

# Perception and Action Planning

Wolfgang Prinz

*Max-Planck-Institute for Psychological Research and Institute for Psychology, University of Munich, Munich, Germany*

A new framework for the understanding of functional relationships between perception and action is discussed. According to this framework, perceived events and planned actions share a common representational domain (common-coding approach). Supporting evidence from two classes of experimental paradigms is presented: induction paradigms and interference paradigms. Induction paradigms study how certain stimuli induce certain actions by virtue of similarity. Evidence from two types of induction tasks is reviewed: sensorimotor synchronisation and spatial compatibility tasks. Interference paradigms study the mutual interference between the perception of ongoing events and the preparation and control of ongoing action. Again, evidence from two types of such tasks is reviewed, implying interference in either direction. It is concluded that the evidence available supports the common coding principle. A further general principle emerging from these studies is the action effect principle—that is, the principle that cognitive representations of action effects play a critical role in the planning and control of these actions.

## INTRODUCTION

In this paper, I present a framework for understanding some aspects of the functional relationships between perception and action. It is mainly a framework for action control and action planning, but to some extent it is a framework for perception as well. Its core contention is that perceived events and planned actions share a common representational domain. This is the *common coding approach* to perception and action. In the following, I try to show that this approach is more powerful than the traditional approach to the relationship between perception and action,

---

Requests for reprints should be addressed to Wolfgang Prinz, Max-Planck-Institute for Psychological Research, Leopoldstrasse 24, D-80802 Munich, Germany.

This paper is based on the Broadbent Lecture delivered at the 8th ESCOP Conference, Rome, September 1995.

which invokes separate rather than common coding. First, I contrast these two approaches and then discuss pertinent evidence from a number of experiments.

The lower part of Fig. 1 (unbroken lines) depicts the major functional components believed to underlie perception and action control. The afferent part is shown on the left-hand side, where events in the environment lead to patterns of stimulation in sense organs. These stimulus patterns generate sensory codes in the brain. In the efferent part on the right-hand side, the activity goes top-down. It starts with motor codes in the brain that again represent certain patterns of excitation in peripheral effector organs and have the power to activate them. If this occurs, the muscles contract or relax and the organism acts in one way or another.

How can sensory codes and motor codes "talk" to each other? It is generally believed that they cannot talk to each other directly and that some transformation or *translation* is required instead. This is because they are incommensurate in terms of their contents. Because sensory codes stand for patterns of stimulation in sense organs and motor codes for patterns of excitation in muscles, there is no obvious way in which these two entities could be compared or matched.

This can be illustrated by a reaction time experiment where the subject has to press one of two keys (e.g. with the left *vs* right hand) in response to one of two lights (e.g. red *vs* green). In this case, the sensory codes

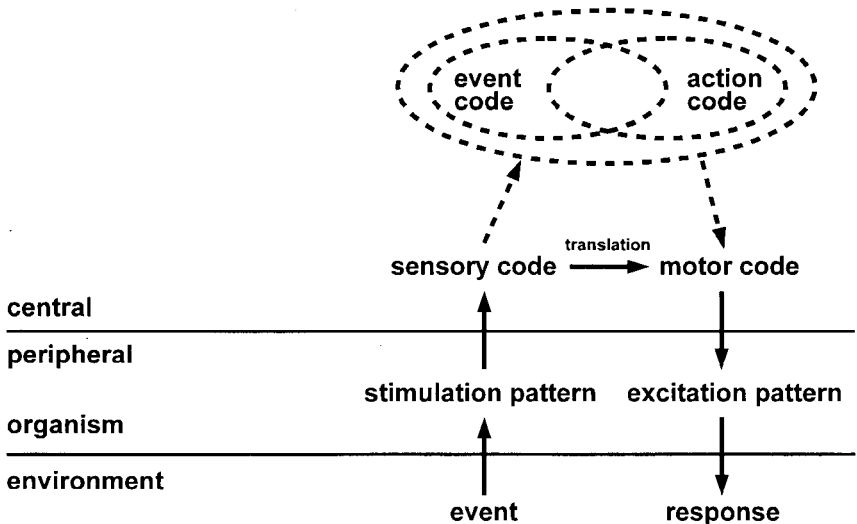


FIG 1. Relationships between perception and action. Lower part (unbroken lines): Separate coding (sensory codes, motor codes and translation between them). Top part (broken lines): Common coding (event codes, action codes and induction between them).

represent two different colours and the motor codes the different hands. A task such as this requires a mapping device that *translates*, as it were, colours into hands (more precisely, codes for colours into codes for the muscles that control the hands). In the rich body of reaction-time research that has emerged over the last 30 years, the concept of translation has indeed become one of the most prominent theoretical notions to account for the internal operations underlying the mapping of responses to stimuli. It was originally introduced by Welford (1968) and later taken up by many other authors, including Sanders (1980) and Massaro (1990).

A beautiful pictorial illustration of this view is available from the *Traitée de l'homme* by René Descartes (1664; see Fig. 2a). Descartes stressed the incommensurability between afferent and efferent processes in two ways. First, he used different metaphors for afferent and efferent conduction. Afferent conduction occurs in terms of mechanical movement (by pulling at certain nerve fibres), whereas efferent conduction occurs in terms of hydraulic pressure (by dispensing neural liquid). Second, to account for the mapping between afference and efference, he used another metaphor that stresses the indirect nature of that mediation. This mapping occurs in the pineal gland, which moves in response to afferent mechanics, thereby discharging a specific pattern of nerve fluid. This pattern, when fed into the efferent hydraulics, eventually causes a specific movement. In a way, then, the pineal gland acts as an interpreter, translating afferent mechanics into efferent hydraulics.

This has since been the standard framework for the linkage between perception and action. It invokes separate coding, because it considers perception and action as two entirely different functions of mental life. On the other hand, nineteenth-century authorities such as Lotze, James and Münsterberg stressed the continuity between the two by pointing to the fact that there are no fundamental differences in their representational contents. This is neatly illustrated by another famous picture from the history of philosophy, which is taken from the *Analyse der Empfindungen* by Ernst Mach (1922; Fig. 2b). In discussing this picture, Mach's main point was that visual perception does not provide us with an *a priori* distinction between body and environment. Rather, the body is perceived in exactly the same way as the environment, and the distinction between the two must be based on a non-visual criterion. Mach's suggestion was that we do it by the criterion of voluntary control, attributing to the body what we can control by will and to the environment what we cannot.

