Ulrich C. Drost

Sensory-Motor Coupling in Musicians

Cuvillier Verlag Göttingen

Sensory-Motor Coupling in Musicians

Inaugural-Dissertation
zur Erlangung des Doktorgrades der Philosophie
an der Ludwig-Maximilians-Universität zu München

vorgelegt von

Ulrich C. Drost
aus München

Cuvillier Verlag Göttingen,
Februar 2005

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.
1. Aufl. - Göttingen : Cuvillier, 2005
Zugl.: München, Univ., Diss., 2005
ISBN 3-86537-363-1

Gutachter:

Prof. Dr. Wolfgang Prinz

PD Dr. Iring Koch

Tag der Disputation: 02.02.2005

Nonnenstieg 8, 37075 Göttingen
Telefon: 0551-54724-0
Telefax: 0551-54724-21
www.cuvillier.de

1. Auflage, 2005
Gedruckt auf säurefreiem Papier

ISBN 3-86537-363-1

Acknowledgements

An extensive project like this dissertation would never have been possible without the support from other people. First, I would like to express my gratitude to Prof. Wolfgang Prinz for enabling me to carry out my dissertation at the Max Planck Institute for Human Cognitive and Brain Sciences (Dept. of Psychology) in Munich. The scientific and social environment he has fashioned at the MPI has been ideal to carry out the research for this dissertation. Then, of course, I would like to thank my direct supervisor, Dr. Martina Rieger, for her continued support throughout the whole project. I have benefited tremendously from her scientific experience, be it in analyzing experimental data, in understanding it, or in writing a comprehensible paper. My further thanks go out to our collaborators in Leipzig, Dr. Marcel Brass and Dr. Thomas Gunter, who have accompanied the project from the distance. I am also very much indebted to Sonja Herberth, who conducted most of the experiments, and to Karl-Heinz Honsberg, who has built most of the ingenious technical devices that were used in the experimental setups. And finally, of course, I am very grateful to my roommate Michael Öllinger and to all my other colleagues at the institute – for many lunches, interesting discussions, and simply for a fascinating and inspiring time.

Munich, September 2004
U. C. D.

Table of Contents

1 Introduction

In the first part of his two-volume magnum opus “The World as Will and Representation” [*Die Welt als Wille und Vorstellung*], which was first published in 1819, the German philosopher Arthur Schopenhauer made the following statement about the acting and perceiving subject: “Every genuine act of his [= the subject’s] will is at the same time and inevitably also an action of his body: he cannot actually will the act without realizing at the same time that it manifests itself as an action of the body. The act of the will and the action of the body are not two objectively discernable, disparate states, which are connected by the bond of causality, their relation is not one of cause and effect; but rather they are one and the same, if only presented in two entirely dissimilar ways: the one quite directly, and the other in the observation of the mind. The action of the body is nothing else but the objectified, that is, the perceivable act of the will. Furthermore it will show that this accounts for all possible actions of the body, not only for those that arise from motives, but also for those that involuntarily arise from simple stimuli, indeed, that the whole body is nothing else but the objectified, that is, the internalized will […]” (Schopenhauer, 1968, pp. 157-8, translation by the author).

This is quite a remarkable statement. Especially so, because for a long time into the 19th century, the scientific and the philosophical world (psychology was to become an independent scientific subject only some time later) had held the belief that, ontologically and epistemically, the mind and the body must be viewed as two absolutely independent entities. This belief had been expressed and made popular by the French philosopher René Descartes in his work

“Meditations on First Philosophy” [*Meditationes de Prima Philosophia*], first published in 1631. Descartes described the world as consisting of two independent entities, the *res extensa* (all things physical, including the body) and the *res cogitans* (the mind). One implication of this “Cartesian” dualism was the belief that the mind is not susceptible to scientific investigation. Accordingly, with the rise of the natural sciences, a more unified view of the world became popular, which tried to reconcile phenomena of the mind with physical principles.

However, apart from such epistemic issues, the considerations of Schopenhauer are also remarkable in another respect. Schopenhauer implies that an “act of the will” is at the same time an “action” of the body. Or, to turn it the other way round, a bodily action (that is, a movement) is nothing else but a “perceivable act of the will”. Other philosophers of the 19th century have expressed similar ideas, as I will show later. In terms of modern cognitive psychology, this conjecture suggests a very close relationship of *sensory* and *motor* processes. In fact, modern cognitive science has gathered a large amount of empirical evidence for such a close relationship. This evidence, as well as the underlying theoretical considerations, will be described in the following sections.

But before I go into the views of modern cognitive psychology in more detail, another quotation. This time, it is not from a philosopher, but from a performer of classical music. In his 1986 book “The Inner Game of Music”, contrabassist Barry Green writes about the performance of music: “When you can hold the sound and pitch of the music clearly in your head […] performing it accurately becomes easier. Your body has a sense of its goal […] Effectively, you are playing a duet between the music in your head and the music you are performing.” (Green & Gallwey, 1986, p. 75).

Note the similarity of the experience of Barry Green and the considerations of Schopenhauer. Both, the subjective experience as well as the analytical philosophical view, express the idea of a very close sensory-motor coupling, which directly manifests itself in the ability of people to control bodily action by means of the voluntary imagination of an intended goal. However, there is also a significant difference. In contrast to Schopenhauer, Green not only speaks of a self-contained action of the body, but also of a goal that lies *beyond* the bodily action. For him as a musician, the goal of his body movements is the production of musical sounds (on the piano, the trumpet, the contrabass,…). However, this ability to perform proper actions on a musical instrument can hardly be supposed to be present *per se*. He as a musician must have *learned* this ability of performing certain movements in relation to certain sounds.

Several questions arise from this point. First, is there empirical evidence for such phenomena of sensory-motor coupling, as presumed generally by Arthur Schopenhauer, and specifically for music by Barry Green? Second, if there is empirical evidence, what might be the underlying structural and functional principles for such effects? And third, what might be the use of such structural and functional principles in the machinery of the mind? The central motivation of the present work has been to investigate the above-mentioned issues in more detail. Thus, the main aim of this dissertation is the investigation of sensory-motor coupling in experienced musicians in a theoretical and empirical manner.

The work consists of three major sections. In the first section, a review of the empirical and theoretical literature on sensory-motor coupling, generally as well as specifically for music, is given. The second section describes and discusses experiments that have examined sensory-motor coupling in musicians in three domains: the harmony dimension of music, the melody dimension of music, and the question of instrument specificity in sensory-motor coupling. In the third and final section these empirical results are discussed more broadly, and are related to the theoretical premises, which have been established in the initial section.

2 Review of Literature

In the following, I will give a review of the relevant empirical and theoretical literature on sensory-motor coupling. The section is ordered into three major parts. In the first part, empirical evidence of sensory-motor coupling is discussed. After a description of a number of exemplary studies, effects of sensory-motor coupling are reviewed in two categories: effects that derive from *hard-wired* sensory-motor linking, and effects that derive from *learned* sensory-motor linking. The latter category mainly focuses on action-effect linking, which is also central for the empirical section of this study. The second part discusses theoretical accounts of sensory-motor coupling. These include accounts based on associative learning theory, computational models, and the ideomotor theory of voluntary action. The third and final part focuses on sensory-motor coupling in musicians. The ability to play an instrument is discussed as a specific form of expertise, which necessarily involves increased sensory-motor coupling. Furthermore, previous empirical studies on sensory-motor coupling in musicians are reviewed.

2.1 Empirical Evidence of Sensory-Motor Coupling

Most studies on sensory-motor integration and interaction have used a *perception-on-action* approach in their experiments: the performance of people in a certain task that involves doing certain actions is examined under varying conditions of perceptual stimulation. Instances of the influence of perception on action have been reported in a number of areas. These include the involuntary imitation of observed movements (Brass, Bekkering, & Prinz, 2001; Brass, Bekkering, Wohlschläger, & Prinz, 2000; Craighero, Bello, Fadiga, & Rizzolatti, 2002), the influence of irrelevant stimulus dimensions on responses (Eriksen & Eriksen, 1974; Kornblum & Lee, 1995; J. R. Simon, 1990; J. R. Simon & Rudell, 1967), and, more recently, the influence of (potential) action effects on action performance (Beckers, De Houwer, & Eelen, 2002; Elsner & Hommel, 2001; Kunde, 2001; Ziessler & Nattkemper, 2001). To a somewhat lesser degree, the case of *action-on-perception* influence has been examined as well, for example in visual discrimination (Müsseler & Hommel, 1997; Müsseler, Steininger, & Wühr, 2001; Wohlschläger, 2000), the perception of apparent motion (Wohlschläger, 2000), and temporal action-effect binding (Haggard, Clark, & Kalogeras, 2002). These latter studies show that, under certain circumstances, action preparation and/or execution can influence perceptual processing.

The studies, which are most relevant for the present work, are discussed in the following. An adapted version of the classical S-R-C model (see Tolman, 1932) may serve as an organizing framework. In the classical S-R-C black box model, the cognitive system is supposed to get input from perceptual *stimuli* (S), which give rise to a *response* (R), which in turn is followed by a certain *consequence* (C). A stimulus directly evokes an associated response. If the response is followed by a consequence of positive valence (a “reward”), the S-R association is strengthened (the “law of effect”, Thorndike, 1911). However, if we concede that people are not simple automatons that follow mechanistic rules (see for example Allport, 1980, for such a model), it is more useful to apply the term *action* (A) instead of *response*. The definition of “action” takes into account the goal-directedness of most human behavior, and emphasizes intrinsic rather than extrinsic control. To give a clear definition, actions are “goal-directed activities that consist of body and/or limb movements" (Magill, 2001, p. 3).

Furthermore, with the *consequence* of an action is usually denoted an event of either negative or positive valence (more explicitly: punishment or reward). However, any event that follows an action (mostly: that is caused by that action) can be perceived as a "consequence" of the action. Therefore, it is more useful to speak of the *effect* (E) of an action. Here, it is important to differentiate between proximal and distal action effects (see Prinz, 1990). *Proximal* effects are sensory events that are immanent to the movements of the acting person. In a way, to the actor a proximal effect is the movement itself. It is represented as kinesthetic, proprioceptive, tactile, and also visual information. As such, an arm movement, for example, is represented as what a person perceives, when he moves his arm. For the most part, proximal information is characterized by the fact that it is accessible only to the acting person (see also Metzinger, 2000, for an analysis of the first-person perspective). Proximal effects can be linked to *distal* effects by way of cause and effect. Distal effects are fed back to a person by far-reaching sensory modalities, like vision and audition. In most cases, the causation of a distal effect is the actual goal of a movement. A pianist, for instance, produces piano tones as distal effects (and even more distal, perhaps, positive emotional responses within the listeners).

If we follow this S-A-E model, sensory-motor interaction can arise from several sources. Perception-on-action effects can arise from perceived sensory events: this is the case when pre-action stimuli (S) or post-action effects (E) influence the control of actions (A). Effects of action-on-perception can arise when action control processes (A) influence the perception of sensory stimuli (S) or of ensuing effects (E). In the following, the major focus will lie on those studies, which have examined the influence of action effects (or stimuli that usually *are* action effects) on action performance. These studies are most relevant for music-related issues, which will become clearer later on.

2.1.1 Examples and Classification

Let me begin with a few exemplary examples that have demonstrated effects of sensory-motor coupling. Some of them directly relate to the present work. A study by Schubö, Aschersleben and Prinz (2001) may serve to illustrate what is meant with sensory-motor interaction. In this study, participants carried out movements on a writing pad, while they observed motions on a computer screen. The observed movement on the computer screen in a trial n was to be carried out on the writing pad in the subsequent trial $n + 1$. The results of this study showed a contrast-like effect of stimulus motion on performed movement: perceiving a small motion while performing a medium-sized movement increased movement size, while perceiving a large motion led to a decrease. Thus, the *perception* of visual events had directly influenced the *execution* of actions. One experiment even showed an opposite effect, that is, an influence of an executed action on visual perception.

Another example for the influence of perception on action is the involuntary imitation of observed actions. This has, for instance, been demonstrated in a study by Brass et al. (2000). In this study, participants carried out finger lifting movements (index or middle finger) in response to symbolic cues. Concurrently with the imperative cues, task-irrelevant picture sequences of finger movements were presented, which could be congruent or incongruent with the required response. Results showed both interference (higher RTs in the incongruent condition) and facilitation (lower RTs in the congruent condition), as compared to a baseline condition with only the imperative cue. When the distractor was made more dissimilar to the response (tapping the finger instead of raising it), interference was reduced, while facilitation disappeared entirely. No interference or facilitation effects were observed, when finger movements were used as imperative stimuli, and symbolic cues as distractors. It seems that the perception of the finger movements *involuntarily* triggered the corresponding movements. Arbitrary symbolic cues (single digits, in the experiments) were not associated with any response from the first, therefore they did not effect in any specific distraction. Similar effects of involuntary imitation have even been found in a simple response task (Brass et al., 2001).

Craighero et al. (2002) were able to demonstrate analogous effects for rotational hand movements. In their study, participants prepared rotational grasping movements, which were to be carried out on presentation of a visual stimulus (showing a mirror image of the hand). Results showed faster responses, the more the mirror image resembled the actual movement in terms of orientation. Even more interesting, these effects were the higher, the more similar the mirror image was to the *final* position of the movement, which was to be carried out. This study primarily illustrates the relevance of the movement goal. It seems that, in involuntary imitation, the most relevant aspect is not the movement itself, but its aimed-for end-state.

How can effects of sensory-motor interaction, as described in the examples above, be accounted for? For such effects to happen, it is necessary that sensory and motor processes are linked by some structural and/or functional principle. There are two possible sources for such sensory-motor linking: it may be hard-wired in the system, or it may be learned. Any *hard-wired* sensory-motor linking has its foundation in the underlying structures of the nervous system. Neurophysiological examples of hard-wired sensory-motor connections include low-level connections, like the monosynaptic reflex circuit (see Chen, Hippenmeyer, Arber, & Frank, 2003), as well as higher-level connections in the brain, like the dorsal visual stream (Goodale, Westwood, & Milner, 2004) or visuomotor neurons (Fadiga, Fogassi, Gallese, & Rizzolatti, 2000). On the other hand, *learned* sensory-motor connections derive from repeated experiences with sensory and motor events. Neuronally, this may result in newly formed connections (see Halsband & Freund, 1993). Note that even the so-called “hard-wired” connections are not present per se, but also have formed during the maturation of the brain. However, given normal developmental circumstances, they are formed similarly in all people.

One might assume that instances of purely hard-wired and purely learned sensory-motor coupling are rather particular extremes, whereas many observable effects might be located on a continuum between those two. For example, hard-wired connections might be modulated by learning experiences. I will focus on these extremes for the sake of clarity, but will also mention those instances, where such a clear ascription has been disputed. In the following two sections, behavioral studies of learned and non-learned sensory-motor relations will be discussed. Since the present study is mainly concerned with *learned* sensory-motor relations, these will be discussed in more detail, particularly in the domain of action-effect coupling.

2.1.2 Non-learned Sensory-Motor Relations

Many of the phenomena of sensory-motor interaction that are reported in the literature are probably due to ‘hard-wired’ characteristics of the cognitive system. Phenomena of this kind can be observed in all people alike, because they do not derive from any specific learning experience. For example, the effects of involuntary imitation in the studies of Brass et al. (2000) and

Craighero et al. (2002), which were described above, do not seem to be specific to any particular group of people.

The first group of such non-learned sensory-motor relations pertains to effects, which derive from pre-action sensory events: S-A relations (classically: S-R relations). First, there are S-R relations that are not only hard-wired, but also decidedly *innate*. To this group belong all the innate reflexes of the human body (for a thorough overview, see Desmedt, 1973). In all the reflexes, a body-related stimulus directly and involuntarily evokes a bodily response. Take, for example, the *blink reflex*: If some object is perceived to (suddenly and rapidly) approach the eye of a person, the person will involuntarily blink. This reaction is automatic and cannot easily be inhibited – it even occurs when a pane of glass is between the object and the eye, or when the "object" is not even solid, but only a moving shadow.

Effects that are not "innate", but probably nonetheless due to hard-wired characteristics of the sensory-motor system, are effects of stimulus-response compatibility (SRC). These were first reported by Paul Fitts in the 1950s (Fitts & Deininger, 1954; Fitts & Seeger, 1953). Fitts and his colleagues showed that response times are the faster, the more similar the spatial alignment of the responses is to the spatial alignment of the stimuli. Other studies showed that this is also the case when compatibility is located on a task-irrelevant dimension. Take, for instance, the well-known SRC effect in the Simon paradigm (J. R. Simon, 1990; J. R. Simon & Rudell, 1967). Here, participants have the task to carry out a spatially aligned response (e.g., a left or right keypress) in response to a certain imperative stimulus feature (e.g., color, auditory pitch). However, the stimulus also has a task-irrelevant spatial feature (e.g., it either appears on the left or the right side), which can be compatible or incompatible with the required response. Results typically show that responses are slower when the task-irrelevant feature is incongruent with the required response (the so-called *Simon effect*). A Simon effect has, *inter alia*, been found for auditory stimuli (J. R. Simon & Rudell, 1967), semantic stimuli (De Houwer, 1998), moving stimuli (Proctor, Van Zandt, Lu, & Weeks, 1993), and even for "stationary moving" stimuli (Bosbach, Prinz, & Kerzel, 2004). However, note that some authors have expressed the opinion that the Simon effect is not based on hard-wired connections, but derives from *learned* sensory-motor relations. For example, in support of this notion, Tagliabue and colleagues have argued that compatible relations are much more frequent in our everyday experience than incompatible relations (Tagliabue, Zorzi, Umilta, & Bassignani, 2000).

Effects of SRC also occur when the task-irrelevant (compatible or incompatible) feature is not even contained within the task-relevant stimulus. One example for this is the so-called *Flanker effect* (Eriksen & Eriksen, 1974). Here, a task-irrelevant "flanker" stimulus appears alongside the imperative stimulus, which contains the compatible or incompatible stimulus feature. For example, the imperative stimulus might be a letter, while the flanker is also a letter, which is either congruent or incongruent. Accordingly, responses are usually slower (and errors more frequent) in the case of incompatibility. Contrary to the Simon paradigm, this paradigm does not examine spatial compatibility effects, but rather effects that derive from visual and/or conceptual compatibility. Flanker effects can also be induced through learning. For example, if the task is to respond with a left keypress to the letters C and H, and with a right keypress to the letters A and F, then a F flanker will have a facilitory effect with an A, while a C flanker will have an interfering effect with an A. Such effects are then due to learned S-A associations.

Effects of stimulus-response compatibility are usually attributed to a principle called *dimensional overlap*, or *feature overlap* (Kornblum, 1992;

Kornblum, Hasbroucq, & Osman, 1990). According to this principle, effects of facilitation or interference occur when the stimulus and the response overlap in a common dimension. For example, if stimuli and responses are aligned on a horizontal spatial dimension (left or right), interference will occur if the S-R relation is incongruent, while facilitation will occur if the S-R relation is congruent (as in the Simon effect). Congruent and incongruent conditions are usually compared to a neutral condition, where stimulus and response do not overlap in any (response-relevant) dimension. Dimensions that can possibly overlap include spatial alignment (horizontal, vertical, depth), identity (same, different), number, semantic dimensions (as in the Stroop effect), and others. The principle of dimensional overlap, or rather feature overlap, is also a central constituent of one of the major theories on sensory-motor coupling, the *Theory of Event Coding* (Hommel, Müsseler, Aschersleben, & Prinz, 2001b), which is described in section 2.2.3.

Effects of compatibility have not only been shown in S-R relations (S-A relations, in our model), but also in A-E relations. In a number of studies it was shown that compatibility (i.e., dimensional overlap) of an action and a forthcoming effect speeds up response times, even when the action effect is irrelevant for the task (Koch & Kunde, 2002; Kunde, 2001, 2003). This could be demonstrated for the perceptual dimensions of location, force/auditory intensity (Kunde, 2001), color (Koch & Kunde, 2002), duration (Kunde, 2003), as well as for compatibility on a purely conceptual level (Koch & Kunde, 2002). These effects were decidedly *not* due to learned action-effect associations, as control experiments showed.

The effects of involuntary imitation in the studies of Brass and Craighero may be viewed as a mixture of both groups: Images of body movements are usually the visual effects of actions. If these serve as pre-action stimuli, as in the experiments, they have a facilitating or interfering influence on action performance: they are involuntarily imitated, as if the actor wishes to produce the perceived movements (or end-states of movements) as action effects.

2.1.3 Learned Sensory-Motor Relations

One may hypothesize that the majority of effects of sensory-motor interaction derive from underlying "hard-wired" connections in the cognitive system. This is most probably the case for the effects in the studies described above. However, it has been observed that effects of sensory-motor interaction can be modulated by specific experience, and that even novel effects can be evoked by learning. Some of these findings are reviewed in the following.

If we follow the S-A-E model, there can be two kinds of learned sensory-motor relations: learned S-A (S-R) relations and learned A-E relations. S-R linking develops, for example, when an "unconditional" stimulus (US), which per se is linked ("hard-wired") to a certain response, is presented repeatedly along with a "conditional" stimulus (CS) ("classical conditioning", see Domjan, 1998). In the classical example, the US is food, which usually leads to the response of salivating, while the CS is a tone. Eventually, the CS alone (the tone) will evoke the same response (salivating, i.e., getting hungry) as the US (the food). Thus, classical conditioning provides the link to non-learned sensory-motor relations, like the reflexes (see above).

S-A/S-R learning can also develop in more intrinsically controlled settings. For example, in any task that involves responses, which follow certain arbitrary S-R rules, people have to *learn* these S-R rules (which, together, compose a so-called *task set*). As such, the linking of stimuli to responses is at first a consciously controlled mapping process, which then becomes increasingly automated (see Nissen & Bullemer, 1987). Eventually, an intrinsically controlled *action* will turn into an extrinsically controlled *response* – a S-R link has formed.

In the context of the current study, the more interesting kind of learned sensory-motor relations is the linking of actions and action effects (A-E linking). Here, people learn to relate body movements to ensuing sensory events. One may speak of a specific A-E linking, when it involves the linking of actions to *distal* action effects. For a given movement (say, the tapping of a finger on a hard surface), its ensuing proximal effects are invariant. However, the resulting distal effect(s) will vary, depending on the current context: It may be a light going on or off, if the movement is carried out on a light switch; it may be a letter appearing on a screen, if it is carried out on a computer keyboard; it may be a piano tone, if the movement it is carried out on a piano; or it may be none at all.

Following their own empirical results, Elsner and Hommel (2001; 2002) have put forward a model for such action-effect linking. This model is also concerned with the question of how such linking may become relevant in the control of goal-directed action. In their study, participants learned to associate left and right keypresses with tones of high and low pitch. Each tone appeared as a contingent effect of a free-choice keypress. This ‘acquisition phase’ (200 trials) was followed by a ‘test phase’, in which the same tones were used as imperative stimuli, requiring a left or a right keypress. Two groups were compared in the test phase: In the non-reversal group, the tone-key mapping was consistent with that of the acquisition phase, while in the reversal group the mapping was inconsistent. Results showed that response times in the non-reversal group were significantly faster than in the reversal group. In a second experiment, the acquisition phase was similar to Experiment 1, while in the test phase participants made free-choice keypresses (left or right) in response to tones of high and low pitch. As the results showed, participants preferably chose the key, which had produced that same tone in the acquisition phase.

