining models allow for. However, this does not mean that there are no phenomena that chaining models capture very well. Did you notice how difficult it is to sing a song or to cite a poem from somewhere in the middle? Did you never catch yourself starting then from the beginning and work through the song or poem until that position "in your mind"? Such observations suggest that chaining models may not be complex enough to serve as a generic model of action sequencing but they may well account for the way we deal with long and highly practiced sequences.

7.2.2 Integrating Action Elements

Some subtle discrepancies notwithstanding, the various versions of the chaining model agree in assuming that sequential action plans emerge through the association between representations of action elements. Practice serves to acquire directed links between these representations, so that they in a sense lose their independence and act as a more comprehensive plan. Human action is often **more flexible** than this perspective suggests however. Consider our example of uttering the word "word." Once we have acquired directed associations between the sounds "w," "o," "r," and "d," it should be easy to say "or" but extremely difficult to say "row"—as this would require working against associations that point into the opposite direction. The fact that we do not experience such difficulties suggests that associations are not fully fixed or permanent but rather depend on the context and the current action goal.

Hull's (1931) idea to connect representations of action sequences with representations of motivations and goals already suggests that considering both action plans and goals makes the former more flexible. Other researchers went even further by assuming that representations are not only associated with goals but **integrated** with various kinds of contextual information. For instance, Schwarz (1933) has suggested that representations of previously independent action elements are transformed into integrated wholes through practice—they become a kind of chunk or action Gestalt. Examples of that are manifold, just think of handwriting. Children are writing words in a letter-by-letter fashion while adults often produce units of larger sizes. Hulstijn and van Galen (1988) had adults handwrite more or less complex letters and unfamiliar, arbitrarily constructed symbols. The reaction time for writing an unfamiliar symbol increased with its complexity, similarly to what we have discussed earlier (Sect. 6.1.3). Familiar letters did not show this effect however; here reaction time was independent from complexity. This implies that the production of a **novel** graphical pattern is planned in terms of a sequence of strokes, so that the reaction time increases as more strokes are to be planned. In contrast, **familiar** letters are produced by an integrated action plan that treats the entire production as one unit (Teulings et al. 1986). Similarly, studies on the acquisition of grammatically structured sequences have revealed that people quickly create **subunits** that chunk together successive elements of a sequence (e.g., Servan-Schreiber and Anderson 1990).

A rather popular integration model has been suggested by Wickelgren (1969). This model assumes that elements of a sequence are represented in an entirely **context-dependent** fashion, in the sense that the representation of each element contains information about the preceding and the following element. According to this logic, the sound sequence "word" would be represented by four chunks (#wo, wor, ord, rd#). The representations of the sounds included in these representations can also be integrated into other representations, such as in the case of "world" (#wo, wor, orl, rld, ld#). On the one hand, this approach still raises the question why for instance people can say "word" instead of "world" even though the first two elements of these two words are identical. It also remains unclear how these kinds of representations allow for the extraction of general word production rules. On the other hand, however, the assumption that representations of action elements and sequences can be contextualized represents a major theoretical step towards more complex models of action sequencing.

7.2.3 Hierarchical Control

Models of hierarchical control assume that actions are planned by creating multilevel representations that divide action elements into functional units and subunits. For instance, preparing a cup of tea may be assumed to comprise a representation of "pouring tea into a cup" that falls into the subunits "lift teapot," "move teapot to cup," and "tilt teapot until tea pours" (Humphreys and Forde 1999). Each of these subunits may of course be further divided into the movement sequences they require, so that for instance "lift teapot" can be subdivided into "approach teapot with dominant hand," "open hand," "grasp handle," and so forth. The assumption of mere chaining makes it difficult to cover the complexity of the entire action. Even the assumption of integrated chunks without assuming representational levels that differ in abstraction does not seem to be sufficient—as Lashley (1951) has pointed out for the planning of verbal actions.

To tackle these challenges, Miller et al. (1960) have suggested a **hierarchical planning model** that comprises several layers. Figure 7.3 illustrates the key assump-

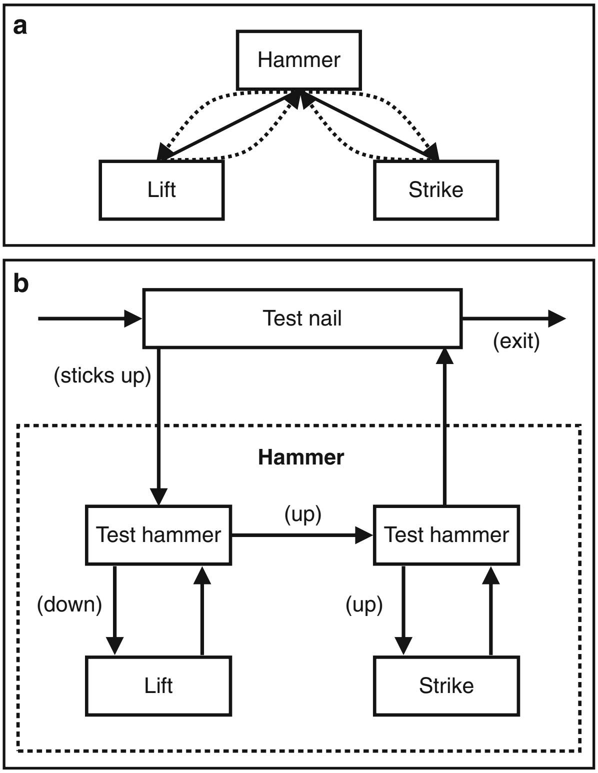


Fig. 7.3 Hierarchical planning model of Miller et al. (1960)

tions of the model for the relatively simple example of hammering a nail into a piece of wood. Figure 7.3a shows the basic idea: the hammering action consists of the components lifting the hammer and striking, where the first component always activates the second (as indicated by the stippled lines). Figures 7.3b makes this grammatical sketch more concrete: the nail will be tested first and, if it still sticks up, the "hammer routine" will be initiated. The nail will be tested again and the whole sequence will be carried out until it no longer sticks out. The logic of this procedure follows the control-theoretical **Test-Operate-Test-Exit (TOTE) principle**, which is here embedded into a hierarchical action plan.

Box 7.1: Evidence for the Hierarchical Planning of Manual Action Sequences

The tendency to plan action sequences in a hierarchical fashion is obvious in rather simple experimental tasks already. Rosenbaum et al. (1983) had participants perform sequences of quick finger movements and measured the initiation time for each single element. Figure 7.4a shows the results for the sequence of movements with the

(continued)

Box 7.1 (continued)

left middle finger → right middle finger → left middle finger → right middle finger →

left index finger → right index finger → left index finger → right index finger.

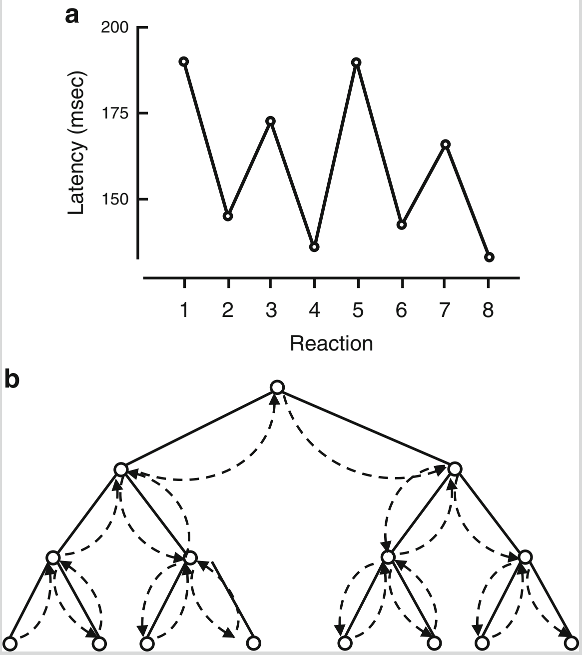


Fig. 7.4 Latencies of initiating individual keystrokes when producing sequences of 1–8 keypresses (redrawn after Rosenbaum, Kenny, & Derr 1983, by permission of the APA)

It is easy to see that the reaction times follow a particular pattern: every second movement is initiated faster than its predecessor. This suggests that the action plan has created chunks integrating two movements each.

Further studies of these authors confirmed that this pattern does not rely on the individual finger movements or the identity of the finger but results from the way the different movements are combined. Very similar patterns were reported by Povel and Collard (1982) for the verbal production of numerical sequences. These authors suggested that action planning makes use of **binary decision trees**. This concept, which was introduced by Greeno and Simon (1974), assumes that sequences are translated into hierarchically structured binary decision trees, and that the production of these sequences is controlled by reading out these binary structures. According to Rosenbaum et al. (1983), the findings shown in Fig. 7.4a could thus imply that action plans consist of

(continued)

Box 7.1 (continued)

binary decision trees (as indicated in Fig. 7.4b), the readout of which would indeed produce the obtained zig-zagging pattern.

Planning and retrieval of an action plan could thus be thought of as a kind of **mental travel** through the branches of decision trees. The planning would start at the highest level, move to the first representation on the left at the lowest level, move to the next on the right of it until the rightmost representation, with the distance covered representing the planning time. Note that this would predict that initiation time would be longer for every first member than every second member of each pair of successive action elements—just as observed by Rosenbaum et al.

The concept of hierarchical action planning has a number of advantages but a number of disadvantages as well. One of the disadvantages is that the **prediction of behavior becomes more difficult**. For instance, it is unlikely that hierarchical decision trees (Fig. 7.4) are necessarily binary in nature. Assume, for instance, you are producing (e.g., verbalizing or writing) the numerical sequence 1-2-3-11-12-13-21-22-23. Would you not be tempted to subdivide the sequence into three chunks comprising 1/2/3, 11/12/13, and 21/22/23? In this example, a particular kind of grouping is very suggestive but how do people structure a shopping tour or a holiday trip? How do they chunk their telephone number?

The concept of hierarchical planning allows the distinction between the elements of a sequence on the one hand and a syntactical control structure on the other. This has considerable advantages for modeling **speech production** (Dell et al. 1997) and the production of **manual sequences**. For instance, Ziessler et al. (1988) had participants perform sequences similar to those studied by Rosenbaum et al. (1983). Before the actual stimulus appeared, which signaled the required movement sequence, participants received pre-information about various aspects of the upcoming action. The research question was whether some kinds of pre-information would be more useful and effective in reducing the reaction time than others. It turned out that pre-information about the general structure of an action (e.g., whether a switch of finger or hand was required at a particular position in the sequence) reduced the reaction time even if the individual action elements were not yet known. This demonstrates that people can plan general characteristics of an action independently from its elements. Similarly, Semjen and Gottsdanker (1991) were able to show that transitions between more forceful and less forceful finger movements could be planned even before the direction of the transition (increase versus decrease of force) was not yet known.

Consistent with hierarchical models of sequential action planning, planners can thus plan content-specific and syntactic characteristics of an action **independently** from each other. This requires separate representations of the sequence elements

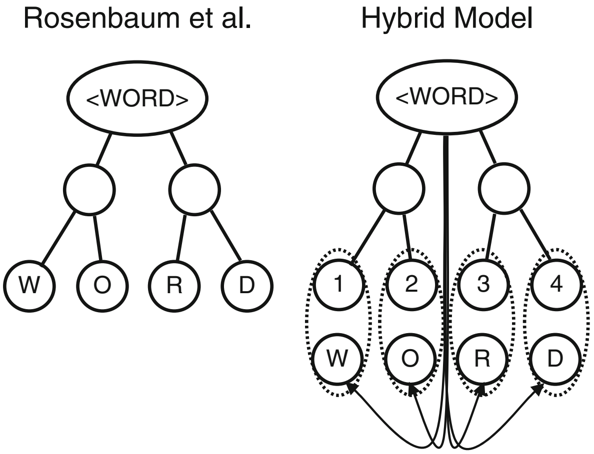


Fig. 7.5 Hierarchical models for the planning of producing the word "word"

proper and the order in which they are performed. How that might work is shown in Fig. 7.5. The left example shows the original concept of binary decision trees as suggested by Rosenbaum et al. (1983). The right example shows a version that is extended by a syntactical level—indicated by the numbers 1–4. The number-labeled representations referred to syntactical slots that stand for the position of an element in the sequence. Each slot is then integrated with a particular sound, so that no direct links between sound representations are necessary. Note that none of the two versions of the model really requires the binary-decision level, which would also make the unreasonable prediction that the pause between producing "o" and "r" would be longer than those between the other elements.

7.3 Planning Extended and Practiced Action Sequences

The available findings on the planning of action sequences do not leave any doubt that people are able to plan multiple elements of a sequence in advance. Some studies demonstrate that action plans include all elements of the respective sequence but one can ask whether this is always the case. We have already discussed some examples where this would be implausible and it is easy to find more, just think of a vacation trip or a theater visit. Whether people do or do not plan their actions has likely to do with the **advantages and the disadvantages of action planning**. Planning reduces the time needed to perform the action but it does require a time investment itself. If thus the duration of the performance matters, such as in a 100 m sprint race, it is better to plan ahead. But if the elements of an action take a lot of time to be carried out, such as on a holiday trip, there is little point in wasting time and energy on preparing more than necessary. Moreover, action plans do not only require time for making them, maintaining a plan has also considerable cognitive costs—demands on working memory capacity in particular. The more extensive the plan, the higher these costs are which, given the severe limitations of working memory, can have serious consequences for other cognitive processes that need to be carried out while maintaining the plan. That does not necessarily mean that longer sequences should not be planned. But the longer a sequence is, the more time it provides for the planning of later elements while executing the earlier ones. In other words, a larger number of sequence elements provides increasing opportunities to distribute planning activities over time.

Consistent with these considerations, Klapp and Wyatt (1976) have concluded from an overview of earlier studies on verbal actions that the **increase of initiation time with the number of sequence elements increases continuously but not linearly**. That is, the longer a sequence the smaller the increase of pre-initiation planning time for each extra element. For manual action sequences, significant increases of initiation time could sometimes be found only up to the second element (Harrington and Haaland 1987) and for longer sequences even a decrease of initiation time with increasing complexity has been observed (Rosenbaum et al. 1987). Planning action sequences ahead is thus **possible** but **not necessary**.

In the following, we will discuss three important factors that determine whether action sequences are planned ahead. We will see that planning ahead is more likely for sequences that are short and not well practiced, especially if they need to be produced quickly. In contrast, if sequences are long and well practiced, if they consist of slow movements, and/or if time pressure is low, people prefer to distribute action planning over time. Moreover, there is evidence that practice changes the way action sequences are planned.

7.3.1 Sequence Length

For reasons of efficiency, early studies on the planning of action sequences have mainly used rather short keypressing or speech sequences. While these studies have commonly shown a linear relationship between initiation time and the number of sequence elements, more recent investigations with longer sequences suggest a more differentiated picture. As mentioned already, there are studies in which participants did not pre-plan more than one or two elements of well-practiced, slow movements (e.g., Harrington and Haaland 1987; Teulings et al. 1986), and little more than six elements with fast, time-critical movements (e.g., Monsell 1986). Longer sequences show a further weakening of the relationship between initiation time and the number of elements (e.g., Garcia-Colera and Semjen 1988).

At least two reasons might be responsible for the weaker impact of the number of elements on the time it takes to initiate longer sequences. For one, this could be a **strategy** (Sect. 7.3.4). People might simply decide to plan no more than a few elements in an attempt to reduce working-memory load or to avoid errors, especially if temporally distributed planning is possible. For another, it is possible that longer sequences provide better opportunities to **organize** the planning process.

Recall the sequence of finger movements in the study of Rosenbaum et al. (1983) (Sect. 7.2.3; Fig. 7.5):

left middle finger → right middle finger → left middle finger → right middle finger →

left index finger → right index finger → left index finger → right index finger.

Such a sequence of eight movements provides more opportunities for the integration (chunking) of its elements. For instance, this sequence might be transformed into an easier and better to recall form such as: 2X (left middle finger → right middle finger) + 2X (left index finger → right index finger). Given that the resulting action plan is more integrated, its demands on time and storage space do not necessarily increase with the number of elements in a linear fashion. For instance, repeating the two stored mini-sequences three rather than two times would not make the plan any more complex, even though it would increase the number of elements by 50 %. Consistent with these considerations, the planning time for longer keypressing sequences in typewriting shows the linear increase of initiation time with the number of elements only if these elements (i.e., the to be typed letters) are randomly combined but not if the sequences contain systematic repetitions that permit the creation of chunks (Kornbrot 1989).

7.3.2 *Movement Duration*

Systematic relationships between initiation time and the length of sequences were mainly observed with the manual production of randomly ordered keypresses in the verbal production of randomly ordered syllables (e.g., Klapp and Wyatt 1976; Sternberg et al. 1978). The elements of the sequences did not take much time to execute, which left little time to plan later components while performing earlier ones. If this is the main reason why planning was carried out before initiation, one would expect that slower movements show a weaker relationship between initiation time and sequence length. Indeed, this relationship is less pronounced with slow manual aiming movements (Harrington and Haaland 1987) and handwriting (Teulings et al. 1986). Even movements that can be carried out quickly in principle hardly show any **sequence-length effects** in the absence of time pressure (Semjen and Garcia-Colera 1986).

7.3.3 *Practice*

The way action sequences are planned is **changing through practice**. While people tried to plan many or all elements of a sequence before initiating the first element, increasing practice leads to a more and more temporally distributed planning mode

(Sect. 7.2.3). Among other things, this is visible in the decreasing strength of the relationship between initiation time and sequence length. As mentioned already, the sequence length effect is obtained for writing movements with novel, arbitrarily constructed symbols but not with familiar letters (Teulings et al. 1986). Likewise, the reaction time costs for the reprogramming of incorrectly prepared letter sequences are very small (Stelmach et al. 1984). Even with sequences of arbitrary movement elements the sequence length effect is visible in the beginning but tends to disappear through training (Klapp and Wyatt 1976).