This criterion applies not only to static objects, but even more to dynamic events. Mach, while lying on his sofa, would perceive an external event when, for example, a bird flew past his window. Yet, when his fingers moved, he would perceive an action of his own. From this

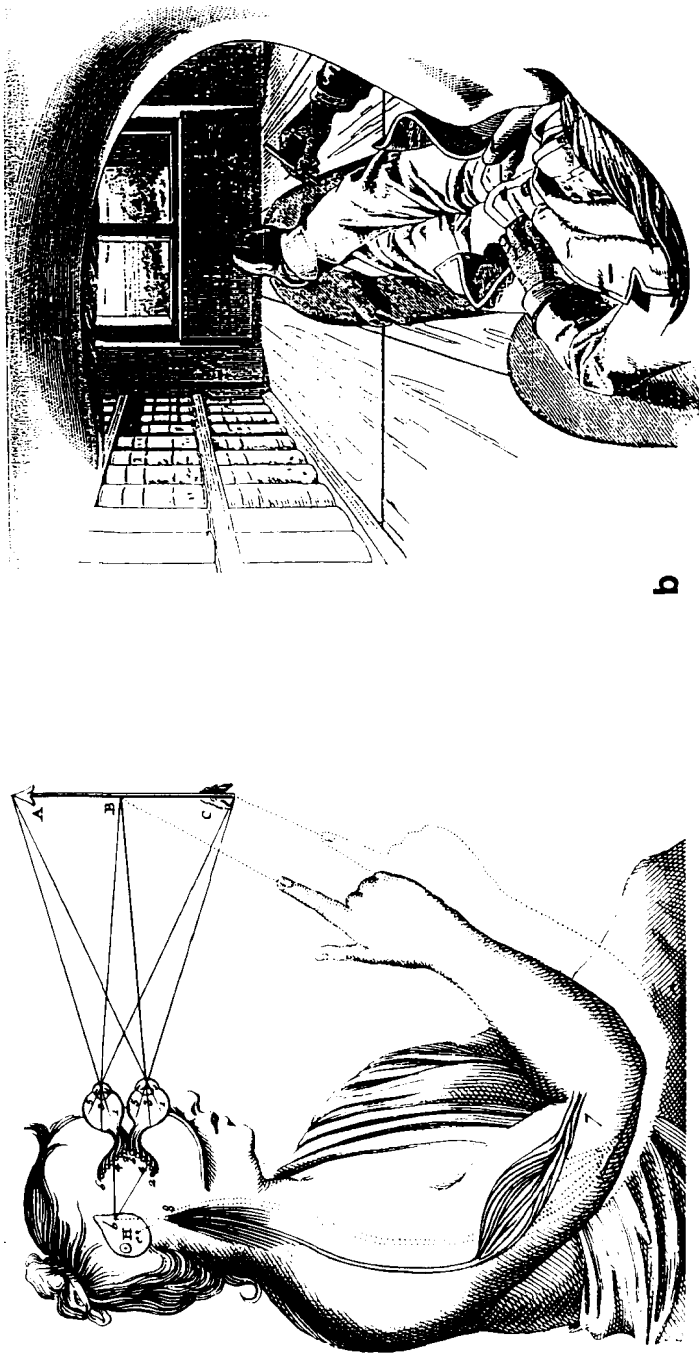


FIG 2. Two classical illustrations of two views of perception and action. (a) Descartes' description of afferent and efferent processes and the central mediation between them implies a separate-coding type of approach (Descartes, 1664). (b) Mach's analysis of the content of conscious experience implies a common-coding type of approach (Mach, 1922).

perspective, actions are represented in the same way as external events, the only difference being that they can be controlled by will. As a consequence, the planning of actions requires no translation between incommensurate entities. Rather, it seems to imply something like conceiving future events that follow—and are continuous with—ongoing events. In this view, events and actions appear to be made of the same stuff.

The *common-coding approach* holds that this continuity applies not only to phenomenal experience but to functional mechanisms as well. This is illustrated by the upper part of Fig. 1 (broken lines), where another level of representation is added to the previous basic scheme. “Event codes” and “action codes” should be considered the functional basis of percepts and action plans, respectively. It is held that they share the same representational domain and are therefore commensurate. As a consequence, they can be more or less similar to each other, as indicated by their graphic overlap.

Related views have from time to time emerged in various fields of study, such as, for example, movement imitation (Koffka, 1925; Meltzoff, Kuhl, & Moore, 1991), language perception and production (MacKay, 1987), speech perception (Liberman, Cooper, Shankweiler, & Studdert-Kennedy, 1967), motion perception (Viviani, 1990) and—perhaps most closely related to the present approach—stimulus–response compatibility (Kornblum, Hasbroucq, & Osman, 1990). In these approaches, the notion of common codes for perception and action has been elaborated in two ways. Motor theories of speech and motion perception try to discover the traces of action in perception. Conversely, theories of imitation and compatibility try to establish the traces of perception in action.

The following experiments were designed in the spirit of this approach. Though they use several different tasks, there is one feature common to all of them: The events presented as stimuli *share certain properties* with the actions to be performed in response to them. The purpose of the experiments was to study to what extent the similarity that arises from these shared properties affects both perception and action planning. Two types of tasks are considered: induction and interference paradigms.

Induction paradigms study the way in which perceptual events *induce*, or *enhance*, actions by virtue of similarity. Induction seems to require some form of common coding, and the study of such tasks should therefore help to elaborate the details of the approach. Two such paradigms will be discussed, one where the supposed common coding refers to the temporal domain, and another where it refers to the spatial domain. Interference paradigms study how perceptual events and actions *interfere with each other* by virtue of similarity. The basic idea here is that some form of interference must result when perception and action make use of the same codes at the same time. Action should then be impaired by

concurrent perception, and perception should be impaired by concurrent action. This prediction is not compatible with separate coding and should therefore provide a strong test of the feasibility of the common coding approach.