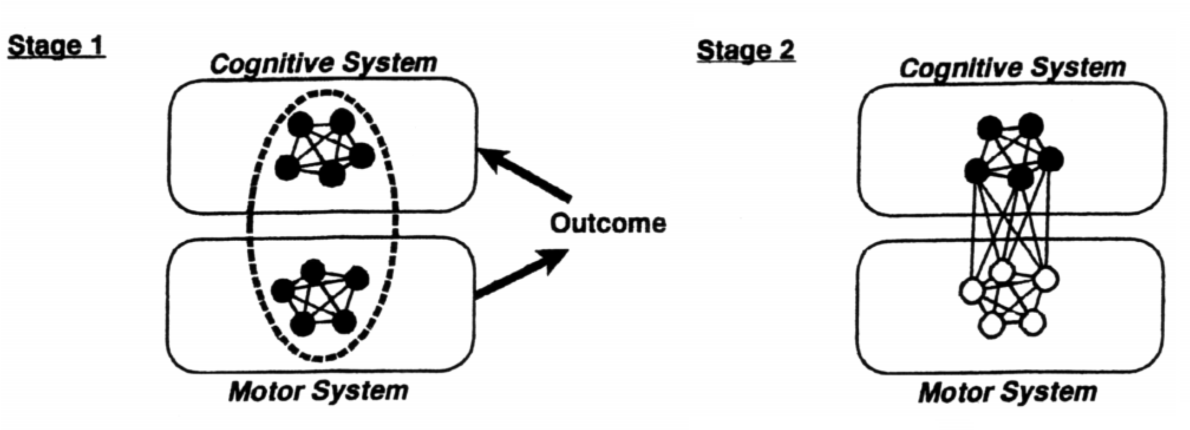


Figure 1. The two-stage model of the acquisition of voluntary action by Elsner and Hommel (2001). In stage 1, motor codes and the perceived sensory outcome (the “action effect”) are connected. At stage 2, the imagination of a potential sensory outcome automatically activates the associated motor code. (Figure adopted from Elsner & Hommel, 2001, p. 230)

The theoretical model of Elsner and Hommel (2001) is based on these results and related data (e.g., Greenwald, 1970a, 1970b; Hommel, 1996), and also draws upon previous theories (Harless, 1861; Lotze, 1852). The model consists of two stages (see Figure 1). In stage 1, sensory and motor codes are connected through learning. For this to happen it is necessary that certain actions (e.g., finger movements) are experienced repeatedly along with certain effects (e.g., tones). The effect must be perceived as being *contiguous* (i.e., close-by in time and space) and *contingent* (i.e., occur in above random frequency) in relation to the action (Elsner & Hommel, 2004). This is in accordance with the principles of associative learning (see section 2.2.1). As a result of this repeated experience, a *bi-directional* association between the action code and the effect code is formed. It is bi-directional, because the activation of either element (the action code or the effect code) will activate the other.

The bi-directional association between action code and effect code can, according to the model, be applied for the voluntary control of behavior. In stage 2, action-effect associations are employed for the selection of appropriate movements in order to reach an intended action goal. The imagination of an intended movement goal (i.e., an intended effect) will activate an action code, which has previously been associated with it. To quote Elsner and Hommel, "…movements are selected by anticipating (i.e., activating the codes of) their consequences" (Elsner & Hommel, 2001, p. 230). It should be noted that such a mechanism only applies for actions, which are not entirely new to the actor, that is, it only applies to actions, in which the actor has a certain amount of "expertise".

In order to examine the brain regions that are involved in the linking of actions and action effects, Elsner and Hommel (2002) have conducted a similar study with positron emission tomography (PET). First, their participants learned to associate self-initiated keypresses with tones. In a test phase, they were then presented tone sequences, which consisted of neutral tones and of the previously learned tones in various ratios. When the frequency of action-effect tones increased, there was increased activity in the caudal supplementary area (SMA) and the right hippocampus. Accordingly, these regions seem to play a crucial part in action-effect coupling. This is in accordance with studies, which have demonstrated the involvement of the SMA in the acquisition of visuomotor associations and motor sequences (Jenkins, Brooks, Nixon, Frackowiak, & Passingham, 1994; Sakai et al., 1999; Toni, Krams, & Passingham, 1997), and the involvement of the hippocampus in associative learning (Henke, Buck, Weber, & Wieser, 1997; Wise & Murray, 1999).

Ziessler and Nattkemper (2002, Experiments 1 and 2) have reported similar effects as Elsner and Hommel, albeit with visual action effects. In an acquisition phase, participants carried out a four-choice response task (keypresses in response to visually presented letters). Each keypress was followed by a contingent visual effect: a letter (taken from the set of stimuli) appeared on the screen. In the subsequent test phase, the letters that previously appeared as effects were used as flankers in the same four-choice task. Results showed an effect of facilitation when the flanker and response were congruent with the previously learned R-E mapping.

A study by Ziessler and colleagues (Ziessler, Nattkemper, & Frensch, 2004) has provided further insight into A-E learning. The study showed that A-E learning is considerably impaired, when a distractor is presented during response preparation (Experiment 1). Ziessler et al. interpret this impairment as evidence for effect anticipation during response preparation. Furthermore, they found evidence that A-E learning can be modulated by intrinsic, volitional factors (similar as in S-R learning, see above). If an action had *two* effects, and participants were instructed to produce *one* of them (Experiment 2), A-E learning only occurred for the instructed (the "intended") effect.

Recently, Hoffmann (2004) has brought forward preliminary evidence, which, in a way, questions the idea of a static A-E linking in the sense of Elsner and Hommel. He conducted an experiment, which was similar to the one of Elsner and Hommel, except that responses were aligned vertically. Similar effects were obtained as in the study of Elsner and Hommel. Then, however, the hand-to-key mapping in the experiment was reversed – and the pattern of results did not change. That is, when in the test phase the upper key was now pressed by the left index finger instead of the right, it was still pressed more often when the tone that corresponded to this key was presented. It seems that participants had not learned associations of specific limb movements and effects, but rather of *response locations* and effects. An implication of this finding may be that specific movements can be dynamically assigned to fit certain targets (response locations, movements goals,…). This is in accordance with findings, which suggest that in the motor and premotor areas of the brain movements are primarily represented in terms of their spatial end-states, which can be attained by a variety of possible movements (Graziano, Taylor, & Moore, 2002).

The studies reported so far were concerned with rather simple actions. But action-effect learning also applies to more complex behavior. Ziessler and Nattkemper (2001) were able to demonstrate that in the learning of event sequences response-effect learning is an important constituent (see also Ziessler, 1994, 1998). In their variant of the serial reaction task (SRT) paradigm, S-S relations and R-S relations were systematically varied, while R-R relations were held constant. Changes in the complexity of the R-S relations had the strongest effect on serial learning. It seems that participants perceived stimuli, which followed a response, as *effects*. In the same vein, the results reported by Hoffmann, Sebald and Stocker (2001) and Stocker, Sebald and Hoffmann (2003) showed that serial learning is considerably improved, when responses are followed by contingent (and contiguous) tone effects. It seems that when movements are followed by tone effects, the internal representation of movements into chunks is facilitated (Stocker & Hoffmann, 2004).

An interesting study by Eenshuistra, Weidema and Hommel (2004) has examined developmental aspects of action-effect learning. Eenshuistra and colleagues could demonstrate that the susceptibility to action induction via learned action-effect associations notably decreases between the age of 4 and 7. One might assume that somewhere at this age children learn to employ integrated action-effect representations in a more intrinsically controlled way, for instance for the voluntary control of movements in the sense of Elsner and Hommel (2001).

Grosjean and Mordkoff (2002) have provided evidence that the size of the Simon effect can be altered (decreased or increased) by the inclusion of post-response stimuli (that is, action effects). This finding is quite interesting, because it shows that processes, which derive from hard-wired sensory-motor coupling (as the Simon effect), can interact with processes, which are not hard-wired. Similarly, Beckers et al. (2002) have found evidence for a learned integration of affective action effect features (namely the positive or negative valence of the effect) in the representation of actions in a variant of the associative Simon task.

If we follow most of the above studies, it becomes clear that there are two possible means of activating the representation of an action code via action-effect coupling: perception and imagination. Obviously, action activation by *perception* is usually not willed by the actor. One may think of exceptions when, coincidentally, the actor intends to perform that same action, or when the actor has voluntarily chosen to expose himself to a certain “stimulating” environment. But usually, the perception of a sensory event, which is a possible effect of an action, will lead to an *involuntary* activation of the action code, as in the studies on involuntary imitation.

For the alleged control of voluntary action, as in the model of Elsner and Hommel (2001), *imagination* of the effect is the method of choice. In this case, the effect code is activated in an “act of the will” in the Schopenhauer sense (see the introductory quotation). Even though the focus of the present study is on the issue of *whether* and *how* people learn to integrate actions and their perceivable effects, it should be noted that there is increasing evidence that supports the idea of a pre-action imagination of forthcoming effects (Kunde, Hoffmann, &

Zellmann, 2002; Kunde, Koch, & Hoffmann, 2004; Ziessler & Nattkemper, 2002). These and other studies are rather interested in the question of the possible *role* of action-effect coupling in human behavior – which is also a major aspect of the theories of sensory-motor coupling, which are described in the following.

2.2 Theoretical Accounts of Sensory-Motor Coupling

Not surprisingly, the extensive empirical data on sensory-motor coupling, some of which has been described above, has entailed attempts to theoretically account for such phenomena. Some theoretical considerations were already mentioned in connection with the related empirical studies – and some of the studies described above were based on theories described below. All theories that try to account for phenomena of sensory-motor interaction purport a close structural and/or functional coupling of perception and action. Denis J. Glencross concurs "…that there is no hard and fast distinction between organization of the sensory and motor systems and that indeed perception and action are tightly coupled" (Glencross, 1995, p. 6), and that "sensory-motor integration is the centre piece of perception-action coupling" (Glencross, 1995, p. 7).

However, the questions of *how*, *where* and *why* such coupling of perception and action may take place is a matter of dispute. Three major lines of theory may be distinguished, which are described in more detail in the following: associative learning theory, computational theory, and ideomotor theory. *Associative learning theory* describes sensor-motor learning as an inevitable forming of associations between concurrently active (sensory and/or motor) representational elements, which may be essential for various functions of the cognitive systems. *Computational theories* describe sensory-motor coupling as a mathematically describable translation of sensory states into motor states (and vice versa), and stress its role in motor control. *Ideomotor theory* (esp. in its most recent form as the ‘Theory of Event Coding’) presupposes a common representational domain for sensory and motor codes, which is hard-wired in the system, but may also adapt through learning. Also, the most important function of sensory-motor coupling is motor control.

2.2.1 Associative Learning Theory

The most straightforward account as to how the coupling of sensory and motor representations might come to pass is *associative learning*. Associative learning theory has derived mostly from classical learning studies (see Domjan, 1998). The central concept of associative learning is that “…the central representations of specified elements can become linked so that activation of one can excite its associate” (Hall, 1991, p. 1). That is, associative learning refers to the bi-directional linking of two cognitive event representations, or, more precisely, the *formation* and subsequent *strengthening* of such associative links.

Hall (1991) has described the essential elements of most theories of associative learning. First, the representations of the events need to be active *simultaneously*. That is, if there is a temporal gap between two (e.g., visual) events, it must not exceed a certain threshold, so that their activation still overlaps. Second, associations are formed between the representations of *events*. Here it should be noted that, traditionally, associative learning has been presumed to be constricted to environmental stimuli (e.g., Roitblat, 1987). However, more recent accounts (Hall, 1996; Pearce, 1997) describe associative learning rather generally as a linking between events – which include *sensory* events, *motor* events, but also memory representations. Third, a connection between elements may be *excitatory* or *inhibitory*. If it is excitatory, activation of the target element increases, if it is inhibitory, activation of the target element decreases. However, not all theories agree with this. Most classical theories only allow for excitatory connections.

A formal description of associative learning is given in the so-called Rescorla-Wagner model (Rescorla & Wagner, 1972). The model describes how the associative strength V between two stimuli (CS and US, see section 2.1.3) increases when the two events appear in contiguity. The change of associative strength within a "conditioning trial" (i.e., when both events are concurrently activated) is given by the following equation:

$$\Delta V_{\mathrm{A}} = \alpha_{\mathrm{A}} \beta (\lambda - \Sigma V),$$

where ΔV_{A} is the increment in associative strength occurring to a given stimulus A (the CS), α_{A} is a learning rate parameter which depends on the intensity, discriminability, or salience of A, β is a learning parameter determined by the nature of the US, λ is the asymptote of associative strength, and ΣV is the summed associative strength of all CSs that are present on that trial and have associations with the US representation. In this model, learning proceeds until $(\lambda - \Sigma V) = 0$, that is, until the asymptote of maximal associative strength is reached. The Rescorla-Wagner model, which is actually quite straightforward and simple, has proven to be very powerful in explaining a wide range of phenomena (see Hall, 1991, chapter 1).

Action-effect coupling may be viewed as a somewhat specific kind of learning, since it involves the learning of causal relationships: an actor learns that his actions (allegedly) *cause* certain proximal and distal effects. Dickinson (2001) has compared the capacity of associative learning theory and computational models in explaining the learning of causal relationships. He comes to the conclusion that a modified version of the standard associative learning theory (see above) can account for causal learning quite well. However, such a model would have to include the capacity to activate elements by memory retrieval (that is, when no external stimulus is present). This is quite similar to the Elsner and Hommel (2001) model, which assumes that an effect representation can be activated by mere imagination (or "anticipation") of a potential sensory effect.

Associative learning is supposed to occur inevitably when two event representations are activated concurrently. The result is an associative network, which is the foundation for a wide range of potential functioning. For example, Heyes (see Heyes, in press) has described how sensory-motor associations may account for effects of involuntary imitation, as they were described in the studies of Brass and Craighero (see section 2.1.1). It seems that such ideas become increasingly popular – again, as must be added. It should be remembered that in the age of behaviorism (that is, from the beginning of the 20th century until into the 1960s) associative learning was supposed to explain *all* human behavior, and that since the "cognitive turn" such ideas have not been very popular, and were, at most, restricted to animal research. The groundwork for a lot of phenomena that are examined today (see, for example, the above section on learned sensory-motor coupling) has already been laid by former behaviorists. For example, the idea of an anticipation of a desired effect (a "reinforcing event"), which is grounded on a R-C association, has already been described by Hull in the 1930s. He proposed that the anticipation already includes something of a reward, a phenomenon that he has termed the "fractional anticipatory goal response" (Hull, 1930, 1931; Spence, 1956).

2.2.2 Computational Models

Like associative learning theory, computational models offer a learning account of sensory-motor coupling. Here, the relations between sensory and motor representations are understood as mathematically describable translation processes. Hence, compuational models are often used for quantitative predictions, and are generally more complex than accounts of associative learning. In the following, I will first give a picture of a typical example of computational models, however without going into the exact details, and then describe more generally how computational models account for sensory-motor coupling.

Ghahramani, Wolpert and Jordan (1997) have described a typical computational model of sensory-motor coupling, or “integration”. According to Ghahramani et al., the sensory-motor integration system can be viewed as “an observer attempting to estimate its own state and the state of the environment by integrating multiple sources of information” (Ghahramani et al., 1997, p. 117). In their view, the most important aspect of sensory-motor integration is the reduction of uncertainty in sensory estimates. For example, the use of multiple sensory modalities (e.g., visual and auditory) may increase the speed and accuracy of reaching movements. All these information must be integrated into a consistent whole, in order to enable an effectual motor system. The model focuses on a formalized description of the integration of sensory modalities, for example of visual and auditory spatial maps. In order to transform, say, two sources X and Y into a common representation, the system must filter information, which is common to both modalities, while it must reject that which is not. In a first step, the signals are transformed into a common coordinate frame. The error noise present in the single measurements is then reduced with a maximum likelihood estimate, assuming that errors are uncorrelated and follow a Gaussian distribution. With the inclusion of a Kalman filter (Kalman & Bucy, 1961), the model is then extended to a dynamical system. With the Kalman filter, each state estimate is calculated from the previous state. The model is also capable of learning, for example, when one source exhibits a constant bias.

In the context of the present study, the most interesting question is how sensory and motor codes are connected. According to Ghahramani et al. (1997), the Kalman filter incorporates an *internal model* of the system dynamics. The notion of internal models has proved to be popular in numerous computational approaches to sensory-motor coupling (Blakemore, Goodbody, & Wolpert, 1998; Jordan & Rumelhart, 1992; Sutton & Barto, 1981; Wolpert & Flanagan, 2001; Wolpert & Kawato, 1998). There are two kinds of internal models: forward models and inverse models (Wolpert & Kawato, 1998). A *forward model* calculates a future sensory state at time $t + 1$ (output), given a current sensory state at time t and motor command (input). For example, its input might be the current position of the arm and an efference copy of an issued motor command, while its output would be the resulting end position of the arm after the movement. Forward models therefore predict the effects of an action. An *inverse model* incorporates the opposite case: it calculates a motor command (output), given a current sensory state at time t and a future sensory state at time $t + 1$ (input). Accordingly, inverse models are particularly suited for motor control, because they estimate motor commands on the basis of desired effects.

Wolpert and Kawato (1998) have proposed a combination of multiple paired forward and inverse models as a possible functional structure for motor control (see also Wolpert, Miall, & Kawato, 1998). Figure 2 depicts a single module within the multiple paired internal model. Each of the n modules incorporates the learned sensory-motor relations for one particular ‘context’ (examples for contexts might include walking, gaze direction, or piano playing). The modules can also be combined for use in novel situations. Thus there is no single controller, which has to adapt to each context change. How does the system know which module to use in a given context? This is administered by a responsibility signal, which is computed from a contextual signal and the prediction of a forward model (see Figure 2, left half). Motor commands are generated by an inverse model, which receives the desired movement trajectory as input, and learns from the motor feedback and the responsibility signal.

Two extremes can be distinguished in the representation of internal models in motor memory (Wolpert, Ghahramani, & Flanagan, 2001): lookup tables and parametric representations. *Lookup tables* store the output for each possible single input setting. Accordingly, lookup tables are prone to require a lot of memory. On the other hand, *parametric representations* of internal models work like an equation. They generalize globally to changes in the parameters (e.g., limb angles). Internal models can also be represented in between those two extremes. These are mappings that generalize within a limited region of the input space.

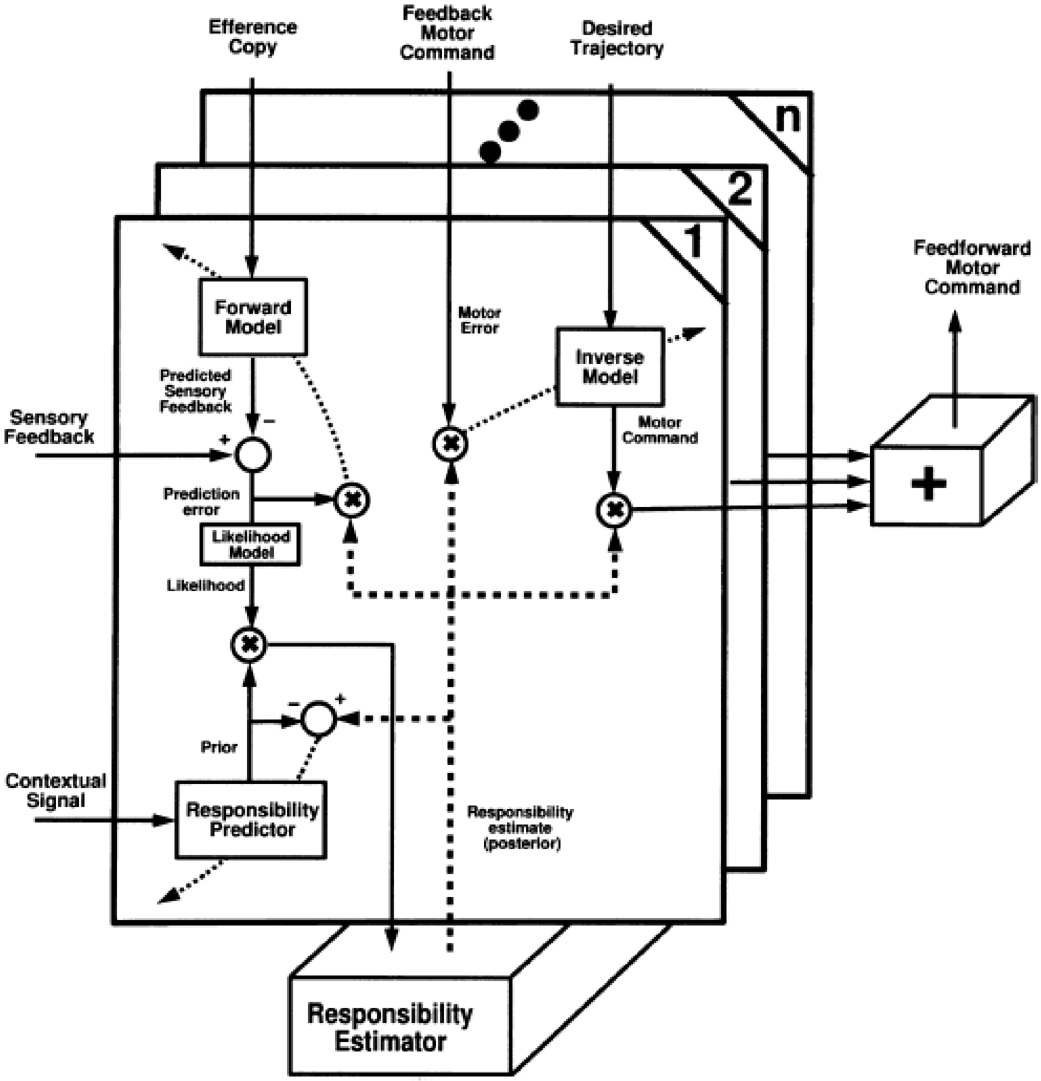


Figure 2. A single module within the multiple paired internal model. Dotted lines represent training signals, × represents signal multiplication. A module consists of three interacting parts. The first two parts (the forward model and the responsibility predictor) determine the responsibility of the module (= the responsibility signal). The smaller the prediction error of the forward model, the more likely the sensory feedback and efference copy are consistent with the context captured by the forward model, and hence the higher the module's responsibility. The higher the prediction error and the responsibility, the more learning occurs. The third part of the module is the controller. It generates a motor command, given a desired trajectory. This is accomplished by an inverse model. (Figure adopted from Wolpert & Kawato, 1998, p. 1325)

Even though internal models are basically computational models, they can be implemented in neural networks (Botvinick & Plaut, 2004; Salinas & Abbott, 1995). And, although the internal models approach derives mainly from computational and neural networks research, there is also neurophysiological evidence. Wolpert and colleagues (see Wolpert et al., 1998) have reviewed the relevant literature, and come to the conclusion that the cerebellum is a most likely candidate to contain an internal model (or models) of the motor apparatus. Other studies have discussed the implementation of internal models in the supra-spinal system for arm movements (Bushan & Shadmehr, 1999), and the role of the prefrontal cortex in internal model learning (Shadmehr & Holcomb, 1997).

2.2.3 Ideomotor Theory

In their Theory of Event Coding (TEC), Hommel, Müsseler, Aschersleben and Prinz (2001b) have outlined a model of sensory-motor integration, which draws on (and also tries to explain) a large amount of empirical research. Most of the studies that have examined the role of action effects (see section 2.1.3) are in support of this theory.

The central tenet of the TEC is an idea, which had already been forwarded by a couple of 19th century philosophers (Carpenter, 1852; Harless, 1861; Herbart, 1816; James, 1890; Laycock, 1840; Lotze, 1852). The idea was rediscovered by Anthony Greenwald in the 1970s (Greenwald, 1970b), and has undergone a lot of theoretical refinement and empirical promotion since then (esp. Hommel et al., 2001b; Prinz, 1990; Prinz, 1997). This idea is called the *ideomotor principle*. A concise overview of the history of the ideomotor principle can be found in a paper by Stock and Stock (2004). The ideomotor principle maintains that the initiation of a movement has its origin in the mental imagination of the movement's sensory feedback. This is in contrast to a *sensorimotor* view on action control, which ascribes the initiation of a movement to external stimuli. For example, according to the ideomotor principle a singer must only imagine the correct pitch in order to produce it (see James, 1890). There is no need to directly control parameters of the movement, like larynx elevation or breathing. Given sufficient experience, these parameters are adjusted automatically.