It seems clear that the way we are planning action sequences is strongly affected by practice. It is less clear how the underlying planning processes are changing however. It is possible that the major change is restricted to the **timepoint of planning**: while the unpracticed novice tends to complete all planning steps before starting the action, the practiced expert might plan later elements while performing the earlier ones, so that performance and planning overlap in time. There is indeed evidence that the time it takes to perform the first element of a sequence changes very little through practice while later elements are performed faster and faster (Portier and van Galen 1992). This suggests that practice moves the planning of later elements to the time the first is executed. Another possibility is that practice allows performing extended action sequences without any planning. Practicing a task could lead to the acquisition of a control structure (that may or may not be hierarchical in nature) that one only needs to retrieve and that then takes over the control of the sequence. This possibility would fit with the observation that practice in typewriting leads to an increasing temporal overlap between keystrokes and corresponding finger movements (Gentner 1983; see Fig. 7.1). A related possibility would be that practice reorganizes the planning process in such a way that different sequencing models (as discussed in Sect. 7.2) apply to different stages in practice; for instance, early stages may involve chaining while later stages rely on hierarchical representations.

References

Dell, G. S. (1986). A spreading activation theory of retrieval in sentence production. *Psychological Review, 93*, 283–321.

Dell, G. S., Burger, L. K., & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review, 104*, 123–147.

Eriksen, C. W., Pollack, M. D., & Montague, W. E. (1970). Implicit speech: Mechanism in perceptual encoding? *Journal of Experimental Psychology, 84*, 502–507.

Estes, W. K. (1972). An associative basis for coding and organization in memory. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 161–190). New York: Wiley.

Fromkin, V. A. (Ed.). (1980). *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand*. New York: Academic.

Garcia-Colera, A., & Semjen, A. (1988). Distributed planning of movement sequences. *Journal of Motor Behavior, 20*, 341–367.

Garrett, M. F. (1982). Production of speech: Observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and pathology in cognitive functions* (pp. 19–76). London: Academic.

Gentner, D. R. (1983). The acquisition of typewriting skill. *Acta Psychologica, 54*, 233–248.

Gentner, D. R., Grudin, J., & Conway, E. (1980). *Finger movements in transcription typing*. La Jolla, CA: University of California, San Diego, Center for Human Information Processing (Technical Report 8001).

Greeno, J. G., & Simon, H. A. (1974). Processes for sequence production. *Psychological Review, 81*, 187–198.

Greenwald, A. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review, 77*, 73–99.

Harrington, D. L., & Haaland, K. Y. (1987). Programming sequences of hand postures. *Journal of Motor Behavior, 19*, 77–95.

Hoffmann, J., & Koch, I. (1997). Stimulus-response compatibility and sequential learning in the serial reaction time task. *Psychological Research, 60*, 87–97.

Hoffmann, J., Sebald, A., & Stöcker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 470–482.

Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review, 38*, 487–506.

Hulstijn, W., & van Galen, G. P. (1983). Programming in handwriting: Reaction time and movement time as a function of sequence length. *Acta Psychologica, 54*, 23–49.

Hulstijn, W., & van Galen, G. P. (1988). Levels of motor programming in writing familiar and unfamiliar symbols. In A. M. Colley & J. K. Beech (Eds.), *Cognition and action in skilled behavior* (pp. 65–85). Amsterdam: North-Holland/Elsevier.

Humphreys, G.W., & Forde, E.M.E. (1999). Category-specific deficits: A major review and presentation of the Hierarchical Interactive Theory. *Aphasiology, 13*, 169–193.

Inhoff, A. W. (1986). Preparing sequences of saccades under choice reaction conditions: Effects of sequence length and context. *Acta Psychologica, 61*, 211–228.

James, W. (1890). *The principles of psychology*. New York: Dover.

Keele, S. W., & Posner, M. I. (1968). Processing of visual feedback in rapid movements. *Journal of Experimental Psychology, 77*, 155–158.

Klapp, S. T., & Wyatt, E. P. (1976). Motor programming within a sequence of responses. *Journal of Motor Behavior, 8*, 19–26.

Kornbrot, D. E. (1989). Organisation of keying skills: The effect of motor complexity and number of units. *Acta Psychologica, 70*, 19–41.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.), *Cerebral mechanisms in behavior* (pp. 112–146). New York: Wiley.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt, Rinehart & Winston.

Monsell, S. (1986). Programming of complex sequences: Evidence from the timing of rapid speech and other productions. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. 72–86). Berlin, Germany: Springer.

Münsterberg, H. (1889). *Beiträge zur experimentellen Psychologie, Heft 1*. Freiburg, Germany: Mohr.

Nattkemper, D., & Prinz, W. (1997). Stimulus and response anticipation in a serial reaction task. *Psychological Research, 60*, 98–112.

Portier, S. J., & van Galen, G. P. (1992). Immediate vs. postponed visual feedback in practising a handwriting task. *Human Movement Science, 11*, 563–592.

Potter, J. M. (1980). What was the matter with Dr. Spooner? In V. A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 13–34). New York: Academic.

Povel, D., & Collard, R. (1982). Structural factors in patterned finger tapping. *Acta Psychologica, 52*, 107–123.

Rosenbaum, D. A. (1991). *Human motor control*. San Diego, CA: Academic.

Rosenbaum, D. A., Hindorff, V., & Munro, E. M. (1987). Scheduling and programming of rapid finger sequences: Tests and elaborations of the hierarchical editor model. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 193–203.

Rosenbaum, D. A., Inhoff, A. W., & Gordon, A. M. (1984). Choosing between movement sequences: A hierarchical editor model. *Journal of Experimental Psychology: General, 113*, 372–393.

Rosenbaum, D. A., Kenny, S. B., & Derr, M. A. (1983). Hierarchical control of rapid movement sequences. *Journal of Experimental Psychology: Human Perception and Performance, 9*, 86–102.

Rumelhart, D. E., & Norman, D. A. (1982). Simulating a skilled typist: A study of skilled cognitive-motor performance. *Cognitive Science, 6*, 1–36.

Schwarz, G. (1933). Über Rückfälligkeit bei Umgewöhnung: 2. Über Handlungsganzheiten und ihre Bedeutung für die Rückfälligkeit. *Psychologische Forschung, 18*, 143–190.

Semjen, A., & Garcia-Colera, A. (1986). Planning and timing of finger-tapping sequences with a stressed element. *Journal of Motor Behavior, 18*, 287–322.

Semjen, A., & Gottsdanker, R. (1991). Structural facilitation of movement sequence planning. In J. Requin & G. E. Stelmach (Eds.), *Tutorials in motor neuroscience* (pp. 397–412). Dordrecht, The Netherlands: Kluwer.

Servan-Schreiber, E., & Anderson, J. R. (1990). Chunking as a mechanism of implicit learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 592–608.

Shaffer, L. H. (1975). Control processes in typing. *Quarterly Journal of Experimental Psychology, 27A*, 419–432.

Stelmach, G. E., Mullins, P. A., & Teulings, H. (1984). Motor programming and temporal patterns in handwriting. *Annals of the New York Academy of Sciences, 423*, 144–157.

Sternberg, S., Monsell, S., Knoll, R. L., & Wright, C. E. (1978). The latency and duration of rapid movement sequences: Comparisons of speech and typewriting. In G. E. Stelmach (Ed.), *Information processing in motor control and learning* (pp. 117–152). New York: Academic.

Terzuolo, C. A., & Viviani, P. (1980). Determinants and characteristics of motor patterns used for typing. *Neuroscience, 5*, 1085–1103.

Teulings, H., Mullins, P. A., & Stelmach, G. E. (1986). The elementary units of programming in handwriting. In G. P. van Galen, H. S. R. Kao, & R. Hoosain (Eds.), *Graphonomics: Contemporary research in handwriting* (pp. 21–32). Amsterdam: Elsevier.

Tubau, E., Hommel, B., & López-Moliner, J. (2007). Modes of executive control in sequence learning: From stimulus-based to plan-based control. *Journal of Experimental Psychology: General, 136*, 43–63.

Verwey, W. B., Shea, C. H., & Wright, D. L. (2015). A cognitive framework for explaining serial processing and sequence execution strategies. *Psychonomic Bulletin and Review, 22*, 54–77.

Wickelgren, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review, 76*, 1–15.

Zießler, M., Hänel, K., & Hoffmann, J. (1988). Die Programmierung struktureller Eigenschaften von Bewegungsfolgen. *Zeitschrift für Psychologie, 196*, 371–388.

Chapter 8
Controlling and Coordinating Multiple Actions

Actions that are used in experimental studies are often much simpler than in everyday life (see Chap. 7). This has many advantages, because simplicity makes it easier to control all the relevant variables and factors, but it is important to ask whether the findings obtained with such tasks can be generalized to reality outside the laboratory. Even strictly controlled laboratory studies should thus keep trying to approach the complexity of everyday life.

Everyday actions often consist of multiple elements that need to be carried out in a particular order to achieve the eventual goal (Chap. 7). But there are other important ways in which they **differ from common laboratory tasks**. One important difference consists in the fact that in everyday life we commonly **do not repeat** the same task hundreds and hundreds of times, as we do in the lab, but switch back and forth between different kinds of action. Just think of preparing a meal in the kitchen or visiting a fun park. The actions we are switching between can be simple or complex, or consist of multiple components. The previous chapters discussed how these actions can be controlled individually, but now we need to ask how we can switch from one of these tasks to another (Sect. 8.1).

Another difference between laboratory studies and everyday action contexts consists in the fact that in everyday life we often carry out **more than one action at a time**. One can ask whether this is actually possible, that is, whether we are truly multitasking or whether we simply shift between different tasks very quickly. Another question is why some actions are easier to combine than others: while it is relatively easy to sing when riding a bike, it is very hard to learn a new language while watching an unrelated movie. Why is that so? Which cognitive processes are responsible for these phenomena? (Sect. 8.2).

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7_8

8.1 Task Switching

Research on how people shift from one task to another took off in the 1920s (Jersild 1927; Zillig 1926). Jersild (1927) presented participants with lists of stimuli to which they were to respond in particular ways. For instance, they could be presented with three lists that all contained random numbers. Participants were to add a constant number to all numbers from the first list, subtract a constant number from all numbers of the second list, and alternate between adding and subtracting with the third list. The time to work through the third list was much longer than the time to work through the other two lists, which suggests that shifting between tasks takes more effort and is more time-demanding. Another important observation of Jersild was that shifting between tasks **did not produce any extra costs** if the stimuli signaled the task they require, such as when alternating between adding a constant to number stimuli and producing rhymes to word stimuli.

While the methods used by Jersild are not without problems (see Box 8.1), his key findings were successfully replicated in numerous studies. We will see that methodological advances have helped to identify and isolate some of the cognitive processes involved in task switching. Before we discuss the most relevant of these processes in more detail, we will first consider how the process of shifting between tasks might be conceptualized, how it might operate in principle, and which cognitive demands it may be associated with.

8.1.1 Theoretical Considerations and Overview

It is our goals that guide and organize our actions. Generally speaking, action goals influence and bias cognitive processing in such a way that actions actually achieve what we want them to achieve, at least in most cases (Chap. 3). How goal representations can fulfill this function is captured in Figs. 3.1 and 3.2.

Naming the color of incongruent color words, such as the word "red" written in green, is so difficult because we have practiced reading words so much more than we have practiced the naming of colors. The model of Cohen et al. (1990) explains how we are able to overcome this practice-induced bias towards reading and are eventually able to name the color nevertheless. The model assumes that **representations of the task context and the current action goal** selectively support the goal-consistent processing route, so that it can outcompete the reading habit. What would be expected if we would ask a participant in a Stroop experiment to read (rather than color-name) the color words? Stroop (1935) could already show that this is easy to achieve. How could that be accounted for in the control model of Cohen et al.? Changing the goal from naming to reading would lead to the deactivation of the naming goal and the activation of the reading goal. As a consequence, the goal representation would no longer support naming but would now support reading, which is rather dominant already. A **task switch** has the purpose of achieving exactly that: the deactivation of old goals and the (re-)activation of the new goal.

Box 8.1 Methodological Advances in the Assessment of Task-Switching Costs

In one of the first studies on task switching, Jersild (1927) had his participants work through lists of random numbers. In some lists, participants were to perform always the same task, such as adding or subtracting a particular number. In other lists, participants were to switch back and forth between two of such tasks. This list method has a number of problems. If participants perform better on lists with just one task than on lists with two tasks, this could be due to a whole number of reasons. It could be that task switching requires more cognitive work, which is the standard interpretation, but keeping in mind two tasks instead of one may also produce higher working memory load, reduce motivation, or induce stress. To avoid such confounds, recent research often uses the task versions developed by Rogers and Monsell (1995) and by Meiran (1996):

- Rogers and Monsell (1995) had participants shift between two tasks in a predictable rhythm (every *n*th trial). The stimuli consisted of letter-digit pairs that were presented in one of four locations marked by boxes on a screen. If the stimulus appeared in one of the top boxes, participants were to classify the letter and if the stimulus appeared in one of the bottom boxes, they were to classify the digit. As the location of the stimulus was predictable (the location rotated in a clockwise fashion), participants were able to prepare the upcoming task. Note that stimulus presentation in the top left and the bottom right box would require a task switch while stimulus presentation in the other two boxes would require a task repetition. The advantage of this so-called **alternate runs paradigm** consists in the fact that task repetitions and alternations occur so fast that changes in motivation and stress are not very likely and that the working memory load is the same for both conditions. One disadvantage is that the sequence of tasks is predictable, so that participants are able to prepare for the upcoming task before the stimulus is presented. However, when they start preparing remains unclear.
- Meiran (1996) used a **task-cuing paradigm** to assess the preparation process more precisely. He randomized the sequence of the tasks, so that participants did not know which task would be required in the upcoming trial. The required task was signaled by means of a task cue that was presented some time before the new stimulus (see also Dixon 1981). Figure 8.1 shows an example of two successive trials in the study of Meiran. The task was to respond to the horizontal or vertical location of the stimuli by pressing a top-left or a bottom-right key. The relevant dimension was indicated in advance—this was the task cue. After a cue-stimulus interval (CSI), the stimulus would appear in one of the four boxes and participants were to respond as fast as possible (in the first trial of the example, they were to press the top-left key to indicate that the stimulus appeared at the top). After

(continued)

Box 8.1: (continued)
the response-cue interval (RCI), the next task cue would appear and indicate the relevant dimension for the next trial, followed by the stimulus.

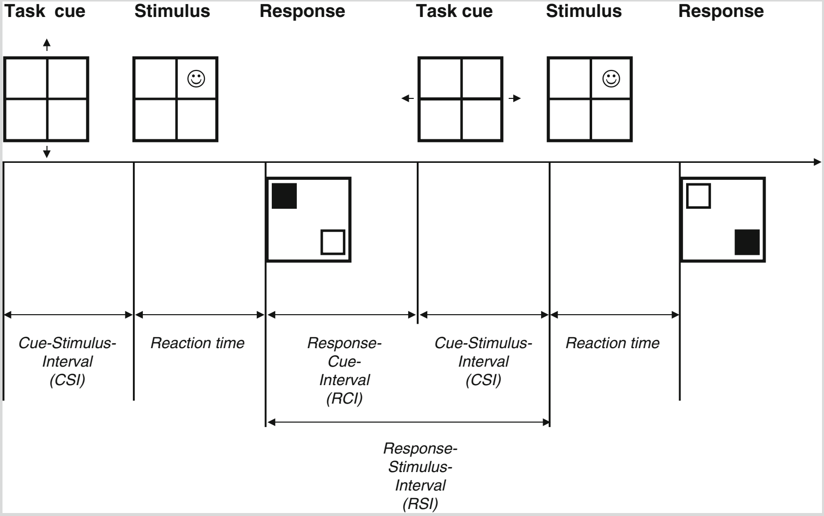


Fig. 8.1 An example of two successive trials in the task-cuing paradigm of Meiran (1996)

In contrast to the alternate-runs method of Rogers and Monsell (1995), Meiran's cuing method allows the separation of some processes that are likely to play a role in preparing for a task switch. The CSI determines how much time participants have to prepare for the upcoming task, so that one can expect that a longer CSI allows for better preparation than a short CSI. The RCI determines how much time is available after the completion of the previous task, and one can expect that possible after-effects of the previous task are getting weaker as this interval increases.

These considerations already suggest a number of requirements, problems, and possible time demands associated with a task switch, which we will then address in more detail in the following sections.

- First, one can shift to another goal only if one knows what this goal is and what it requires. This means that shifting between tasks is likely to put demands on **working memory** and that it requires the **coordination** of the elements of the new task (Sect. 8.1.2).
- Second, retrieving a new goal and activating its representation is likely to **take time** (Sect. 8.1.3). According to the ECTVA model of Logan and Gordon (2001; see Box 3.1), preparing a task involves the specification of a number of task parameters that control attention and response selection—which amounts to a

more detailed version of Cohen et al.'s goal representation. Specifying these parameters requires the retrieval of information, which also is likely to take some time. How can we measure these time demands? Best suited is the interval between the task cue, which signals the upcoming task goal, and the actual stimulus on which the new task is to be applied (the CSI; see Box 8.1). It is this interval that reflects the time a participant has to prepare for the task, including the retrieval of the task goal and the specification of the required parameters. The CSI does not measure these demands directly, but one can estimate the demands by systematically manipulating the length of the CSI. An extremely short CSI would not allow any preparation, so that the new task would need to be prepared after the stimulus is presented, which in turn would increase reaction time. The longer the CSI will be, the more time would be available for task preparation, so that preparatory processes should be less and less reflected in reaction time.