## INDUCTION PARADIGMS

Two major issues will be raised with respect to tasks where perception induces action by virtue of similarity: (1) Which components of actions are critical for similarity-based induction effects to occur? (2) What kinds of shared properties are involved and what is their representational status?

### Sensorimotor Synchronisation

For synchronisation tasks, subjects are required to perform certain acts in synchrony with certain external events. In the simplest version of the task, subjects are presented with a sequence of isochronous auditory clicks and are required to tap with their index fingers on a silent electrical contact plate, such that tap onsets coincide with click onsets. Performing this task is a child's game. Action is immediately captured by perception, and everybody believes that he or she is perfect. In fact, however, hardly anybody is. Instead, tap onsets tend to be ahead of click onsets by about 30–50 msec. This effect is called “negative asynchrony” and it has been observed in a number of previous studies (Aschersleben & Prinz, 1995; Dunlap, 1910; Fraisse, 1980; Kolers & Brewster, 1985; Vos, Mates, & van Kruysbergen, 1995).

The negative asynchrony raises the question of what exactly it is in the tap that gets synchronised with the click. To obtain possible candidates one may consider the functional components involved in individual clicks and taps. A click's functional history is very simple. It starts with a “click” in the ear and ends with the formation of a sensory code in the brain. A tap's functional history is somewhat more complex. It starts with a motor code in the brain, which activates the muscles, so that the finger's “tap” is actually performed. Finally, this event gets fed back to the brain, where a sensory feedback code is formed, on which the tap's perception is based.

Given these components, the occurrence of negative asynchronies could suggest that taps get synchronised with clicks in terms of the sensory codes that represent their perceptual effects. This presupposes that one can reasonably argue that the delay at which these codes lag behind the respective stimulus events is larger for taps than for clicks, such that to

let the taps' sensory effects coincide with the clicks' sensory effects, the taps must lead the clicks by a certain amount of time. An argument such as this can indeed be established in two different versions: code generation and onset computation (Fig. 3).

The Code-generation Hypothesis, which was first proposed by Paillard (1949), relies on the times required for generating brain codes from sensory stimulation (Fig. 3a). Paillard reasoned that the brain might synchronise taps with clicks by way of superimposing the two sensory codes in time. Furthermore, he argued that when tap codes coincide with

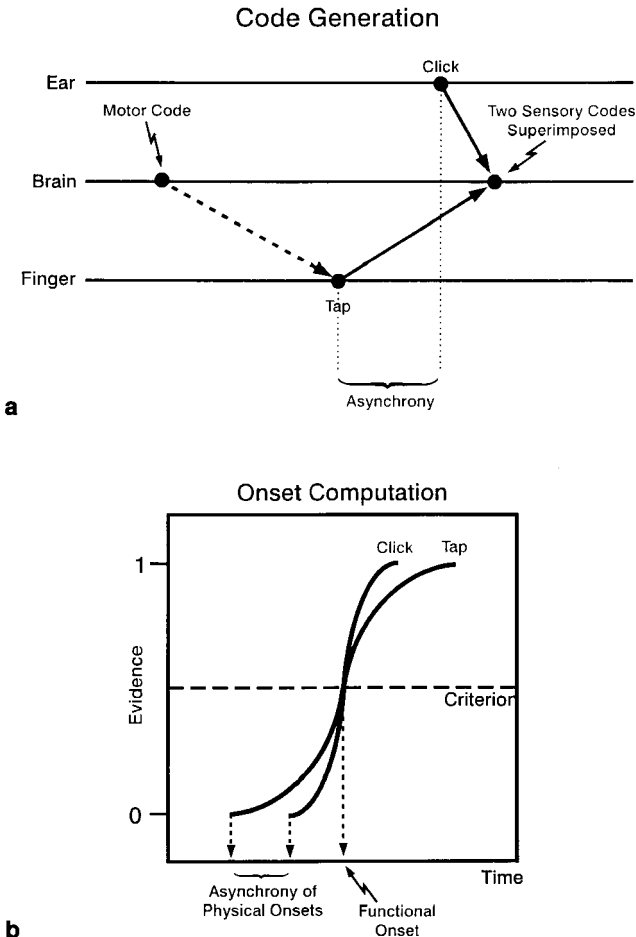


FIG 3. Two explanations for the occurrence of negative asynchronies in sensorimotor synchronisation tasks. (a) Code-generation Hypothesis; (b) Onset Computation Hypothesis (see text for details).

click codes in the brain, the taps and the clicks themselves cannot coincide in the world, because the times required to generate the two codes are different for the two modalities, as is indicated in Fig. 3a by the different lengths of the two converging arrows. As the pathway between finger and brain is longer than between ear and brain, the tap must lead the click *in the world* to warrant coincidence of the two codes *in the brain*.

The Onset Computation Hypothesis is closely related to the notion of perceptual centres, as introduced by Morton and his co-workers in the area of speech processing (Morton, Marcus, & Frankish, 1976). It holds that the functional dating of events is actively computed on the basis of event-inherent features such as intensity, duration, onset abruptness, and so forth, rather than being passively derived from central arrival times of their codes. When one adopts this view, taps may differ from clicks in several respects. First, their perceived onsets are less clearly defined. For instance, as the taps' tactile onsets are preceded by the fingers' downward movements, they may appear less abrupt compared to the onsets of auditory clicks. Second, tap duration is usually much longer compared to click duration (usually between 50 and 200 msec as compared to 10 msec). Third, taps and clicks may differ in intensity or size of the population of cortical neurons activated by them: Taps might be weaker and/or activate a smaller neuron population than clicks. As a consequence, taps may be dated later than clicks (relative to their physical onsets), and again tap onsets would have to lead click onsets in the world to achieve perceived synchrony.

This is illustrated in Fig. 3b in terms of two normalised accumulation functions that represent the build-up of sensory codes over time—a steep one for clicks and a more gradual one for taps. Each of these functions has a true onset (where the accumulation starts) and a computed onset (where the evidence accumulated reaches a preset criterion). To achieve synchrony, they must intersect in their functional onsets. This, in turn, requires that taps must lead clicks in terms of their true onsets. Furthermore, if differences in code generation times play no role, this must apply to both codes in the brain and events in the world.