In the TEC, the ideomotor theory of voluntary action is integrated with a conception of sensory-motor coupling, which has come to be called the *common coding* approach of perception and action (Prinz, 1990). The central conjecture of this conception is that, on some level of the cognitive system, sensory and motor codes share one representational domain (see also Prinz, 1997). That is, on some structural level of the cognitive system sensory and motor codes are not represented as distinct entities, but in one *common* code. As such, these representations are neither purely sensory nor purely motor codes, but representations of *events*. An event can either be *perceived* (the input into the common representation then comes from a sensory channel), or a person can have the intention to *produce* the event himself (then it must be transformed into a motor command). Here, the relation to the ideomotor principle becomes obvious. If the event representation comprises both sensory and potential motor events, then the imagination (the "idea") of a desired event (an intended goal) can (and, to some extent, involuntarily *will*) directly initiate an appropriate motor event. This is exactly what is meant with ideo-motor control of action.

It is important to note that commonly represented events are always *distal* events. That is, they are events that occur in the "outside" world, and are not contained within, say, body movements. In experiments, such distal events may be a light going on or off, or a "beep" sound. Outside the lab, one may think of the direction where to one steers the own car, or of the tones of a musical instrument. According to the TEC, these distal events are represented as composites of elementary *features*. Those features, which are part of a (perceived or to-be-produced) event, are weighted and temporarily conjoined by a binding mechanism. Features might be rather low-level, like relative position (RIGHT, LEFT,…), shape (LARGE, SMALL, RECTANGULAR,…), or color, but they might also include higher-level features, like time and change.

The coding of events in terms of their distal features is central to the explanation of many effects of sensory-motor interaction. Take the Simon effect, for example. Experiments have shown that there is a Simon effect independent of whether the responses (left or right) are executed normally or with hands crossed (e.g., Wallace, 1971). If one thinks of the response location as a distal code, this effect can be easily explained, because the feature for location (LEFT or RIGHT) does not change when one crosses hands. But why is there a Simon effect in the first place? According to the TEC, the perceived stimuli, as well as the planned responses, are represented as distal events in the common domain. If the location feature of the perceived stimulus has the content RIGHT, this will at the same time activate an involuntary *response* tendency with the feature RIGHT. If the location feature of the planned response has the content LEFT, there is a conflict: the response will be executed with a delay – and sometimes even with an error (here, a 'left' response).

In the present theoretical overview, the 'Theory of Event Coding' has been included as the most elaborated and most typical example of a theory that ascribes sensory-motor coupling to *structural* characteristics of the cognitive system. It should be noted that other theories have expressed similar conceptions. For example, the *ecological theory* by James Gibson (see Gibson, 1979) illustrates how the visual system may be grounded on the motor system:

according to this theory, the visual scenery is perceived according as to how the perceiving person can possibly *interact* with it. A chair, then, is perceived as a possible object for the action of *sitting*. In the terminology of the ecological theory, the chair has a so-called “affordance” for the action of sitting (Gibson speaks of the “sit-able-ness” of the chair). The affordances of the environment directly activate according response tendencies within the perceiving person. In contrast to the TEC, which tries to give an integrated view of perception and action, the ecological theory is mainly an account of perception. As such, both theories are not necessarily in contradiction. Hommel and colleagues have argued that ecological theories emphasize the *What*, while theories like the TEC emphasize the *How* of information use (see Hommel, Müsseler, Aschersleben, & Prinz, 2001a).

Finally, it should be noted that structural theories like the TEC are not necessarily in contradiction to accounts that ascribe sensory-motor coupling to learning, like the associative and computational models described above. In fact, these approaches may even profit from each other: theories like the TEC provide the framework for sensory-motor coupling which is hard-wired and structurally immanent to the system, while theories of sensory-motor learning may account for phenomena of environment adaptation and skill acquisition, which have their foundation in specific experience. Alternatively, theories like the TEC might be understood as describing the state of the system, while other theories describe how this state is acquired.

2.3 Sensory-Motor Coupling in Music

In the above sections, a general account (empirically and theoretically) of sensory-motor coupling has been given. Most of the empirical data in the above studies had been obtained from people, who had no specific training in the examined domain whatsoever. Accounts of learned sensory-motor integration (especially of action-effect learning) have thus derived from explicitly implemented learning phases within the experiments. Many of these studies have shown that it is possible to establish effects of sensory-motor coupling within a relatively short time (Eenshuistra et al., 2004; Elsner & Hommel, 2001; Ziessler & Nattkemper, 2002; Ziessler et al., 2004). But what about people, who have been subject to such specific learning over months, years – or even a whole lifetime?

The present study is an examination of learned sensory-motor integration in such a specific group of people: musicians. Why choose musicians? There are several reasons. First, musicians usually have year-long training on their respective instrument. Evidently, the expertise of a musician cannot easily be compared to situations where a person has learned the performance of a task in a single session of maybe half an hour. It is simply not *possible* to learn to play the piano after half an hour. Second, this year-long training is usually *explicit* and *deliberate*. That is, it takes place in a controlled setting, has well-defined learning goals, and incorporates explicit feedback on the learning progress. And third, instrument playing allows for a clear distinction of sensory and motor events. *Motor* events are all the movements that are carried out in relation to the instrument (e.g., finger movements on the piano). *Sensory* events are, especially, the distal auditory effects of these movements (the sounds of the instrument).

In the following, it is first described what, from the perspective of cognitive science, makes musicians different to other people, and how such differences have developed. This is mainly an account of *expertise*. Afterwards follows a review of previous studies that have examined sensory-motor integration in musicians. It will show that, for the most part, these were studies that have used neurophysiological methodology.

2.3.1 Constituents of Musical Expertise

What makes a musician different from other, from 'normal' people? Let us first look at the phenomenon of expertise in general. Expertise is usually considered as superior performance ability in a specific domain, be it in the arts, sports, games, or other fields (see Ericsson, 1999). Such superior performance ability has been attributed to the experts' capability to represent a large amount of complex domain-specific (declarative and procedural) knowledge patterns in terms of informational chunks (H. A. Simon & Chase, 1973). Note that the dichotomy of experts versus non-experts is, in a way, artificial. It is generally agreed that expertise forms a continuum of varying grades between novices and experts. Furthermore, expertise is domain-specific. Probably all people have some degree of expertise in certain domains, be it the own job, sports, or one's own language (Sloboda, 1991).

Musicians belong to a specific subgroup of experts: they are *motor experts*. Simply spoken, a motor expert is "someone who's very good at doing something motoric" (Starkes, 1993, p. 3). Obviously, in musicians this "something motoric" pertains to the specific movements that handle the musical instrument. Take, for example, piano playing. If we follow the classification of motor skills by Magill (2001), piano playing involves movements that are *fine* (they require precise control of small muscles), *continuous* (they are repetitive movements with arbitrary beginning and end points), and *closed* (they are performed in a stable and predictable environment). As has already been mentioned above, these movements in musicians are not self-contained, but serve the goal to produce musical sounds. Accordingly, perhaps, it would be more appropriate not to speak of musicians in terms of motor experts, but in terms of "sensory-motor" experts.

A controversial issue concerning the foundation of expertise is the respective import of *practice*, as specific ongoing training and experience, and innate *talent*, as defining the starting point, possible speed, and ultimate boundary for learning. Most often, and this applies especially to music, non-scientific folk psychology attributes outstanding ability and creative experience to talent (see Sloboda, Davidson, & Howe, 1994). Such beliefs are promoted by legends of child prodigies (Mozart, Menuhin, Arrau, etc.), whose abilities were seemingly present from birth. However, the scientific research on expertise has come to a quite contrary conclusion: Expertise is the product of *specific practice and learning* (Charness, Krampe, & Mayr, 1996; Ericsson, Krampe, & Tesch-Römer, 1993; Howe, Davidson, Moore, & Sloboda, 1995; Sloboda, 1996; Starkes, Deakin, Allard, Hodges, & Hayes, 1996).

Karl Anders Ericsson and colleagues, for example, found that the critical difference between expert musicians who differed in the level of attained solo performance could be ascribed to the amounts of time they had spent in solitary practice during their music development (Ericsson et al., 1993). At the age of twenty, time of practice totaled around ten thousand hours for the best professionals, around five thousand hours for the least-accomplished professionals, and only two thousand hours for serious amateur pianists. As a general rule it can be said that, in order to reach a high level of expertise, at least ten years of deliberate practice are necessary (Ericsson, 1996; Ericsson et al., 1993; H. A. Simon & Chase, 1973). Besides the time span, it is this *deliberate* practice, which seems to be of specific import. The more controlled certain factors are in the learning situation, the better the learning results. In their thorough review of skill learning studies, Ericsson et al. (1993) come to the conclusion that effective learning requires a well-defined task with an appropriate difficulty level for the particular individual, informative feedback, and opportunities for repetition and correction of errors (see also Lehmann, 1997).

Expertise in music, then, is a product of *specific learning*. But what constitutes this learning? What has changed in the cognitive system of a musician throughout his learning history? In their three-stage model of motor skill learning, Fitts and Posner (1967) describe how sensory-motor information processing is altered during learning. The initial stage of skill acquisition, the *cognitive stage*, is characterized by a direct control of movements (limb positions, etc.), a conscious processing of feedback, and a large number of errors. Gradually, this stage turns into the *associative stage*. Now the person has learned to associate certain environmental cues with the movements that are required to achieve the goals of the skill. Progress within this stage is characterized by a “refining” of the skill. Error rate and performance variability decreases. Eventually, the learner reaches the so-called *autonomous* or *automatic stage*. Now, the skill has become almost automatic, or habitual. The skill (like piano playing) can be performed without conscious control of the movements. Sensory and motor information are accurately related. One may now speak of actual “expertise” in the skill.

Two points should be noted about this widely acknowledged model. First, people have to learn the execution of the movements themselves, in order to properly perform the skill. For example, the (bimanual) finger movements in piano playing are quite specific, and are usually not contained within the movement repertoire of a beginner. In this movement learning, deliberate practice is most important (Lehmann, 1997). Second, the model not only describes the development of the movements, but also how these movements relate to the perceived environment. That is, the pianist learns his movements in relation to the keyboard, the piano, and the resulting sounds. All this is integrated into a consistent representation in a gradual learning process – in a process of *sensory-motor integration*.

This premise, namely that the acquisition of expertise in music performance is a process of sensory-motor integration, also follows, implicitly or explicitly, from most theories on motor control. For example, *closed-loop* models of motor learning stress the role of sensory feedback in the movement control process (e.g., Adams, 1971). Here, motor learning is described as a transition from the reliance on externally-provided feedback to internally-derived feedback. This part of the theory is closely related to the two-stage model of Elsner and Hommel (2001, see above). In contrast, *open-loop* models of motor control (where sensory feedback is irrelevant) can hardly account for such complex movements of musicians, as when they perform a piece of music. At best, open-loop movement control might become relevant in very fast movement sequences, as in trills or arpeggios (Palmer, 1997). Other theories that emphasize the role of sensory information in motor control are *motor programming* theories (e.g., Schmidt, 1975, however, note that this is not typical for most motor programming theories), and *dynamical systems* theories (Erlhagen & Schöner, 2002; Kelso, 1995; Turvey & Carello, 1995). The latter also provide quantitative accounts on how sensory and motor information are possibly conjoined. For example, Camurri (1997) has described how a particular kind of dynamical systems, a *multimodal environment*, can serve as a model for motor control and music.

To sum up, expertise in musicians may be viewed as an acquired skill that has developed in year-long training, and that, from a cognitive perspective, necessarily involves an increased coupling of sensory and motor processes. Complex movements are learned in relation to resulting sounds of the instrument. This latter point has, for instance, been explicitly stated by Baker (2001), who writes that guitarists "are likely to associate heard chords with particular finger configurations on their instrument" (Baker, 2001, p. 251). In the following, I will discuss the existing empirical evidence for such sensory-motor coupling in musicians.

2.3.2 Empirical Evidence

There are some studies, though not very many, that have empirically investigated sensory-motor coupling in the domain of music. Obviously, most effects of sensory-motor integration in music are not hard-wired, but learned. There may be a few exceptions. For example, there is evidence that the perception and production of rhythm (or, more precisely, *meter*) are closely connected to the functioning of the motor system (see Clarke, 1999). For instance, a study by Gabrielson using factor analysis and multidimensional scaling showed that listeners rated perceived rhythms (in similarity ratings and descriptive adjective ratings) primarily in terms of their 'movement character' rather than in structural or emotional dimensions (Gabrielson, 1973). In fact, it has been proposed that quite generally the temporal control in behavior might be regulated by some kind of internal clock (e.g., Luce, 1972), which would make meter a necessary constituent of movement itself. Also, think of the close relations of rhythm and movement in dancing – here, almost everybody can experience effects of "action induction" through music. However, besides such a presumably *universal* coupling of rhythm and movement, there is also evidence for *specifically* learned sensory-motor coupling in musicians themselves.

First empirical evidence for sensory-motor coupling in musicians was brought forward in 1975 by Eugene Holdsworth in an (otherwise unpublished) dissertation (see Holdsworth, 1975). He posed the following question: "As individuals see and hear musical stimuli for which they are known to possess performance skills, can covert bioelectrical neuromuscular activity be observed in the musculature associated with the performance of those skills?" (Holdsworth, 1975, paragraph 1). Holdsworth used electromyography (EMG) to measure covert bioelectrical neuromuscular activity in eighteen trumpet players, while presenting to them melodic excerpts in different conditions. Results showed significant EMG activity in the according arm musculature. There were no overt responses, and mostly participants were not aware of the covert muscular activity.

In a more recent study with pianists, Haueisen and Knösche (2001) have demonstrated quite similar effects. They used magnetencephalography (MEG) to measure motor activity in the primary motor cortex (M1). In their experiment, pianists and non-pianists were presented monophonic piano sequences of well-known piano pieces. Their task was to detect a certain piece of music, and press a button when this piece contained a wrong note. The MEG measures during the task revealed an involvement of the M1 in the perception of the piano sequences. Even more, the individual contralateral M1 areas of thumb and little finger were activated specifically, when the presented sequence would have required those fingers in its performance. This activation was only present in pianists, and not in non-pianists. According to Haueisen and Knösche, this M1 activation was *involuntary*, because the participants' attention was focused primarily on the detection task. Also, there was no activation of the supplementary motor area (SMA) and the premotor cortex (PMC), which was probably owing to the fact that, due to the pianists' high expertise, *automatized* processes were predominant in the motor and premotor areas.

The reverse case, that is, the activation of auditory representations through movement, has been reported as well. In a study with functional magnetic resonance imaging (fMRI) by Scheler and colleagues (Scheler et al., 2001), eight professional violinists and eight amateurs tapped out the first 16 bars of a Mozart violin concerto. The expert performers revealed significant activity in primary auditory regions, which was not present in the amateur group.

The studies reported so far provide no within-group evidence that the effects of sensory-motor coupling that they describe have derived from specific musical experience. In a longitudinal learning study using electroencephalography (EEG) methodology, Bangert and Altenmüller (2003) have investigated the acquisition of sensory-motor coupling (see also Bangert, Haeusler, & Altenmüller, 2001). Two groups of non-musicians received piano training (with auditory feedback) in ten single 20-minute sessions over a period of six weeks. In the 'map' group a conventional key-to-pitch and force-to-loudness assignment was used on the piano, while in the 'no-map' group the five relevant tones were randomly reassigned to the five relevant keys after each training trial. EEG data was acquired in two different probe conditions during the course of learning: *passive listening* (purely auditory) and *silent finger movement* (purely motor). First alterations in cortical activity already occurred during the first training session. Then, during the course of learning, came to pass a remarkable effect: The EEG patterns in the listening and the movement condition became increasingly similar. This was only the case in the 'map' group, where the motor and auditory events were contingently mapped. It seems that the activation of either the auditory representation or the motor representation led to a co-activation of the other. There was no such effect in the ‘no-map’ group. However, other than in the Holdsworth study, EMG revealed no covert muscular activity during listening. Presumably, such activity can only observed after a much longer learning phase.

The above studies, which all used neurophysiological methodology (EMG, EEG, MEG, fMRI), have provided evidence for sensory-motor coupling in experienced musicians. But what might be the benefits of such coupling? Steven Finney and Caroline Palmer have conducted a study, where they examined the role of auditory feedback in the memorizing of musical sequences (Finney & Palmer, 2003). In various experiments, auditory feedback was either present or absent during the learning and the subsequent retrieval of an unfamiliar piece. Results showed that the presence of auditory feedback during learning significantly improved later retrieval. However, the presence or absence of auditory feedback during retrieval did not significantly affect retrieval performance. It seems that in the process of learning participants had somehow integrated the auditory feedback into their movement representation of the piece, and then had used this integrated information during the actual performance. They must have used an *internal* representation of the auditory events, because actual auditory feedback during retrieval had no significant influence. To use the terminology of Elsner and Hommel (2001), these pianists had encoded their actions along with their distal effects, and had used these action-effect associations in the control of movements.

2.4 Summary

As this review showed, close coupling of sensory and motor processes has been demonstrated in a large number of empirical studies. Such coupling has shown to manifest itself in the *interaction* of sensory and motor processes. Interaction has been demonstrated in both ways: as the influence of sensory processes upon motor processes (“perception-on-action”), and as the influence of motor processes upon sensory processes (“action-on-perception”). It was argued that such effects may arise from two sources: they may be grounded in the hard-wired structural characteristics of the cognitive system, or they may arise from specific learning. This distinction is also reflected in the theories of sensory-motor coupling. Theoretical approaches like the ‘Theory of Event Coding’ describe the possible nature of a *structural* sensory-motor coupling, while theories of associative learning and computational models illustrate how specific experience might give rise to effects of *learned* sensory-motor coupling. It showed that musicians might be an exemplary group to exhibit learned sensory-motor coupling. First evidence for this assumption can be derived from a number of neurophysiological studies.

3 Empirical Part: Overview

In the following overview of the empirical part are first described the aims of the present study. Here are propounded the specific research questions that established the basis for the experiments. Afterwards follows a general description of the interference paradigm, which was used in the experiments.

3.1 Aims and Structure

In the previous section musicians were identified as an exemplary group to exhibit close coupling of sensory and motor processes. It was argued that such coupling has its foundation in the specific learning history of musicians. As has been shown, first empirical evidence for specific sensory-motor coupling in musicians was brought forward in a number of previous studies. All these studies were interested in the question of how such coupling manifests in certain neurophysiological parameters. Among the parameters that were examined were covert muscular activity (Holdsworth, 1975), activity of the primary motor cortex (Haueisen & Knösche, 2001), activity of the primary auditory cortex (Scheler et al., 2001), and the correlation of sensory and motor activation in the cortex (Bangert & Altenmüller, 2003).

The aim of the present study is to investigate *behavioral* effects of sensory-motor coupling in musicians. That is, to what extent does sensory-motor coupling affect the actual execution of movements in a musician? Hence, in contrast to the aforementioned neurophysiological studies, the focus of the present study is not on underlying, non-observable processes, but on actual, observable behavior. As such, the present work is directly related to the studies of Elsner and Hommel (2001) and Brass et al. (2000). Elsner and Hommel have described how the repeated experience of concurrent sensory and motor events leads to the formation of bi-directional associations. If such associative structures have developed in experienced musicians, instrument-related movements should be directly coupled with corresponding tones. Accordingly, the perception of these tones should activate associated movement representations (see also the theoretical section). This activation should occur in a similar way as in the imitation study of Brass and colleagues, with the difference that the "imitation" is not related to the observed movements (proximal visual effects), but to the perceived sounds (distal auditory effects).

The experiments of the present study, which are described in the following section, were designed to elucidate a number of questions. First, can behavioral evidence for sensory-motor coupling in musicians be brought forward at all? If this were the case, musicians should exhibit different effects than non-musicians in specifically designed experiments. Second, taken that such coupling can be observed, on what levels in the cognitive system does it take place? For example, are tones directly linked to certain movement representations on a rather low level? In this case, the perception of a certain sound would directly activate an associated action. Or, is there also an involvement of higher cognitive levels, which might, for example, represent abstract musical features like major-minor mode? In this case, the perception of a sound would also activate abstract information. Third, is there an influence of elementary musical structure? Musical events may be separated into the dimensions of "harmony" and "melody" (Spitzer, 2002). These dimensions might also be observable in the coupling of sounds to movements. And fourth, is sensory-motor coupling in musicians specific for the own instrument? If it is not, one might expect effects of generalization onto other, maybe similar, instruments.

These research questions were dealt with in a series of experiments, which are described in the following section. The section is ordered into three parts. The first part examines sensory-motor coupling in the harmony dimension of music. Here, pianists are compared to non-musicians (Experiment 1.1), and it is investigated to what extent the perception of musical chords activates motor and abstract representations (Experiment 1.2 and 1.3). The second part examines the melody dimension of music. Here, it is investigated in how far the perception of two-tone sequences can directly induce corresponding movements in pianists (Experiment 2.1 and 2.2). Again, these results are compared to a group of non-musicians (Experiment 2.3). The third part is concerned with the specificity of sensory-motor coupling in musicians. Here, a group of pianists (Experiment 3.1) is compared to a group of guitarists (Experiment 3.2) in the perception of different instrument timbres.

3.2 Experimental Paradigm

All of the following experiments are based on an interference paradigm, which is somewhat similar to the paradigm of Brass et al. (2001). In the study of Brass and colleagues, participants had the task to perform finger movements in response to visually presented stimuli. Concurrently with these imperative stimuli, distractor stimuli were presented, which could be congruent, incongruent or neutral in relation to the required response. The rationale of this paradigm is quite similar to the flanker paradigm (see section 2.1.2). In the experiments of the present study, participants had the task to carry out movements on their instrument (mostly a piano-like keyboard) in response to visual stimuli. Responses included the playing of chords and the playing of two-note sequences. In contrast to the study of Brass et al. (2001), the task-irrelevant distractors were not visual, but auditory stimuli. Each distractor was presented concurrently with the imperative stimulus. The relation of distractor and response could either be congruent, incongruent, or neutral.

With this paradigm, behavioral effects of sensory-motor coupling can be investigated quite efficiently. If musicians have learned to associate sounds and movements, then the perception of a sound should involuntarily activate a corresponding movement representation, even if the sound is not relevant in the current task context. Accordingly, these auditory distractors might be termed as “potential” action effects, because for a musician they are usually the perceived effects of instrument-related movements. An involuntarily activated movement representation might be either congruent or incongruent with the movement representation, which is part of the planned response in the task. This can effect in either *interference* (incongruent condition) or *facilitation* (congruent condition) of the response. Sensory-motor coupling should thus become observable in comparatively slower (incongruent condition) or faster (congruent condition) responses, and possibly in more (incongruent condition) or less (congruent condition) task errors.

4 Empirical Part I: Harmony Dimension

Two questions define the outset for the first series of experiments. First, is there evidence for sensory-motor coupling in musicians in dimension of harmony? And second, what levels of the cognitive systems may be involved in this putative coupling?

The conception of music as consisting of the two, more or less independent, dimensions of *harmony* and *melody* has, for example, been maintained by Spitzer (2002). These dimensions may also be described as vertical (harmony) and horizontal (melody) temporal dimensions. *Harmony* means the concurrent layering of tones at a time t. In most Western music, this layering is composed of tones of the diatonic scale, and results in a chord with a defined keynote. For example, the tones F, A, and C form a “major” chord with the keynote F. On the other hand, *melody* means the sequence of tones in subsequent time steps t_1, t_2,…t_n. In polyphonic music, there is a progression of two or more voices, each with its own “melody”. Typically, the combined result of such a layering and succession of tones is a harmonic progression, which follows a certain musical logic. Note that *rhythm* may be considered as a third elementary dimension of music. Rhythm, however, is not concerned with tonal relations, but with the temporal aspects of musical sequences. Rhythm is not examined in the present study. The examination of rhythm requires a quite different experimental approach. Besides, it has already been studied extensively, and also incorporates non-learned sensory-motor aspects (see also section 2.3.2).

In instrument playing, harmony and melody are directly related to certain bodily actions. Take, for example, piano playing. Here, harmony is produced by certain finger configurations on the keyboard, while melody is produced by finger movements on the keys. Thus, motor processes (finger configurations and movements) and sensory processes (layers and sequences of tones) are closely related. The aim of the first series of experiments is to investigate whether this relationship is reflected in a learned coupling of these sensory and motor processes in experienced pianists.