- Third, given that our brain is a biological organism but no digital computer, one can imagine that no longer needed goal representations cannot simply be switched off, but that they linger on for some more time before decaying (Sect. 8.1.4). If you kept yourself busy with a particular thought or topic, you will see that moving to a new thought or topic may sometimes not be sufficient to get rid of the old mindset, which may sometimes lead to rumination. Phenomena of that sort are known as "**proactive interference**" (from the old on the new representation) in memory psychology. If we assume that there are capacity limitations with regard to the number of goals we can keep active at the same time, it is to be expected that effects of proactive interference make switches to a new task more difficult.
- Fourth, it is possible that switching to some tasks is **easier** than switching to others. Jersild (1927) for instance found no switching costs if the stimuli were such that they would indicate the corresponding task in an unequivocal fashion. This suggests that some goals do not require a lot of cognitive effort to implement them, it rather seems that **stimuli** may be able to activate some goals directly (Sect. 8.1.5).

8.1.2 Task Coordination

Did it ever happen to you that you found yourself standing in a supermarket having forgotten what you actually came for? Shopping is a rather typical combination of different kinds of tasks that have to be executed in a particular sequence: first you need to go to the shop of your choice, then you need to find the products you intended to buy (in no particular order), then you need to pay, to put your products into your basket or car, and then go back home. Each step in the sequence comprises a number of more specific actions, but let us ignore this for a moment. The fact that we sometimes lose track of the required sequence of actions or of which step is the next suggests that our **working memory** plays an important role in **coordinating** tasks and subtasks in complex actions. In studies on task switching, this important role is not always sufficiently appreciated and may thus be easily overlooked.

Consider the study of Jersild (1927). In this study, participants were to apply particular cognitive operations on lists of stimuli. In one list, a constant would be added to each stimulus number; in another list, a constant would be subtracted from each stimulus number; and yet another list required continuous switching between adding and subtracting. These three conditions correspond to the examples 1–3 in Table 8.1. We have already learned that Jersild found the slowest performance for the third list, but let us consider what these observed costs might represent.

For one, this third list required participants to frequently switch between two different cognitive operations: adding and subtracting. It makes sense to assume that each cognitive operation needs to be implemented in the cognitive system before it can be applied, and that this implementation takes some time. If so, it would be appropriate to attribute the slower performance on the third list to the sum of all **switching costs**, that is, the extra time needed to implement another cognitive operation after each stimulus. But there is another possibility. Even though participants apply only one cognitive operation on each stimulus on a list, they cannot afford to forget about the other cognitive operation that they are to apply on the next stimulus. Moreover, participants must not forget to monitor the ongoing operation, so to make sure they keep switching as required. Both of these extra requirements are not related to the actual switching process and do thus not represent switching per se but rather reflect a mixture between memory and organizing processes, which we in the following will refer to as **mixing costs**. Note that Jersild's third list differed from the other two in terms of both switching costs and mixing costs, which means that these two sources of cognitive demands were confounded (see Table 8.1).

Meanwhile, methodological advances (see Box 8.1) have provided tools to deconfound these two sources, such as with Rogers and Monsell's (1995) alternate runs paradigm or Meiran's (1996) cuing technique. The purpose of these variants consists in measuring the time demands of mixing tasks and of switching from one task to another separately as indicated in the two lowermost rows in Table 8.1. Note that the two conditions are equivalent with respect to mixing costs but differ with respect to switching costs only. That is, participants are to keep in mind the two tasks under both conditions but they repeat the task in one condition and switch to a new task in the other.

Numerous observations suggest that the difference between mixing costs and switching costs is important. For instance, Kray and Lindenberger (2000) found that aging strongly increases mixing costs while having hardly any impact on switching

Table 8.1 Relations between costs of task mixing and costs of task switching

Task sequence	Condition	Mixing costs	Switching costs
Add→Add→Add→Add→Add…	Pure list	–	–
Sub→Sub→Sub→Sub→Sub…	Pure list	–	–
Add→Sub→Add→Sub→Add…	Switch list	+	+
Add→Add→Sub→Sub→Add…	Mixed list	+	+
– Add – Sub –	Task repetition	+	–
– Sub – Add	Task switches	+	+

costs. This suggests that age-related cortical decline impairs the coordination of multiple tasks while leaving the actual switch from one task to another unaffected. Conversely, children with attention disorders show considerably increased switching costs but rather normal mixing costs. Hence, the two kinds of costs can be empirically dissociated, which suggests that they rely on **different functional and neural mechanisms**. The same conclusion is suggested by the findings of Rubin and Meiran (2005), who reported that mixing costs are obtained only if the stimulus fails to specify the corresponding task. If a given stimulus is exclusively associated with one task, however, mixing costs do not occur. This suggests that mixing costs reflect the difficulty to identify the required task and rely on processes that are concerned with determining which task is to be carried out on which stimulus (see Sect. 8.1.5).

Demands on task coordination can thus impair performance and produce mixing costs. In addition to such **quantitative effects**, there is also evidence for more **qualitative effects** on task coordination. Van Duren and Sanders (1988) had participants name digits that were presented either with high visual clarity or masked by a cloud of dots. Not surprisingly, unmasked digits were identified faster than masked ones. More interestingly, however, this masking effect was more pronounced if masked and unmasked digits were presented in different blocks than if trials with masked and with unmasked digits were intermixed. The authors attribute this observation to a simplification strategy: if all stimuli in the block are easy to see, they may be identified more or less automatically and, thus, with very high efficiency. In contrast, if all stimuli in the block are masked, an additional information-extraction process is necessary, which increases identification time. However, if masked and unmasked stimuli appear in random sequence, it might be more efficient to run the information-extraction process in each trial rather than checking first and running this process only if necessary. As a consequence, reaction time is unchanged for the masked trials (in which the information-extraction process was run in the blocked condition as well) but increased for the unmasked trials (in which now the information-extraction process is also running), so that the reaction time difference between masked and unmasked trials is reduced.

Duncan (1978) has applied a similar reasoning to explain his findings on stimulus-response compatibility. As others before him, Duncan observed that performance is better if the location of a response corresponds to the location of the stimulus signaling it (Sect. 4.2.1). More interestingly, the benefit for compatible stimulus-response pairings disappeared when they were combined with incompatible stimulus-response pairings (e.g., stimulus-response correspondence for two fingers and stimulus-response non-correspondence for two other fingers) in the same block. As Duncan suggests, having only compatible pairings in a block may allow participants to rely on the automatic activation of responses through the stimuli while mixing compatible and incompatible pairings may suggest the strategy to rely on controlled stimulus-response translation for both compatible and incompatible pairings. As a consequence, compatible pairings are processed more slowly than they would need to be, which reduces their benefit over incompatible pairings but avoids the need to identify in each trial whether the pairing is compatible or incompatible. As the latter can be seen as a coordination cost, **simplification strategies** could be thought as aiming to avoid unnecessary coordination processes. Hence, in a sense simplification strategies and mixing costs are two sides of the same coin: while mixing costs represent the unavoidable extra demands on coordination induced by task switching, simplification strategies can be seen as an attempt to avoid coordination where possible.

8.1.3 Switching Between Goal Representations

Switching to a new task requires a change of the relevant goal representation (Sect. 8.1.1). The new goal supports task-related associations, activates the required cognitive operations, and specifies their parameters, so that stimuli can be translated into responses according to the task requirements. The functional and neural network controlling performance of a particular task is commonly called a **task set**—a not particularly well-defined concept that derives from the German concept **Einstellung** (Gibson 1941).

Studies on task switching are based on the idea that the time demands of implementing a new task set can be measured by **subtracting reaction time for task repetitions from the reaction time for task alternations**. Particularly well suited for comparing task repetitions and task alternations are paradigms in which the sequence of tasks varies randomly, as in the cuing design developed by Meiran (1996). If the CSI in a task-switching trial is short, so that the new task is signaled only briefly before the new stimulus is presented, participants first need to implement the new task set before they can appropriately respond to the stimulus. This should increase the reaction time in comparison to trials in which the task is repeated, so that no new task set needs to be implemented. In contrast, a longer CSI would leave the participant more time to implement the new task set before the stimulus appears, so that the reaction time increase in switch trials should become less and less pronounced. Both predictions could be confirmed in numerous experiments, such as in the study of Meiran (1996). As shown in Fig. 8.2, reaction times are longer for task alternations than for repetitions, but this effect is much stronger for short than for long CSIs.

More findings could demonstrate that people can use precues to prepare the upcoming task, especially if they have sufficient time (1 s or more). What exactly produces this preparation effect is not entirely clear however. There are two obvious possibilities:

- First, it could be that successful preparation consists in the **retrieval and implementation of task-specific rules**, such as stimulus-response translation rules. Rubinstein et al. (2001) and Mayr and Kliegl (2000) have suggested that people first activate the new goal representation, which in turn specifies and implements all task-relevant rules. Consistent with this possibility are the observations that switching costs increase with the complexity of task-relevant rules (Rubinstein et al. 2001) and with the difficulty to retrieve task-relevant information (Mayr and Kliegl 2000).

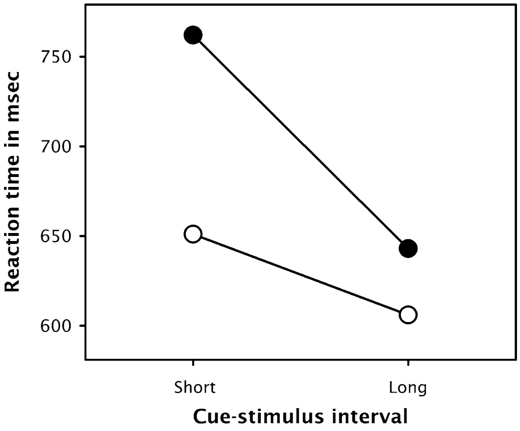


Fig. 8.2 Reaction times are longer for task alternations than for repetitions, the more so the shorter the CSI

The more information a task set comprises, so one could reason, the longer the activation and implementation process takes.

- Second, it could be that preparation consists in the **adjustment of attentional settings** (Meiran 2000; Meiran et al. 2000; cf., Logan and Gordon 2001). Shifting to a new task often involves shifting attention to other stimulus locations and/or feature dimensions. For instance, if one task requires responses to the horizontal, and the other to the vertical location of stimuli (Fig. 8.1), preparing for a new task will require a change in the spatial dimension one is attending, so that the task-relevant spatial feature of the new stimulus can properly be identified. Consistent with the attentional hypothesis of task preparation is the observation that the relationship between CSI and switching costs disappears if the stimuli of the new task do not require a reinterpretation of their features (Meiran 2000).

You may have noticed a particularity of the outcome pattern shown in Fig. 8.2: the switching costs with a long CSI are much smaller than with a short CSI, but they did not disappear entirely. It is possible that even the longer CSI was insufficient to complete the preparation, so that extending the CSI even further might eliminate the switching costs. However, studies have demonstrated that even very long CSI's are insufficient to eliminate the switching costs. Why is that so? We will devote the following two sections to two possible explanations for this observation. For one, the old, no longer used task set may still linger on for some time and impair the implementation or application of the new set (Sect. 8.1.4). And, for another, the new stimulus itself could lead to a reactivation of the old abandoned task set (Sect. 8.1.5). In addition to these two possibilities, participants may also be **insufficiently motivated** to fully prepare in each single trial, perhaps because preparation is effortful and, thus, aversive to some degree. Consistent with this failure-to-engage hypothesis (De Jong 2000; De Jong et al. 1999), switching costs can become larger with longer blocks, which are likely to lead to less motivation. However, it may also be possible that longer blocks lead to a stronger activation of the competing goal representations, which may also account for stronger switching costs.

8.1.4 Proactive Interference

Switching to a new task does not only require the activation of the new task set but also the sufficient de-activation of the old set. Hence, if you for instance shift from naming the color of Stroop stimuli to reading the color words, you do not only need to increase the activation of the representation of the reading task but also reduce the activation of the representation of the naming task. As long as the old set is not sufficiently deactivated, it is likely to compete with the new set and interfere with engaging in the new task (Allport et al. 1994). In our Stroop example, it is possible that the naming set still lingers on for some time, which would induce the tendency to name the color of the stimuli rather than reading them. If we assume that the activation of a no longer needed set **decays over time**, one would expect that the impact of the old set becomes smaller the more time one has to prepare for the new task and the longer one is busy with it.

Shifting costs indeed tend to decrease as the time between the last response associated with the previous task in the stimulus for the new task increases (Allport et al. 1994; Rogers and Monsell 1995). However, as these observations stem from studies in which the sequence of tasks was known to the participants, they could just as well reflect **preparation effects**. Fortunately, preparation effects can be disentangled from aftereffects of the old set by using a task-cuing paradigm (Meiran 1996), in which the sequence of tasks varies randomly (see Box 8.1). This technique allows the separation of the time between the last reaction and the next stimulus for the new task into two intervals: the interval between the reaction and the task cue (RCI) and the interval between the task cue and the stimulus (CSI). The RCI gives the old set an opportunity to decay even though the new task is not yet known, while the decay can be accompanied (or perhaps even sped up) by the implementation of the new set.

To study **proactive interference** from the previous task set, Meiran et al. (2000) have kept the CSI constant while the RCI was systematically varied either in different blocks (Groups 1 and 3) or randomly within blocks (Group 2). As shown in Fig. 8.3, the switching costs decreased with increasing RCI in all three groups. Consistent with the assumptions of Allport et al. (1994), having engaged in another task impairs the performance in another, succeeding task, and it does so the shorter the time between the tasks. A comparison between Groups 1 and 3 reveals that this effect is rather passive: in Group 1 were task repetitions and task alternations equally likely, while in Group 3 repetitions were much more likely than alternations. The main effect of group shows that switching costs decrease with the probability of switching, which makes sense because one can assume that greater likelihood of switching makes good preparation more likely. Interestingly, however, this effect does not interact with RCI—the slope of the RCI function is about the same in all three groups. If we assume that the probability (Group) effect reflects active preparation, the independence of this effect from RCI implies that the task-set decay in this study was preparation-independent.

In addition to this demonstration of passive decay, there is also some evidence suggesting that active **inhibition** may play a role under some circumstances. Allport et al. (1994) for instance presented Stroop stimuli that were to be read or named in

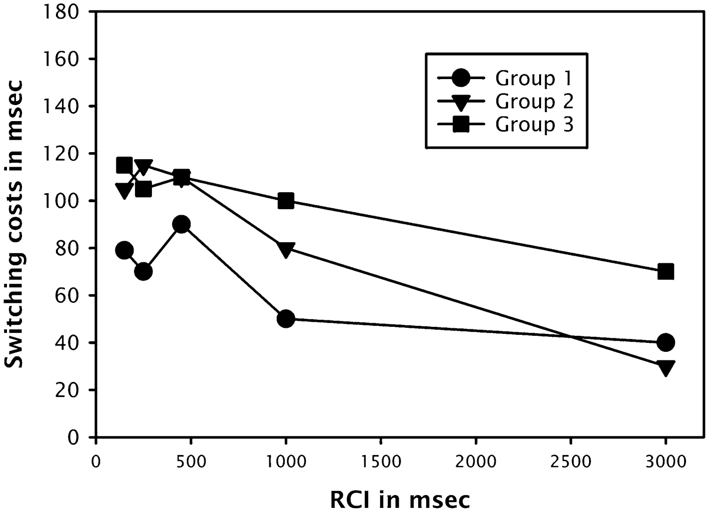


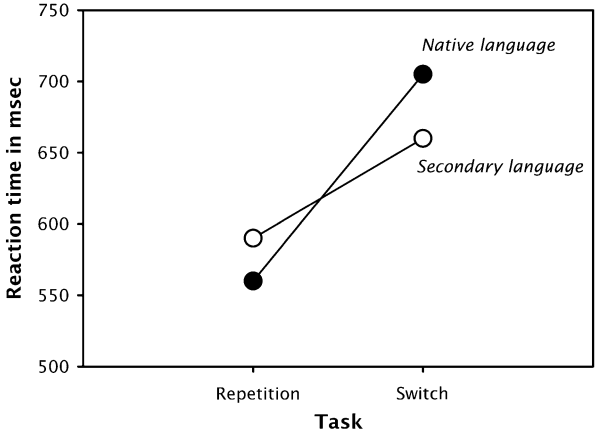
Fig. 8.3 Costs of task switching dependent on RCI (redrawn after Meiran et al. 2000)

alternation. Surprisingly, participants found it more difficult to switch to reading, even though this should be the more automatized and, thus, easier task. Similarly, Meuter and Allport (1999) observed that participants found it more difficult to switch to their native language than to a foreign language (see Fig. 8.4). A possible explanation for these seemingly paradox findings could be that easy, and presumably **dominant** tasks are more difficult to suppress when moving to a less dominant task. If you for instance spontaneously name a picture, you are more likely to use your native language, suggesting that the use of this language is more dominant. Preventing yourself from engaging in such a dominant task may require more inhibition, so that switching to a foreign language would require a particularly strong inhibition of your native language. If you then are to switch back to your native language, the respective representations may still suffer from this strong inhibition and need more time to recover—thus producing stronger switching costs.

Support for such an inhibition mechanism was also reported by Mayr and Keele (2000). They had participants randomly alternate between three tasks (A, B, and C). Interestingly, performing one of these tasks was easier as the time since this task was performed last increased. That is, switching to task A was easier in the sequence <u>A</u>BC<u>A</u> than in the sequence B<u>A</u>C<u>A</u>. This might suggest that switching to a new task required the inhibition of the old task, and that the degree of this inhibition **decreased over time**. If so, reactivating a task would indeed be easier if the time since their last performance is longer.