It should be noted, however, that there is a common principle underlying both code generation and onset computation that we call the "action-effect principle". For the task at hand, it implies that it is the taps' sensory or perceptual effects that get synchronised with the sensory or perceptual effects arising from clicks.

*Auditory Feedback.* In one of the experiments conducted by Aschersleben (1994; Aschersleben & Prinz, 1995), a short auditory feedback tone was administered on each tap. This tone was always triggered by the



tap's physical onset. In this case, each tap is represented by *two* feedback signals: the *intrinsic* tactile/kinaesthetic signal and the additional *extrinsic* auditory signal.

Consider code generation first. A strong version of this hypothesis would predict that it is now the sensory code of the extrinsic feedback tone that takes the role of the tap code that gets superimposed on the click code—simply because it arrives earlier in the brain than the intrinsic feedback signal, due to its shorter pathway. In this case, the asynchrony should completely disappear because the generation of the auditory tap code should require exactly the same time as the generation of the auditory click code. For a weaker version of the hypothesis, one could also think of the possibility that the two tap codes enter into a joint event code whose effective onset is later than the early component, but earlier than the late component. In this case, one would expect that additional auditory feedback serves to reduce the size of the asynchrony without bringing it down to zero.

The same result can be expected under onset computation. Adding a short extrinsic feedback tone to the intrinsic kinaesthetic and tactile feedback signal should alter the pattern of the tap-related feedback signal. For instance, if one assumes that the different sources of feedback are entered into a joint multimodal code, it is reasonable to expect that this joint code exhibits a steeper accumulation function compared to the control task, where no extrinsic feedback is provided. This should also reduce the asynchrony.

This is what we observed in several experiments. When we provided additional auditory feedback, the size of the asynchrony was clearly reduced but—no less clearly—it did not go down to zero. In further experiments, we imposed some delay on the delivery of the auditory feedback tone. The prediction is straightforward for both code generation and onset computation. Any delay in the delivery of the extrinsic signal should also delay the onset of the tap's functional representation in the brain relative to its physical onset in the world. This should, in turn, lead to an increase in the negative synchrony, provided that the delay does not exceed the integration zone within which the feedback components can be integrated into a joint code.

This is exactly what the data showed (see Fig. 4; Aschersleben & Prinz, 1997). The delivery of the auditory feedback signal could be delayed up to 70 msec without the subjects becoming aware of it, suggesting that the zone of feedback integration is at least that large. Within this range, a linear relationship between the size of the asynchrony and the delay of feedback delivery was obtained. This function suggests a linear combination of the two feedback components, with the slope indicating their relative weights.

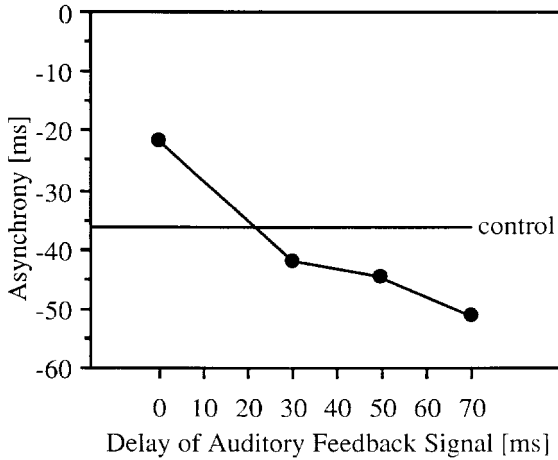


FIG 4. Results from a synchronisation experiment with delayed auditory feedback (Aschersleben, 1994). The horizontal line indicates a control condition where the auditory feedback signal was absent. When the delay is zero, the amount of asynchrony is reduced relative to this control, but it increases monotonically with increasing delays.

*Tap Pressure.* In a recent experiment, Gehrke (1996) explored the role of the dynamics of the tapping movement (see also Gehrke, Aschersleben, & Prinz, in prep.). This was done in an attempt to distinguish between code generation and onset computation. Consider what must be expected when the force that the tapping exerts on the contact plate is systematically manipulated. As far as code generation is concerned, it is difficult to see how a manipulation such as this could affect the time for the generation of central codes. For example, as the evidence available from studies with event-related potentials (ERPs) indicates, the latency of somatosensory-evoked brain potentials (SEPs) is not substantially affected by stimulus intensity. Studies with electrical stimulation of the median nerve have shown that SEPs remain invariant with stimulus intensity (Beric, 1992; Lesser, Köhle, & Lüders, 1979) and that stimulation of the posterior tibial nerve yields a slight decrease in the latency of early SEP components of about 10 msec (Rappaport, Portillo, & Leonard, 1992). For mechanical stimulation of the skin, SEP latencies also tended to decrease with increasing stimulus intensity (Franzén & Offenloch, 1969; Hashimoto, Gatayama, Yoshikawa, & Sasaki, 1992a, b; Johnson, Jürgens, Kongehl, & Kornhuber, 1975). It is generally believed that these slight effects are due to the speeding up of peripheral receptor processes rather than an increase in conduction speed.

Conversely, as far as onset computation is concerned, the computation of functional tap onsets should clearly be affected by this manipulation.

Increasing tap force should increase the intensity of the tactile feedback signal, which should, in turn, increase the size of the population of neurons involved in its central representation. This is in line with ERP studies indicating that stimulus intensity is reflected in the amplitude of cortically evoked potentials (Beric, 1992; Franzén & Offenloch, 1969; Hashimoto et al., 1992a, b; Johansson, 1979).

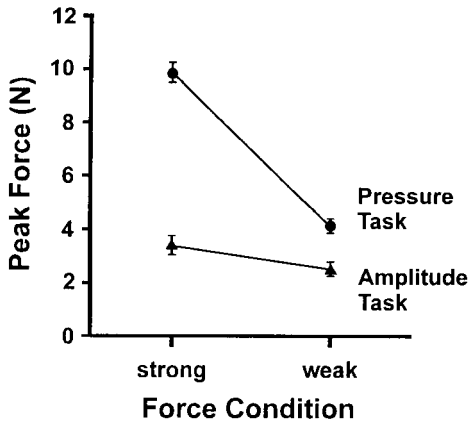
According to the logic of onset computation, this should reduce the delay at which functional onsets lag behind physical onsets. As a consequence, the size of the asynchrony should decrease when tap pressure increases. In principle, it might even go down to zero, or become positive, when the tactile feedback signal equals or exceeds the auditory pacing signal in terms of intensity and/or abruptness.