4.1 Experiment 1.1 – Experts and Novices

Does the perception of a "potential" action effect involuntarily evoke a corresponding, learned action representation in experienced pianists? Experiment 1.1 was designed to address this issue. Participants were required to play one of four triad chords, which varied in major-minor mode (major, minor) and pitch (C, F). The chord to be played was specified by a visual text stimulus (e.g., "F major"). Concurrently with each visual stimulus, a task-irrelevant auditory distractor was presented in piano timbre. The auditory distractor was either congruent (the same chord), incongruent (another chord) or neutral (a non-tonal sound, or no sound at all) with respect to the imperative stimulus.

A group of experienced pianists (expert group) was compared to a group of non-musicians (novice group). If pianists have acquired enduring integrated representations of actions (e.g., finger movements) and effects (piano sounds), then the perception of a task-irrelevant auditory stimulus should give rise to faster responses in the congruent than in the incongruent condition, because an associated action representation will be activated. In contrast to the piano experts, novices should not be systematically affected in their performance by either congruent or incongruent stimuli.

4.1.1 Method

4.1.1.1 Participants

16 experienced pianists (12 female; 1 left-handed; age in years: *M* = 23.5, *SD* = 3.8; years of practice: *M* = 12.1, *SD* = 3.7) formed the expert group. They were piano students from the Richard-Strauss-Conservatory, Munich. The novice group consisted of 16 participants (9 female; all right-handed, age: *M* = 22.4, *SD* = 4.0), who did not play any instrument. 20 Euros were paid for participation.

4.1.1.2 Material and Equipment

Participants responded on a YAMAHA CBXK2 MIDI keyboard. The experiment was conducted in a dimly lit, soundproof room. Visual stimuli were presented on a computer screen, which was positioned behind the keyboard. Auditory stimuli were played over headphones. The four visual stimuli consisted of black text on white background (Times New Roman, visual angle approximately 1.43° horizontally and 0.48° vertically), each denoting a triad chord ("C major", "C minor", "F major" or "F minor"). The five auditory stimuli were recorded piano sounds of the same four triad chords, and a non-tonal "chimes" sound (i.e., it did not consist of distinctly perceivable notes), taken from a sound library.

4.1.1.3 Procedure

Each trial started with the presentation of a black fixation cross on white background (500 ms). After that, the imperative visual stimulus (visible for 150 ms) and irrelevant auditory stimulus (muted at response onset, maximal duration 3000 ms) were presented concurrently, that is, with the same onset time. In the expert group 48 practice trials were carried out. Participants of the novice group were instructed in a longer demonstration phase (96 trials) which keys to press on the keyboard in response to the stimuli presented, just like in an ordinary RT experiment. 1200 experimental trials were conducted. The relation between the irrelevant auditory and the imperative visual stimulus was either *congruent* (the same chords, e.g., both F major), *incongruent* (different major-minor mode and/or pitch, e.g., imperative C major, auditory F minor) or *neutral* (chimes sound or no sound). These conditions were randomly intermixed. Participants were instructed to respond as fast and as accurately as possible by playing the visually specified triad chord with their right hand. They were told to ignore the sounds they heard over the headphones. There was no auditory effect/feedback of the response. Feedback whether the response was correct was given visually (“Ok” or “Error”). After playing the chord, participants were instructed to leave their hand in the current horizontal position, lifted slightly above the keyboard. Therefore, when playing a chord, participants were required to either move their hand away from the previous position (e.g., from the C to the F position) or to remain in this position (*Shift* vs. *Tarry*).

4.1.1.4 Data Analysis

Analysis of results was based on the mean response time (RT) of the three triad keys from valid trials. A trial was considered invalid if (1) at least one key in the response was wrong, *key error*, (2) the keys were not pressed synchronously, *sync error* (more than 150 ms between the fastest and the slowest keypress), or (3) responses occurred outside the response window (above 3000 ms), *time error*. Both control conditions were combined as "neutral" for analysis, since there was no significant RT difference between them (paired t test, novices: $p > .30$, experts: $p > .50$). RTs and error rates for key and sync errors were entered into a 2 x 3 x 2 mixed factor analysis of variance (ANOVA) with the variables Group (Experts, Novices), Congruency (Congruent, Neutral, Incongruent) and Movement (Tarry, Shift).

4.1.2 Results

Table 1 displays the RTs and percentages of key and sync errors obtained in the different conditions of Experiment 1.1. There was a main effect of RT for Group, $F(1, 30) = 50.34$, $p < .001$, indicating that RTs in the expert group (823 ms) were shorter than in the novice group (1281 ms) and a main effect for Movement, revealing that responses were slower in the Shift condition (1128 ms) than in the Tarry condition (968 ms), $F(1, 30) = 49.16$, $p < .001$. In addition, there was a significant interaction of Group x Movement, $F(1, 30) = 21.20$, $p < .001$: Hand shifts in the expert group (55 ms difference between Tarry and Shift) were much faster than in the novice group (266 ms). The main question of the study concerned Congruency. There was a main effect for Congruency, $F(2, 60) = 4.95$, $p < .05$, and a significant interaction of Congruency x Group, $F(2, 60) = 3.93$, $p < .05$. Post-hoc comparisons with the Least-Significant-Difference (LSD) test showed that within the expert group the Congruent condition was faster than the Neutral condition, and the latter in turn was faster than the Incongruent condition (all $p < .05$). Thus, both interference and facilitation were obtained. In the novice group, none of the conditions different from each other (all *p* > .40). No further effects were significant in the ANOVA.

Time errors (overall rate, experts: 0.18%, novices: 1.35%) were considered as outliers and were not analyzed further. Key error (experts: 2.6%, novices: 5.1%) and sync error rates (experts: 0.18%, novices: 1.35%) are also displayed in Table 1. The ANOVAs on each error type revealed only a main effect for Group (key errors: *F*(1, 30) = 7.21, *p* < .05, sync errors: *F*(1, 30) = 7.20, *p* < .05). That is, experts generally made fewer errors than novices. There were no further significant main effects or interactions.

In order to examine in which of the dimensions of pitch (C, F) and major-minor mode (major, minor) the congruency manipulation led to interference in the expert group, a further ANOVA on RT with the variables Movement (Tarry, Shift), Pitch (Congruent, Incongruent) and Mode (Congruent, Incongruent) was computed. First, there was a main effect for Movement, *F*(1, 15) = 12.12, *p* < .01, which reflects the fact that responses in the Shift condition were slower than in the Tarry condition (49 ms). The ANOVA further yielded a main effect for Pitch, *F*(1, 15) = 15.77, *p* < .01. Responses in the Pitch Incongruent condition were 26 ms slower than in the Pitch Congruent condition. There was also a main effect for Mode, *F*(1, 15) = 12.86, *p* < .01. Mode Incongruent condition were 23 ms slower than in the Mode Congruent condition. There was no interaction. Thus, the dimensions of pitch and mode were both effective in producing interference when they were incongruent.

	Congruent	Neutral	Incongruent
RT			
Experts			
Tarry	763.0 (148.5)	782.5 (138.6)	816.9 (168.8)
Shift	830.3 (190.6)	838.1 (187.3)	859.0 (186.2)
Novices			
Tarry	1146.7 (179.0)	1137.2 (157.6)	1159.0 (181.4)
Shift	1420.8 (263.4)	1410.2 (259.2)	1409.3 (269.3)
Key			
Experts			
Tarry	2.4 (3.3)	2.6 (2.1)	2.4 (2.0)
Shift	2.1 (2.3)	2.9 (3.2)	3.4 (2.5)
Novices			
Tarry	4.6 (3.5)	5.5 (2.9)	4.6 (2.8)
Shift	5.0 (4.2)	5.3 (4.2)	5.5 (3.8)
Sync			
Experts			
Tarry	1.6 (3.6)	1.4 (2.1)	1.8 (2.7)
Shift	1.8 (3.2)	1.8 (3.1)	1.8 (3.1)
Novices			
Tarry	8.5 (11.1)	8.7 (10.3)	8.9 (8.1)
Shift	7.5 (8.6)	6.9 (8.9)	7.6 (9.2)

Table 1. Mean Response Times in Milliseconds, key error and sync error rates in percent (with Standard Deviations) for Experiment 1.1 for experts and novices in conditions with congruent, neutral, and incongruent irrelevant auditory potential action effects for Tarry and Shift conditions.

4.1.3 Discussion

The results of the experiment support the hypothesis that in experienced pianists the perception of a potential action effect automatically produces interference with, or facilitation of, to-be-performed actions. Responses of experienced musicians were slowest when they were presented with incongruent potential effects, and fastest when they were presented with congruent potential effects. This was not the case in non-musicians. This suggests that perception of potential auditory effects led to an involuntary, automatic activation of associated action representations.

If we look closer into the design of the experiment, we see that interference could have taken place at two possible levels in the task. It could have taken place (a) between the perception of the auditory distractor and the perception of the imperative stimulus, or (b) between the perception of the auditory distractor and the selection of the response. For instance, it might be that participants automatically named the chords they heard (Segalowitz, Bebout, & Lederman, 1979), which might have led to interference with the imperative text stimulus that was read. Furthermore, imperative text stimuli and auditory chord stimuli are similar in an abstract sense – they designate same or different major-minor mode and key (C, F). In order to argue in favor of action-effect coupling, it needs to be assured that interference occurred between the auditory stimulus and the response, and not between the auditory stimulus and the imperative stimulus. To this end, Experiment 1.2 was designed, where different types of imperative stimuli were employed that could be more or less related to the auditory distractor.

4.2 Experiment 1.2 – Types of Imperative Stimuli

Imperative stimuli provide decisive information for response selection in a task. Stimulus features can overlap more or less with the features of a to-be-performed response (Kornblum & Lee, 1995). They can provide information which is directly motor-related (which keys to press), or which pertains to a more abstract level (e.g., major-minor mode). In Experiment 1.2, the types of imperative stimuli were varied to represent those different levels. Stimuli were pictures of marked keys, notes, note characters (e.g., “c – e – g”), text (e.g., “C – Major”), or colored squares (conditions: Keys, Notes, Char, Text, Color; see Figure 3 for examples of the stimuli). Each stimulus type contained the necessary information to select the appropriate response (in the case of color stimuli the mapping was given in the instructions).

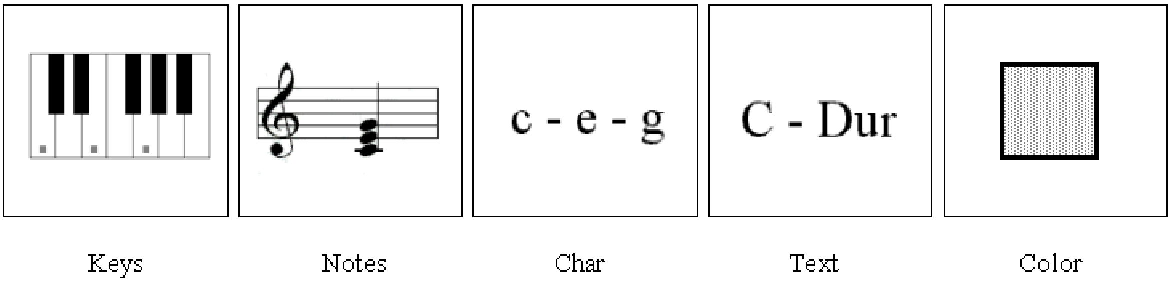


Figure 3. Examples of imperative stimulus types in Experiment 1.2 and 1.3. From left to right: Keys, Notes, Char, Text, Color. Text stimuli were in German (“C – Dur” is German for “C – major”). Color stimuli were either blue or red.

The general task of playing chords was retained from Experiment 1.1, while task complexity was reduced to the selection of either a C major or a C minor chord. This was done for the reason of experiment length, owing to the additional manipulation of stimulus type. Also, the major-minor mode manipulation in Experiment 1.1 has proved to be effective in pianists. Only pianists were examined, because the focus of the experiment was on the specific question of where interference occurred in experts. Furthermore, Experiment 1.1 has provided clear evidence that there was no interference in non-musicians.

If there is interference between the auditory stimulus and the response, it should occur in all conditions. In this case, different types of imperative stimuli should not have a modulating effect. However, if interference occurs between the imperative stimulus and the auditory stimulus, it should differ, depending on the type of imperative stimulus.

Marked keys directly convey information about the movement to be produced. In this condition, therefore, there should be no interference between processing of the auditory distractor and the imperative stimulus. If participants automatically name the chords they hear, and this leads to interference with the reading of the imperative stimulus, there should be greater interference in the Char and Text conditions than in the Keys condition, since both Char and Text consist of verbal codes. One might also expect higher interference in the Notes condition than in the Keys condition. Notes have to be read, too, and there are several similarities between music reading and reading verbal text (Nakada, Fujii, Suzuki, & Kwee, 1998). However, there are also differences between these forms of reading (Nakada et al., 1998; Schön, Anton, Roth, & Besson, 2002). For example, in the case of the present experiment, text and notes both signify a certain quality of sound, while text also designates the abstract major-minor mode category. Therefore, naming the imperative stimulus might not necessarily result in the same kind of interference with musical notes as with verbal stimuli, because of the probable differences in their cognitive representation. If there is interference between the imperative stimulus and the auditory distractor on an abstract level (because they designate the same or different major-minor mode), there should be greater interference in the Text condition than in the Char condition, because only the former directly refers to the major-minor mode concept. In the Color condition, information about the chord to be played has to be extracted first. If there is interference between the imperative stimulus and the auditory distractor on an abstract level (i.e., on a level of categorization), this condition could also yield heightened interference, because participants presumably quite consciously associate color stimuli with the mode categories „major“ or „minor“.

4.2.1 Method

4.2.1.1 Participants

16 experienced pianists (12 female; all right handed; age in years: *M* = 23.9, *SD* = 4.1; years of practice: *M* = 12.9, *SD* = 5.7) served as participants. They were piano students from the Richard-Strauss-Conservatory, Munich. None of these pianists participated in any of the other experiments. 20 Euros were paid for participation.

4.2.1.2 Material and Equipment

The setting and the equipment used was the same as in Experiment 1.1. The two auditory stimuli were recorded piano sounds of triad chords (C major, C minor). The ten visual stimuli coded either a C major or a C minor chord. There were five sets (see also Figure 3): (1) Pictures of marked keyboard keys (Keys), (2) musical notes (Notes), (3) single characters specifying the notes (Char, e. g., “c

– e – g”), (4) text stimuli (Text, e.g., “C – Major”), and (5) a blue or red square denoting either C major or C minor (Color).

4.2.1.3 Procedure

The procedure was the same as in Experiment 1.1, with the following differences: The experiment consisted of 40 practice and 1000 experimental trials. Visual stimuli were visible for 300 ms, in order to make the more complex ones still discernible.

4.2.1.4 Data Analysis

Invalid trials were discarded from the analysis of RT. RTs and error rates were entered into a 2 x 5 repeated measures ANOVA with the variables Congruency (Congruent, Incongruent) and Imperative Stimulus (Keys, Notes, Char, Text, Color).

4.2.2 Results

Table 2 displays the RTs and percentages of key and sync errors obtained in the different conditions of Experiment 1.2. Table 2 also shows the differences between Incongruent and Congruent conditions (IC-Diff) and corresponding effect sizes for dependent samples (ES, see Cohen, 1988). Effect sizes are discussed in the Results and Discussion sections of Experiment 1.3. The ANOVA for RT revealed a significant main effect for Imperative Stimulus, $F(4, 60) = 27.46$, $p < .001$. As the post-hoc LSDs showed, RTs for Notes and Keys did not differ significantly (all $p > .80$), but are successively longer for Char, Text, and Color (all *p* < .05). This effect is in accordance with higher dimensional overlap between keys and notes and the required response (Kornblum & Lee, 1995).

	Keys	Notes	Char	Text	Color
RT					
Congruent	522.7 (65.2)	522.1 (76.7)	561.1 (78.4)	590.5 (61.8)	648.4 (100.8)
Incongruent	540.0 (67.0)	536.8 (77.7)	579.5 (75.9)	627.4 (70.6)	665.3 (117.4)
IC-Diff	17.3 (18.9)	14.7 (23.4)	18.4 (19.6)	36.9 (29.6)	16.9 (32.2)
Effect size	0.92	0.63	0.94	1.25	0.53
Key errors					
Congruent	1.3 (1.1)	2.2 (2.6)	2.5 (3.2)	4.2 (3.6)	5.9 (4.0)
Incongruent	2.3 (2.0)	3.8 (3.7)	3.7 (3.8)	7.9 (5.2)	8.0 (5.4)
Sync errors					
Congruent	1.3 (1.4)	1.8 (2.0)	1.2 (1.1)	1.3 (1.0)	1.9 (2.2)
Incongruent	0.8 (1.2)	1.9 (1.8)	1.1 (1.3)	2.0 (1.2)	2.0 (2.2)

Table 2. Mean Response Times in Milliseconds, IC-Diffs (RT difference between Incongruent and Congruent conditions), effect sizes, and key and sync error rates in percent (with Standard Deviations) for expert pianists in Experiment 1.2 in conditions with congruent and incongruent irrelevant auditory potential action effects for imperative stimuli of the types Keys, Notes, Char, Text, and Color.

In addition, there was a significant main effect for Congruency, *F*(1, 15) = 31.33, *p* < .001. RTs in the Incongruent condition were longer than in the Congruent condition. Furthermore, there was significant interaction of Imperative Stimulus x Congruency, *F*(4, 60) = 2.53, *p* < .05. To investigate the interaction, the IC-Diff for each imperative stimulus was analyzed separately. Post-hoc LSD tests revealed that, although with all types of stimuli a significant IC-Diff was obtained (all *p* < .05, see Table 2), the effect was higher for the Text condition (37 ms) than for the other conditions (*M* = 17 ms, *p* < .05).

Time errors (0.08%) were considered as outliers and were not analyzed further. Key error (4.2%) and sync error rates (1.6%) are also displayed in Table 2. The ANOVA for key errors revealed a significant main effect for Congruency, *F*(1, 15) = 18.75, *p* < .001. More errors were made in the Incongruent condition. Also, there was a main effect for Imperative Stimulus, *F*(4, 60) = 11.89, *p* < .001. As the LSD test shows, significantly more errors were made in the Text and Color than in the other conditions (all *p* < .01). There was no interaction. The ANOVA for sync errors yielded a main effect for Imperative Stimulus, *F*(4. 60) = 2.63, *p* < .05. The LSD test did reveal no clear pattern as to the source of this effect. There was no interaction.

4.2.3 Discussion

As the response time results show, there was significant interference in all conditions. Let us first look at the Notes, Keys and Char conditions. Here, the three relevant motor components, that is, the three single fingers or tones, are directly specified. These are situations of basic sensory-motor compatibility (Kornblum & Lee, 1995). Therefore, it is most likely that in these conditions interference occurred on the motor level (Elsner et al., 2002; Haueisen &

Knösche, 2001). According to this account, the perception of the auditory stimulus directly activated certain motor representations.

How can the larger effect in the Text condition, as compared to the other conditions, be accounted for? It is unlikely that this effect was due to interference between naming the auditory stimulus and the reading of the imperative stimulus. If this were the case, there should have been higher interference in the Char condition than in the Key condition, because the Char condition also relies on verbal coding. It seems more likely that interference between the irrelevant auditory stimulus and the imperative stimulus occurred on an abstract level, because only the Text condition exhibited a heightened interference effect. The higher RT level in the Text condition cannot account for the higher IC-Diff, since RTs in the Color condition were even higher, but the IC-Diff was smaller.

The key characteristic of interference on an "abstract" level is that it denotes interference on a processing level where discriminations between major and minor chords are made, that is, where major-minor mode categorization takes place. This categorization may involve (a) auditory representations, and/or (b) categorical knowledge of the major-minor mode concept. In the first case, perception of text stimuli would activate an auditory representation of the perceptual major or minor mode sound quality. Evidence for such auditory activation by symbolic stimuli has been reported by Widmann and colleagues (see Widmann, Kujala, Tervaniemi, Kujala, & Schröger, 2004). In the second case, discrimination would take place on a level of abstract knowledge, that is, on a level, which does not rely on auditory representations. The abstract categorical information of the major-minor mode feature would then have to be extracted from the auditory stimulus. This would be in accordance with Segalowitz et al. (1979). From the present data, it cannot be decided between these possibilities. However, in order to distinguish this effect from interference on the motor level, I will use the term "abstract" interference in the following. Abstract interference always implies interference which involves major-minor mode categorization, however it actually comes to pass.

It might seem surprising that the color stimulus did not induce a higher interference effect – interference was not higher than with the motor-related stimuli, and lower than in the Text condition. There are three possible explanations for this. The instruction was to play either a C major or C minor chord in response to the color of this stimulus. Hence, the major-minor mode category must have been activated in one way or another. However, this may not have been the strategy participants used. Instead, they might have directly translated the visual stimuli into motor representations, for instance by associating color with a certain response, without referring to abstract properties. A second explanation could be that abstract information is extracted in parallel to response selection. This process takes time, and congruency or incongruency might only be effective in the early stages of response selection. The critical feature for interference to occur between the auditory and the imperative stimulus would then be, whether abstract information is directly given by the imperative stimulus (the text stimlus directly conveys the major-minor mode category), or has to be extracted by the participants themselves. There is also a third possiblity to explain the lower effect in the Color condition. I have argued above that abstract interference might be due to an activation of auditory representations by the perceived imperative stimuli, or it might be due to an activation of categorical knowledge. It might be that categorical knowledge is activated by color stimuli, but that it does not play a decisive role for interference. According to this explanation, color stimuli are less effective in activating auditory representations of major-minor mode than text stimuli, which results in smaller interference.

A significant interference effect was obtained with all different types of imperative stimuli. According to the predictions, this leads to the conclusion that, for the most part, interference occurred between the auditory stimulus and the response. That is, potential auditory effects automatically activated associated actions. However, one might still argue that abstract information was always activated by the different imperative stimuli, but to different degrees in the Text and the other conditions. Therefore, it might still be that the observed effects were not due to interference between the irrelevant auditory stimulus and the response, but to interference between the auditory stimulus and the imperative stimulus. Therefore, Experiment 1.3 was conducted as a further control.

4.3 Experiment 1.3 – Types of Responses

In Experiment 1.3, Experiment 1.2 was replicated, but with a decisive difference: participants did not respond by playing the chord on a keyboard, but by deciding with simple keypresses whether the imperative stimulus designated a C major or a C minor chord. Responses thus were dissimilar to the motor pattern pianists usually perform to produce chords.

Again, interference should occur in the Text condition, because of interference between the auditory stimulus and the imperative stimulus on an abstract level. Interference between the irrelevant auditory stimulus and the response should not be expected, because the responses were dissimilar to the responses that are usually performed to produce the auditory stimuli as effects. The perception of the auditory stimulus might still activate an associated action on the motor level, but this activation would not interfere with the to-be-performed response. Thus, if the observed effects in the other conditions in Experiment 1.2 had really been due to interference between the auditory stimulus and the response, there should be no interference, or smaller interference, in these conditions in Experiment 1.3.

4.3.1 Method

4.3.1.1 Participants

16 experienced pianists (12 female; one ambidextrous, all other right-handed; age in years: *M* = 24.6, *SD* = 4.4; years of practice: *M* = 15.6, *SD* = 6.0) served as participants. They were piano students from the Richard-Strauss-Conservatory, Munich. None of these pianists participated in any of the other experiments. 20 Euros were paid for participation.

4.3.1.2 Material and Equipment

Setting, stimuli, and material were the same as in Experiment 1.2, with one difference: The response device consisted of two custom-built response keys on which responses were made with the index fingers of both hands.

4.3.1.3 Procedure

The procedure was the same as in Experiment 1.2, with the following difference: Participants were instructed to respond to the visual stimulus by deciding whether it designated a C major or a C minor chord. To do so, they were required to press either the left or the right response key. The chord-key mappings were balanced between participants.

4.3.1.4 Data Analysis

Data analysis was the same as in Experiment 1.2, except that RT was now derived from the single keypress. Thus, there were no sync errors. Invalid trials were discarded for the analysis of RT.