Taken altogether, the available evidence suggests that the activation of no longer needed task sets decreases over time; hence, one increasingly forgets the previous task. Under some circumstances, such as when switching between tasks that differ substantially in terms of their dominance, the more dominant task may require **more inhibition**, which makes it more difficult to reactivate later on. The available

Fig. 8.4 Latencies for switching between two languages (native vs. foreign language) (redrawn after Meuter & Allport 1999)



research does not yet allow predicting under which circumstances task sets decay passively and under which they are to be inhibited. It is possible that this depends on the exact task requirements, so that active inhibition comes into play if the speed and accuracy of task switching matters a lot.

8.1.5 Stimulus-Driven Goal Activation

We have mentioned already that task switching does not always produce measurable costs. It was Jersild (1927) who reported first that shifting to a new task is easy if this task is unequivocally specified by the stimulus that the task needs to be applied to. Some objects afford all sorts of actions while others afford only a few. Colored color words can be named or read, which creates uncertainty if you are frequently shifting between naming and reading. If you would be presented with colored dots for naming and with black words for reading, it would be much easier to switch back and forth. How can we explain this **dependency of switching costs** on the relationship between stimuli and tasks?

A first hint emerged from the study of Wylie and Allport (2000). Participants were presented with to-be-named colored dots, to-be-read words, and with the typical Stroop stimuli—color words in incongruent ink that were to be named or read. The researchers were interested in the difficulty to switch from color naming to word reading, that is, the costs to shift to a dominant task. One question was whether it mattered if the stimuli in the reading task would be univalent or multivalent, that is, whether the to-be-read words were colored or presented in black. It turned out that this played no role, suggesting that multi-valence per se is not a problem. More important was whether the multivalent Stroop stimuli were previously presented in the color-naming task or not. If they had been presented in the naming task before, switching costs were increased considerably.

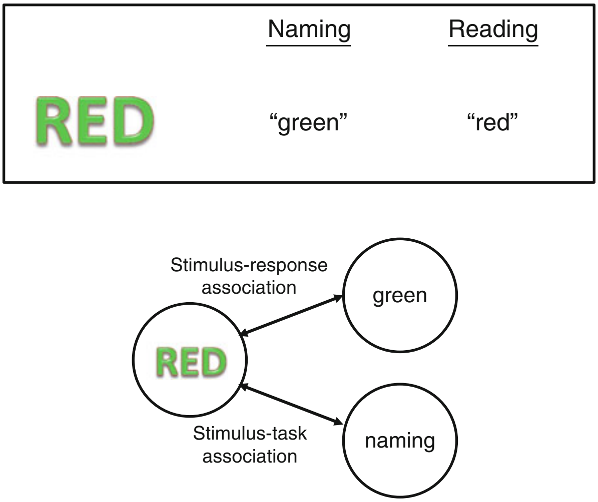


Fig. 8.5 Two types of associations when performing a task

Allport and Wylie (2000) explain this observation with the assumption that the performance of a response to a stimulus creates an **association between the stimulus and the response representation**—an event file in the terms of Hommel (1998b). If you thus respond to the word "red" written in green ink by saying "green," the stimulus will be associated with the response "green" (Fig. 8.5). If then the same stimulus appears again in the word-reading task, this creates a **conflict**: you now should say "red" to the same stimulus that you just called "green"!

The assumption of associations between stimuli and responses provides the opportunity to explain at least a substantial portion, if not all of switching costs. But there is another possibility. If a stimulus appears and is processed in the context of a particular task, this could create an association between the stimulus and the task set. If you for instance encounter a word, this could directly trigger the tendency to read it, at least if reading is an option in the current context, while encountering a number might trigger numerical operations. Applied to the example depicted in Fig. 8.5, this might suggest that processing a Stroop stimulus in the context of a naming task creates an association between the stimulus and the corresponding naming set. If you then later encounter the same stimulus again in the reading task, this could create a **task conflict**, as the stimulus may reactivate the currently inappropriate naming task.

To investigate this possibility, Waszak et al. (2003) presented participants with Stroop-like incongruent combinations of pictures and words, such as the picture of a church with the word "factory" printed on it. Similar to the study of Wylie and Allport (2000), participants had more difficulty to switch to the reading task if the current stimulus had appeared in the naming task before. This was only observed in task-switching trials but not in task-repetition trials. This confirms that encountering a stimulus in a particular task does create an association between representations of the two. If then the same stimulus appears in another task later on, it tends to activate the task it had been associated with. If that happens when the new task had been repeated sometime already, this does not create a problem but it does so in task-switching trials, when people are uncertain about which task set to implement. In other words, under task uncertainty, the **stimulus-driven reactivation of previous task sets** can create conflict between task sets.

We can summarize that the processing of a stimulus in the context of a particular task leads to an association between the corresponding stimulus, response, and task representations. Encountering the same stimulus in another task tends to reactivate the previously associated response and the previously associated task. Under laboratory conditions, this can create **task conflict**. Under more natural conditions, however, the relationships between stimuli and responses are often **less arbitrary**, so that stimuli are likely to activate the task they belong to. Associations between stimuli, responses, and tasks will thus often lead to the activation of the appropriate task, and thereby reduce the demands on working memory. This does not exclude **action errors** in particular situations, such as if you entered your bedroom for changing your dress but then find yourself lying down for a nap. Fortunately, however, such errors are rare, so that we can often rely on our associative machinery.

8.2 Multitasking

Psychological studies often employ singular, isolated actions because they allow a high degree of experimental control and facilitate the analysis of performance. In everyday life, however, we rarely engage in just one action at a time: we think while walking, listen to music while riding a car, and make phone calls while cooking. This poses the question whether and to what degree temporally overlapping actions interfere with each other, such as when using a mobile phone while driving, how such interference can be explained, and how it can be minimized or avoided. Before addressing questions of this sort, we need to consider whether people actually can perform more than one action at a time and in which sense that may be the case. The **conceptual problem** this question refers to has to do with the difficulty to define what may count as “one” action and as “at the same time.”

Consider a person who is talking while walking. Both talking and walking are intentional actions, but it is not self-evident whether combining them leads to **multitasking** rather than to the performance of one single integrated talking-while-walking action. Or consider driving a car. At the beginning of your driving lessons, driving a car was undoubtedly a multitasking challenge and the driving process consisted of multiple, party overlapping action elements. For a skilled driver, however, driving becomes a much more **integrated action** in which the multiple elements have become part of a more coherent action plan. What then is “an action”? It is difficult to see how this problem can be tackled conceptually; it rather needs an empirical answer. Psychological studies often try to avoid such interpretational difficulties by using arbitrary combinations of rather artificial tasks and testing unpracticed participants. However, we must not forget that separating different actions is much more difficult in everyday life.

Even if we could be sure that we are dealing with two separate actions, rather than one integrated action, we need to define under which circumstances we refer to the performances of these two actions as occurring **concurrently**. At a coarse, abstract level of analysis, it is certainly easy to agree that people can carry out tasks that overlap in time. But what does that mean for more detailed levels of analysis? While our home PCs give us the **illusion** that they run more than one program at a time, the underlying operations do not really run in parallel. Rather, the control system switches so quickly between different operations that we do not notice their sequential nature. The same could be true for the human cognitive system, which means that being busy with more than one task at a time does not necessarily imply parallelism between the underlying cognitive processes. This question also requires an empirical answer: if two given tasks can be combined without any costs at any level of measurement, it is at least a reasonable assumption that the underlying processes run in parallel. In contrast, the more costs the combination of two tasks creates, the more sense it makes to assume that the combination produces an **information-processing bottleneck**.

8.2.1 Task Coordination and Resource Allocation

The probably most obvious method to assess performance in dual tasks is based on a **comparison** of three measures. Let us assume we are interested in the two tasks A and B, and in the effects of combining these two tasks. We could have participants perform task A and measure their average performance, do the same for task B, and then measure performance on A and on B while the two tasks are performed together.

Figure 8.6 shows, for a hypothetical example, the performance measures we may have obtained for A and B when performed in isolation. The theoretical point P_i reflects and integrates these two measures. If participants were perfect dual-taskers, their performance on A and B would be exactly the same if the two tasks would be carried out at the same time. Accordingly, the corresponding point P_c would fall together with P_i. This is commonly not the case however: performance on one or both of the tasks will suffer if they are performed in combination. This means that P_c will typically lie under and/or to the left of P_i, and the distance between P_c and P_i can be taken as an indicator for the **efficiency of dual-tasking**—the shorter the distance the greater the efficiency.

The observation that the total performance in dual tasks is reduced as compared to performance in isolation has led to the question whether people can strategically influence which of the two tasks suffers more, that is, whether **the efficiency of one action can be optimized at the expense of the other**. To test that, participants were to perform two tasks and instructed or rewarded to distribute their cognitive resources over the two tasks in a particular ratio (e.g., 70:30, 50:50, or 30:70). Such manipulations turned out to be very effective, in the sense that the relative performance in the two tasks mirrored the instruction or reward structure rather accurately (for overviews, see Gopher and Sanders 1984; Wickens 1984). This has been interpreted as suggesting that people are equipped with **limited mental resources** of a

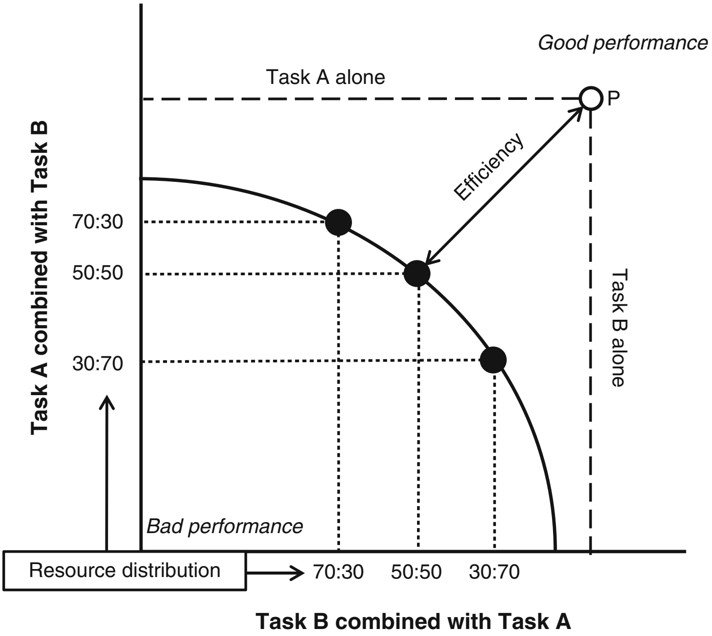


Fig. 8.6 Tool for estimating effects of dual-tasking on performance (redrawn after Wickens 1992)

particular amount, which they can strategically allocate to particular tasks or processes. The more resources are allocated to the processing of a particular task, the fewer resources are left for other tasks, which would indeed result in the negative relationship between performance in the two tasks indicated in Fig. 8.6.

This resource interpretation has been criticized however. If people would really have one unitary kind of mental resource at their disposal, then every increase in the difficulty of one task should reduce the total performance. Various studies could not find such a reduction however. Later approaches have tried to account for these observations by distinguishing between **different kinds of resources**, such as between activation and arousal (Sanders 1983), or between resources that are specific for perception and action processes, respectively (Wickens 1980, 1984). For instance, the approach of Wickens assumes that two tasks are more likely to make use of the same resources the more similar they are with respect to the modality of the stimuli, the cognitive operations they involve, and the modality of their actions. According to this approach, combining tasks that involve very different components is more likely to produce efficient performance than combining tasks with very similar components.

The **advantage** of resource approaches consists in their intuitive plausibility and the fact that they have been very functional in applied human factors studies (see e.g., Wickens 1984). Their **disadvantage** consists in their descriptive nature that does not allow many predictions and does not afford a good mechanistic understanding of possible bottlenecks. Without knowing of what kind the hypothetical mental resource may be, whether its amount is really fixed and its capacity is really limited (or whether it can increase by investing more effort), and how it can be measured independently from the performance it is assumed to explain, a better theoretical understanding of the cognitive processes underlying dual task performance is difficult to obtain. Moreover, the commonly used **comparison between single-task and dual-task performance** is methodologically questionable. This comparison confounds a whole number of other variables: As compared to single-task situations, dual-task conditions require participants to keep in mind more instruction-relevant information in working memory, to process a larger number of stimuli and responses, and they are likely to exhibit a different level of motivation and stress. Any difference between single-task and dual-task performance may be affected by these factors, which limits their interpretation in terms of resource-related effects (see the very similar problems that we have discussed for list-based task-switching paradigms; Sect. 8.1, Box 8.1).

Considering these theoretical and practical problems, it is not surprising that resource models do not play a major role in the current discussion. More successful have been **information-processing models** that attribute dual-tasking costs to particular processing bottlenecks. They assume that some processing stages can be used by multiple tasks in parallel, while other processing stages are serial in nature, so that they can be used by only one task at a time. Particularly important for the success of information-processing models have been the development of the **locus-of-slack method** (**LoSM**; see Box **8.2**). By using this method, processing bottlenecks can be identified and located in the processing stream through systematic manipulation of experimental factors. The application of this technique has revealed three bottlenecks that we will discuss in the following.

Box 8.2: Locus-of-Slack Method

An important tool for localizing processing bottlenecks under dual-task conditions is the **locus-of-slack method (LoSM)**, which is related to the so-called PRP paradigm. Studies using this paradigm have participants perform two tasks T1 and T2 that are performed with different degrees of temporal overlap. Let us assume that T1 requires a left or right key press (response R1) signaled by the high or low pitch of a tone (S1). Let us further assume that T2 requires a verbal naming response (R2) to a visual stimulus (S2), such as the letter O vs. X. If the temporal interval between S1 and S2 (the so-called **stimulus onset asynchrony** or **SOA**) is very long, the two tasks are carried out more or less sequentially. The shorter the SOA, however, the more the two tasks overlap, thus creating a real dual-task condition.

If we were perfect dual-taskers, it should not matter whether we perform tasks sequentially or at the same time, which means that performance in the two tasks should not depend on the length of the SOA. Conversely, we can say that the more performance on T2 suffers as the SOA decreases, the more we can speak of **dual-task costs**. Telford (1931) has attributed such costs to a hypothetical **Psychological Refractory Period (PRP)**, which is why the corresponding experimental design with two overlapping tasks and systematically manipulated SOA has been coined **PRP paradigm**.

(continued)

Box 8.2: (continued)

The concept of a PRP implies that some stages in the processing of information represent a processing bottleneck, in the sense that they can only be used by one task at a time. That processing bottlenecks exist is uncontroversial but the question is which cognitive processes exhibit the assumed bottleneck characteristic. Based on Sternberg's (1969) **additive factors method**, the LoSM has been developed to localize possible bottlenecks in the processing stream (for overviews, see Pashler and Johnston 1989; Schweickert 1978).

The basic idea is sketched in Fig. 8.7. It is assumed that a task requires information to pass along a series of processing stages. Some of these stages may have a bottleneck characteristic, which means that they are available for information from only one task at a time (=bottleneck stages), while other stages before (pre-bottleneck stage) or after a bottleneck (post-bottleneck stage) can be accessed by multiple tasks in parallel. If pre- and post-bottleneck stages can process information from both tasks, the operation of these stages can overlap in time. With respect to Fig. 8.7, this means that the pre-bottleneck stage of T1 can fully overlap with the pre-bottleneck stage of T2, and the same is true in principle for the post-bottleneck stages. This is not true for the actual bottleneck stages, however: If we assume that the bottleneck stage for T1 starts first (which need not be the case, but which is likely if S1 is presented before S2, as common in PRP designs), the bottleneck stage for T2 has to wait until the bottleneck stage for T1 is completed. In the figure, this leads to a delay of T2—the PRP effect.

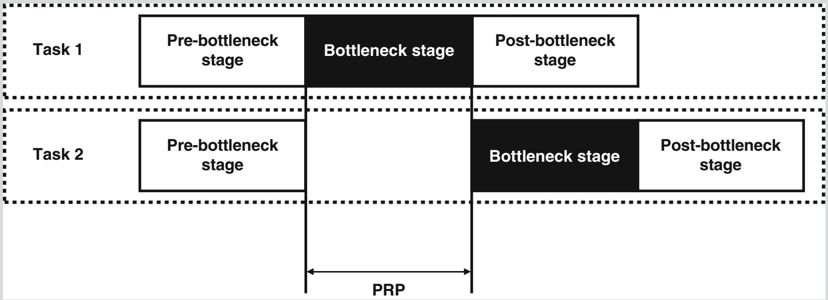


Fig. 8.7 The rationale of the *locus-of-slack method* (LoSM) for localizing processing bottlenecks under dual-task conditions (redrawn after Pashler & Johnston 1989)

But how do we find out which cognitive processes exhibit the sought-for bottleneck characteristics? Which stage is the bottleneck stage? Is there more than one? To address these questions, LoSM suggests to systematically vary the duration of the assumed processing stages. To interpret the outcomes of variations of that sort, LoSM offers four particularly diagnostic

(continued)

Box 8.2: (continued)
predictions, of which two apply to manipulations targeting T1 and two to manipulations targeting T2.