Gehrke manipulated tap pressure in two ways. First, subjects were instructed to tap with small versus large amplitudes (amplitude task). This instruction can easily be followed after some practice. Due to the close coupling between kinematics and dynamics, the difference in amplitude reflects the corresponding difference in the resulting force exerted on the contact plate. This force is higher for large than for small amplitudes.

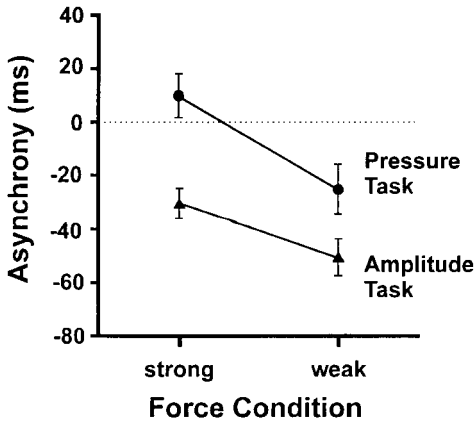
Second, a pressure task was administered to the same subjects. This was an isometric tapping task, where the finger was required to press on the contact plate without moving at all and to modulate that pressure periodically, such that peak pressure coincided with the auditory pacing stimuli. The pressure task was introduced because it allows for substantially higher peak forces compared with standard tapping. Subjects had to tap with high versus low pressure in two conditions of the task, which was again easy to do after some practice.

Figure 5a shows the mean peak forces obtained in the four conditions of the two tasks. As can be seen, peak forces are substantially larger in the pressure than in the amplitude task. Moreover, practice and instructions were effective in both tasks. Under high-amplitude and high-pressure instructions, peak force was larger than with low-amplitude and low-pressure instructions. The critical results are the asynchronies observed in the same two tasks. They are shown in Fig. 5b. First, in the amplitude task, we see negative asynchronies throughout, as well as a modulating effect of peak force that points in the expected direction. With strong force, the amount of asynchrony is smaller than with weak force. Second, in the pressure task, we see much smaller asynchronies, ranging around zero. Again, they are strongly modulated by instructions: they are negative with weak force and become positive with strong force.

This picture is incompatible with code generation. There is no way that code generation can account for the positive asynchronies in the pressure task. It is, however, compatible with onset computation. The fact that the



a



b

FIG 5. Results from an experiment with two manipulations of tap pressure (amplitude task and pressure task). (a) Peak forces obtained in the two force conditions of the two tasks; (b) asynchronies obtained in the same four conditions. Data from Gehrke (1996).

pattern of asynchronies mirrors exactly that of peak forces suggests that we should prefer onset computation to code generation.

*Summary.* In summary, these results support the action-effect hypothesis. They support the notion that synchrony is established in a common-coding domain where clicks are represented by sensory codes and taps by sensory feedback or effect codes. Moreover, the last experiment suggests discarding code generation and relying on onset computation. It seems

that synchronisation does not rely on the passive recording of arrival times, but rather on the active computation of functional event onsets.

## Stimulus-Response Compatibility

Everybody knows that choice reaction times are fast when the spatial arrangement of the response keys corresponds to the spatial arrangement of the stimuli to be responded to, compared to control conditions where no such correspondence exists. These compatibility effects tend to be substantial and stable and can be observed in each subject.

Various models have been suggested to account for these effects, and most of them draw, in one way or another, on the spatial similarity between stimulus and response arrangements. For instance, in Kornblum's terms, compatibility effects arise to the extent that stimulus and response sets share some features, so that their cognitive representations overlap (Kornblum et al., 1990; Kornblum, 1992, 1994; Kornblum & Lee, 1995). In the case of compatible stimulus-response (S-R) assignments, stimuli will prime the correct responses, whereas in incompatible assignments, this priming will always refer to the incorrect response and therefore lead to interference.

This view implies a common representational domain, and it invokes common codes for perception and action. It can readily be applied to cases where the similarity between stimuli and responses is obvious. It is, however, less clear how it can account for tasks that show clear preferences for one over the other assignment, though there is no obvious dimensional overlap between stimulus and response sets. Two such examples will be discussed. The first demonstrates that common coding may not be restricted to physical features but may include more abstract properties of actions as well. The second example shows that common coding may not only apply to actions themselves, but to action goals as well.

*Orthogonal Dimensions.* A puzzling task, where stable preferences for certain S-R assignments arise in the absence of any obvious similarity between stimulus and response sets, has recently been studied by Lippa (1996a,b). This task requires the mapping of two stimuli onto two responses. On each trial the stimulus can appear to the left or right of the fixation point. To respond, subjects have to press a response key located above or below a central home key from which the response starts.

The puzzling finding Lippa replicated from an earlier study by Bauer and Miller (1982) is that the two hands "prefer", as it were, different S-R assignments. The left hand is fast when the task requires an upward response to a left stimulus and a downward response to a right stimulus.

The same hand is about 40 msec slower with the reverse assignment of the same responses to the same stimuli. Conversely, for the right hand, this reverse assignment, which requires going down when the stimulus appears to the left and up when the stimulus appears to the right, is faster than its counterpart. This puzzling pattern applies not only to response times but to error rates as well.

At first glance, it is difficult to see how dimensional overlap could account for this effect, because stimulus and response dimensions are literally orthogonal under all conditions. As a possible solution, Lippa suggested that the hand's resting position on the home key might serve as a reference that generates an implicit left–right coding of the response set. As is illustrated in Fig. 6, an upward move goes to the right for the right hand and to the left for the left hand, whereas a downward movement goes left for the right hand and right for the left hand. When such “referential coding” of the response set is taken into account, we should indeed expect compatibility effects due to dimensional overlap that must be countercurrent for the two hands—exactly the pattern the data show.