4.3.2 Results

Table 3 displays the RTs and percentages of key and sync errors obtained in the different conditions of Experiment 1.3. The ANOVA for RT revealed a significant main effect for Imperative Stimulus, $F(1, 15) = 8.02$, $p < .05$. This effect mainly derives from the Color condition, in which RTs were significantly longer than in the other conditions, except Text (LSD test, all $p < .05$). There was a main effect for Congruency, $F(4, 60) = 3.39$, $p < .05$, and an interaction of Congruency x Imperative Stimulus, $F(4, 60) = 2.77$, $p < .05$. The IC-Diff for each Imperative Stimulus condition was computed, and it was tested whether these differences differed significantly from zero (*t* test). There was only a significant effect for the Text condition, $t = 2.68$, $p < .05$. Thus, only in the Text condition evidence for interference on an abstract level was obtained.

Time errors (0.21%) were considered as outliers and were not analyzed further. Key error rates (4.8%) are displayed in Table 3. The ANOVA for key errors revealed no significant effect.

	Keys	Notes	Char	Text	Color
RT					
Congruent	683.1 (207.4)	665.6 (199.3)	706.0 (200.1)	690.7 (191.6)	753.1 (250.6)
Incongruent	693.9 (196.1)	669.5 (200.1)	698.2 (187.6)	726.4 (206.9)	756.3 (247.0)
IC-Diff	10.8 (28.8)	3.9 (26.8)	-7.8 (36.2)	35.7 (48.6)	3.2 (41.4)
Effect size	0.37	0.15	-0.21	0.73	0.08
Key errors					
Congruent	3.6 (4.4)	4.1 (4.7)	3.5 (2.9)	4.0 (3.6)	6.5 (6.0)
Incongruent	5.1 (7.7)	3.7 (3.4)	5.4 (4.1)	5.0 (4.1)	6.3 (4.8)

Table 3. Mean Response Times in Milliseconds, IC-Diffs (RT difference between Incongruent and Congruent conditions), effect sizes, and key error rates in percent (with Standard Deviations) for expert pianists in Experiment 1.3 in conditions with congruent and incongruent irrelevant auditory potential action effects for imperative stimuli of the types Keys, Notes, Char, Text, and Color.

A comparison of the IC-Diffs of Experiment 1.2 and Experiment 1.3 in one ANOVA revealed a main effect for Experiment, $F(1, 30) = 5.60$, $p < .05$, reflecting that IC-Diffs are higher in Experiment 1.2, and a main effect for Imperative Stimulus, $F(4, 120) = 4.74$, $p < .01$, reflecting that the IC-Diff is higher for the Text condition, but no interaction of Experiment x Imperative Stimulus, $F(4, 120) = .68$, $p = .61$. The non-existence of an interaction indicates that the pattern of effects is the same in Experiment 1.2 and Experiment 1.3, though it seems that the effects in Experiment 1.3 lack interference between auditory stimulus and the responses, as was the case in Experiment 1.2.

Although the interference effect in the Text condition (36 ms) is numerically about the same size as in Experiment 1.2 (37 ms), a calculation of the effect size shows that the effect is about half a standard deviation smaller for Experiment 1.3 ($ES = 0.73$) than for Experiment 1.2 ($ES = 1.25$).

4.3.3 Discussion

Let us first look at the differences between the experiments in the Text condition. The results indicate that the effect for the Text condition was in fact smaller in Experiment 1.3 than in Experiment 1.2. It may be concluded that the effect for the Text condition in Experiment 1.2 was due to interference (a) between the auditory stimulus and the response, and (b) between the auditory stimulus and the imperative stimulus, whereas it was only due to interference between the auditory stimulus and the imperative stimulus in Experiment 1.3.

In both experiments, 1.2 and 1.3, there was a marked difference between the Text condition and the other conditions. This difference can be explained by interference occurring between the processing of the auditory stimulus, and the direct activation of abstract information by the text stimulus. Additionally, there were significant effects for the other stimulus types in Experiment 1.2, while there were none in Experiment 1.3. Most probably, this reflects automatic response activation by the auditory stimuli on the motor level, which led to interference with the response that had to be carried out in Experiment 1.2. In Experiment 1.3, this response activation did not lead to interference, because participants carried out quite dissimilar responses.

4.4 General Discussion Part I

Two questions were posed at the outset for this first series of experiments. First, is there evidence for sensory-motor coupling in musicians in dimension of harmony? And second, what levels of the cognitive systems may be involved in this coupling? The first question can be answered in the positive. As Experiment 1.1 showed, auditory ”potential action effect” stimuli activated representations that interfered with an ongoing keyboard playing task. This effect occurred in pianists, but not in non-musicians – which is evidence for acquired A-E coupling in pianists. Experiments 1.2 and 1.3 served as control experiments. Here, the possibility that these effects were merely brought about by interference between the auditory and the imperative stimulus was ruled out.

Experiments 1.2 and 1.3 also served to investigate the question of the representational levels involved. The results of these experiments showed that when text stimuli were used as imperative stimuli, interference indeed occurred between the auditory and the imperative stimulus, which was probably owing to an activation of abstract categorical representations by both of these stimuli. This kind of interference partly contributed to the observed effects. However, the effects which were obtained when other imperative stimuli were used can clearly be attributed to interference between the auditory stimulus and the response, because the effects were only present when the response was similar to the one, which is usually performed to produce these chords as effects (that is, when they were similar to piano playing). The effects in the Text conditions indicate that the potential auditory effects did not only activate features on a sensory-motor level. Rather, it seems that, in these conditions, the potential action effect was also processed on an abstract categorical level. This is accordance with previous findings with verbal material, where it was shown that action effects can evoke abstract information regarding word meaning (Koch & Kunde, 2002).

5 Empirical Part II: Melody Dimension

Experiments 1.1 to 1.3 have shown that pianists, as compared to non-musicians, have learned to associate actions with piano sounds in the domain of chords, that is, in the “harmony” dimension of music. This second empirical part examines sensory-motor coupling in the domain of tone sequences, that is, in the “melody” dimension of music.

Effects in the previous part were only observable in response times and not in error rates. However, if perceived potential action effects automatically activate corresponding actions, they should have the capability to actually *induce* specific actions. Interference of potential action effects with an action should become observable in specific errors, that is, errors that correspond to presented potential action effects. Thus, the second aim of Part II is to investigate action and action effect coupling in an experimental design, which makes it more likely to observe erroneous movement induction.

Induction of actions via perceived events, albeit not action effects, has been reported in the study by Schubö et al. (2001), which has been described in section 2.1.1. Many effects on errors in interference paradigms could probably be understood as induction errors (Eriksen & Eriksen, 1974; J. R. Simon, 1990; J. R. Simon & Rudell, 1967). However, most of these studies used two-choice paradigms, and it is thus impossible to distinguish between induction errors and other errors. To investigate the issue of whether perceived action effects actually *induce* the corresponding action, one needs a task, which allows more than two response alternatives and has a certain likelihood of inducing errors. I decided that a sequence-playing task with four possible response alternatives, as described below, would suit those needs.

Hence, the aim of Part II is to investigate whether the perception of a heard piano tone sequence could directly activate a corresponding movement representation in pianists. I wanted to (a) replicate the results of Part I and extend them to another domain of actions and (b) find a design, which allows the direct examination of action induction. The task in the following three experiments consisted of playing two-tone sequences (intervals) on a keyboard in response to visual stimuli. Task-irrelevant auditory distractor intervals, which could be congruent or incongruent with the required response, were presented concurrently with the imperative visual stimuli. Participants were experienced pianists in Experiment 2.1 and 2.2, and non-musicians in Experiment 2.3.

5.1 Experiment 2.1 – Note Stimuli

The simplest kinds of tone sequences that can be thought of are generic intervals, that is, sequences of two tones. In order to produce intervals on the piano, two keys have to be pressed in succession. This results in two corresponding piano tones, which are unambiguously mapped to the keys. Positive evidence for action-effect coupling would be provided if the mere perception of a piano tone sequence activates a corresponding action in pianists. It should be noted that for such an effect to happen, the starting note of the sequence must be known. This was ensured by the experimental design.

The experiment required participants to play intervals on a MIDI keyboard. After pressing an initial key (E) with the middle finger, a second target tone was to be pressed (either C, D, F, or G). Concurrently with the presentation of the visual imperative stimuli, which consisted of notes, task-irrelevant auditory distractors were presented. These also consisted of intervals from E to one of the four possible target tones. Thus, the interval to be played and the auditory interval presented could either be congruent (both were the same) or incongruent (they were different).

If auditory stimuli have the capability of inducing associated actions, interference effects should be observable in the incongruent conditions. That is, responses should be slower and error rates should be higher in conditions, where the auditory interval was different from the interval to be played, as compared to those conditions, where they were the same. If auditory stimuli indeed “induce” certain actions, this should become apparent in the distribution of errors in the incongruent condition. *Induction errors*, that is, responses where participants erroneously play the second tone they are presented auditorily, should be more likely to occur than other errors. Such a direct induction of erroneous responses should also be reflected in faster responses in induction error trials.

5.1.1 Method

5.1.1.1 Participants

16 experienced pianists (8 female; 1 ambidextrous, all other right-handed; age in years: *M* = 23.1, *SD* = 3.6; years of practice: *M* = 11.9 years, *SD* = 5.1) participated in the experiment. They were piano students from the Richard-Strauss-Conservatory, Munich. None of the participants reported to have perfect pitch. Ten Euros were paid for participation.

5.1.1.2 Materials and Equipment

Responses were acquired with a YAMAHA CBXK2 MIDI keyboard. The experiment was conducted in a dimly lit, soundproof room. Participants were seated in front of a computer screen (screen size 19”), on which the visual stimuli were presented. Auditory stimuli were presented over headphones. There were five different visual stimuli. They all consisted of black musical notes and staves on white background (visual angle approximately 2.14° horizontally and 1.05° vertically). One depicted a single E_4 note, the other four depicted intervals of two successive notes: E_4-C_4, E_4-D_4, E_4-F_4 and E_4-G_4. The five auditory stimuli consisted of single piano tones: C_4, D_4, E_4, F_4 and G_4. They were sampled from a real piano.

5.1.1.3 Procedure

Participants were instructed to play the notes, which appeared on the screen in front of them, with their right hand on the keyboard. They were told to play the note sequences *legato*, that is, to keep the first key pressed until the onset of the second key. Participants were instructed to place their fingers on the keyboard such that the thumb was on the C_4 key, and the following fingers were placed on the subsequent keys. This position was to be retained throughout the experiment. Participants were told to ignore the tones they heard over the headphones.

The course of events within a trial is shown in Figure 4. At the beginning of each trial, a black fixation cross was presented on white background for 500 ms at the centre of the screen. It was followed by a visual stimulus, which always showed a single E_4 note. When participants had correctly pressed the E_4 key on the keyboard, immediate auditory feedback was given (an E_4 tone was played). After a fixed interval of 500 ms, a second note was added to the first visual note stimulus. It depicted one of the four possible target notes. The visual stimulus disappeared with the onset of the target response. Concurrently with the second note (the target stimulus), a second auditory stimulus was presented. The second auditory stimulus was one of four tones (C_4, D_4, F_4 or G_4). It was muted at the onset of the target response. Thus, the resulting auditory interval was either congruent or incongruent with the interval to be played. Immediately after participants had responded by playing the target note, they were given visual feedback as to whether the response given was correct ("Ok" or "Error", visible for 300 ms). There was no auditory feedback from the second response. The subsequent trial started after an interval of 500 ms.

At first, a practice block of 32 trials was carried out. The one experimental block consisted of 640 trials (4 target stimuli x 4 irrelevant auditory stimuli x 40 trials each). The sequence of conditions was randomly intermixed. There was a short break after 400 trials. The session lasted about 45 minutes.

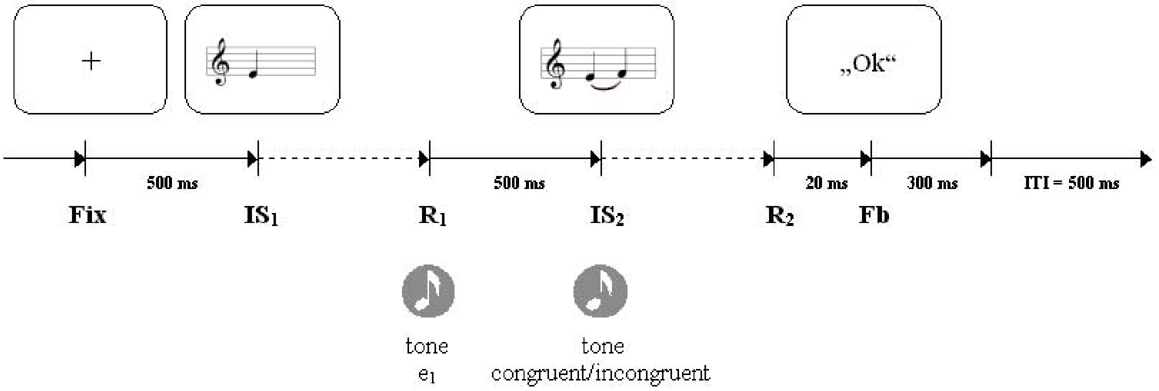


Figure 4. Course of events within a trial in Experiment 2.1. The upper part depicts examples of visual events as shown on the screen, the lower part illustrates the occurrence of auditory events. (Fix = Fixation, IS_1 = Imperative Stimulus 1, R_1 = Response 1, IS_2 = Imperative Stimulus 2, R_2 = Response 2, Fb = Feedback, ITI = Inter-Trial-Interval)

5.1.1.4 Data Analysis

Responses that were outside the response window (RT greater than 1000 ms or smaller than 200 ms) were considered as outliers, and were discarded from further analysis. Percentages of responses and response times (RTs) in trials with correct and erroneous responses were analyzed as dependent variables. RT was measured as the difference between the onset time of the target note stimulus and the onset time of the target response. *T* tests were computed to compare congruent and incongruent conditions.

In order to examine action induction, errors in the incongruent condition were specifically analyzed. In erroneous trials in the incongruent condition participants had (per definition) not pressed the key that was specified by the visual stimulus, but one of the three others instead. In case the auditory stimulus was incongruent with the required response, two different types of errors could be made. Either the error was such that participants had incorrectly pressed the key corresponding to the tone that they heard, or they had pressed any of the two remaining incorrect keys. In the following, cases in which participants erroneously played the tone they heard are labeled *induction errors*. I am aware that in a strict sense one can only speak of induction errors when the observed rate of those errors is above the expected rate. However, for reasons of convenience, I decided to use the term „induction error“ for this type of error at all times. If errors were randomly distributed, the relative induction error rate (i.e., the erroneous playing of the heard tone) should be at 33.3% of all errors in the incongruent condition. The observed rate of induction errors was statistically compared with this expected rate. Furthermore, the RTs of induction errors were compared to the RTs of other errors in the incongruent condition.

Main analysis		
	Congruent	Incongruent
RT	424.9 (66.2)	443.7 (62.0)
Percentage of Errors	2.21 (2.29)	4.13 (4.26)
Errors in the Incongruent condition		
	Induction Errors	Other Errors
Percentage	1.82 (2.24)	2.31 (2.18)
RT	396.0 (71.9)	466.6 (128.2)

Table 4. Response times in milliseconds and percentages of errors (with standard deviations) for congruent and incongruent conditions, and percentages and response times in the incongruent condition for induction errors and other errors in Experiment 2.1 (N = 16).

5.1.2 Results

Table 4 displays the RTs and error rates in the different conditions in Experiment 2.1. For non-erroneous trials, responses in the incongruent condition (444 ms) were about 19 ms slower than in the congruent condition (425 ms; two-tailed t test, $t = 6.91$, $p < .001$). Furthermore, the percentage of errors was higher in the incongruent condition (4.13%) than in the congruent condition (2.12%; two-tailed t test, $t = 2.26$, $p < .05$). Incongruency thus led to slower responses and a higher percentage of errors.

Table 4 also displays the rate of induction errors in the incongruent condition. Their percentage in relation to all the errors made in the incongruent condition was 44.17%. If errors were equally distributed, one would expect an induction error rate of about 33.3%. The observed rate of induction errors was significantly higher than this expected rate (one-tailed t test, $t = 1.84$, $p < .05$).

Thus, induction errors were more frequent than one would expect from a random distribution of errors. Responses in trials with induction errors (396 ms) were significantly faster than in trials with other errors (467 ms; one-tailed t test, $t = 2.47, p < .05$).

5.1.3 Discussion

The aim of the experiment was to find out whether the perception of incongruent potential action effects could interfere with an ongoing interval-playing task in experienced pianists. That is, does the perception of an incongruent potential action effect activate an associated action representation? The results provide evidence that this was the case. Responses were slower and errors more frequent in the incongruent condition, as compared to the congruent condition. These results replicate those of the previous part and extend them to another domain of piano playing action, that is, the playing of short sequences.

The most convincing evidence for an actual activation of actions by auditory stimuli on the motor level derives from the rate of induction errors. A perceived incongruent auditory interval frequently led participants to actually play this perceived interval – although the task required them to play another interval. This direct activation hypothesis is corroborated by the fact that responses were significantly faster in induction error trials than in the other error trials.

Thus, the results point to the capacity of potential action effects (auditory interval stimuli) to *induce* the actions by which they are usually produced. However, as in Part I, there is also an alternative account of these results. If one looks closer into the design of the experiment, it becomes apparent that interference could have taken place at several levels in the task. Interference could either have taken place (a) between the perception of the irrelevant auditory stimulus and the selection of the response, which is the present focus of interest, but also (b) between the perception of the irrelevant auditory stimulus and the perception of the imperative stimulus. Furthermore (c) the use of note stimuli as imperative stimuli could have biased the congruency effect, either because of automatic response activation by note stimuli, or because of dimensional overlap between the note stimuli and responses.

Interference between the irrelevant auditory stimulus and the imperative stimulus might follow from dimensional overlap between the visual stimulus display and the pitch of the auditory stimulus. That is, interference might take place in the abstract category of "high" and "low" which is present in pitch (high and low tones) as well as in musical notation (notes are placed on the upper or lower part of the staves).

Even if one concedes that interference occurred between the auditory stimuli and the responses, this effect might still have been modulated by the imperative note stimuli in two ways. First, for musicians notes are usually directly associated with the playing of the instrument. That is, music reading already involves sensory-motor translation of notes into adequate responses (see Schön et al., 2002; Stewart et al., 2003). Such direct, automatic activation of certain responses by notes might thus counteract effects of response activation by auditory stimuli. Second, automatic activation of responses might have been engendered by effects of dimensional overlap of stimuli and responses (S-R compatibility). Notes represent pitch on a vertical spatial dimension (higher pitch being represented by more upward notes), while keys on the keyboard are also aligned spatially, namely in the horizontal dimension (higher pitch produced by keys on the right hand side). Activation of "right" responses through upwardly oriented stimuli has been observed in a number of experiments (for a concise overview, see Cho & Proctor, 2003). Also, there is evidence that pitch is generally associated with vertical and horizontal spatial dimensions (Mudd, 1963). Such spatial mapping may also be learned by musicians (Stewart, Walsh, & Frith, 2004; Zakay & Glicksohn, 1985). Thus, it might be that the design of the experiment even led to an underestimation of the influence of potential action effects on actions.

Hence, interference between auditory stimuli and responses might have been modulated in several ways by the use of notes as imperative stimuli in the current experiment. To control for these possible biases, I tried to replicate Experiment 2.1 in a second experiment with colored squares as imperative stimuli, because color codes responses in a rather arbitrary way.

5.2 Experiment 2.2 – Color Stimuli

In order to rule out (a) that interference might occur between the perception of the auditory distractors and the perception of the imperative stimuli, and (b) that interference between responses and action effects might be modulated by the imperative stimuli, I used a different set of imperative stimuli in Experiment 2.2. Instead of notes, the required responses were now coded by colored squares. Two possible sources of confound from Experiment 2.1 are thus reduced. First, it is unlikely that musicians have acquired any S-R relationships between colors and keypresses on the piano. Second, there is no possible spatial compatibility of imperative stimuli and responses and/or auditory stimuli as in Experiment 2.1. Apart from the different types of stimuli, design and procedure of Experiment 2.2 were similar to Experiment 2.1.

If the observed interference effects in Experiment 2.1 had been due to interference between the imperative stimulus and the irrelevant auditory stimulus, there should be no effect in the current experiment. If, on the other hand, the interference effects in Experiment 2.1 had been due to interference between the auditory stimulus and the response, they should be replicated in Experiment 2.2.

5.2.1 Method

5.2.1.1 Participants

16 experienced pianists (10 female; 1 ambidextrous, 1 left-handed, all other right-handed; age in years: *M* = 23.3, *SD* = 3.5; years of practice: *M* = 13.8 years, *SD* = 5.8) participated in the experiment. They were piano students from the Richard-Strauss-Conservatory, Munich. None of them had participated in Experiment 2.1. None of the participants reported to have perfect pitch. Ten Euros were paid for participation.

5.2.1.2 Materials, Equipment, Procedure, and Data Analysis

The material and equipment used were the same as in Experiment 2.1. Stimuli were different from those in Experiment 2.1 in the following way: The keys to be pressed were not coded by notes, but by squares in arbitrary colors (visual angle of each square approximately 1.62° horizontally and vertically). There were five color stimuli: one showing a single, black-colored square, which represented the E_4 note. In the four other stimuli, a second square was placed besides the black one in either of the following colors: brown, green, red or blue. The black square appeared to the left, and the second square appeared to the right of the centre of the screen, separated by a gap of about 0.43° visual angle. Each color represented one of the notes C_4, D_4, F_4 or G_4. The color-note mapping was balanced between subjects. The five auditory stimuli used were the same as in Experiment 2.1.

The procedure was similar to Experiment 2.1, with the following difference: The demonstration block consisted of 64 trials, in order for the participants to get used to the color-key mappings. A sheet of paper was placed beside the screen in the demonstration block, where the five colored squares were depicted in the order corresponding to the assigned piano keys from left to right. The instruction was to press the middle key, when the square with the color in the middle appeared on the screen, to play the first key when the square on the utmost left appeared, and so on. Data analysis was analogous to Experiment 2.1.

5.2.2 Results

The RT results and error rates for the different conditions in Experiment 2.2 are displayed in Table 5. RTs in trials with correct responses were about 41 ms higher in the incongruent condition (677 ms) than in the congruent condition (637 ms; two-tailed *t* test, $t = 6.08$, $p < .001$). The comparison of the percentages of errors between congruent and incongruent condition yielded no significant difference ($p = .15$). Thus, an effect of incongruency was only observable in response times, but not in error rates.

However, participants still made enough errors to allow a comparison of induction errors and other errors in the incongruent condition (see Table 5). Their percentage in relation to all the errors made in the incongruent condition was 40.25%. Again, in an equal distribution of errors, one would expect an induction error rate of about 33.3%. The observed rate of induction errors was significantly higher than this expected rate (one-tailed *t* test, $t = 3.29$, $p < .01$).

Thus, induction errors were more frequent than one would expect from a random distribution of errors. As in Experiment 2.1 responses in trials with induction errors (638 ms) were significantly faster than in trials with other errors (692 ms; one-tailed *t* test, *t* = 1.79, *p* < .05).

Main analysis		
	Congruent	Incongruent
RT	636.7 (93.0)	677.3 (101.8)
Percentage of Errors	4.31 (3.89)	5.29 (2.61)
Errors in the Incongruent condition		
	Induction Errors	Other Errors
Percentage	2.13 (1.11)	3.16 (1.62)
RT	638.3 (140.1)	692.1 (107.0)

Table 5. Response times in milliseconds and percentages of errors (with standard deviations) for congruent and incongruent conditions, and percentages and response times in the incongruent condition for induction errors and other errors in Experiment 2.2 (N = 16).

5.2.3 Discussion

Experiment 2.2 was conducted to replicate the results Experiment 2.1, and to exclude possible biases of imperative stimuli. These possible biases concerned learned relationships of notes and responses, and compatibility between notes and auditory stimuli and/or responses, as they were possibly present in Experiment 2.1. To rule out such effects, arbitrary color stimuli were used in Experiment 2.2. As in Experiment 2.1, responses in the incongruent condition were slower than in the congruent condition. Also, the rate of induced incorrect responses was above the expected rate, and responses in induction error trials were faster than those in other error trials. Thus, the effects of Experiment 2.1 could be replicated, while unwanted sources of interference were controlled for.