Let us assume we make the **identification of S1** more difficult, for instance by using two very similar tones. This manipulation should increase the reaction time for T1, but the duration of which stage did we prolong? We do not yet know, but which options do we have? If it is a pre-bottleneck stage or the bottleneck stage, we would expect that reaction times do not only increase for T1 but for T2 as well. In both cases, the completion of the bottleneck stage for T1 would be delayed, which would also delay the bottleneck stage for T2. Now we can turn around this logic to identify the affected processing stage: if reaction times of both tasks were equally delayed, especially if the SOA is very short (so that the two tasks overlap considerably), our manipulation must have impacted either the pre-bottleneck stage or the bottleneck stage of T1 (LoSM **prediction #1**). It could not have been the post-bottleneck stage of T1, as that should not have affected performance in T2. Again, you can turn around the logic of this assumption: if a manipulation targeting T1 does not affect performance in T2, the manipulation must have impacted a post-bottleneck stage (LoSM **prediction #2**).

Let us now assume that we manipulate factors affecting T2. For instance, we may make a **response selection** in this task more difficult by introducing more response alternatives. Should that have affected the pre-bottleneck stage of T2, we can make a counterintuitive prediction: If the SOA is so long that the two tasks hardly overlap, this manipulation should of course increase the reaction time for T2 without affecting T1. If the SOA is short, however, then the bottleneck stage for T2 has to wait until the bottleneck stage for T1 is completed. During that time, the bottleneck stage for T2 cannot start anyway, so that any pre-bottleneck stage of T2 has more time to process. This means that factors that delay the pre-bottleneck stage of T2 do *not* necessarily delay reaction time for T2. And, turning around the logic: any factor that delays the reaction time of T2 with a long SOA but not (or not that much) with short SOA can be assumed to impact a pre-bottleneck stage of T2 (LoSM **prediction #3**). Another possibility is that our manipulation affects the bottleneck or a post-bottleneck stage. This should delay reaction times in T2 without affecting T1. In contrast to manipulations targeting the pre-bottleneck stage of T2, this delay should be the same for short and long SOAs. If we turn around the logic: manipulations that have the same impact on T2 for long and short SOAs can be assumed to affect the bottleneck stage or post-bottleneck stages (LoSM **prediction #4**). Try playing with these predictions and applying them to real-world dual-task situations. Even though this will sometimes require some simplification of the tasks, you will see that the LoSM allows for interesting and specific predictions.

8.2.2 Working Memory

Researchers have distinguished between various kinds of **memory systems**, such as between short-term and long-term memory, semantic and episodic memory, between modality-specific systems, and more. There is hardly any evidence that the encoding or retrieval of information represents a true structural processing bottleneck. In stark contrast, phenomena like the Stroop effect suggest that retrieval is actually often too little constrained, so that seeing a word can activate word-reading tendencies and even representations of the word's meaning. Hence, the problem seems to be a too much but not a too little of memory access. This suggests that it is not the availability of information that creates bottlenecks under dual-task conditions but the organization and the goal-directed processing of this information. Their organization and the handling of memory information are commonly attributed to working memory. The concept of a working memory comprises slave systems serving for the maintenance of information and a control system that takes care of the way the maintained information is organized (Baddeley 2003). Even though research on dual-tasking has mostly focused on response selection as the key bottleneck (Sect. 8.2.3), more recent findings suggest that the recruitment of working memory can lead to processing bottlenecks and dual-tasking costs.

Evidence for the limiting contribution of working memory has been provided by studies on tasks that require the encoding of information for later report. For instance, Jolicœur and Dell'Acqua (1998) had participants encode briefly presented letters for later report (T1) and carry out manual binary-choice reactions to tones (T2). As shown in Fig. 8.8, encoding in T1 affected the reaction time in T2. There

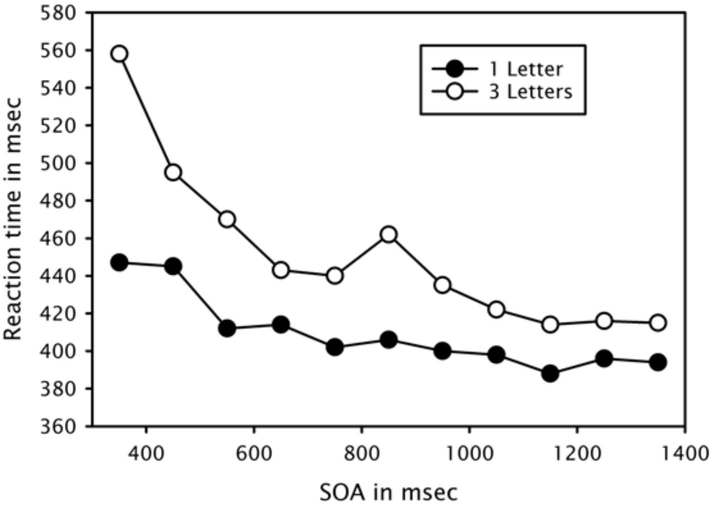


Fig. 8.8 Latencies for the second of two combined tasks (redrawn after Jolicoeur & Dell'Acqua 1998, by permission of Elsevier)

were three interesting findings: First, reaction times in T2 increased with the number of encoded stimuli in T1. Second, the reaction times increased more with short than with long SOAs (see Box **8.2**). Third, the dual-tasking costs were particularly pronounced if SOA was short (so that the two tasks overlap strongly) and the number of encoded items was maximal. This suggests that the temporal demands of the **encoding process** increase with the number of encoded items, and that the duration of this encoding process determines the delay of T2.

In addition to encoding, the **retrieval** of information from working memory can also represent a processing bottleneck. For instance, Jolicœur (1999) combined a manual choice-reaction task with a memory-search task, in which participants were to judge whether a presented stimulus was part of a previously shown, memorized stimulus set. A typical outcome of such tasks consists in the observation that reaction time increases with the set size. If the memory task was carried out first, reaction times in the choice-reaction task increased with set size. If the choice-reaction task was carried out first, the set size effect was independent of task overlap (SOA). According to LoSM, this means that memory search represents a bottlenecks stage that cannot be started before response selection in T1 is completed.

The selective, goal-directed **retrieval of information from long-term memory** seems to represent another bottleneck. Carrier and Pashler (1995) had participants learn word pairs before presenting them with a dual task. T1 required a choice reaction to a tone. In T2, participants were presented with one member of a word pair and they were to reproduce the other member. The difficulty of this task was manipulated by having some pairs learned better than others. As it turned out, the temporal overlap between the two tasks had no impact on the difficulty effect, suggesting that the retrieval of information could not start before the bottlenecks stage of T1 was completed.

These and other observations suggest that the handling of memory information represents an **effective processing bottleneck under dual-task conditions**. This is remarkable because the stimulus-driven retrieval of memory information seems to proceed rather automatically, which may sometimes create functional problems but does not point to a structural limitation. This suggests that it is the intentional, perhaps even conscious access to memory content that represents the actual bottleneck (Jolicœur et al. 2002).

8.2.3 Stimulus-Response Translation and Response Selection

Since the beginnings of systematic research on multitasking, researchers have suspected that the selection of responses might represent a bottlenecks stage and it is since Welford (1952) that **response-selection-bottleneck** (**RSB**) **models** dominate the research field (for overviews, see McCann and Johnston 1992; Pashler 1994; Pashler and Johnston 1989). This dominance results from the multitude of hints from studies using the LoSM. Let us reconsider the four key predictions of the LoSM approach (see Box **8.2**) under the assumption that response selection indeed represents a bottleneck stage (and, for the sake of simplicity, that it is the only

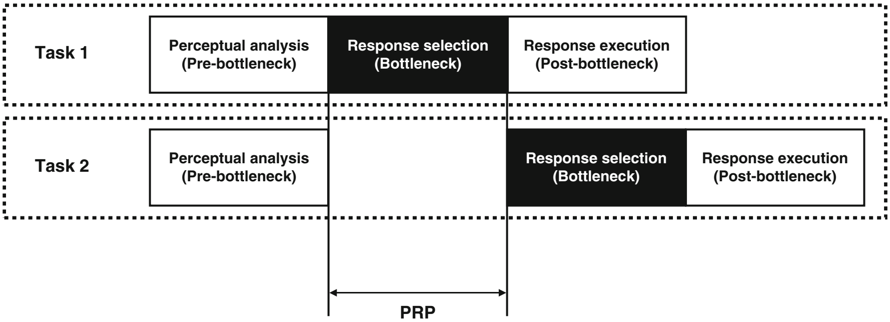


Fig. 8.9 Applying the LoSM for modeling the flow of processing

bottleneck stage). If response selection is indeed the bottleneck, stages preceding response selection, like stimulus processing, could be considered a pre-bottleneck stage while stages following response selection, like response execution, could be considered a post-bottleneck stage (see Fig. 8.9).

According to **prediction #1**, this would suggest that manipulations of **response-selection** difficulty in T1 affect performance in both tasks. One way to increase this difficulty is to increase the number of response alternatives in T1, and this manipulation has indeed been reported to increase reaction times in both tasks (Karlin and Kerstenbaum 1968). Making response choice in T1 more complicated thus seems to delay response selection in T2, as RSB models would suggest. If so, we can conclude that **selecting responses in more than one task at a time is not possible**.

According to **prediction #2**, manipulations targeting a post-bottleneck stage in T1 should impact performance in T1 but not in T2. There is indeed evidence that increasing the complexity of the required movements in T1 (a manipulation that is likely to increase the time it takes to execute T1 responses) leads to an **increase of reaction times in T1 but not, or hardly so, in T2** (Pashler and Christian 1996).

According to **prediction #3**, making a pre-bottleneck stage in T2 more difficult should impair performance in T2 if the SOA is long but not, or hardly so, if the SOA is short. An example of this so-called underadditivity of the difficulty manipulation is shown in Fig. 8.10. The data patterns stem from a study of Pashler and Johnston (1989), who manipulated the intensity of visual stimuli in T2. We can see that reaction times in T1 are unaffected by this manipulation, as the LoSM would predict. We also see a typical PRP effect, which shows that performance in T2 gets worse as the SOA gets shorter, suggesting that the temporal overlap between the two tasks impairs performance in T2. Furthermore, we see that less intensive stimuli in T2 delay reaction times in T2 for long SOAs, that is, if the two tasks are executed more or less sequentially. This means that the intensity manipulation was successful. Most interestingly, however, the intensity effect disappears entirely for the shortest SOA, which statistically means that the effects of SOA and intensity combine in an **underadditive** fashion (i.e., the reaction time for the combination of

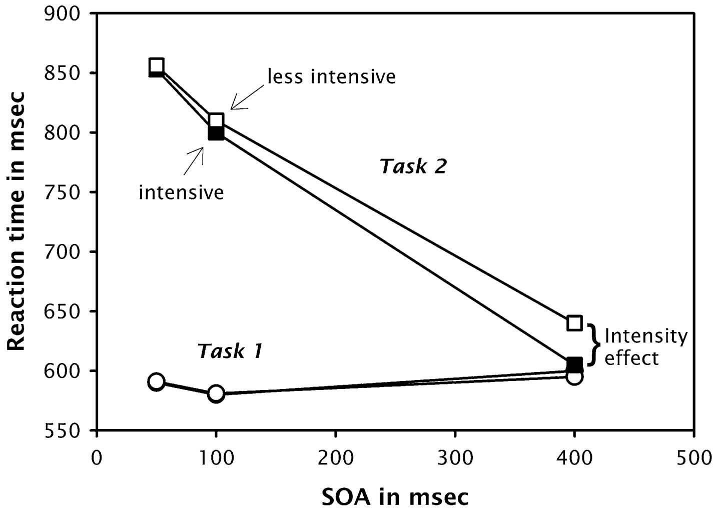


Fig. 8.10 Latencies for the first and second of two combined tasks (redrawn after Pashler & Johnston 1989, by permission of Taylor and Francis Group)

less intensive stimuli and the shortest SOA is smaller than adding the SOA effect and the intensity effect would suggest). This is exactly what the RSB model would predict for the shortest SOA: As T2 has to wait for the completion of response selection in T1 anyway, the extra available time can be devoted to identifying the stimulus for T2, so that the extra demands imposed by the lesser intensity are no longer visible in the reaction time.

According to **prediction #4**, manipulations targeting the bottleneck stage of T2 should affect performance in T2 but not in T1, independently from SOA. Examples for this pattern have indeed been reported: Reducing stimulus-response compatibility in T2 (McCann and Johnston 1992), a manipulation that should delay response selection in T2, **increases reaction times in T2 but not in T1**.

These are just a few examples of the many findings supporting the assumption that response selection represents a processing bottleneck in dual-tasking. This means that RSB models have a very solid empirical basis. And yet, there are **three reasons** why traditional RSB models have been criticized recently.

8.2.3.1 Multiple Bottlenecks

The first reason is implicit in our discussion of bottleneck stages so far: Even if **response selection** is certainly an important bottleneck stage, it does not seem to be the only one. As we have discussed, access to **working memory** is likely to represent another bottleneck and, as we will see in the following, the **execution of particular actions** can create capacity problems. This poses the question whether all these bottlenecks represent independent characteristics of information processing or whether they reflect a general principle of cognitive functioning that we do not yet fully understand.

For instance, it might be possible that processing bottlenecks emerge whenever local processing is insufficient so that information from different local systems has to be **integrated** (e.g., Baars 1980). This may explain why consolidation into memory represents a bottleneck, because this requires the integration of features of the to-be-consolidated stimulus event. It may also explain why response selection represents a bottleneck, because selecting a response requires relating stimulus information to response information according to a particular instruction. Finally, it may account for observations that response execution sometimes represents a bottleneck (see Sect. 8.2.4), because execution sometimes requires extra checking whether the environmental conditions for the to-be-executed actions are adequate.

8.2.3.2 Conceptual Issues

The second reason why RSB models have been criticized has to do with their **conceptual fuzziness**. These models use terms like stimulus-response translation and response selection synonymously, as they (explicitly or implicitly) assume that responses are selected by applying a stimulus-response rule. Let us assume that T1 requires a response R1 to stimulus S1 while T2 requires a response R2 to stimulus S2. Presenting S1 would then activate the representation of that stimulus and the instructed stimulus-response rule would be applied to activate the representation of R1. If it is this process that represents response selection, the traditional RSB model would suggest that the representation of R2 cannot be activated before this process is completed.

One possibility to test this prediction of the RSB model is to **vary the relationship between R1 and R2** systematically. Let us for instance assume that the two responses are compatible in some trials (e.g., because R1 consists in pressing a left key while R2 consists in saying the word “left”) but incompatible in other trials (e.g., because R1 consists in pressing a left key while R2 consists in saying the word “right”). Would you expect that the compatibility of this relationship affects performance in T1? According to the RSB model, it shouldn’t, because R2 is not yet known while R1 is being selected. However, several studies have demonstrated that selecting R1 is facilitated if it is compatible with R2 (Hommel 1998a; Logan and Schulkind 2000). This suggests that the representations of R2 were activated before the selection of R1 was completed—how else could the selection of R1 have been affected? But this implies that activating response representations for multiple tasks can occur in parallel, suggesting that response activation does not represent a processing bottleneck. If we thus would stick with the suggestion that response selection is capacity limited, the selection process cannot consist in translating stimulus information into response activation—a process that does not seem to be limited. It is possible that the selection process proper does not involve **activation** of response representations but rather the **decision** which of the activated response representations should actually be executed (Hommel 1998a). In other words, the actual purpose of the selection process seems to consist in checking activations against the instruction and the stimulus. If so, we can conclude that actions are selected in two phases: one that serves to **translate** stimulus information into codes of response representations and another that selects and perhaps integrates these codes into a coherent action plan (Sect. 6.4).

While the first phase does not seem to underlie any capacity limitations, the second phase seems to be devoted to only one task at a time.

8.2.3.3 Practice

A third reason that has raised skepticism against the traditional RSB model is motivated by observations suggesting that processing bottlenecks can be "**trained away**." The idea of a processing bottleneck does not necessarily require the assumption that it relies on structural limitations (see Box **8.3**), but many authors explicitly or implicitly assume that dual-tasking bottlenecks are structural in nature and, thus, immutable. One possible objection is that many psychological studies use rather artificial tasks and participants are often not provided with a lot of practice (Meyer and Kieras 1997a, b). It would thus be possible that massive practice reduces or perhaps even eliminates the assumed bottlenecks. Some training studies have indeed provided evidence that dual-task costs get smaller through practice and sometimes disappear altogether (e.g., Van Selst et al. 1999).

From a practical standpoint, such training effects are of enormous importance, as they demonstrate that **multiple tasks can be performed in parallel** in principle. From a theoretical standpoint, however, the available findings are not easy to interpret. It is certainly possible that practice eliminates processing bottlenecks entirely, but one needs to ask how this works in detail. Alternatively, practice may merely reduce the duration of bottleneck stages, which would not necessarily eliminate the bottleneck but reduce or even eliminate its temporal overlap with other bottleneck stages. If so, it would be an important aspect of training regimes to optimize the way multiple tasks are **organized** in time, and to reschedule processing stages in such a way that temporal overlap between bottleneck stages is minimized (Ruthruff et al. 2001).

8.2.4 Response Initiation

Even though this topic did not attract a lot of attention so far, there is evidence that dual-tasking can encounter processing bottlenecks with respect to the **implementation and initiation** of actions. For instance, Logan and Burkell (1986) had participants prepare and execute a response for T1, but sometimes presented a stop signal requesting the immediate abortion of the response preparation. Performance in T2

Box 8.3: Structural and Functional Limitations of Information Processing

Psychological research often studies the limitations of human information processing: How many objects can we attend simultaneously? How many events can be keep in memory? How many actions can we prepare concurrently? In discussions of possible limitations, researchers often distinguish between **structural** and **functional** limitations and bottlenecks. Even though

(continued)

Box 8.3: (continued)
it is not always easy to distinguish the two underlying concepts, they do have very different implications for how we can deal with the respective limitation under real-life conditions.