To test this account, Lippa ran two further experiments. In the first experiment, subjects were instructed to hold their responding hands *in line with* the vertically oriented response keys. Under this condition, where the hands' movements are restricted to the up/down dimension, it

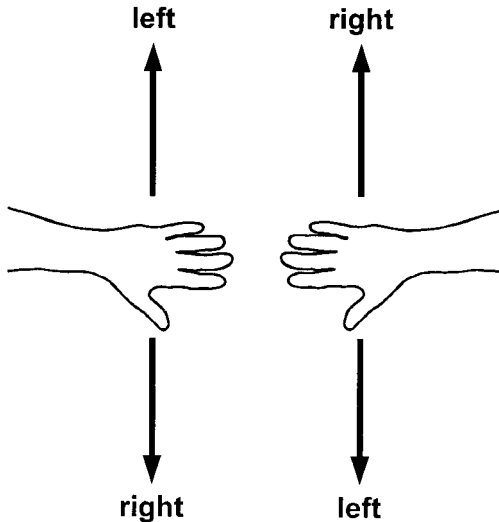


FIG 6. Referential coding. The coding of upward/downward movements may be different for the two hands. For the left hand, an upward movement goes to the left and a downward movement to the right; for the right hand, the reverse holds (after Lippa, 1996a,b).

is not obvious how the referential coding of responses on the left/right dimension could be preserved. Therefore, the difference between the preferences of the two hands should appear. This was, in fact, observed—identical results were obtained for the two hands.

In the second experiment, the response panel was located either to the left or to the right, relative to the subject's sagittal midline and the responding hand was held at a 45–90° angle to the line of the response keys. With this arrangement, referential coding should be identical for the two hands. Instead, the location of the response panel should now determine which of the two assignments is compatible versus incompatible, and this should hold for both hands. Again, this is exactly the pattern the data showed.

In summary, Lippa's findings suggest (1) that compatibility effects can indeed be accounted for in terms of dimensional overlap (i.e. shared features between stimulus and response sets), and (2) that these effects may not arise at the level of basic sensory features of stimulus and response sets, but rather apply to high-level or cognitive representations of stimulus events and actions. It is only at such higher levels of representations that referential coding can come into play, because it requires coding the spatial locations of events in a way that takes the position of the acting limb into account.

*Action Goals.* In Lippa's experiments, it remains open which aspects of the actions are critical for the compatibility effects to occur. What exactly is it in the action that gets coded as left or right? There are various candidates, such as the relative position of the perceived target key, the direction of the planned movement, or the intended movement's goal location. Some years ago, Hommel ran another series of experiments pertaining to this issue (Hommel, 1990, 1993). They concerned the role of intentions in the mechanisms underlying the Simon effect.

In the Simon task, there are two response keys that are assigned to the left and the right hand. Two possible stimuli can occur (e.g. a high-pitched or a low-pitched tone) and the task requires the subject to press one of two keys in response to one stimulus and the other key in response to the other. Stimulus identity (high vs low pitch) is therefore the relevant dimension on which the stimuli vary. In addition to this, stimulus position is varied as a further dimension (the tone coming from a loudspeaker on the left- vs the right-hand side). This dimension is irrelevant because subjects are instructed to ignore it completely. Yet it is usually observed that performance is clearly superior in the case of correspondence between stimulus and response position, as compared to the case of non-correspondence. This is the Simon effect. Obviously, one can again think of the effect in terms of shared properties on the afferent and

the efferent side. Responses are fast (and correct) when stimulus and response share the same spatial position, and slow (and less correct) when positions are different.

There are various candidate levels at which such correspondence could be effective. First, when, for example, the left-hand key is pressed in response to a stimulus in the left visual field, the response comes from an effector attached to the same body hemisphere to which the stimulus is presented. This may be summarised under the notion of “anatomical correspondence”. Second, correspondence may be claimed between the location at which the stimulus is presented and the location at which the action is performed. This is called “locational correspondence”.

In the standard version of the task, locational correspondence is completely confounded with anatomical correspondence. There is, however, an easy way to unconfound them. Consider what must be expected when subjects perform the task with their hands crossed (Simon, Hinrichs, & Craft, 1970; Wallace, 1971). If anatomical correspondence was the critical factor, one would have to expect that the effect is then completely reversed. In contrast, it should be unchanged if locational correspondence counts. The results of the experiments are clearly in favour of locational correspondence, implying that the effect is grounded in a high-level coding domain where actions and events are localised in extracorporeal space.

Hommel's suggestion was that one might even go one step further and distinguish between the location of the action itself and the location of the action's goal, that is, its intended effect. A related notion had already emerged in an experiment by Riggio, Gawryszewski and Umiltà (1986), in which subjects operated response keys with crossed sticks, to the effect that the action and its effect could be separated. In the standard paradigm where the action is a simple keypress, there is again no way to separate the two, because the action (pressing the key) shares the same location with its intended effect (the key being pressed).

What happens when one dissociates the two? In this case, “intentional correspondence” might be effective as well. This notion refers to the correspondence between the locations of the stimulus and of the response's intended effect. In Hommel's experiments, the hands were left uncrossed. What was crossed instead were the relations between the locations at which the two actions were performed and the locations of two lights that were switched on by these actions (Fig. 7). For example, when the left-hand response key was pressed in response to a low-pitched tone, light *l* went on on the right-hand side. Conversely, light *h* went on on the left-hand side when the right-hand response key was pressed in response to a high-pitched tone. The task was a standard Simon task,