It can be ruled out that interference took place between the imperative stimulus and the auditory stimulus in the present experiment. The imperative stimuli were absolutely arbitrary, with respect to their mapping to tones (however, note that Hasbroucq and Guiard, 1991, have proposed a stimulus-identification account in which the arbitrary stimuli are presumed to take on the meaning of the responses to which they are assigned). Instead, the observed effects must have been due to interference between the auditory stimulus and the response. This interference could (a) reflect learned action-effect associations (as hypothesized), but it might (b) also reflect spatial relationships between pitch and response location. To ensure that these results can be interpreted as being due to learned action-effect associations, I conducted a third experiment, in which Experiment 2.2 was repeated with non-musicians.

5.3 Experiment 2.3 – Non-Musicians

Experiment 2.1 and 2.2 showed that experienced pianists were systematically affected by perceived auditory distractor intervals, while they played intervals on a keyboard. I interpreted this as evidence for learned A-E associations in pianists. In order to show that the effects obtained were really due to such learned associations, Experiment 2.2 was repeated with a control group of non-musicians. Just like the pianists in Experiment 2.2 they were told which keys to press in response to which color. If there is an effect in non-musicians, this would be evidence for dimensional overlap of auditory stimuli and responses, which cannot be attributed to specific experience. If, as compared to pianists, there is a smaller effect in non-musicians, or even no effect at all, this would be evidence for acquired action-effect associations in pianists.

5.3.1 Method

5.3.1.1 Participants

16 non-musicians (10 female; 1 ambidextrous, 1 left-handed, all other right-handed; age in years: *M* = 23.3, *SD* = 2.0) participated in the experiment. None reported to play, or to have played, any musical instrument. Ten Euros were paid for participation.

5.3.1.2 Materials, Equipment, Procedure and Data Analysis

Material and equipment, procedure, and data analysis were similar to Experiment 2.2.

5.3.2 Results

The RT results and error rates for the different conditions in Experiment 2.3 are displayed in Table 6. There was neither a significant difference between congruent and incongruent conditions in response times ($p > .50$), nor in error rates ($p > .60$). Thus, no significant effect of incongruency was observable. The mean level of RT was the same as in Experiment 2.2 (about 667 ms in both experiments).

The relative rate of induction errors (see Table 6) was 29.35%. It was not statistically different from the expected rate of 33.3% (one-tailed t test, $p > .05$). Also, the RT level in induction error trials did not significantly differ from the RT level in other error trials (*p* > .50). Thus, no evidence for involuntary action induction could be observed in Experiment 2.3.

Main analysis		
	Congruent	Incongruent
RT	664.1 (100.2)	667.3 (95.7)
Percentage of Errors	5.29 (3.74)	5.56 (3.23)
Errors in the Incongruent condition		
	Induction Errors	Other Errors
Percentage	1.57 (0.88)	3.99 (2.50)
RT	701.8 (160.7)	690.7 (118.2)

Table 6. Response times in milliseconds and percentages of errors (with standard deviations) for congruent and incongruent conditions, and percentages and response times in the incongruent condition for induction errors and other errors in Experiment 2.3 (N = 16).

5.3.3 Discussion

Contrary to the similar experiment with pianists (Experiment 2.2), non-musicians did not show any effect of incongruency. It did not matter whether they were presented congruent or incongruent distractor intervals – their response times and error rates were the same. In particular, there was no evidence for induction in incongruent trials, neither in a higher rate of specific induction errors, nor in response times. From these results I conclude that all the effects in Experiment 2.2 can be attributed to specifically learned associations of actions and action effects in pianists.

5.4 General Discussion Part II

Part II investigated two major questions. First, the question of whether experienced pianists have learned to associate movements and sounds in the domain of two-tone sequences (intervals). Second, the question of whether potential auditory action effects have the capability to directly *induce* associated movements. This issue was investigated in three experiments. The results of Experiment 2.1 support the hypothesis of learned action-effect associations in the “melody dimension” in pianists. Experiment 2.2, which was carried out in order to exclude possible influences of the stimulus material, replicated these findings. Experiment 2.3 demonstrated that these effects were due to the pianists’ expertise, because no such effects could be observed in non-musicians.

There is strong evidence that interference actually occurred between the perceived auditory stimuli and the selection of the responses. The alternative account (interference between auditory stimuli and visual stimuli) is contradicted by the results from Experiment 2.2. Here, dimensional overlap of auditory and visual stimuli was excluded by using arbitrary color stimuli without any spatial properties. Also, the use of arbitrary stimuli ruled out learned S-R relationships of notes and responses. Furthermore, it can be ruled out that the observed effects were due to non-learned spatial compatibility between tones and finger movements (Experiment 2.3). Hence, it seems that the auditory distractors interfered quite systematically and specifically with the ongoing response processes. Further support for this is provided by the results on induction errors. The auditory intervals specifially induced erroneous actions more often than it would be expected if errors were randomly distributed across all response alternatives.

6 Empirical Part III: Instrument Specificity

Imagine two musicians attending a piano recital. One is a professional guitarist, and the other is a professional pianist. Both will listen to the same music in the concert. But will they also process the music that they hear in the same way? There is reason to assume that this is not the case. The guitarist knows piano sounds only from listening to them, and at best may have a vague abstract understanding of how such sounds are produced (Rosch, Mervis, Gray, Johnson, & Boyes-Braem, 1976). But, he has never acquired the motor skills to actually produce them by means of a piano. On the other hand, a pianist is certainly able to relate piano sounds to actual motor representations (see the pervious two sections). Hence, it seems reasonable to assume that the guitarist and the pianist will process perceived piano music differently – although both are listening to the same auditory events. The aim of Part III is to examine this issue experimentally.

It has been said that events, like a piano tone or the movement of a finger, are believed to consist of a set of distinctive features (Hommel et al., 2001b, see section 2.2.3). In the case of musical events, they include low-level features, like the frequency spectrum and the amplitude pattern of a sound, or the muscle forces and torques in an elementary movement. Additionally, there are higher-level features, like the categorization of a sound into a certain timbre type (e.g., "string instrument", see Palmer, Jones, Hennessy, Unze, & Pick, 1989), or the major-minor mode category of a chord (see Part I). The question arises, which representational levels are involved in the coupling of events, like in A-E coupling in musicians. Is it that a certain action is linked to a certain frequency and amplitude pattern? A-E associations would then be constricted to a purely sensory-motor level. Or, is it that features like consonance and dissonance or major-minor mode are part of this coupling? This would imply an involvement of higher representational levels.

One may reformulate this question with regard to learning specificity. In case A-E coupling is rather specific, it is constrained to exactly the learned action and the learned effect. On the other hand, this linking could also generalize over similar actions and/or action effects. Generalization phenomena are well known from classical conditioning (Domjan, 1998). It might be the case, for example, that the perception of a sound of similar timbre activates a movement representation, although such an association has never been learned (e.g., guitar and harp have quite similar timbres). Or, alternatively, the perception of a sound, which is similar in other, abstract musical features (e.g. major-minor mode) may engender such activation. Tones of any instrument might activate corresponding actions in a pianist.

In the previous two parts, I have provided evidence for acquired action-effect coupling in pianists. Part III examines whether A-E associations in musicians have been learned specifically for sounds of the own instrument. In order to examine this issue, pianists and guitarists were tested in two experiments. To test for specificity, the auditory distractors were now presented in different instrument timbres.

6.1 Experiment 3.1 – Pianists

Experiment 3.1 was carried out in order to find out whether a similar interference effect as described in Part I could be found for different types of instrument timbre. As in the experiments of Part I, participants were required to play major and minor chords on a keyboard in response to visual stimuli. The distractor stimulus was presented in one of the following timbres: piano, organ, guitar, flute, or voice. Participants were experienced pianists. They had no particular experience or practice on any of the other instruments that were presented auditorily.

Which instrument timbres will cause interference in this experiment? This depends on the specificity of A-E coupling, and on the representational levels involved. If there is a specific sensory-motor coupling of certain sounds to certain movements, significant interference should only occur in the piano condition, because piano is the participants' own instrument. If, on the other hand, this coupling involves the processing of higher-level features of the sound (e.g., the major-minor mode category of the sound), a significant, if not equivalent effect should occur in the other timbre conditions, too. Furthermore, categorization on an abstract level might involve the processing of instrument categories (Palmer et al., 1989). Two could lead to two possible effects. First, piano belongs into the category of keyboard instruments. This categorization is based on the motor requirements of playing the piano. This category also incorporates other keyboard instruments. Hence, interference might also occur the organ condition. Second, piano belongs into the category of string instruments. This categorization is based on features of sound generation and sound similarity. Therefore, one might expect interference in the guitar condition, since the guitar is also a string instrument. Such an effect might be due to categorization (string instrument category) or perceptual generalization (sound similarity).

6.1.1 Method

6.1.1.1 Participants

20 experienced pianists (10 female; 1 ambidextrous, 1 left-handed, all other right-handed; age in years: *M* = 23.7, *SD* = 2.7; years of practice: *M* = 13.7 years, *SD* = 5.1) participated in the experiment. They were piano students from the Richard Strauss Conservatory in Munich. None of them had any particular training on either organ, guitar, or flute. 15 Euros were paid for participation.

6.1.1.2 Material and Equipment

Participants responded on a YAMAHA CBXK2 MIDI keyboard. The keys used in the experiment were C_4, Eb_4, E_4 and G_4. The experiment was conducted in a dimly lit, soundproof room. Participants were seated before the keyboard, with their right hand above the keys. Visual stimuli were presented on a computer screen, which was positioned behind the keyboard. Auditory stimuli were played over headphones.

The two visual stimuli consisted of black musical notes on white background (visual angle approximately 1.84° horizontally and 1.05° vertically) in the violin clef, each denoting a triad chord (C major, C minor). Ten auditory stimuli were used. They consisted of C major and C minor chords in five timbres: piano, organ, guitar, flute and voice. Each major/minor chord consisted of the same three notes (C major: C_4, E_4, G_4; C minor: C_4, Eb_4, G_4). The sounds were recorded from real instruments. Flute and voice chords were recorded as ensembles. The duration of each auditory stimulus was 3000 ms.

6.1.1.3 Procedure

At the beginning of each trial, a black fixation cross on white background was shown for 500 ms at the center of the screen. It was followed by the presentation of the imperative visual stimulus (C major or C minor chord, visible for 200 ms). Concurrently, that is, with the same onset time as the imperative stimulus, an irrelevant auditory stimulus (either C major or C minor) in one of the five timbres was played over headphones. It was muted at response onset, and its maximal duration was 3000 ms. Imperative and auditory stimulus could either be *congruent* (e.g., both C major) or *incongruent* (e.g., imperative: C major, auditory: C minor) in major-minor mode. Participants were instructed to respond to the visual stimulus by playing the designated chord on the keyboard. They were told to ignore the sounds they heard over headphones. There was no auditory effect/feedback to the response. Visual feedback as to whether the response was correct was given immediately after the response ("Ok" or "Error", visible for 300 ms). The inter trial interval was 500 ms.

The experiment started with a block of 40 demonstration trials. In the experimental block, each combination of visual imperative and irrelevant auditory stimulus (20 combinations) occurred 50 times. Thus, there were a total of 1000 trials in the experiment. Experimental conditions were randomly intermixed. There was a break of about five minutes after 600 trials. The experimental session lasted about one hour.

6.1.1.4 Data Analysis

Analysis of response times was based on the mean response time (RT) of the three triad keys from valid trials. A trial was considered invalid if (1) at least one key in the response was wrong, *chord error*, (2) the keys were not pressed synchronously (more than 150 ms between the fastest and the slowest keypress), *sync error*, and (3) responses occurred outside the response window (less than 200 ms or more than 3000 ms), *time error*.

Sync errors and time errors were not analyzed further. RTs and chord error rates were entered into 2 x 5 repeated measures analyses of variance (ANOVAs) with the variables Congruency (Congruent, Incongruent) and Timbre (Piano, Organ, Guitar, Flute, Voice).

	Piano	Organ	Guitar	Flute	Voice
RT					
Congruent	441.2 (60.5)	447.8 (68.8)	444.1 (58.7)	457.2 (64.2)	454.2 (68.6)
Incongruent	452.0 (70.6)	456.7 (73.7)	448.6 (59.4)	453.2 (67.6)	457.2 (70.0)
IC-Diff	10.8* (17.1)	8.9* (14.4)	4.5 (10.6)	-4.0 (16.4)	3.0 (11.6)
Chord Errors					
Congruent	1.4 (1.4)	1.1 (1.7)	1.3 (1.5)	1.0 (1.0)	1.1 (2.1)
Incongruent	2.5 (3.0)	1.7 (2.3)	1.9 (2.6)	1.1 (2.2)	1.7 (2.6)

Table 7. Mean response times in milliseconds and chord error rates in percent (standard deviations in parenthesis) for experienced pianists in Experiment 3.1 in conditions with congruent and incongruent auditory stimuli, and IC-Diff values (RT difference between Incongruent and Congruent conditions) for auditory stimuli of the types Piano, Organ, Guitar, Flute, and Voice (* $p < .05$).

6.1.2 Results

Table 7 displays the RTs obtained in the different conditions of Experiment 3.1. The ANOVA on RT revealed a main effect for Timbre, *F*(4, 76) = 6.90, *p* < .01, deriving from RT level differences in the various Timbre conditions. There was also a main effect for Congruency, *F*(1, 19) = 5.95, *p* < .05, and a significant interaction of Congruency x Timbre, *F*(4, 76) = 4.06, *p* < .01. To determine the source of the interaction, the IC-Diff for each Timbre condition was computed separately (see Table 7 for IC-Diff values). It was then tested whether the IC-Diff in each Timbre condition differed significantly from zero (*t* test). A significant IC-Diff was obtained in two conditions: Piano, *t* = 2.76, *p* < .05, and Organ, *t* = 2.67, *p* < .05 (all other *p* > .05).

There was a rate of 0.06% (of all trials) for time errors and a rate of 1.54% for sync errors. The ANOVA on chord errors revealed a tendential effect for Congruency, *F*(1, 19) = 4.12, *p* = .06. It seems that participants generally produced more errors when the auditory chord was incongruent. Also, there was a main effect for Timbre, *F*(4, 76) = 2.84, *p* < .05. Overall, more errors were made in the Piano condition (2.0%) than in all other conditions (mean 1.4%, post-hoc LSD test, all *p* < .05). There was no interaction of Congruency x Timbre (*p* > .60).

6.1.3 Discussion

A significant interference effect in RT occurred in two timbre conditions: piano and organ. Most errors were made in the piano condition. Hence, the first thing to note is that incongruency in the major-minor mode dimension was in itself not sufficient to bring about a significant interference effect in RT. There were no significant effects in the guitar, flute and voice conditions. Thus, the results speak against a timbre-independent processing of the major-minor mode category.

Interference occurred when sounds of the participants' own instrument (piano) served as a distractor stimulus, and also when organ sounds did. None of the participants actually played organ, and had hardly any experience on this instrument. How can this result be interpreted? According to the predictions, one would expect an "organ effect", if instrument sounds were categorized with respect to the "keyboard instrument" category. It seems that this was the case. Palmer and colleagues (1989) have argued that the categorization into instrument categories follows a principle, which is similar to Gibson's notion of affordances (Gibson, 1979). If we follow this assumption, it seems reasonable to assume that organ sounds might have an "affordance" for pianists, because, concerning hand movements, an organ is played in the same way as a piano. There was no effect in the guitar condition, which would have been evidence for an influence of the sound-related category (string instruments are similar in sound, and/or sound generation, respectively).

However, one may also think of an alternative explanation for the effects of this experiment. Organ and piano sounds may have intrinsic qualities that make them easier to be recognized and differentiated than other instruments, especially regarding the recognizability of major-minor mode. One cannot decide between these explanations with the results of the present experiment. To exclude this possibility, Experiment 3.2 was conducted.

6.2 Experiment 3.2 – Guitarists

A perceptual account for the results of Experiment 3.1 might argue that piano and organ sounds are perceptually more salient, and that their features (especially major-minor mode) can be discriminated more easily. If this were the case, it should not matter whether the perceived sounds have timbres of the own or of other instruments: piano and organ sounds should lead to interference in any musician, who performs a similar task.

Experiment 3.2 was designed to clarify this issue. Its design was similar to Experiment 3.1. However, instead of piano players, participants were now guitarists, who performed the task on a guitar. Instead of chords in the C key, chords in the A key were presented to and played by participants, because this was technically easier to accomplish. If the results in Experiment 3.1 had been due to perceptual characteristics of the presented instrument timbres, interference should occur in the piano and organ conditions in Experiment 3.2 as well. If, on the other hand, these effects derived from learned, instrument-specific representations, this should not be the case. Rather, the most pronounced interference effect should be observed in the *guitar* condition. Together with the results of Experiment 3.1, this would be evidence for a double dissociation of instrument specificity.

6.2.1 Method

6.2.1.1 Participants

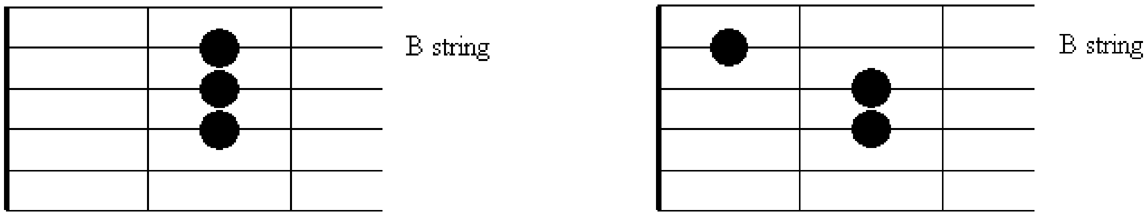
20 experienced guitarists (5 female; one ambidextrous; age in years: *M* = 27.9, *SD* = 8.9; years of practice: *M* = 12.4 years, *SD* = 7.6) participated in the experiment. The group consisted of guitar students from the Richard-Strauss-

Conservatory, Munich, and of professional guitar teachers. None of them had any particular experience on either piano, organ, or flute. 15 Euros were paid for participation.

6.2.1.2 Material and Equipment

For the experiment, a customary steel string guitar was modified such as to allow response time measurements. Below the B string on each of the two upmost frets, a response button was attached, which was activated when the string was pressed down. This setup still allowed the guitar to be played in a normal way. Two kinds of responses could occur: A minor and A major (see Figure 5). Since in the A minor chord the B string is pressed on the first fret, and in the A major chord the B string is pressed on the second fret, this setup enabled the measurement of which of the two chords was grasped. Due to technical constraints it was not possible to control for the positions of all three fingers with this setup. I think this was no problem for the experiment, since the decisive measure lay in the differentiation of A major and A minor chords. As a control, the experimenter checked during the demonstration phase that participants actually grasped the whole chord with all three fingers.

The experiment was conducted in a dimly lit, soundproof room. Visual stimuli were presented on a computer screen, and auditory stimuli were played over headphones. Visual and auditory stimuli were similar to Experiment 3.1, except that they consisted of major and minor chords in the A key (A minor: A_3, C_4, E_4; A major: A_3, $C\#_4$, E_4).



(a) A major chord

(b) A minor chord

Figure 5. Finger positions on the guitar for two chords (a) A major (b) A minor. Sensors on the first two frets below the B string enabled the measurement of the chord that was grasped.

6.2.1.3 Procedure

The experimental procedure was similar to Experiment 3.1, with the following adjustments: Instead of C chords, A chords were used (see above). Participants were instructed to grasp the visually specified chord but not to pluck the strings, so that there was no auditory feedback. After each trial, participants moved their hand into a neutral position above the strings.

6.2.1.4 Data Analysis

Due to technical constraints, only the response time of the B string was measured. Analysis of results was based on RTs from valid trials. A trial was considered invalid if (1) the wrong chord was grasped, *chord error*, and (2) the response occurred outside the response window (less than 200 ms or more than 3000 ms), *time error*. Time errors were not analyzed further. RTs and chord error rates were entered into a 2 x 5 repeated measures analysis of variance (ANOVA) with the variables Congruency (Congruent, Incongruent) and Timbre (Piano, Organ, Guitar, Flute, Voice).

6.2.2 Results

Table 8 displays the RTs obtained in the different conditions of Experiment 3.2. The ANOVA neither revealed a main effect for Congruency, nor for Timbre (all $p > .15$). However, there was a significant interaction of Congruency x Timbre, $F(4, 76) = 2.59$, $p < .05$. Again, to determine the source of the interaction, each IC-Diff value was tested whether it differed significantly from zero (*t* test). Only the IC-Diff in the Guitar condition (20 ms) turned out to be significant, $t = 2.49$, $p < .05$ (all other $p > .15$).

	Piano	Organ	Guitar	Flute	Voice
RT					
Congruent	738.8 (216.5)	725.0 (206.2)	723.6 (215.2)	719.9 (207.5)	727.0 (221.8)
Incongruent	720.5 (215.1)	716.7 (211.3)	743.8 (228.6)	719.0 (203.7)	734.6 (207.9)
IC-Diff	-18.3 (53.7)	-8.3 (38.4)	20.1* (35.2)	-0.9 (36.5)	7.6 (52.5)
Chord Errors					
Congruent	4.0 (5.8)	3.9 (5.3)	4.6 (7.0)	4.7 (6.5)	4.8 (8.1)
Incongruent	4.6 (6.9)	5.2 (9.0)	4.3 (7.1)	5.1 (8.3)	4.6 (7.0)

Table 8. Mean response times in milliseconds and chord error rates in percent (standard deviations in parenthesis) for experienced guitarists in Experiment 2 in conditions with congruent and incongruent irrelevant auditory stimuli, and IC-Diff values (RT difference between Incongruent and Congruent conditions) for auditory stimuli of the types Piano, Organ, Guitar, Flute, and Voice (* $p < .05$).

There was an overall time error rate of 3.17%. The ANOVA on chord error rates (see Table 8) revealed no significant main effects, and no interaction (all $p >$ .30).

6.2.3 Discussion

Experiment 3.2 examined whether evidence for instrument-specific A-E coupling could be obtained in guitarists. The results of the experiment are clearly in favor of this hypothesis. Only in the guitar condition, incongruent auditory stimuli led to significant interference. Thus, a double dissociation of instrument-specific A-E representations was obtained: whereas pianists in Experiment 3.1 exhibited significant interference with piano and organ sounds, but not with guitar sounds, guitarists in Experiment 3.2 exhibited interference with guitar sounds, but not with piano or organ sounds.

This result rules out a perceptual account of the results from Experiment 3.1. If piano and organ timbres had intrinsic qualities that would enable people to better recognize and differentiate these sounds, effects in these conditions should have turned up in the present experiment with guitarists, too. But, in guitarists congruency was only effective with *guitar* sounds – that is, the sound of the participants' own instrument. It seems that action-related representations are only activated by sounds of the own instrument. That is, only sounds of the own instrument are "potential" action effects. However, note that there was no condition comparable to the "organ" condition of Experiment 3.1 in the present experiment. That is, no instruments were presented that were within a motor-related category as guitar. Such a condition was not included, because I wanted to make the experiment as similar as possible to Experiment 3.1.

6.3 General Discussion Part III

The aim of Part III was to examine the specificity of A-E coupling in experienced musicians. Auditory distractors of different instrument timbre were presented to pianists and guitarists, while they performed a chord-playing task on their own instrument. Experiment 3.1 showed that in pianists interference occurred only when piano or organ sounds were presented, but not with sounds in other instrument timbres. Experiment 3.2 showed the analogous result pattern for guitarists: interference only occurred when guitar sounds were presented.

What do these results tell us about the specificity of A-E associations and the involvement of different representational levels? Let us first look at the differences between pianists and guitarists. Participants of each group exhibited strong interference effects when sounds of their own instrument were presented to them as distractors. That is, the perception of sounds of the own instrument triggered processes, which, when they were incongruent, interfered with ongoing response processes. The coupling of action and effect codes thus seems to contain a component, which is markedly *instrument-specific*.