The concept of **structural limitations** commonly refers to some upper limit of processing capacity, irrespective of task, context or kind of information. For instance, it has been assumed that we cannot keep more than about seven independent events in short-term memory (Miller 1956) and no more than about four objects in visual short-term memory (Luck and Vogel 1997). These assumptions refer to structural limits which might show some intraindividual variability and that may be circumvented by means of tricks and strategies (such as chunking information), but the limitation itself remains.

The concept of **functional limitations** refers to processing bottlenecks that result from a less effective combination of cognitive operations or interactions between their outcomes. For instance, let us assume that your short-term memory is unlimited and you could keep traces of as many events as you want. Each single memory would entertain hundreds of associations with other events in long-term memory (having been in Rome reminds you of red wine, the Catholic Church, the Coliseum, and the airport, which reminds you of other airports, and so forth), so that keeping 20 items in short-term memory would be likely to activate large portions of your long-term memory. This would create so much noise that it would not be particularly practical—the noise would overshadow the signal. Note that the problem here does not refer to a limitation, not to a too little but to a too much of information. Hence, the problem is not structural but functional in nature.

This example also shows that the distinction between structural and functional limitations may be clear in some examples but more difficult to draw in others. As our example suggests, the seemingly structural limitation of our short-term memory may have functional reasons: a much greater capacity may tend to activate so many long-term traces that we could not keep our thoughts straight. How might our brain solve such a functional problem? It could try to organize the stored information, so that only thematically relevant associations are currently enabled while others are inhibited. Such an integration strategy may be realized through neural synchronization, in such a way that all neurons representing currently relevant associations are firing in the same rhythm (Raffone and Wolters 2001). To properly separate the phases of the rhythms representing different associative structures, only phases that are sufficiently different may be permitted. Is this a functional or structural limitation? The answer is difficult, suggesting that at least some of the empirical observations that we attribute to structural limitations may actually represent functional limitations and, more importantly, that at a certain level of analysis the distinction may no longer be meaningful. Nevertheless, structural and functional limitations often call for different ways to tackle them, which means that the distinction can still have some practical value.

was particularly impaired if the response in T1 could not be suppressed. One possible interpretation of this observation is that the initiation of a response can impair the selection or initiation of other responses.

These considerations would fit with findings reported by Ivry et al. (1998). These authors studied possible processing bottlenecks under dual-task conditions in J.W., a split-brain patient that previously underwent commissurotomy. The authors consider the possibility that having two independent cortical hemispheres might permit the execution of two tasks (one in each hemisphere) at the same time without any costs. To test that, the stimuli for the two tasks were presented to two different visual hemifields, so that they could be processed in different hemispheres. This did eliminate the typical processing bottleneck associated with response selection: Normally, performance in dual tasks is better if the tasks are performed with different effectors (e.g., with hand vs. mouth as compared to the left vs. right hand), but this effect of effector modality did not play a role for J.W. Likewise, performance in dual-task conditions are commonly strongly dependent on the consistency of the stimulus-response rules and response–response compatibility in the two tasks, but that was not the case for J.W. As effector-modality and rule-compatibility effects are likely to be associated with response selection, these findings suggest that J.W. could indeed select responses for the two tasks in parallel and without mutual interference. And yet, performance was worse with short SOAs, which shows that temporal overlap between the two tasks did have an impact. If we exclude response selection as the bottleneck stage, this suggests that the impairment must have to do with the **initiation of actions**.

References

Allport, D. A., Styles, E. A., & Hsieh, S. (1994). Shifting intentional set: Exploring the dynamic control of tasks. In C. Umiltá & M. Moscovitch (Eds.), *Attention and performance* (Conscious and nonconscious information processing, Vol. XV, pp. 421–452). Cambridge, MA: MIT Press.

Allport, D. A., & Wylie, G. (2000). 'Task-switching', stimulus-response bindings and negative priming. In S. Monsell & J. Driver (Eds.), *Attention and performance* (Control of Cognitive Processes, Vol. XVIII, pp. 421–452). Cambridge, MA: MIT Press.

Baars, B. J. (1980). Eliciting predictable speech errors in the laboratory. In V. A. Fromkin (Ed.), *Errors in linguistic performance: Slips of the tongue, ear, pen, and hand* (pp. 307–318). New York: Academic.

Baddeley, A. D. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839.

Carrier, L. M., & Pashler, H. (1995). Attentional limits in memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1339–1348.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.

De Jong, R. (2000). An intention-activation account of residual switch costs. In S. Monsell & J. Driver (Eds.), *Attention and performance* (Control of cognitive processes, Vol. XVIII, pp. 357–376). Cambridge, MA: MIT Press.

De Jong, R., Berendsen, E., & Cools, R. (1999). Goal neglect and inhibitory limitations: Dissociable causes of interference effects in conflict situations. *Acta Psychologica, 101*, 379–394.

Dixon, P. (1981). Algorithms and selective attention. *Memory & Cognition, 9*, 177–184.

Duncan, J. (1978). Response selection in spatial choice reaction: Further evidence against associative models. *Quarterly Journal of Experimental Psychology, 30*, 429–440.

Gibson, J. J. (1941). A critical review of the concept of set in contemporary experimental psychology. *Psychological Bulletin, 38*, 781–817.

Gopher, D., & Sanders, A. F. (1984). S-Oh-R: Oh stages! Oh resources! In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 231–253). Berlin, Germany: Springer.

Hommel, B. (1998a). Automatic stimulus-response translation in dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 1368–1384.

Hommel, B. (1998b). Event files: Evidence for automatic integration of stimulus-response episodes. *Visual Cognition, 5*, 183–216.

Ivry, R. B., Franz, E. A., Kingstone, A., & Johnston, J. C. (1998). The psychological refractory period effect following callosotomy: Uncoupling of lateralized response codes. *Journal of Experimental Psychology: Human Perception and Performance, 24*, 463–480.

Jersild, A. T. (1927). Mental set and shift. *Archives of Psychology, 14*, 89.

Jolicœur, P. (1999). *Capacity demands of accessing short-term memory*. Paper presented at the Ninth Annual Meeting of the Canadian Society for Brain, Behaviour, and Cognitive Science, June 18–19, Edmonton, AB.

Jolicœur, P., & Dell'Acqua, R. (1998). The demonstration of short-term consolidation. *Cognitive Psychology, 36*, 138–202.

Jolicœur, P., Tombu, M., Oriet, C., & Stevanovski, B. (2002). From perception to action: Making the connection. In W. Prinz & B. Hommel (Eds.), *Attention and performance* (Common mechanisms in perception and action, Vol. XIX, pp. 558–586). Oxford: Oxford University Press.

Karlin, L., & Kerstenbaum, R. (1968). Effects of number of alternatives on the psychological refractory period. *Quarterly Journal of Experimental Psychology, 20*, 167–178.

Kray, J., & Lindenberger, U. (2000). Adult age differences in task switching. *Psychology and Aging, 15*, 126–147.

Logan, G. D., & Burkell, J. (1986). Dependence and independence in responding to double stimulation: A comparison of stop, change, and dual-task paradigms. *Journal of Experimental Psychology: Human Perception and Performance, 12*, 549–563.

Logan, G. D., & Gordon, R. D. (2001). Executive control of visual attention in dual-task situations. *Psychological Review, 108*, 393–434.

Logan, G. D., & Schulkind, M. D. (2000). Parallel memory retrieval in dual-task situations: I. Semantic memory. *Journal of Experimental Psychology: Human Perception and Performance, 26*, 1072–1090.

Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature, 390*, 279–281.

Mayr, U., & Keele, S. W. (2000). Changing internal constraints on action: The role of backward inhibition. *Journal of Experimental Psychology: General, 129*, 4–26.

Mayr, U., & Kliegl, R. (2000). Task-set switching and long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition, 26*, 1124–1140.

McCann, R. S., & Johnston, J. C. (1992). Locus of the single-channel bottleneck in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 471–484.

Meiran, N. (1996). Reconfiguration of processing mode prior to task performance. *Journal of Experimental Psychology: Learning, Memory and Cognition, 22*, 1423–1442.

Meiran, N. (2000). Reconfiguration of stimulus task-sets and response task-sets during task-switching. In S. Monsell & J. Driver (Eds.), *Attention and performance* (Control of cognitive processes, Vol. XVIII, pp. 377–399). Cambridge, MA: MIT Press.

Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology, 41*, 211–253.

Meuter, R. F., & Allport, A. (1999). Bilingual language switching in naming. Asymmetrical costs of language selection. *Journal of Memory and Language, 40*, 25–40.

Meyer, D. E., & Kieras, E. D. (1997a). A computational theory of executive cognitive processes and multiple task performance: Part 1, Basic mechanisms. *Psychological Review, 104*, 3–75.

Meyer, D. E., & Kieras, D. E. (1997b). A computational theory of executive cognitive processes and multiple-task performance: Part 2. Accounts of psychological refractory-period phenomena. *Psychological Review, 104*, 749–791.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*, 81–97.

Pashler, H. (1994). Dual-task interference in simple tasks: Data and theory. *Psychological Bulletin, 116*, 220–244.

Pashler, H., & Christian, C. (1996). *Bottlenecks in planning and producing manual, vocal, and foot responses*. Unveröffentlichtes Manuskript.

Pashler, H., & Johnston, J. C. (1989). Chronometric evidence for central postponement in temporally overlapping tasks. *Quarterly Journal of Experimental Psychology, 41A*, 19–45.

Raffone, A., & Wolters, G. (2001). A cortical mechanism for binding in visual working memory. *Journal of Cognitive Neuroscience, 13*, 766–785.

Rogers, R. D., & Monsell, S. (1995). Cost of a predictable switch between simple cognitive tasks. *Journal of Experimental Psychology: Human Perception and Performance, 124*, 207–231.

Rubin, O., & Meiran, N. (2005). On the origins of the task mixing cost in the cuing taskswitching paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 1477–1491.

Rubinstein, J., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 763–797.

Ruthruff, E., Johnston, J. C., & Van Selst, M. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance, 27*, 3–21.

Sanders, A. F. (1983). Towards a model of stress and human performance. *Acta Psychologica, 53*, 61–97.

Schweickert, R. (1978). A critical path generalization of the additive factor method: Analysis of a Stroop task. *Journal of Mathematical Psychology, 18*, 105–139.

Sternberg, S. (1969). The discovery of processing stages: Extensions of Donders' method. *Acta Psychologica, 30*, 276–315.

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

Telford, C. W. (1931). The refractory phase of voluntary and associative responses. *Journal of Experimental Psychology, 14*, 1–36.

Van Duren, L. L., & Sanders, A. F. (1988). On the robusmess of the additive factors stage structure in blocked and mixed choice reaction designs. *Acta Psychologica, 69*, 83–94.

Van Selst, M., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the psychological refractory period? *Journal of Experimental Psychology: Human Perception and Performance, 25*, 1268–1283.

Waszak, F., Hommel, B., & Allport, A. (2003). Task-switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology, 46*, 361–413.

Welford, A. T. (1952). The "psychological refractory period" and the timing of high-speed performance—a review and a theory. *British Journal of Psychology, 43*, 2–19.

Wickens, C. D. (1980). The structure of attentional resources. In R. Nickerson (Ed.), *Attention and performance* (Vol. VIII, pp. 239–257). Hillsdale, NJ: Erlbaum.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). Orlando, FL: Academic.

Wylie, G., & Allport, A. (2000). Task switching and the measurement of "switch costs". *Psychological Research, 63*, 212–233.

Zillig, M. (1926). Die Übung der Umstellbarkeit. *Zeitschrift für Psychologie, 100*, 18ff.

Chapter 9
Action Monitoring

Up to this point, we have discussed how people plan and execute intentional, context-adjusted actions successfully. But some of our actions fail. For one, this raises the question when and how action errors are actually generated. Even though we will not be able to address this issue exhaustively, we will discuss some considerations and conclusions from research on action errors in Sect. 9.1. For another, if we do commit errors, one can ask why we do not commit more of them. The relationship between correctly and incorrectly executed actions is indeed of considerable scientific interest: if we interpret correctly performed action as the successful avoidance of action errors, a better theoretical understanding of the causes underlying action errors will obviously be beneficial for our insight into successful action control. We will therefore also discuss whether and how people monitor action execution and identify action errors (Sect. 9.2), how they avoid errors (Sect. 9.3), and how they learn from them (Sect. 9.4).

9.1 Action Errors

The analysis of human action errors has a long tradition in psychology (see Sect. 6.1.2). Among others, Freud (1904) has broadly discussed various kinds of (mainly verbal) action errors and their underlying psychological causes in his book on "The Psychopathology of Everyday Life." More recently, action errors have been investigated mainly in applied fields, such as human performance. Reason (1979) has analyzed systematic diaries to inventarize and classify a large number of everyday errors. Rasmussen (1980) suggested an influential category system to analyze errors. The system distinguishes between errors on three different levels of cognitive processing:

- **Knowledge-based errors** occur in novel and unfamiliar situations, in which actions need to be planned ad hoc. Such situations, such as an outbreak of fire, can lead to stress and, as a consequence, panic-induced, not particularly goal-directed actions.

B. Hommel et al., *Human Action Control:* DOI 10.1007/978-3-319-09244-7_9

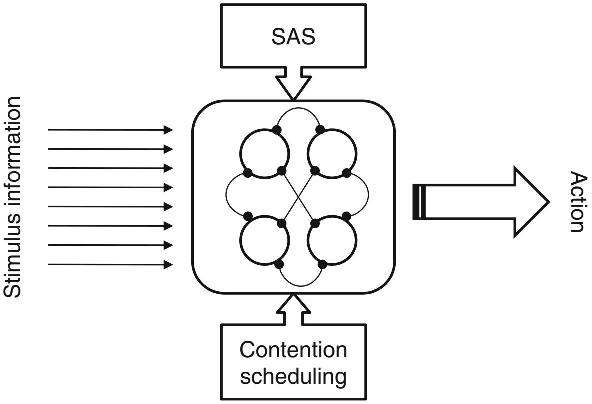


Fig. 9.1 The SAS-model (redrawn after Norman & Shallice 1986, by permission of Plenum Press)

- **Rule-based errors** occur in situations in which cognitive if-then rules are available and used, but not appropriately applied. An example would be a pedestrian crossing a street at red stoplight when another nearby traffic light goes green. The rule "if green then walk" is appropriate in principle but the perception of the trigger condition went wrong.
- **Skill-based errors** occur through the faulty functioning of cognitive processes that are adequate in principle. A common example are action slips caused by insufficient attention, such as when bumping into another person on the street.

These considerations fit well with the theoretical approach of Norman and Shallice (1986; Fig. 9.1). These authors assume that incoming stimulus information activates so-called **schemata**, overlearned stimulus-response rules that trigger a corresponding action. These schemata compete for action control and inhibit each other. If, thus, a processed stimulus activates only one single schema, the corresponding action can be immediately performed. If stimuli activate multiple schemata, however, as in the case of a Stroop stimulus, a paralyzing conflict situation is created.

According to the model, such conflict situations can be **resolved** through the influence of two factors. For one, the person might have learned to carry out the competing actions not at all or in a particular sequence, or to prefer one over others. These kinds of learning-dependent conflict regulations proceed automatically and are referred to as "contention scheduling." For another, the person could devote more attention to the task. The regulation of endogenous attention is taken care of by the **Supervisory Attentional System** (SAS). The role this system plays in the approach of Norman and Shallice is comparable to the role of action goals as discussed in Chap. 3 (Sect. 3.1) and the structure of the SAS model is very similar to the contextual-control model of Cohen et al. (1990; Sect. 3.1). Also, the way the SAS model is applied to explain action errors is highly consistent with other considerations. According to the model, action errors occur through:

- New situations for which appropriate schemata are not yet available, so that the SAS (which is particularly vulnerable against stress and other kinds of workload) needs to resolve the conflict.

- The erroneous activation of the wrong schema.
- The lack of sufficient monitoring through the SAS.

It is easy to see that these alternatives fit nicely with the three error categories suggested by Rasmussen: **knowledge-, rule-, and skill-based errors**, respectively.

9.2 Action Monitoring and Error Detection

The long-standing interest of psychological research in the analysis of human errors derives from the fact that in everyday life errors play a much more important role than the mainly theoretically interesting and often tiny reaction time differences that psychological experiments focus on. It is easy to imagine various situations where a delay of 20 ms is much less dramatic than making a mistake, such as in traffic. However, this research line has commonly defined errors in an objective fashion and derived the existence of an error from overt behavior. Whether and how the acting person **processes and perceives an error** was not in the focus. This latter question has attracted increasing attention in the recent years however. One reason for this increasing interest was the observation that acting people are surprisingly good and reliable in assessing and identifying their own errors. This is obvious from informal observations in psychological experiments, where they are often visibly and audibly frustrated when pressing the wrong key, but also from numerous empirical studies.

9.2.1 Behavioral Responses to Errors

Rabbitt (2002) is one of the pioneers investigating the **cognitive processing of errors** empirically. He had for instance participants indicate each committed error they noticed in a task by pressing a key, which they were able to do very reliably. Participants in choice-reaction time tasks often **spontaneously react** to their own mistakes and they tend to **correct** action errors. For instance, they often press the actually correct key after having pressed the wrong one, even if they were explicitly instructed not to do so (Rabbitt and Rodgers 1977). Such corrections are extremely fast and can occur about 250 ms after the wrong reaction. This suggests that the correction of action errors is based on very **rapid, automatically operating processes**.