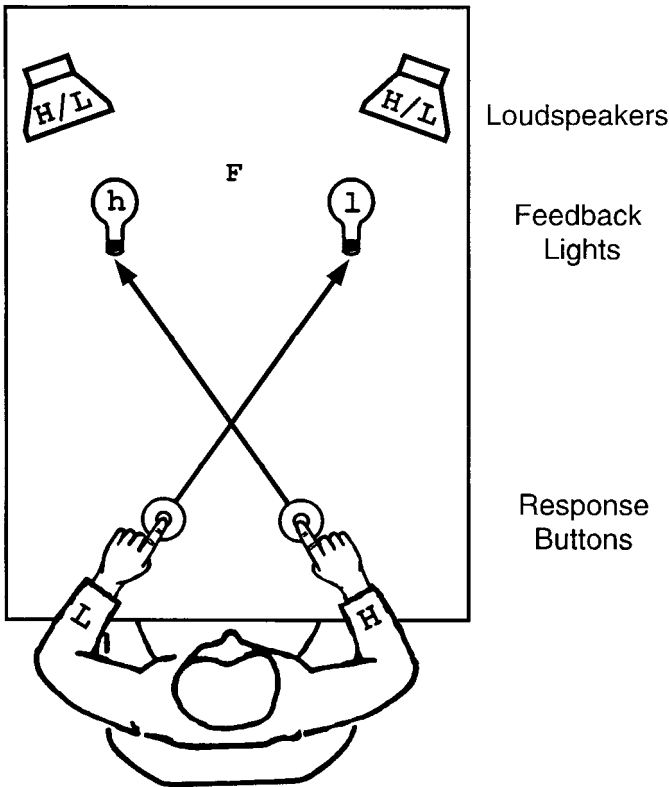


FIG 7. Hommel's modification of the Simon paradigm (Hommel, 1990, 1993). The basic task is a Simon task in which high- and low-pitched tones (H/L) can be presented on a loud-speaker on the left- or right-hand side (relative to the fixation point F). In the example shown, the left hand responds to low-pitched tones and the right hand to high-pitched tones, and each response turns on a light on the opposite side.

with pitch as the relevant stimulus dimension and location as the irrelevant stimulus dimension.

There were two groups of subjects. In the control group, subjects were instructed to ignore the lights and press the left key in response to low-pitched tones and the right key in response to high-pitched tones. In this group, the normal Simon effect was obtained. In the experimental group, the same task was administered under different instructions. This time, subjects were instructed to illuminate as quickly as possible light *l* on the right-hand side in response to low-pitched tones and light *h* on the left-hand side in response to high-pitched tones. Under these instructions, the response keys were simply not mentioned. The idea was to implement

two different intentional sets: a set that refers to the action itself (keypress) in the control group and one that refers to the action's remote effect (light onset) in the experimental group.

The results showed that the location of the action itself became irrelevant under this latter intentional set. Here, the critical factor was the correspondence between the location of the stimulus and the action's intended effect: responses were fast when the stimulus and the light shared the same location (despite the fact that the location of the action itself was always on the opposite side). Conversely, responses were slow when the stimulus and the light appeared in different locations (despite the fact that in this case the location of the action itself was always on the same side as the stimulus).

*Summary.* In summary, our observations from spatial compatibility tasks suggest two related conclusions. First, it seems that the observed effects do not arise at the level of low-level sensory features, but at a higher perceptual level of representation, where they emerge on the basis of shared spatial properties of perceived events and planned actions. Second, by demonstrating that compatibility may be effective not only between stimuli and responses but also between stimuli and response effects, we obtain strong further support for the action-effect principle.

## INTERFERENCE PARADIGMS

The interference paradigms discussed in the following section have been designed to study possible conflicts between perception and action. Unlike separate coding, common coding leads one to expect such conflicts. Given the functional logic of common coding, interference paradigms study what happens when perception and action make use of the same codes at the same time. The basic idea is that a code that is busy with perception cannot at the same time be functional in action planning and action control, and the same should hold in the reverse direction. This is called the Interference Hypothesis for perception and action.

More specifically, we can distinguish between a general and a specific version of the hypothesis. The general version relies on the assumption that the involved codes exhibit some degree of refractoriness such that they need some time to recover after each activation. The specific version rests on the belief that the system must have a way to distinguish between the use of codes for perception and action and, more importantly, a way to protect one from the other. For instance, when a given

code is being used for action control, this activity should not alter the perception of ongoing related events. Conversely, the perception of events should not have any effect on the preparation and control of ongoing related actions. Yet, irrespective of this difference, both versions join in the prediction that similarity-based interference must be obtained in the case of conflict between perception and action.

## Perception Impairs Action Planning

What happens when a stimulus activates a given code at a moment when this code is busy with action planning? The solution to this conflict might depend on the relative importance of the two. If the action is important and the new information is not, it can perhaps be ignored altogether. However, if the new information cannot be ignored, action planning should be impaired. Therefore, when a particular code is busy with action planning and a compatible stimulus that addressed this code is presented at the same time, we should expect interference, due to the similarity between the stimulus event and the ongoing action. The action must be delayed, because the code that controls it is, for a short time, required by the perceptual side.

This conflict can be seen in a paradigm developed by Müsseler (1995). In this task, each trial has two parts: memorisation and execution. In the memorisation part, an array of four arrows is presented for 3 sec. The arrows point to the left or the right, specifying a corresponding sequence of left and right keypresses, to be performed by the subject later on, following the arrows from left to right. Subjects are instructed to memorise the arrows and prepare the corresponding action sequence as efficiently as they can. After 3 sec the instruction stimulus disappears and subjects can start the action sequence whenever they want. Once they have started, they are required to work through it as fast as possible.

The critical feature of the task is the additional presentation of another arrow. This presentation is triggered by the first keypress. The additional arrow cannot be ignored because the instructions require the subject to respond to it with an additional keypress. This additional keypress has to be attached as the fifth link to the chain of the four keypresses programmed earlier.

The question was whether the processing of the additional arrow would interfere with the ongoing planning of the second keypress and whether that interference would depend on the compatibility between the arrow and the action. Somewhat paradoxically, we must expect that interference arises in the compatible but not in the incompatible condition, yielding an “inverted compatibility effect”.