The effect in the organ condition in Experiment 3.1 is evidence for an additional involvement of higher representational levels. A perceptual account for this effect (as well as for the piano effect) can be excluded, because it was not present in Experiment 3.2. Organ sounds do not seem to possess any intrinsic qualities (perceptual features like discriminability) that make them especially suited for inducing interference. Interference in the organ condition cannot draw on low level action-effect coupling, but must involve abstract knowledge. Previous studies have shown that musicians organize information about musical instruments in movement-related categories (Palmer et al., 1989; Rosch et al., 1976). It seems that instruments with similar "affordances" are represented within the same category: “…the manner of playing pianos is like the manner of playing other members of the keyboard family […] Musical instruments within families have similar affordances for playing” (Palmer et al., 1989, pp. 19 and 36). The perception of a sound from that category may thus act as an “affordance” for a possible action, by which it can be produced.

7 General Discussion

The experiments, which were described in the three empirical parts, were all aimed to investigate behavioral effects of sensory-motor coupling in musicians. A number of guiding questions had been posed at the outset of the empirical section. These were, first: Can behavioral evidence for sensory-motor coupling in musicians be brought forward at all? Second, taken that such coupling can be observed, on what levels in the cognitive system does it take place? Third, is there an influence of elementary musical structure? And fourth, is sensory-motor coupling in musicians specific for the own instrument? In the following, I will first give a brief descriptive summary of the main results of the experiments. Afterwards, I will try to answer the four guiding questions under separate headings. Finally, I will make the attempt to relate the empirical results of the present study to the previous empirical and theoretical works, which have been described at the beginning, and will discuss open questions and possible future research.

7.1 Summary of Results

The first two empirical parts addressed the question of whether the two structural dimension of music, harmony and melody, are reflected in sensory-motor coupling in musicians. Part I addressed the harmony dimension. In all the experiments, participants were presented task-irrelevant auditory chord stimuli while they played, or decided between chords, which were congruent or incongruent with the auditory distractor stimulus. Experiment 1.1 showed that, in comparison to a control condition, pianists responded slower when the distractor was incongruent, and faster when the distractor was congruent. No such effect was observable in non-musicians. Experiment 1.2 showed that interference occurred in all kinds of imperative stimulus type conditions, independent of whether they coded the required response in a motor-related or an abstract way. This independence of the imperative stimulus is evidence that interference took place between the auditory stimulus and the response. However, the experiment also showed an increased effect in the ‘Text’ condition, where the imperative stimulus coded the response with plain text. It seems that when the task-relevant category (the major-minor mode category, in this case) was directly coded in the imperative stimulus, interference increased. In Experiment 1.3, responses were made dissimilar to chord playing on a keyboard. That is, responses were such that they would not produce chords as auditory effects under normal circumstances. Results showed that, except for the ‘Text’ condition, there was no interference effect at all. Taken together, the three experiments of Part I showed that the perception of auditory chords directly activated corresponding motor representations in pianists. Thus, positive evidence for sensory-motor coupling could be demonstrated for the harmony dimension in pianists.

Part II had two major aims: First, to extend the results of Part I to the melody dimension, and second, to investigate the direct activation of responses (“action induction”) by auditory stimuli. In the experiments, pianists had the task to play two-tone sequences while they were presented auditory distractor sequences, which were either congruent or incongruent with the ones to be played. Experiment 2.1 demonstrated that a similar RT effect as in the experiments with chords (Part I) could be observed. Responses in the incongruent condition were significantly slower than in the congruent condition.

Additionally, the experiment provided direct evidence for action induction by auditory stimuli: induction errors occurred significantly above chance level. These results were replicated in Experiment 2.2. This experiment precluded possible spatial compatibility biases from the visual imperative stimuli, as they were possibly present in Experiment 2.1. Experiment 2.3 replicated Experiment 2.2 with non-musicians. There was no RT effect, and no evidence for action induction. Hence, the effects in pianists can be attributed to their specific expertise. In sum, Part II provided further behavioral evidence for sensory-motor coupling in pianists, and extended the results of Part I to the melody dimension of music. The results on action induction, especially, provided direct evidence for sensory-motor coupling in pianists.

Part III examined the question of instrument specificity in sensory-motor coupling in musicians. To this end, a group of guitarists was compared to a group of pianists. The experimental paradigm was similar to that in Part I, except that the auditory distractor chords were presented in varying instrument timbre. In Experiment 3.1, pianists exhibited significant interference effects for piano and organ sounds. In order to exclude a perceptual account of this result, and to obtain a "double dissociation" of the effect, the experiment was replicated in Experiment 3.2 with guitarists. Here, significant interference could be observed only for guitar sounds. The results are evidence for instrument-specific sensory-motor coupling, which seems to involve processing of the instrument category. Instrument sounds from the same motor category seem to act as "affordances" for musicians, who have the ability to play an instrument from that category.

7.2 Evidence for Sensory-Motor Coupling

The first guiding question of the empirical part of this study was: Can behavioral evidence for sensory-motor coupling in musicians be brought forward at all? An answer to this question can be given on the basis of the comparison of musicians and non-musicians. This comparison was made in Part I and II. In both parts, the group of musicians consisted of pianists, while the group of non-musicians consisted of participants who had never learned to play a musical instrument. Both groups performed similar experiments. Of course, the pianists already knew how to play chords and intervals on a keyboard. Non-musicians were trained the correct responses in a longer demonstration phase before the actual experiment. The rationale behind the experiments was as follows: If pianists have acquired coupled representations of sensory and motor events during their learning history, then perceived auditory events should have the capability to directly activate corresponding action representations. Hence, the experiments required participants to play chords or intervals on a keyboard, while at the same time task-irrelevant auditory distractor stimuli were presented. Positive evidence for sensory-motor coupling would be given, if interference effects could be observed.

The results of the comparisons of pianists and non-musicians can be summarized quite shortly: Pianists exhibited significant interference effects, while in non-musicians there were no effects at all. That is, pianists responded slower when they were presented incongruent auditory distractors before they made their responses. Non-musicians were not affected by either congruent or incongruent auditory distractors. The differences between both groups can be attributed to the pianists' specific learning history. It seems that, in pianists, the perception of auditory stimuli of their instrument involuntarily activated action representations that interfered with ongoing action. Somehow, the perceived auditory stimuli are linked to action-relevant processing. The first question can thus be answered in the positive: there is behavioral evidence for sensory-motor coupling in musicians.

It should be noted that other, especially neurophysiological studies, have also looked at differences between musicians and non-musicians, albeit not under the topic of sensory-motor coupling. Gaser and Schlaug (2003) have examined morphological differences between brains of musicians (keyboard players) and non-musicians. They found significant grey matter differences in motor, auditory, and visual-spatial brain regions, attributable to long-term skill acquisition. In these areas, grey matter was more pronounced in musicians. Koeneke and colleagues examined cerebellar processing in pianists and non-musicians during finger movements (Koeneke, Lutz, Wustenberg, & Jancke, 2004). Cerebellar activity in musicians was significantly less pronounced than in non-musicians, which is interpreted by the authors as evidence for decreased neuronal "effort" for trained movements. This is to say that sensory-motor coupling is just one, though perhaps major factor, which marks the difference between motor experts and non-experts.

7.3 Abstract and Motor Representations

The second guiding question that was posed at the outset of the experiments was as follows: Taken that sensory-motor coupling can be observed in musicians, on what levels in the cognitive system does it take place? Positive evidence for sensory-motor coupling in musicians *per se* was derived from the comparisons of musicians and non-musicians (see above). But what are the characteristics of this coupling? It may be that the observed effects had been due to coupling on an abstract level. In this case, perceived potential auditory effects would have led to an activation of abstract representations. For example, in the case of major and minor chords, the central abstract category is the major-minor mode category. On the other hand, if this coupling had taken place on a sensory-motor level, then a perceived potential auditory effect would have directly activated an action representation on the motor level.

A first answer to this issue may be gained from the control experiments in Part I and II. In both experiments, 1.1 and 2.1, there are several possibilities where interference might have occurred. Three events must be taken into account in both cases: (a) the imperative visual stimulus, (b) the task-irrelevant auditory distractor, and (c) the response. The assertion that there is direct motor activation can only be made if interference took place between the auditory stimulus and the response. In Part I, this was controlled for by varying the types of imperative stimuli (Exp. 1.2), and by using responses that are dissimilar to piano playing (Exp. 1.3). These control experiments showed that the imperative stimuli only had a modulating effect, if they directly addressed the abstract response-relevant category (the major-minor mode category). However, the differences between Exp. 1.2 and Exp. 1.3 are clear evidence for direct motor activation by the auditory stimuli: when the task required responses that are dissimilar to piano playing, the interference effect vanished. In Part II, the task-relevant dimension was, in a sense, spatial. Here, up and down movements had to be carried out on the keyboard. Since this up and down category was present in all three major events, results may again have been biased. This was controlled for in Exp. 2.2, where the imperative stimuli were absolutely arbitrary in this respect. Since in this experiment the interference effects were replicated, there is strong evidence that interference actually occurred between the auditory stimulus and the response, just like in Part I. This is also corroborated by the null effects in non-musicians in Exp. 2.3, which exclude the possibility that the effects in musicians derived from spatial overlap of the auditory stimulus and the response, and not from learned associations.

The most convincing evidence for direct motor activation by auditory "potential action effects" is the results on induction errors in Part II. These results show that interference was not unspecific, but *specific*. The experiments of Part I provided evidence that interference mainly occurred on the motor level. However, this interference could also have been unspecific, in the sense that processing on the motor level may have been unspecifically disturbed. Evidence for specific interference comes from the induction error results in Part II (Exp. 2.1 and 2.2). These results show that auditory stimuli directly activated certain actions, namely those actions that can potentially produce the auditory events as action effects in an actual piano playing situation. This is clear evidence that certain auditory representations are directly coupled with certain motor representations.

In addition to this direct coupling of actions and action effects, evidence for interference on a level of abstract processing was observed as well. It should be noted that interference on the abstract level of major-minor mode categorization in Part I was modulated by the imperative stimuli, and not by the auditory stimuli or the response. This abstract interference only occurred when text stimuli were used. These stimuli directly referred to the task-relevant category. Thus, this effect does not allow any inferences on sensory-motor coupling, because it derived from the task context. However, it allows inferences on the interference process. It seems that, in this special case, there was not only interference between the auditory stimulus and the response, but also between the imperative stimulus, and either the auditory stimulus or the response. The question of whether there was interference between the imperative stimulus and the auditory stimulus or between the imperative stimulus and the response in the Text condition cannot really be resolved from the experiments. However, one aspect seems to advocate interference between the imperative text stimulus and the auditory stimulus: the effect in the Text condition in Exp. 1.3, where responses were dissimilar to piano playing. Since there were no effects in the other conditions, it seems quite implausible to assume direct response activation only for text stimuli. Rather, there must have been interference between the representation of the text stimulus and the representation of the auditory stimulus. This interference could have occurred in two ways. Either, the perception of the text stimulus led to an imagined auditory representation (major or minor), which interfered with the actual auditory input. This would be in line with Widmann et al. (2004), who demonstrated auditory activation by abstract stimuli. Or, the perception of the auditory stimulus led to an abstract representation in the major-minor mode category, which interfered with the representation of the imperative text stimulus that was also coded in this abstract category. This would be in line with Segalowitz et al. (1979), who demonstrated abstract activation by reading. From the current data it cannot be decided, which of the two alternatives is correct.

7.4 The Dimensions of Harmony and Melody

The third question was: is there an influence of elementary musical structure? To answer this question, sensory-motor coupling was examined separately for the dimensions of harmony and melody. Evidence for learned sensory-motor coupling was obtained for both dimensions. The results of Part I and II both suggest direct motor activation by perceived auditory events in pianists. It seems that, in the experiments, perceived chords have activated representations of corresponding finger configurations, and perceived intervals have activated representations of corresponding finger movements. The fact that these activations actually occurred on the motor level is supported by the control experiments in both parts (see section 7.3). Hence, harmony and melody both seem to be reflected in corresponding sensory-motor representations in pianists. The *sensory* part of these representations is the actual musical output from the instrument. These are the sounds of chords and intervals, that is, layers and sequences of tones. The *motor* part of these representations consists of the procedural knowledge as to how these sounds can be produced by adequate body movements.

It must remain open, in how far the dimensions of harmony and melody are actually *independent* of each other. From a musical perspective they seem to be, because they reflect the temporal concurrence and succession of musical events. In a musical piece, both dimensions can be analyzed independently. But can such an independence also be assumed for the representation of these dimensions in musicians? According to Magill (2001), one possibility to classify motor skills is to assume a continuum between discrete and continuous motor skills. A *discrete* motor skill has specific beginning and end points. Magill mentions hitting a piano key as one example. In *continuous* motor skills, beginning and end are arbitrary, and movements are repetitive. It has been said that, in piano playing, harmony is produced by certain finger configurations, and melody is produced by finger movements. In a way, this fits into the classification of discrete and continuous motor skills. This is an indication for a differentiation of harmony and melody on the motor side. However, according to Magill, discrete and continuous skills constitute the extreme ends of a continuum on one dimension, rather than two independent dimensions. Hence, from a temporal definition (concurrence and succession) the assumption of independence may also make sense for the motor side, but from a descriptive view this assumption is rather inadequate.

It has been said that the dimensions of harmony and melody may be viewed as the elementary structural dimensions of music (with rhythm as a possible third dimension). In the experiments of the present study, harmony was examined with chords in the major and minor mode, while melody was examined with small generic intervals. Note that these chords and intervals are of the most basic and simplest kind. Any "real" piece of music is much more intricate and complex. Therefore, the results of the present study cannot simply be transferred into such more complex contexts. The aim of the present study was to examine learned sensory-motor coupling in its most basic forms. The examination of actual musical performance would require a much more complicated experimental setup. In the end, however, any musical piece can be considered as consisting of generic elements, like chords and intervals.

7.5 The Question of Instrument Specificity

The fourth question I have posed, was: is sensory-motor coupling in musicians specific for the own instrument? This question was dealt with in Part III of the empirical section. From the perspective of associative learning, specificity indicates in how far associations generalize over non-learned elements. If the range of generalization is rather narrow, this coupling is constricted to the specific elements that were part of the learning procedure. Sensory-motor coupling in musicians would then be constricted to specifically learned movements and sounds of the own instrument. To some extent, this view can be verified from musicians' subjective experience. For example, pianists, who have learned to play the piano on a conventional upright piano, usually have problems at first in the one-to-one transfer of their skill onto a grand piano. Keypresses on a grand piano require slightly different muscle forces, and have different playing characteristics. On the other hand, generalization may also be quite broad. This would allow for an easy transfer of motor skills to instruments with similar motor requirements. Many examples of such transfer can be found in instrument playing, especially in players of wind instruments, like saxophone or clarinet.

Concerning sensory-motor coupling, the question of instrument specificity may also be viewed from a different angle. The coupled elements may be processed in various “elaboration depths” (in the sense of Craik & Lockhart, 1972), or levels of symbolic abstraction. If this processing does not exceed the lowest perceptual level, the information of a sound, for instance, would be constricted to a certain frequency and amplitude pattern. If there is further processing on higher cognitive levels, then features like major-minor mode would be involved. These features are not part of the sound *per se*, but must be extracted by additional cognitive processes, including memory retrieval. If there is processing of such categorical information, it would be reasonable to assume that it would occur for all kinds of sounds, even if they are not from the own instrument.

Perception is thought to be a process that runs through these two levels, or stages (Eimas & Miller, 1978; Pitt, 1995; Sussman, 1993). Low-level features are integrated into higher-level representations. In the literature, support can be found for the involvement of both higher and lower levels in associative learning. For example, Paivio’s dual coding theory (Paivio, 1979, 1990) assumes two kinds of memories, albeit for non-auditory material. One memory stores low-level visual information, the other (verbal) memory stores semantic information. It is also known that the involvement of higher (alternatively: deeper) levels of processing significantly influences learning. Associative learning is more effective, when the material to be learned is interpreted as being meaningful (Lai, 2000).

In music, there is also evidence for a two-level type of organization. For example, the perceptual processing of instrument timbre seems to involve a sensory and a "central" level (Pitt, 1995). Timbre discriminations, which involve abstract acoustic information, are primarily made at the central level. This is the case in musicians as well as in non-musicians. A similar organization has been observed in active instrument performance, namely in piano playing (Chaffin & Imreh, 2002). Both motor features (e.g., finger movements) and abstract features (formal features of a piece) are involved in the learning of a new piece (Palmer & Meyer, 2000). The formal structure of a piece is then used in performance as a retrieval scheme.

In Experiment 3.1 and 3.2, instrument specificity in sensory-motor coupling was examined with respect to the sound of the instrument. The results of both experiments demonstrate a clear double dissociation of an instrument-specific effect. In pianists, interference occurred only with piano and organ sounds. In guitarists, interference occurred only with guitar sounds. This shows that, first, sensory-motor coupling in musicians involves a markedly instrument-specific component. Incongruent sounds of non-related instruments did not activate response-relevant representations. It seems reasonable to assume that this specific component is due to a rather low-level coupling of specific sounds and actions. However, the fact that pianists exhibited significant interference with organ sounds points to an additional involvement of higher-level information. This effect can only be explained by processing beyond a simple sensory level. The most obvious explanation for the observed effect is that abstract information of the movement-related instrument category was processed. This explanation is in line with the results of Palmer (1989), who found that musicians represent musical sounds with respect to instrument categories with similar “affordances”. Since piano and organ are both keyboard instruments, they are played in roughly the same way. For the player of a keyboard instrument, sounds of instruments of this category have similar “affordances”. This explanation also fits to the results of Pitt (1995), who found that timbre discriminations are primarily made on a central level, and not on a purely sensory level. Other potential outcomes of Experiment 3.1 were disqualified by the experimental results. In particular, there was no influence of the sound-related category, which might also have been due to perceptual characteristics. In this case, a significant effect should have occurred in the guitar condition. However, this was not the case.

7.6 Theoretical Perspectives

In the review of literature at the beginning of this dissertation, a wide range of empirical studies and theoretical approaches on sensory-motor coupling had been presented. How do the results of the present study relate to the results of previous studies and to underlying theories? The first thing to be noted is that the present study has examined one specific type of sensory-motor coupling, namely that which derives from learning experiences. In the experiments of the current study it has been attempted to exclude any influences of non-learned sensory-motor relations. For example, the results of Experiment 2.1 might have been biased by influences of dimensional overlap in the sense of Kornblum (1992). Therefore, Experiment 2.2 was conducted in order to exclude such possible biases. However, this example also demonstrates that learned and non-learned sensory-motor relations might interact, and that their respective influence might sometimes be hard to tell apart.

At the beginning, I have argued that effects of perception on action may derive from two sources: they may derive from pre-action stimuli, or from post-action effects (see section 2.1). The interference paradigm that was employed in the present study has made use of the influence of task-irrelevant pre-action sensory stimuli on action. However, the actual focus of interest was on the learned coupling of actions and post-action effects. The distractor stimuli that were used in the experiments were stimuli, which, at least for musicians, usually are the sensory *effects* of actions. Hence, the present study is to be viewed in relation to those studies, which have examined learned relationships of actions and action effects (e.g., Elsner & Hommel, 2001, 2004; Hoffmann, 2004; Ziessler et al., 2004). These studies have focused on general mechanisms of action-effect coupling. In contrast, the present work has examined action-effect coupling in a specific domain of sensory-motor expertise. The main aim of the present work was to demonstrate learned action-effect coupling after year-long training. In the experiments of Elsner and Hommel (2001), for example, A-E associations were explicitly trained in the experimental session. The present work demonstrates that in a specific group of experts such coupling has developed over years of training.

The general mechanisms of sensory-motor coupling, as they were examined in previous studies and described in the framework theories, certainly also apply to action-effect learning in musicians. Since the underlying mechanisms of this learning were not examined in the present study, it must remain open, which of the framework theories are best at describing sensory-motor coupling in general. Associative learning theory, as well as the computational models approach, would predict action activation, as it was observed in the current study. However, it is to be assumed that the situation would be different in more complex settings. Strict associative learning implies that there is a non-ambiguous mapping of the associated elements. However, in instrument playing this mapping is often ambiguous. Certain tone sequences can usually be played in more than one way. For example, Sloboda (1998) has described how pianists choose certain finger configurations, depending on the current musical context within a piece. That is, a certain auditory action effect (say, an interval) can be achieved by a variety of different actions. In the experiments of the current study, this was resolved by instructing the participants to use specified finger configurations. Dynamic finger-to-key mappings, as they are required in real music performance, are probably best described by computational models or dynamical systems theories (e.g., Erlhagen & Schöner, 2002; Kelso, 1995; Turvey & Carello, 1995). The results of Hoffmann (2004, see section 2.1.3) are also in favor of a rather dynamic linking of actions to effects.

7.7 Open Questions and Future Research

The present work has examined the basics of sensory-motor coupling in musicians. Four elementary questions concerning this coupling were dealt with (see above). Naturally, the answers to those few questions give rise to a lot of further questions that might be interesting to do research upon. I will describe four of them, which concern sensory-motor coupling itself, but also its possible function in the cognitive system.

First, as has already been remarked above, it is unclear whether pianists really associate movements with tones. In the experiments of Hoffmann (2004), effects of action activation were not dependent of the effector, but dependent of the present effector location. In the same vein, one might assume that pianists associate key locations with tones, rather than movements with tones. In order to examine this issue, one would have to dissociate finger locations and key locations in a series of experiments, just like in the Hoffmann study. For example, in an interval playing task as in Part II, participants might be instructed to use different fingering. If this would lead to similar effects, this would be evidence against a static finger-to-tone linking, but evidence for a dynamic assignment of fingers to keys.

Second, the question of higher-level processing of abstract, categorical characteristics of auditory musical action effects would present an interesting topic to examine in more detail. This accounts especially for the alleged processing of the movement-related instrument category. Do pianists really "play" internally when they hear organ music? Such an effect should vanish, if the pianist knew that the sounds are produced on the *pedals* and not on the keys – as a pianists, he only knows how to play the keyboard, and not how to play the pedals. Furthermore, it should be possible to produce effects of action activation by subtle experimental manipulations. For example, if a pianist were made to believe that the sounds he hears were produced with a keyboard, these sounds might develop an affordance-like quality – and also lead to action activation of keyboard playing.

Third, it might be interesting to examine the question of instrument specificity in a brain imaging study. Just like in the study of Haueisen and Knösche (2001), the perception of sounds of the own instrument (category) should result in specific M1 activation, while sounds of an unrelated instrument (category) should not. Such specific cortical activation for sounds of the own instrument has already been found by Pantev and colleagues (see Pantev, Roberts, Schulz, Engelien, & Ross, 2001). However, this study only examined activation of the auditory cortex, and not of motor areas. Furthermore, one might assume that the perception of motor-related instruments from the same category as the own instrument might lead to an activation of additional, possibly prefrontal areas.

As a final point it should be noted that I have examined the existence of sensory-motor coupling in musicians, but not its possible function in the cognitive system. Theoretical approaches like the *Theory of Event Coding* (Hommel et al., 2001b) or the idea of paired internal models (Wolpert & Kawato, 1998) assume that the major function of this coupling is in motor control (see sections 2.2.2 and 2.2.3). Their common assumption is that an actor must only activate a representation of the desired movement goal (the action effect), which, because of the acquired coupling, will activate a corresponding action representation. The representation of the desired effect is thought to be a voluntary imagination, or “anticipation”, of that effect. It may seem tempting to assume such a mechanism for music performance: a trained musician would only have to imagine a certain melody, and then could “automatically” perform it on his instrument. Remember the statement of contrabassist Barry Green: “When you can hold the sound and pitch of the music clearly in your head […] performing it accurately becomes easier. Your body has a sense of its goal […]” (Green & Gallwey, 1986, p. 75). Similarly, think of jazz musicians, who improvise seemingly new melodies on the spur of a moment. Nonetheless, it should be noted even great soloists can never play a difficult, unknown piece from scratch. Intricate parts with high motor demands almost always require prior training, and need to be automatized to a certain degree. And if we take a closer look at improvisations of jazz musicians, it becomes obvious that they largely consist of iterative elements (like scales and certain harmonic progressions) that are fittingly combined into a current musical context. Hence, the automatization of certain motor elements almost always plays an undeniable part in music performance. Even so, the alleged role of effect anticipation in music performance would present an interesting topic for further experimental research. It is to be assumed that it would not only comprise such elements as the mapping of fingers to tones, but also such rather subtle elements as the fine-control of accentuation or timbre.