To account for these observations, Rabbitt and colleagues assume that the eventual selection of a motor response results from the continuous **collection of evidence** for one or the other of the available action alternatives. Once sufficient evidence favoring one alternative is available, the reaction is triggered. This does not stop the evidence-collection process, so that it can happen that more evidence favoring the other alternative is available while the previously favored reaction is still under way or even after it has been executed. If the now available evidence is favoring the alternative to a sufficient degree, it is then carried out as well (Rabbitt et al. 1978; Rabbitt and Vyas 1981).

If we consider that the two hypothetical evidence-collection processes (one supporting the erroneous response and the other supporting the actually correct response) operate within splits of a second, it becomes clear why correction reactions are so **difficult to suppress**. On the one hand, such corrections can be considered automatic, given that they are difficult to suppress but, on the other hand, they do make sure that the actual action goal is eventually reached. In that sense, corrections represent what we have earlier coined "**conditional automaticity**" (see Sect. 5.3)—they operate automatically and yet reflect the current action goal. These considerations suggest that errors are corrected more likely as more time is available, as this implies that more evidence can be collected. To test that, Rabbitt and Vyas (1981) have analyzed the probability of performing a correction as a function of the duration of stimulus presentation. As expected, more corrections were observed when the stimulus was presented for a longer time.

These studies demonstrate that acting people detect their errors spontaneously. This raises the question **whether and how they monitor their actions** to check whether any errors might occur. Note that the studies of Rabbitt and colleagues do not necessarily suggest a monitoring process. What they show is that people carry out two different actions if the first of them is an error. But whether they carry out the second action **because** the first was an error is not clear. It could just as well be that both actions collect evidence in their favor simultaneously and independently from each other, and are then triggered if this evidence surpasses a certain threshold. Even the fact that they can later report that they made an error does not necessarily imply that they had registered the error at the time point they made it and that they had to do so even without instruction to report it later.

Alternatively, it might be that analyses of conscious reactions to errors actually **underestimate** the cognitive system's capacity to discover errors. It is possible that errors **need not be consciously detected** to trigger corrections and other repair operations. Testing this possibility requires the employment of measurement tools that allow us to analyze unconscious processes as well. As we will see in the next section, electrophysiological methods turned out to be very useful for such analyses.

9.2.2 Electrophysiological Correlates of Error Detection

The quest for possible mechanisms of error detection have revealed a diagnostic component in the event-related potential obtained in response to errors. The component comes under two different names: it was coined N_e by Falkenstein et al. (1991) when discovering it but is now better known as **Event-Related Negativity** (**ERN**; Gehring et al. 1993). A brief look at Fig. 9.2 explains both labels. The figure shows two event-related potentials, one for correct performances of an action and another for incorrect performances, measured from the onset of the reaction (indicated by the Y axis). Note the course of the two potentials over time and the relationship between the two lines. They diverge at the beginning already, with the function of the erroneous action going way more negative than that of the correct action. It is

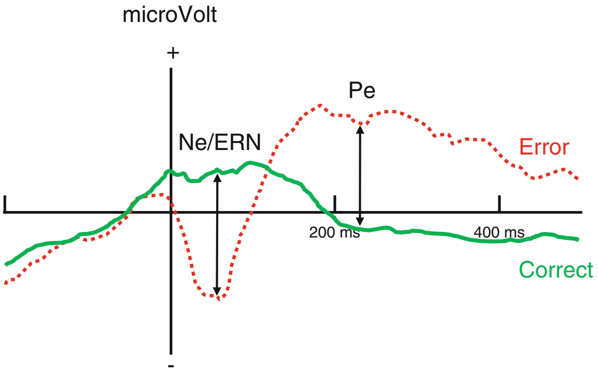


Fig. 9.2 Error-related components of the EEG (redrawn after Yeung et al. 2004, by permission of Elsevier, coloring is ours)

this stronger negativity associated with the erroneous response that the two labels refer to. Somewhat later, the two functions converge but only to diverge again. This time inversed in sign: the error-related function now gets much more positive than the one for the correct reaction, which is why Falkenstein et al. (1991) have coined the error-related part of the component P_e (standing for error-related positivity).

The N_e/ERN is commonly observed in **three situations**: if one is committing an error in a choice-reaction time task, if one responds too late, or if one receives feedback about the action's accuracy. The component is assumed to reflect activity of neural populations in the **anterior cingulate cortex (ACC)**, a cortical region that is assumed to play an important role in the monitoring of actions (see Sect. 2.6.3). The N_e/ERN is not only very fast but also independent from conscious awareness, as it even occurs when the acting person **did not notice** to have committed an error (Nieuwenhuis et al. 2001). In contrast, the P_e seems to be more strongly associated with the conscious awareness of the error, but we do not yet know much about this component. In any case, the existence of both components suggests that our cognitive system is continuously busy with **monitoring** the progress and goal-related utility of ongoing actions.

Recent studies have indicated that monitoring processes do not only automatically check for success or failure of actions but they also seem to register the occurrence of cognitive conflict. In his **theory of cognitive dissonance**, Festinger (1957) has emphasized the particular role of conflict for the organization of decision-making. These considerations have informed modern **models of error/conflict monitoring**, which we will discuss in the next section. Even the component N_e/ERN has been suspected to indicate not only the occurrence of **errors** but the occurrence of **conflict** as well. Indeed, significant N_e/ERNs can be obtained even in correct actions if they are likely to involve decision-making conflict, such as in Stroop-like situations. It is thus possible that the component reflects any kind of cognitive problem during processing, such as conflict between response alternatives, even if the problem is eventually resolved before execution (so that no error occurs). Or it reflects a process that continuously compares expected outcomes of the action against intended outcomes of the action. Both kinds of processes would be of particular importance for the identification of errors, but they may also inform cognitive processing about conflict as such.

9.2.3 Neural Correlates of Error and Conflict Detection

As mentioned already, it has been hypothesized that the N_e/ERN component reflects the activity of the ACC, which is a part of the medio-frontal cortex. This hypothesis has received ample support from fMRI studies. In a particularly influential study, Botvinick et al. (1999) had participants perform a flanker task in an fMRI scanner. The stimuli consisted of rows of arrowheads, in which the central arrowhead was the target, with its direction signaling a left or right keypress. The target was flanked by other arrowheads that would point into the same direction (e.g., >>≥>>) or into another direction (e.g., <<≥<<), with the assumption that the latter condition would create more response conflict. The authors found stronger activity of the ACC in the conflict-inducing conditions, even in trials in which participants responded correctly. Comparable findings were reported for other conflict-inducing tasks, like those using the Stroop effect. An increase of activation in the ACC has also been observed when participants carried out uncommon and infrequent actions or when deciding between equally appropriate action alternatives (e.g., Thompson-Schill et al. 1997).

These and other findings have motivated the development of the **conflict monitoring theory** suggested by Botvinick et al. (2001). According to this theory, the ACC monitors the cognitive system for the occurrence of cognitive conflict, such as between response alternatives. If a conflict is registered, the ACC reports to systems responsible for cognitive control, which results in the immediate increase of top-down control. This increase can be fast enough to resolve the conflict before an error occurs, even though this challenges the cognitive system more than processing in the absence of conflict.

This theory has **two attractive features**. For one, it proposes a mechanism that is both simple and efficient. Even though the hypothesized function of the ACC is not very complicated, it does increase the quality of action control enormously (via the loop "conflict detection → increase of control → conflict resolution") while playing the role of a whistleblower that allows the cognitive control system to disengage whenever tight control is unnecessary. Given that control is capacity limited and takes substantial effort, it makes sense to spare control whenever possible, especially given that control problems are in a sense reporting themselves.

For another, the theory fits very well with other theoretical considerations. Consider, for instance, the control model suggested by Cohen et al. (1990; Fig. 3.1). Conflict theory simply adds a further module to this model, a module that merely receives information about conflict and outputs activating signals to the goal representation. The conflict theory is also very similar to the SAS model of Norman and Shallice (1986; Fig. 9.1), who assume that conflict between competing schemata brings the SAS into play—exactly as suggested by the conflict theory.

9.3 Action Regulation and Error Avoidance

The available evidence provides solid support for the hypothesis that the ACC plays an important role in the monitoring of actions and action-related conflict. According to conflict monitoring theory, this role is rather passive, however: the ACC merely serves to translate signals indicating conflict into alerts sent to control systems without actually being part of the control process. Is the role of the ACC really that passive and, if it is, how can we characterize the actual control process?

To tackle these questions, MacDonald et al. (2000) had participants perform a task with Stroop stimuli in an fMRI scanner. In a given trial, participants were either to name the color of a congruently or incongruently colored color word or to read that word, as signaled by a task cue presented several seconds before the stimulus. The main question was whether and to which degree these conditions would activate the ACC and the dorsolateral prefrontal cortex (DLPFC), which is assumed to represent the current action goal and to orchestrate goal-related cognitive-control processes (Sect. 2.6.1). The authors observed that DLPFC was mainly activated during the preparation of the new upcoming task, and that this activation was particularly strong while preparing for the more difficult color-naming task. In contrast, the activation of the ACC during the same time interval was independent from the prepared task and its difficulty. This relationship was exactly reversed after the stimulus was presented. Now the ACC was particularly active, especially if the word and the color implied different responses, while the activation of DLPFC was independent from response conflict. This means that ACC and DLPFC play **dissociable roles** in action regulation, and they do so in the way implied by conflict monitoring theory: the DLPFC establishes a particular task set in the cognitive system and controls the flow of information through it, while the ACC monitors this flow and alerts the DLPFC in the case of problems.

The findings of McDonald and colleagues (2000) are clear in assigning separable roles to DLPFC and ACC, but they are not particularly telling with respect to how these two systems **interact**. More informative with regard to this interaction is a study conducted by Kerns et al. (2004). If ACC does not only detect conflict but also increase control exerted by DLPFC, so the authors reasoned, then this increases should have a beneficial impact on conflict processing in the following trial. Hence, the occurrence of cognitive conflict, and the resulting upregulation of control, should make it easier to deal with the next upcoming conflict. This would imply that the conflict-induced effects obtained in classical conflict tasks (like the Stroop task, Simon task, or flanker task) should get **smaller** right after a conflict trial. This pattern has indeed been often observed, for the first time in a flanker task: After a trial in which target stimulus and flanker were related to different responses, the flanker-compatibility effect was smaller than after a trial in which target stimulus and flanker were related to the same response (Gratton et al. 1992).

Like others before them, Kerns et al. (2004) found the same kind of effect in a Stroop task: the Stroop effect was smaller after trials with incompatible word-color pairings than after trials with compatible pairings. ACC was again more activated with conflict-inducing word-color pairings, but this activation was dampened right after a conflict trial. It is thus possible that activating the ACC in the face of conflict (in the previous trial) has strengthened cognitive control, so that the conflict in the following trial was less pronounced. This scenario has received support from another observation of Kerns et al.: the more strongly ACC was activated in the first of two successive conflict trials, the smaller was the conflict-induced reaction-time delay in the second trial. Moreover, the more ACC was activated when encountering the first conflict, the more DLPFC was activated when encountering the second conflict—a pattern that supports the hypothesized **interaction between ACC and DLPFC** in conflict-induced action control.

These findings point to the existence of an adaptive control loop, in which the ACC registers conflict and DLPFC resolves it. Up to now, we have focused on the prospective effects of this **control loop**, that is, on the consequences of current adjustments for future control operations. In laboratory tasks, when the same task is carried out hundreds and hundreds of times, this is of course particularly adaptive, as it helps to keep the performance level at a rather high level. In everyday life, however, we do not repeat our actions very often, at least not immediately, so that **prospective** effects of control are likely to have disappeared before they can affect the next repetition. More useful would be adjustments that immediately affect ongoing processes, so that the occurrence of conflict can contribute to resolving itself. The ACC/DLPFC control loop could manage this in principle, but is it really that fast?

A particularly clever fMRI study of Egner and Hirsch (2005) suggests that it is. In this study, participants were again presented with a conflict-inducing Stroop-like task. The stimuli consisted of response-compatible or response-incompatible combinations of faces and names of famous actors and politicians. In the most relevant condition, participants were to decide whether the presented face would show an actor or a politician while the names should be ignored. Unsurprisingly, performance was better if face and name were response-compatible. For instance, participants were faster to categorize the face of Robert de Niro as belonging to an actor if the name “Jack Nicholson” rather than the name “Bill Clinton” was printed over it. Like in the study of Kerns et al. (2004), this compatibility effect was smaller after incompatible trials than after compatible trials. This suggests that **control was adjusted** from one trial to the next, but was there also evidence for adjustment effects for the ongoing trial? And how did adjustment work?

To test that, Egner and Hirsch (2005) measured the activation of DLPFC and of the cortical area that is involved in the processing of faces—the fusiform face area (FFA). Two observations were particularly diagnostic: just-experienced response conflicts (induced by response-incompatible stimulus combinations) increased the activation of the FFA in the next trial and this increase in activation was apparently caused by the DLPFC (as suggested by so-called connectivity analyses). Given that the FFA is coding for the task-relevant stimulus information, this observation suggests that the experience of a response conflict leads to an immediate adjustment, which consists in the increase of cortical sensitivity for the relevant information. In other words, conflict leads to a stronger focus on the action goal, which in turn guides attention to focus on the task-irrelevant stimulus information.

9.4 Error-Based Learning

We learn from our mistakes. Indeed, continuously monitoring our actions for possible errors does not only allow immediate and short-term adjustments of action control but also temporally much more extended learning processes. We do not repeat unsuccessful actions over and over again but try to increase the efficiency of our actions every time we perform them. Learning to select the appropriate action has been associated with processes that are informed by reward and punishment, success and failure, in the sense that we tend to repeat actions that have positive consequences and avoid actions leading to negative consequences. External reward and punishment is not always available, however, and the fine-tuning of actions needs much more detailed information about action outcomes than reward and punishment can provide. For instance, if you are in the process of learning to ski downhill, it is not particularly helpful if all your teacher does is to shout "no good, no good," as correct as that may be. In situations like this you do not only need information regarding **whether** you made a mistake but also information about **what** went wrong in particular. Where do you get this information from? How can you learn from experience even without external feedback from others?

Box 9.1: Why We Can't Tickle Ourselves

Our considerations on the reafference principle (Box 9.2) in the role that predictions of action consequences play in the detection of errors are perhaps easier to understand in the context of an everyday phenomena: the observation that you can't tickle yourself. Why is that so? Blakemore et al. (1998) explain this observation by assuming that predicted consequences of our actions are experienced less intensively than unexpected consequences. The underlying reasoning is captured in their model as sketched in Fig. 9.3.

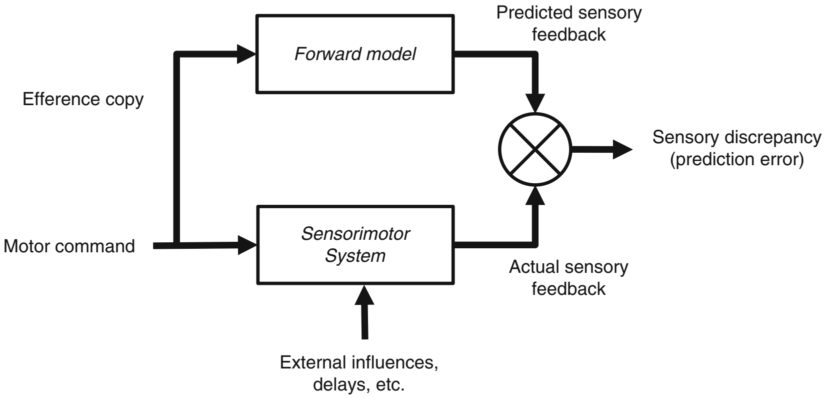


Fig. 9.3 Model for understanding why predicted consequences of our actions are experienced less intensively than unexpected consequences (adapted from Blakemore et al. 1999, by permission of MIT Press)

(continued)

Box 9.1: (continued)

If you are in the process of tickling yourself, for instance by touching the palm of your left hand with the fingers of your right, you are sending a corresponding motor command to your sensorimotor system. Executing this command (which controls the movements of your right hand) produces the intended sensorimotor feedback, as shown in the lower loop in the model. At the same time, a copy of the command—the efference copy—is used to compute the expected sensory feedback, as shown in the upper loop of the model. Finally, the system compares the expected and the actual feedback by subtracting the former from the latter, so to detect possible discrepancies, which in turn would indicate possible action slips or errors. If you intend to tickle yourself, and everything goes as intended, the resulting discrepancy is close to zero, which means that you hardly perceive anything. However, if you are tickled by another person, you will be unable to predict the expected feedback accurately, which creates discrepan-

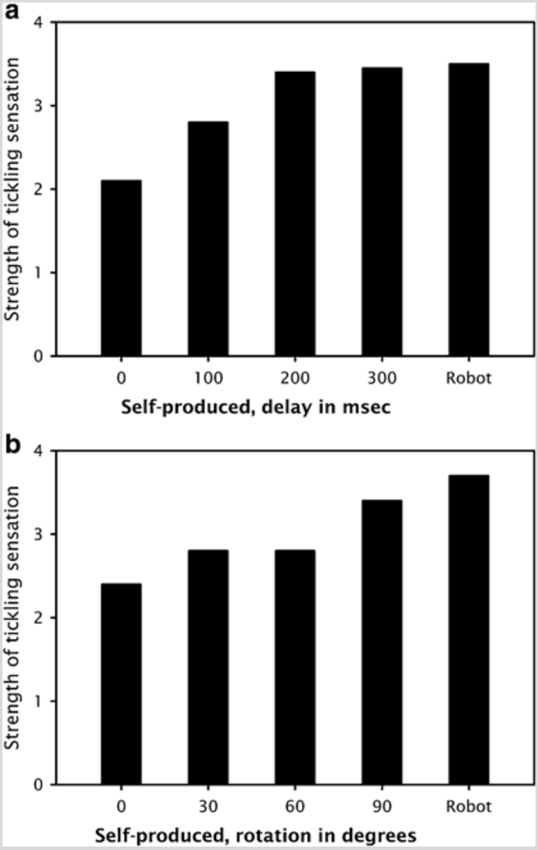


Fig. 9.4 The intensity of the experienced tickling depends on the similarity between temporal (**a**) and spatial (**b**) parameters of the self-produced and the robot-generated movement effects (redrawn after Blakemore et al. 1999)

(continued)

Box 9.1: (continued)
cies, which in turn correspond to the perception of tickling. So far the theorizing.

To test this interpretation, Blakemore and colleagues (1999) asked participants to tickle themselves. They did not do this directly, but by controlling a robot that tickled the palm of their left hand by means of a piece of rubber foam. In another condition, the robot was tickling the left hand of the participant autonomously, that is, without any contribution of the participant. As predicted, participants experienced the tickling much more intensively if they were tickled by the autonomous robot than if they produced the tickling themselves (Fig. 9.4a). In another study, participants also tickled themselves by controlling the robot, but the robot's tickling response was delayed by 100, 200, or 300 ms. The authors assumed that delays between the movement command in the occurrence of the actual feedback would induce some degree of discrepancy between expectations and actual feedback. This should increase the intensity of tickling, which indeed was the case (Fig. 9.4a).

Blakemore et al. also manipulated the similarity between expected and actual sensory consequences of the participants' actions. In particular, they varied the spatial relationship between the movement participants carried out to control the robot and the tickling movement of the robot, which were either aligned or rotated by 30°, 60°, or 90°. As expected, the intensity of the tickling perception increased with the degree of rotation, that is, with increased discrepancy (Fig. 9.4b).

In addition to subjective experience of participants, Blakemore et al. (1998) also analyzed the corresponding neural activities. FMRI analyses showed that intentionally triggered tickling movements produce significantly less activity in the somatosensory cortex and the ACC than tickling movements that the robot autonomously produced. This supports the assumption that the successful prediction of the sensory consequences of one's action reduces the intensity with which these consequences are perceived. This in turn suggests that the comparison between the expected and the actual feedback takes place before this information reaches the sensory areas representing it. Moreover, these observations suggest that the ACC is involved in detecting discrepancies between expected and actual feedback, which is consistent with the conflict theory of Botvinick et al. (2001; see Sect. 9.2.3).

Box 9.2: The Reafference Principle

The reafference principle, which assumes that errors are detected by comparing expected and actual feedback, is a control principle that derives from cybernetics. Originally, this principle entered psychology as an account for why we experience our visual world as stable even though we move our eyes about four times per second. Each movement shifts the part of the environment that is projected on our retina rather dramatically, so why do we not have the impression that we are living in a floating world?

To explain this phenomenon, von Holst and Mittelstaedt (von Holst 1954) have suggested that the control of eye movements is based on the **comparison of the visual changes that a given eye-movement is expected to produce and the actual changes**: the reafference principle. At each single eye-movement, so the idea, a copy of the motor command that is moving the eye (the efference copy) is sent to neural structures that calculate how the movement will change the sensory projection on the retina. If then the eye movement is carried out, the visual information about the resulting change is compared against the predicted change. If the discrepancy between the two is close to zero, the sensory signal is canceled, so that the perceiver is not aware of the change. A movement is perceived only if the discrepancy is considerable or if one of the two signals is missing, as if the eye has been moved non-intentionally or if the expected change does not occur, such as in studies where the eye has been experimentally paralyzed. You can easily test this prediction by moving one of your eyes by hand (through the eyelid, be careful)—which should give you the impression that your visual world is moving. It is this experience of motion that the reafference principle aims to prevent.

For quite some time, all we had was good theoretical reasons to assume that the reafference principle takes care of the **stability** of our visual world, but recent studies in monkeys have provided strong empirical support as well. Sommer and Wurtz (2006) have trained two monkeys to fixate a light dot on a monitor and to perform saccades towards a visual target whenever it appeared on the screen (Fig. 9.5a). At different points in time a probe stimulus would appear either close to the fixation point or in the neighborhood of the target (Fig. 9.5b). Microelectrodes measured the activities of neurons in the frontal eye fields (FEFs) of the monkeys, especially in those neurons the receptive fields of which would cover the area in which the probe was presented (Fig. 9.5c).

(continued)

Box 9.2 (continued)

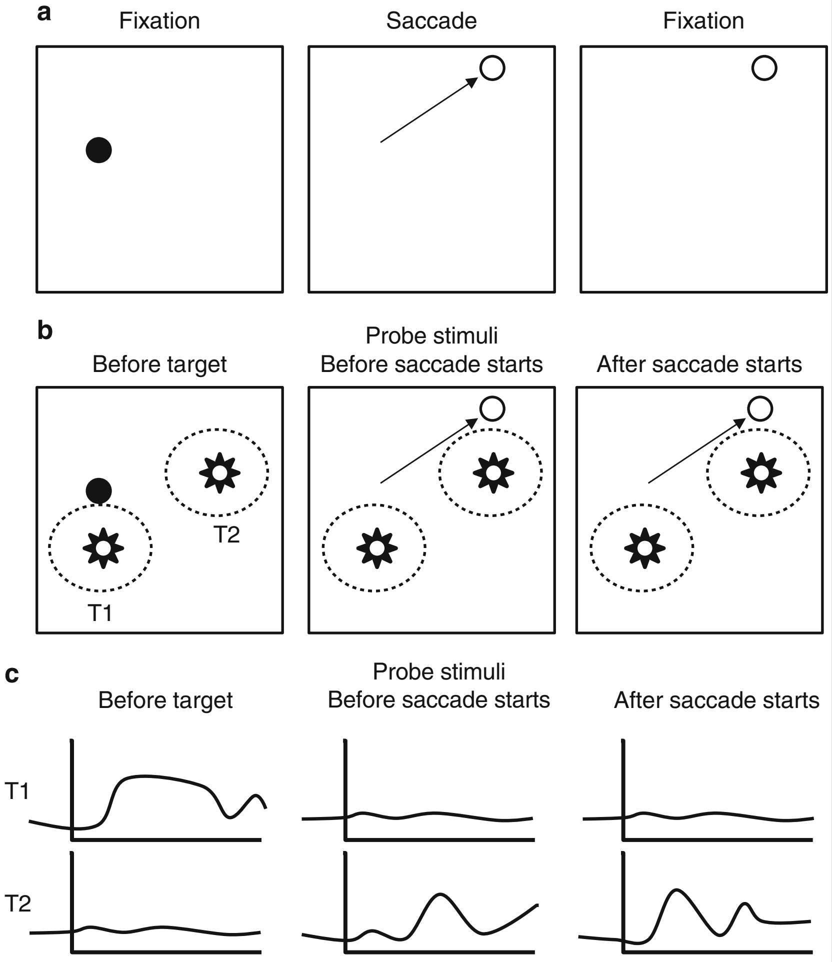


Fig. 9.5 Activity of FEF neurons (in monkeys) when executing saccadic eye movements from a fixation point to a peripheral target (adapted from Sommer & Wurtz 2006), by permission of MacMillan Publishers Ltd.)

Before the target appeared (see leftmost column), only neurons with receptive fields in the neighborhood of the fixation point reacted (T1 but not T2, see panels b and c), as one would expect. When a saccade was executed (rightmost column), these neurons were no longer active (see T1 in panel c) while those with receptive fields in the neighborhood of the target would get into action (see T2 in panel c), just as one would expect as well. This suggests that the monkeys' focus of attention **moves from the fixation point to the new target**. What was interesting, however, was that the neurons with receptive

(continued)

Box 9.2 (continued)

fields close to the target started firing before the eye movement was executed (see center column, T2 in panel c). The focus of attention is thus faster than the actual movement and it anticipates the new location. The authors suggest that neurons of the frontal eye fields move their receptive fields in the direction of the new location before the eye movement begins. As a consequence, the movement-induced changes of the retinal input are compensated even before the actual movement starts, so that the neurons meet basically the same visual information before and after the movement. In a sense, the actual change only affects the retina while higher-order systems of visual information processing get the impression that nothing has changed. This is the logic underlying the reafference principle.

One important principle that explains how errors cannot only be identified qualitatively but also described quantitatively, which is of particular importance for adaptive learning, stems from **cybernetic systems theory**. As touched upon in the section on the functioning characteristics of the cerebellum (Sect. 2.4) and the TOTE model (Sect. 7.2.3), cybernetic control loops operate by comparing intended and actual outcomes of particular processes. Your central heating system is a good example. If you for instance choose 21 °C as the room temperature you prefer, the system compares this intended value against the actual value. If the actual value is lower than the intended one, the room is heated up, but if the actual value is higher than the intended, the room is cooled down. In this example, the sign of the outcome of the comparison can be considered to inform about success or failure but the comparison also provides information about the degree of the discrepancy between ideal and actual outcome. Learning to adjust the details of our actions may thus be conceived of as aiming to **reduce the discrepancy** between intended action and actual action. Given that this discrepancy can often be calculated without any external teacher, comparison-based learning provides the learner with a high degree of autonomy.

This kind of learning is commonly referred to as "supervised learning," which follows a different logic and seems to have a different neural basis than other forms of learning. According to Doya (2000; see Fig. 9.6), supervised learning can be distinguished from "unsupervised learning" and "reinforcement learning." Unsupervised learning relies on the neocortex, which is highly specialized in picking up all sorts of relations between events irrespective of reward, punishment, or particular expectations. Hebb (1949) had suggested the existence of such processes already (often referred to as **Hebbian learning**) and developed a corresponding learning algorithm, which still plays an important role in neural-network modeling ("what fires together wires together"; see Sect. 4.3.1). In contrast, **reinforcement learning** relies on neural structures that connect the neocortex with the basal ganglia. It is there where affective implications of actions (i.e., of reward and punishment) are coded, which in turn inform and guide the acquisition of contextually adequate actions.

Supervised learning, which is based on errors, relies on the cerebellum and its interactions with neocortical structures (see Sect. 2.4). As shown in Fig. 9.6, this kind

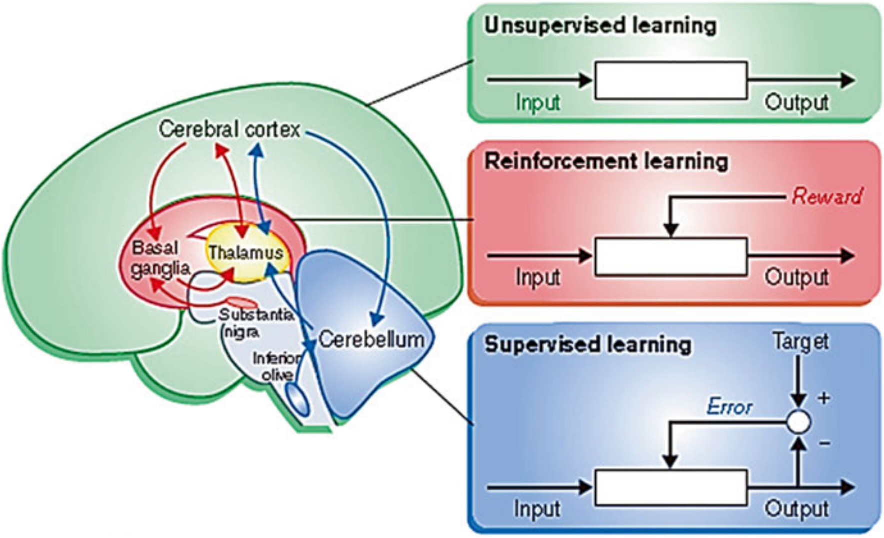


Fig. 9.6 The neural networks underlying supervised learning, unsupervised learning and reinforcement learning, respectively (adapted from Doya 2009, by permission of Elsevier)

of learning makes use of the cybernetic principle of discrepancy reduction by comparing intended outcomes against actual outcomes of actions. The intended outcomes can be specified by the learner herself or by another person, as in learning through imitating the teacher. Supervised learning helps to acquire so-called **forward models** and **inverse models**, which play an important role in movement control (see Sect. 2.4 and Box 9.2). Forward models serve to predict the sensory consequences of actions, they in a sense "look forward." The critical comparison is then between the predicted consequences and the actual consequences. The larger the discrepancy between the two, the more the learner adjusts her forward model, so that predictions become more and more precise over time (the process of "feedback error learning," Kawato 1990). Inverse models integrate information about which effects can be obtained by which action under which circumstances, which makes these models particularly important for the selection of actions and movements (Chap. 5).

Acquiring such models is difficult, because there is **no learning signal** at the beginning of the learning process: If one doesn't know what one could do/achieve and how one could do/achieve it, it is hard to develop realistic expectations. But without such realistic expectations, matching expected against actual consequences—which in supervised learning generates the most crucial information—is impossible. Assume, for instance, you are busy with learning basketball and now you want to learn how to throw the ball into the basket from different distances. You need to find out which motor commands are generating body movements that make you reach the intended goal. But this is particularly difficult because success depends on, and varies with various not yet fully appreciated factors, such as the position of the basket in the room, the distance between thrower and basket, and the thrower's current body position.

All the absolute beginner has at her disposal is **trial and error**. You just throw and look what happens. But with each try you are building up a sensory motor knowledge base. Each movement produces a multitude of body- and environment-related sensory impressions: you perceive how your arms and hands are moving, you register (typically unconsciously) which combinations of force in joint postures are associated with successful throws, how these combinations predict how the ball moves in time and space, which combinations lead to failure, and so forth. All this information provides the opportunity to build up **associations** between motor patterns and representations of these patterns' sensory effects, which in turn allows building up expectations about to-be-produced effects that are increasingly realistic (see Sects. 3.3 and 4.3.1). The acquisition of these associations is easier if complex skills are learned in a stepwise fashion, if you acquire one component after the other. Particularly obvious is this divide-and-conquer strategy when learning to drive a car. Shifting into another gear is first achieved by performing all the component actions separately: lifting off your foot from the gas pedal, moving it to the clutch pedal, press the pedal down, move the lever, and so forth. However, the more often you practice this sequence, the smoother everything goes, suggesting that the elements have been integrated into a unitary action plan (see Sect. 7.2). Now the integrated pattern cannot only be performed with much less effort, it can also become part of a more complex skill through supervised learning.

References

Blakemore, S.-J., Frith, C. D., & Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience, 11*, 551–559.

Blakemore, S.-J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience, 1*, 635–640.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review, 108*, 624–652.

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature, 402*, 179–181.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review, 97*, 332–361.

Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10*, 732–739.

Egner, T., & Hirsch, J. (2005). Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nature Neuroscience, 8*, 1784–1790.

Falkenstein, M., Hohnsbein, J., Hoorman, J., & Blanke, L. (1991). Effects of crossmodal divided attention on late ERP components: II. Error processing in choice reaction tasks. *Electroencephalography and Clinical Neurophysiology, 78*, 447–455.

Festinger, L. (1957). *A theory of cognitive dissonance*. Evanston, IL: Row & Peterson.

Freud, S. (1904). *Zur Psychopathologie des Alltagslebens*. Berlin, Germany: Karger.

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*, 385–390.

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation and responses. *Journal of Experimental Psychology: General, 4*, 480–506.

Hebb, D. O. (1949). *The organization of behavior*. New York: Wiley.

Kawato, M. (1990). Feedback-error-learning neural network for supervised learning. In R. Eckmiller (Ed.), *Advanced neural computing* (pp. 365–372). Amsterdam: North-Holland.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., et al. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science, 303*, 1023–1026.

MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science, 288*, 1835–1838.

Nieuwenhuis, S., Ridderinkhof, K. R., Blom, J., Band, G. P. H., & Kok, A. (2001). Error-related brain potentials are differentially related to awareness of response errors: Evidence from an antisaccade task. *Psychophysiology, 38*, 752–760.

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behaviour. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York: Plenum.

Rabbitt, P. M. A. (2002). Consciousness is slower than you think. *Quarterly Journal of Experimental Psychology, 55A*, 1081–1092.

Rabbitt, P. M. A., Cumming, G., & Vyas, S. M. (1978). Some errors of perceptual analysis in visual search can be detected and corrected. *Quarterly Journal of Experimental Psychology, 30*, 319–332.

Rabbitt, P. M. A., & Rodgers, B. (1977). What does a man do after he makes an error? An analysis of response programming. *Quarterly Journal of Experimental Psychology, 29*, 727–743.

Rabbitt, P. M. A., & Vyas, S. M. (1981). Processing a display even after you make a response to it. How perceptual errors can be corrected. *Quarterly Journal of Experimental Psychology, 33A*, 223–239.

Rasmussen, J. (1980). What can be learned from human error reports? In K. D. Duncan, M. M. Gruenberg, & D. Wallis (Eds.), *Changes in working life* (pp. 97–113). Chichester, England: Wiley.

Reason, J. (1979). Actions not as planned: The price of automatization. In G. Underwood & R. Stevens (Eds.), *Aspects of consciousness* (Psychological issue, Vol. 1, pp. 51–74). London: Wiley.

Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature, 444*, 374–377.

Thompson-Schill, S. L., D'Esposito, M., Aguirre, G. K., & Farah, M. J. (1997). Role of left inferior prefrontal cortex in retrieval of semantic knowledge: A reevaluation. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 14792–14797.

Von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *The British Journal of Animal Behavior, 2*, 89–94.

Index

B. Hommel et al., *Human Action Control*, DOI 10.1007/978-3-319-09244-7