8 Conclusion

The present work has investigated sensory-motor coupling in experienced musicians in a theoretical and empirical manner. I believe that this line of research – that is, the examination of the relation of sensory and motor processes in motor experts – is a valuable approach in order to understand the fundamentals of cognitive representations and refined movement control. In the present work, I have tried to investigate basic issues of learned sensory-motor coupling in musicians. Such coupling revealed itself in direct motor activation under sensory stimulation, while several results suggested additional higher-level processing. Motor expertise in general – indeed, any learned movement skill – may be grounded on such coupling. Needless to say that this is just a speculation and needs to be shown in further research.

I wish to end this dissertation, just as I began it, with a quotation by Arthur Schopenhauer. Somewhere (I can't remember where exactly) he wrote, frustrated with the experiences of his time: „Where to thinking leads without experimenting, this was demonstrated in the middle ages; but the current century shows us, where to experimenting leads without thinking" (translation by the author). For my part, I hope to have attained an adequate balance between experimenting and thinking, and hereby conclude this work.

References

Adams, J. A. (1971). A closed-loop theory of motor learning. *Journal of Motor Behavior, 3*, 111-150.

Allport, A. (1980). Patterns and actions: Cognitive mechanisms are content-specific. In G. Claxton (Ed.), *Cognitive Psychology: New directions* (pp. 26-64). London: Routledge & Kegan Paul.

Baker, J. M. (2001). The Keyboard as Basis for Imagery of Pitch Relations. In R. I. Godoy & H. Jorgensen (Eds.), *Musical Imagery* (pp. 251-269). Lisse: Swets & Zeitlinger.

Bangert, M., & Altenmüller, E. O. (2003). Mapping perception to action in piano practice: a longitudinal DC-EEG study. *BMC Neuroscience, 4*, 26-36.

Bangert, M., Haeusler, U., & Altenmüller, E. O. (2001). On practice: How the brain connects piano keys and piano sounds, *Biological Foundations of Music* (Vol. 930, pp. 425-428). New York: New York Acad Sciences.

Beckers, T., De Houwer, J., & Eelen, P. (2002). Automatic integration of non-perceptual action effect features: the case of the associative affective Simon effect. *Psychological Research, 66*, 166-173.

Blakemore, S. J., Goodbody, S. J., & Wolpert, D. M. (1998). Predicting the consequences of our own actions - the role of sensorimotor context estimation. *Journal of Neuroscience, 18*, 7511-7518.

Bosbach, S., Prinz, W., & Kerzel, D. (2004). A Simon effect with stationary moving stimuli. *Journal of Experimental Psychology: Human Perception & Performance, 30*, 39-55.

Botvinick, M., & Plaut, D. C. (2004). Doing without schema hierarchies: A recurrent connectionist approach to normal and impaired routine sequential action [Review]. *Psychological Review, 111*, 395-429.

Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica, 106*, 3-22.

Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition, 44*, 124-143.

Bushan, N., & Shadmehr, R. (1999). Computational nature of human adaptive control during learning of reaching movements in force fields. *Biological Cybernetics, 81*, 39-60.

Camurri, A. (1997). Network models in motor control and music. In P. Morasso & V. Sanguineti (Eds.), *Self-Organization, Computational Maps, and Motor Control* (pp. 311-355). Amsterdam: North-Holland.

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

Chaffin, R., & Imreh, G. (2002). Practicing perfection: Piano performance as expert memory. *Psychological Science, 13*, 342-349.

Charness, N., Krampe, R. T., & Mayr, U. (1996). The Role of Practice and Coaching in Entrepreneural Skill Domains: An International Comparison of Life-Span Chess Skill Acquisition. In K. A. Ericsson (Ed.), *The Road to Excellence* (pp. 51-80). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Chen, H. H., Hippenmeyer, S., Arber, S., & Frank, E. (2003). Development of the monosynaptic stretch reflex circuit. *Current Opinion in Neurobiology, 13*, 96-102.

Cho, Y. S., & Proctor, R. W. (2003). Stimulus and response representations underlying orthogonal stimulus-response compatibility effects. *Psychonomic Bulletin & Review, 10*, 45-73.

Clarke, E. F. (1999). Rhythm and timing in music. In D. Deutsch (Ed.), *The Psychology of Music (Second Edition)* (pp. 473-500). San Diego: Academic Press.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Erlbaum.

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002). Hand action preparation influences the responses to hand pictures. *Neuropsychologia, 40*, 492-502.

Craik, F. I., & Lockhart, R. S. (1972). Levels of Processing: A Framework for Memory Research. *Journal of Verbal Learning and Verbal Behavior, 11*, 671-684.

De Houwer, J. (1998). The Semantic Simon Effect. *Quarterly Journal of Experimental Psychology A, 51*, 683-688.

Desmedt, J. E. (Ed.). (1973). *Human Reflexes, Pathophysiology of Motor Systems, Methodology of Human Reflexes*. Basel: Karger.

Dickinson, A. (2001). Causal Learning: Association Versus Computation. *Current Directions in Psychological Science, 10*, 127-132.

Domjan, M. (1998). *The Principles of Learning and Behavior (Fourth Edition)*. Pacific Grove: Brooks/Cole.

Eenshuistra, R. M., Weidema, M. A., & Hommel, B. (2004). Development of the acquisition and control of action-effect associations. *Acta Psychologica, 115*, 185-209.

Eimas, P. D., & Miller, J. L. (1978). Effects of selective adaption of speech and visual patterns: Evidence for feature detectors. In H. L. Pick & R. D. Walk (Eds.), *Perception and experience* (pp. 307-345). New York: Plenum.

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. (2004). Contiguity and contingency in action-effect learning. *Psychological Research, 68*, 138-154.

Elsner, B., Hommel, B., Mentschel, C., Drzezga, A., Prinz, W., Conrad, B., & Siebner, H. (2002). Linking actions and their perceivable consequences in the human brain. *Neuroimage, 17*, 364-372.

Ericsson, K. A. (1996). The Acquisition of Expert Performance: An Introduction to Some of the Issues. In K. A. Ericsson (Ed.), *The Road to Excellence*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Ericsson, K. A. (1999). Expertise. In R. A. Wilson & F. C. Keil (Eds.), *The MIT Encyclopedia of the Cognitive Sciences* (pp. 298-299). Cambridge, Massachusetts: MIT Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363-406.

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

Erlhagen, W., & Schöner, G. (2002). Dynamic field theory of movement preparation. *Psychological Review, 109*, 545-572.

Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (2000). Visuomotor neurons: ambiguity of the discharge or 'motor' perception? *International Journal of Psychophysiology, 35*, 165-177.

Finney, S. A., & Palmer, C. (2003). Auditory feedback and memory for music performance: Sound evidence for an encoding effect. *Memory & Cognition, 31*, 51-64.

Fitts, P. M., & Deininger, 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., & Posner, M. I. (1967). *Human Performance*. Belmont, CA: Brooks/Cole.

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

Gabrielson, A. (1973). Similarity ratings and dimension analyses of auditory rhythm patterns. II. *Scandinavian Journal of Psychology, 14*, 244-260.

Gaser, C., & Schlaug, G. (2003). Brain structures differ between musicians and non-musicians. *Journal of Neuroscience, 23*, 9240-9245.

Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1997). Computational Models of Sensorimotor Integration. In P. Morasso & V. Sanguineti (Eds.), *Self-Organization, Computational Maps, and Motor Control* (pp. 117-147). Amsterdam: Elsevier.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston MA: Houghton Mifflin.

Glencross, D. J. (1995). Motor Control and Sensory-Motor Integration. In D. J. Glencross & J. P. Piek (Eds.), *Motor Control and Sensory Motor Integration*. Amsterdam: Elsevier.

Goodale, M. A., Westwood, D. A., & Milner, A. D. (2004). Two distinct modes of control for object-directed action. *Progress in Brain Research, 144*, 131-144.

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

Green, B., & Gallwey, W. T. (1986). *The inner game of music*. New York: Doubleday & Company, Inc.

Greenwald, A. G. (1970a). A choice reaction time test of ideomotor theory. *Journal of Experimental Psychology, 86*, 20-25.

Greenwald, A. G. (1970b). Sensory feedback mechanisms in performance control: with special reference to the ideo-motor mechanism. *Psychological Review, 77*, 73-99.

Grosjean, M., & Mordkoff, J. T. (2002). Post-response stimulation and the Simon effect: Further evidence of action-effect integration. *Visual Cognition, 9*, 528-539.

Haggard, P., Clark, S., & Kalogeras, J. (2002). Voluntary action and conscious awareness. *Nature Neuroscience, 5*, 382-385.

Hall, G. (1991). *Perceptual and Associative Learning* (Vol. 18). Oxford: Clarendon Press.

Hall, G. (1996). Learning about associatively activated stimulus representations: Implications for acquired equivalence and perceptual learning. *Animal Learning & Behavior, 24*, 233-255.

Halsband, U., & Freund, H.-J. (1993). Motor learning. *Current Opinion in Neurobiology, 3*, 940-949.

Harless, E. (1861). Der Apparat des Willens [The apparatus of the will]. *Zeitschrift fuer Philosophie und philosophische Kritik, 38*, 50-73.

Hasbroucq, T., & Guiard, Y. (1991). Stimulus-Response Compatibility and the Simon Effect - Toward a Conceptual Clarification. *Journal of Experimental Psychology: Human Perception & Performance, 17*, 246-266.

Haueisen, J., & Knösche, T. R. (2001). Involuntary motor activity in pianists evoked by music perception. *Journal of Cognitive Neuroscience, 13*, 786-792.

Henke, K., Buck, A., Weber, B., & Wieser, H. G. (1997). Human hippocampus establishes associations in memory. *Hippocampus, 7*, 249-256.

Herbart, J. F. (1816). *Lehrbuch zur Psychologie*. Königsberg; Germany: Unzer.

Heyes, C. (in press). Imitation by association. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation*. Cambridge, MA: MIT Press.

Hoffmann, J. (2004). *Ideo-Motorik: Alte Antworten und neue Fragen [Ideomotor theory: old answers and new questions]*. Paper presented at the 46. Tagung experimentell arbeitender Psychologen (TeaP), Giessen.

Hoffmann, J., Sebald, A., & Stocker, C. (2001). Irrelevant response effects improve serial learning in serial reaction time tasks. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 470-482.

Holdsworth, E. I. (1975). Neuromuscular Activity and Covert Musical Psychomotor Behavior: An Electromyographic Study. *Dissertation Abstracts International, XXXVI.*

Hommel, B. (1996). The cognitive representation of action: Automatic integration of perceived action effects. *Psychological Research, 59.*

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001a). Codes and their vicissitudes. *Behavioral & Brain Sciences, 24*, 910-937.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001b). The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences, 24*, 849-937.

Howe, M. J. A., Davidson, J. W., Moore, D. G., & Sloboda, J. (1995). Are there early childhood signs of musical ability? *Psychology of Music, 23*, 162-176.

Hull, C. L. (1930). Knowledge and purpose as habit mechanisms. *Psychological Review, 30*, 511-525.

Hull, C. L. (1931). Goal attraction and directing ideas conceived as habit phenomena. *Psychological Review, 38*, 487-506.

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

Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S., & Passingham, R. E. (1994). Motor sequence learning: A study with positron emission tomography. *Journal of Neuroscience, 14*, 3775-3790.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science, 16*, 307-354.

Kalman, R., & Bucy, R. S. (1961). New results in linear filtering and prediction. *Journal of Basic Engineering (ASME), 83D*, 95-108.

Kelso, J. A. S. (1995). *Dynamic patterns: the self-organization of brain and behavior*. Cambridge, MA: MIT Press.

Koch, I., & Kunde, W. (2002). Verbal response-effect compatibility. *Memory & Cognition, 30*, 1297-1303.

Koeneke, S., Lutz, K., Wustenberg, T., & Jancke, L. (2004). Long-term training affects cerebellar processing in skilled keyboard players. *Neuroreport, 15*, 1279-1282.

Kornblum, S. (1992). Dimensional overlap and dimensional relevance in stimulus-response and stimulus-stimulus compatibility. In G. E. Stelmach

& J. Requin (Eds.), *Tutorials in Motor Behavior II* (pp. 743-777). Amsterdam, Netherlands: Elsevier Science Publishers.

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

Kornblum, S., & Lee, J. W. (1995). Stimulus response compatibility with relevant and irrelevant stimulus dimensions that do and do not overlap with the response. *Journal of Experimental Psychology: Human Perception & Performance, 21*, 855-875.

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

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

Kunde, W., Hoffmann, J., & Zellmann, P. (2002). The impact of anticipated action effects on action planning. *Acta Psychologica, 109*, 137-155.

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

Lai, S.-L. (2000). Increasing associative learning of abstract concepts through audiovisual redundancy. *Journal of Educational Computing Research, 23*, 275-289.

Laycock, T. (1840). *A treatise on the nervous diseases of women; comprising an inquiry into the nature, causes, and treatment of spinal and hysterical disorders*. London: Longmans.

Lehmann, A. C. (1997). The acquisition of expertise in music: Efficiency of deliberate practice as a moderating variable in accounting for subexpert performance. In I. Deliège & J. Sloboda (Eds.), *Perception and Cognition of Music* (pp. 161-187). Hove: Psychology Press Ltd.

Lotze, R. H. (1852). *Medicinische Psychologie oder die Physiologie der Seele [Medical psychology or physiolgy of the soul]*. Leipzig: Weidmann'sche Buchhandlung.

Luce, G. G. (1972). *Body time*. London: Temple Smith.

Magill, R. A. (2001). *Motor learning: concepts and applications* (6th ed.). Boston: McGraw-Hill.

Metzinger, T. (2000). The subjectivity of subjective experience: A representationalist analysis of the first-person perspective. In T. Metzinger (Ed.), *Neural Correlates of Consciousness: Empirical and Conceptual Questions*. Cambridge, MA: MIT Press.

Mudd, S. A. (1963). Spatial Stereotypes of 4 Dimensions of Pure-Tone. *Journal of Experimental Psychology, 66*, 347-352.

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

Müsseler, J., Steininger, S., & Wühr, P. (2001). Can action affect perceptual processing? *Quarterly Journal of Experimental Psychology A, 54*, 137-145.

Nakada, T., Fujii, Y., Suzuki, K., & Kwee, I. L. (1998). 'Musical brain' revealed by high-field (3 Tesla) functional MRI. *Neuroreport, 9*, 3853-3856.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*, 1-32.

Paivio, A. (1979). *Imagery and Verbal Processes*. Hillsdale, New Jersey: Lawrence Erlbaum.

Paivio, A. (1990). *Mental Representation: A Dual Coding Approach*. New York: Oxford University Press.

Palmer, C. (1997). Music performance. *Annual Review of Psychology, 48*, 115-138.

Palmer, C., Jones, R. K., Hennessy, B. L., Unze, M. G., & Pick, A. D. (1989). How is a trumpet known? The "basic object level" concept and perception of musical instruments. *American Journal of Psychology, 102*, 17-37.

Palmer, C., & Meyer, R. K. (2000). Conceptual and motor learning in music performance. *Psychological Science, 11*, 63-68.

Pantev, C., Roberts, L. E., Schulz, M., Engelien, A., & Ross, B. (2001). Timbre-specific enhancement of auditory cortical representations in musicians. *Neuroreport, 12*, 169-174.

Pearce, J. M. (1997). *Animal learning and cognition: an introduction* (2nd ed.). Hove: Psychology Press.

Pitt, M. A. (1995). Evidence for a central representation of instrument timbre. *Perception & Psychophysics, 57*, 43-55.

Prinz, W. (1990). A common coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action* (pp. 167-201). Berlin Heidelberg New York: Springer.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology, 9*, 129-154.

Proctor, R. W., Van Zandt, T., Lu, C. H., & Weeks, D. J. (1993). Stimulus-response compatibility for moving stimuli: Perception of affordances or directional coding? *Journal of Experimental Psychology: Human Perception and Performance, 19*, 81-91.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In

A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning, II: Current research and theory* (pp. 64-99). New York: Appleton-Century-Crofts.

Roitblat, H. L. (1987). *Introduction to comparative cognition*. New York: Freeman.

Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*, 382-439.

Sakai, K., Hikosaka, O., Miyauchi, S., Sasaki, Y., Fujimaki, N., & Putz, B. (1999). Presupplementary motor area activation during sequence learning reflects visuo-motor association. *Journal of Neuroscience, 19*, RC1-6.

Salinas, E., & Abbott, L. F. (1995). Transfer of Coded Information from Sensory to Motor Networks. *Journal of Neuroscience, 15*, 6461-6474.

Scheler, G., Lotze, M., Braitenberg, V., Erb, M., Braun, C., & Birbaumer, N. (2001). Musician's brain: balance of sensorimotor economy and frontal creativity [abstract]. *Society for Neuroscience Abstracts, 27*, 14.

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

Schön, D., Anton, J. L., Roth, M., & Besson, M. (2002). An fMRI study of music sight-reading. *Neuroreport, 13*, 2285-2289.

Schopenhauer, A. (1968). *Die Welt als Wille und Vorstellung [The world as will and representation]*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Schubö, A., Aschersleben, G., & Prinz, W. (2001). Interactions between perception and action in a reaction task with overlapping S-R assignments. *Psychological Research, 65*, 145-157.

Segalowitz, S. J., Bebout, L. J., & Lederman, S. J. (1979). Lateralization for reading musical chords: disentangling symbolic, analytic, and phonological aspects of reading. *Brain & Language, 8*, 315-323.

Shadmehr, R., & Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. *Science, 277*, 821-825.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394-403.

Simon, J. R. (1990). The effects of an irrelevant directional cue on human information processing. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus-response compatibility: an integrated perspective* (pp. 31-86). Amsterdam: North-Holland.

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

Sloboda, J. (1991). Musical expertise. In K. A. Ericsson & J. Smith (Eds.), *Towards a general theory of expertise* (pp. 153-171). Cambridge: Cambridge University Press.

Sloboda, J. (1996). The Acquisition of Musical Performance Expertise: Deconstructing the "Talent" Account of Individual Differences in Musical Expressivity. In K. A. Ericsson (Ed.), *The Road to Excellence*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Sloboda, J., Clarke, E. F., Parncutt, R., & Raekallio, M. (1998). Determinants of finger choice in piano sight-reading. *Journal of Experimental Psychology-Human Perception and Performance, 24*, 185-203.

Sloboda, J., Davidson, J., & Howe, M. J. A. (1994). Is everyone musical? *The Psychologist, 7*, 349-354.

Spence, K. W. (1956). *Behavior theory and conditioning*. New Haven, CT: Yale University Press.

Spitzer, M. (2002). *Musik im Kopf [Music in the head]*. Stuttgart: Schattauer.

Starkes, J. L. (1993). Motor experts: opening thoughts. In J. L. Starkes & F. Allard (Eds.), *Cognitive Issues in Motor Expertise* (pp. 3-16). Amsterdam: North-Holland.

Starkes, J. L., Deakin, J., Allard, F., Hodges, N. J., & Hayes, A. (1996). Deliberate Practice in Sports: What Is It Anyway? In K. A. Ericsson (Ed.), *The Road to Excellence* (pp. 81-106). Mahwah, New Jersey: Lawrence Erlbaum Associates.

Stewart, L., Henson, R., Kampe, K., Walsh, V., Turner, R., & Frith, U. (2003). Brain changes after learning to read and play music. *Neuroimage, 20*, 71-83.

Stewart, L., Walsh, V., & Frith, U. (2004). Reading music modifies spatial mapping in pianists. *Perception & Psychophysics, 66*, 183-195.

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

Stocker, C., & Hoffmann, J. (2004). The ideomotor principle and motor sequence acquisition: Tone effects facilitate movement chunking. *Psychological Research, 68*, 126-137.

Stocker, C., Sebald, A., & Hoffmann, J. (2003). The influence of response-effect compatibility in a serial reaction time task. *Quarterly Journal of Experimental Psychology A, 56*, 684-702.

Sussman, J. E. (1993). Focused attention during selective adaption along a place of articulation continuum. *Journal of the Acoustical Society of America, 93*, 488-498.

Sutton, R. S., & Barto, A. G. (1981). Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review, 88*, 135-170.

Tagliabue, M., Zorzi, M., Umilta, C., & Bassignani, F. (2000). The role of long-term-memory and short-term-memory links in the Simon effect. *Journal* *of Experimental Psychology: Human Perception & Performance, 26*, 648-670.

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

Tolman, E. C. (1932). *Purposive Behavior in Animals and Men*. New York: Appleton-Century-Crofts.

Toni, I., Krams, M., & Passingham, R. E. (1997). A comparison of externally and internally guided motor learning. *NeuroImage, 5*, 252.

Turvey, M. T., & Carello, C. (1995). Some Dynamical Themes in Perception and Action. In R. Port & T. van Gelder (Eds.), *Mind as Motion: Explorations in the Dynamics of Cognition*. Cambridge MA: MIT Press.

Wallace, R. A. (1971). S-R compatibility and the idea of a response code. *Journal of Experimental Psychology, 88*, 354-360.

Widmann, A., Kujala, T., Tervaniemi, M., Kujala, A., & Schröger, E. (2004). From symbols to sounds: Visual symbolic information activates sound represent. *Psychophysiology, 41*, 709-715.

Wise, S. P., & Murray, E. A. (1999). Role of the hippocampal system in conditional motor learning: Mapping antecedents to action. *Hippocampus, 9*, 101-117.

Wohlschläger, A. (2000). Visual motion priming by invisible actions. *Vision Research, 40*, 925-930.

Wolpert, D. M., & Flanagan, J. R. (2001). Motor prediction. *Current Biology, 11*, R729-R732.

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning [Review]. *Trends in Cognitive Sciences, 5*, 487-494.

Wolpert, D. M., & Kawato, M. (1998). Multiple Paired Forward and Inverse Models for Motor Control. *Neural Networks, 11*, 1317-1329.

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

Zakay, D., & Glicksohn, J. (1985). Stimulus congruity and S-R compatibility as determinants of interference in a Stroop-like task. *Canadian Journal of Psychology, 39*, 414-423.

Ziessler, M. (1994). The Impact of Motor Responses on Serial-Pattern Learning. *Psychological Research, 57*, 30-41.

Ziessler, M. (1998). Response-Effect Learning as a Major Component of Implicit Serial Learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 24*, 962-978.

Ziessler, M., & Nattkemper, D. (2001). Learning of event sequences is based on response-effect learning: Further evidence from a serial reaction task. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 27*, 595-613.

Ziessler, M., & Nattkemper, D. (2002). Effect anticipation in action planning. In W. Prinz & B. Hommel (Eds.), *Common Mechanisms in Perception and Action* (pp. 645-672). Oxford: Oxford University Press.

Ziessler, M., Nattkemper, D., & Frensch, P. A. (2004). The role of anticipation and intention in the learning of effects of self-performed actions. *Psychological Research, 68*, 163-175.

Der Autor

Name:	Ulrich Christian Drost
Geboren:	27. September 1976 in Erlangen, Bayern

Promotion

01 / 2002 bis 02 / 2005	Promotion bei Prof. Wolfgang Prinz am Max-Planck-Institut für Kognitions- und Neurowissenschaften, München

Studium

09 / 1997 bis 11 / 2001	Studium im Fach Psychologie (Diplom) an der Friedrich Alexander-Universität, Erlangen ▪ Schwerpunkte: Diagnostik, Organisationspsychologie, Rechtspsychologie ▪ Thema der Diplomarbeit: „Control of Multiple Actions“

Wehrdienst

09 / 1996 bis 06 / 1997	Wehrdienst bei der FlaRak-Gruppe 23 (Manching) ▪ Letzter Dienstgrad: Obergefreiter

Schulausbildung

09 / 1987 bis 06 / 1996	Ehrenbürg-Gymnasium, Forchheim (Abitur)
09 / 1983 bis 07 / 1987	Grundschule in Igensdorf, Bayern