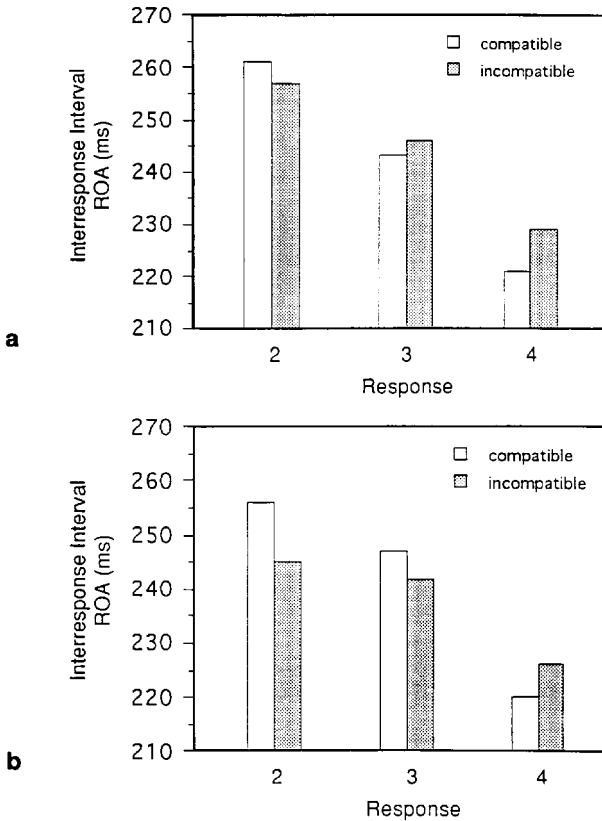


FIG 8. Inter-response intervals (response onset asynchronies: ROAs) preceding the second, third and fourth responses in Müsseler's interference task (Müsseler, 1995). In one of the experiments, the action sequence is determined by arrows alone (a); in another experiment, it is determined by arrows plus tones attached to them (b).

This counterintuitive prediction was in fact warranted by the data. Figure 8a shows the inter-response times, that is, the intervals elapsing between response onsets in the action sequence. As can be seen, the inter-response times preceding the second response were slightly longer when the additional arrow was compatible with that response than when it was incompatible.

It can also be seen that the effect was reversed in the following two intervals where compatible pairings of arrows and actions were faster than incompatible ones, yielding a normal compatibility effect. This is exactly what we had expected. In these later intervals, no intake of new information is required and the inverted compatibility effect should therefore disappear. Instead, a normal compatibility effect should arise, due to

the fact that an action code for the additional response has been built up in the meantime, which serves to speed up or slow down the third and fourth responses, depending on compatibility.

This pattern was more clearly demonstrated in a further experiment, where a short tone was added to each arrow—low-pitched for left and high-pitched for right arrows. Basically, this manipulation should have two effects. First, the tone sequence that now accompanies the arrow sequence can be maintained in auditory working memory. This should support the memorisation of the sequence. Second, as the fifth stimulus now comes as an arrow with the tone attached to it, its attention-demanding value is increased. This should increase its power to impair or interrupt ongoing action preparation and action control. Therefore, the inverted compatibility effect should now be more pronounced than the tiny effect in the first experiment.

In this experiment, the four arrows were presented in the same display locations as before, but they now appeared in succession and were accompanied by corresponding tones (at a rate of two stimuli per second). The fifth stimulus, which was another pairing of an arrow and a tone, was again triggered by the first response. As can be seen, the inverted compatibility effect was now more pronounced for the first inter-response interval, and again it gradually changed into an ordinary compatibility effect over the following intervals (Fig. 8b). Somewhat unexpectedly, there was still a slight inverted effect in the second inter-response interval. This is probably due to the fact that the fifth stimulus was now preserved somewhat longer in working memory, so that it impaired the common code over a longer period.

In summary, the inverted compatibility effect is another important piece of evidence in support of the notion of common coding. First, it shows that common coding is not restricted to temporal and spatial similarity between perception and action, but covers symbolic equivalence as well, thereby corroborating our previous conclusion about the representational status of the codes involved. Second, it suggests that these codes cannot be used for both perception and action planning at the same time. Rather, impairment of performance will arise when action planning is switched in and perception is called for.

## Action Planning Impairs Perception

Would the reverse be true as well? Would perceptual performance be impaired by concurrent action planning and would that impairment depend on compatibility? This issue has recently been addressed by Müsseler and Hommel (1997). In their paradigm, the basic task was to identify a briefly presented stimulus while a particular response was being

prepared. Again, the stimuli were arrows pointing to the left or the right and the actions were left and right keypresses. On each trial, the subjects had to perform two tasks that were interlocked in a particular way.

The first task was a speeded action task, shown in the upper part of Fig. 9. The imperative stimulus for this task was an arrow pointing to the left or right. The required action consisted of two responses in a speeded sequence: an unspecific double keypress first and a specific single keypress second. The unspecific double keypress was always the same, regardless of which imperative stimulus was presented. The specific single keypress had to be left for left arrows and right for right arrows (compatible S-R assignment throughout). The task was speeded in the sense that the two responses had to follow each other as quickly as possible. There was, however, no speed requirement for starting the sequence as a whole: the subjects could start responding whenever they wished.

The second task was an unsped perceptual task (lower part of Fig. 9). It also required the subjects to identify an arrow and to indicate its direction by a corresponding keypress. This second arrow was masked, however, and its presentation time was individually adjusted such that an error rate between 10 and 40% was obtained for each subject. The response to this arrow was unsped in the sense that subjects could press the corresponding key at leisure.

The two tasks were interlocked as follows: Each trial started with the presentation of the imperative stimulus for the speeded action task (unmasked arrow). This was then followed (at a point in time chosen by each subject) by the sequence of the two speeded responses. The first of these responses (the unspecific double keypress) served to interlock the two tasks (indicated by the dotted circle in Fig. 9). It triggered the immediate presentation of the masked arrow, to the effect that this stimulus was presented at a time when the specific single keypress was being prepared.

The question was whether identification accuracy of the masked arrow would depend on the compatibility between that arrow and the response being prepared at the time of its presentation, and whether that dependency would again take the form of inverted compatibility. In a series of experiments, Müsseler and Hommel (1997) observed that this was indeed the case: When left or right responses were being prepared, identification accuracy for compatible arrows pointing in the same direction was lower than for incompatible arrows pointing in the opposite direction. This inverted compatibility effect was observed in a series of five experiments.

Some of these experiments were conducted to meet obvious objections. These objections refer to the possibility of purely visual interactions between the two stimuli or purely response-related interference between

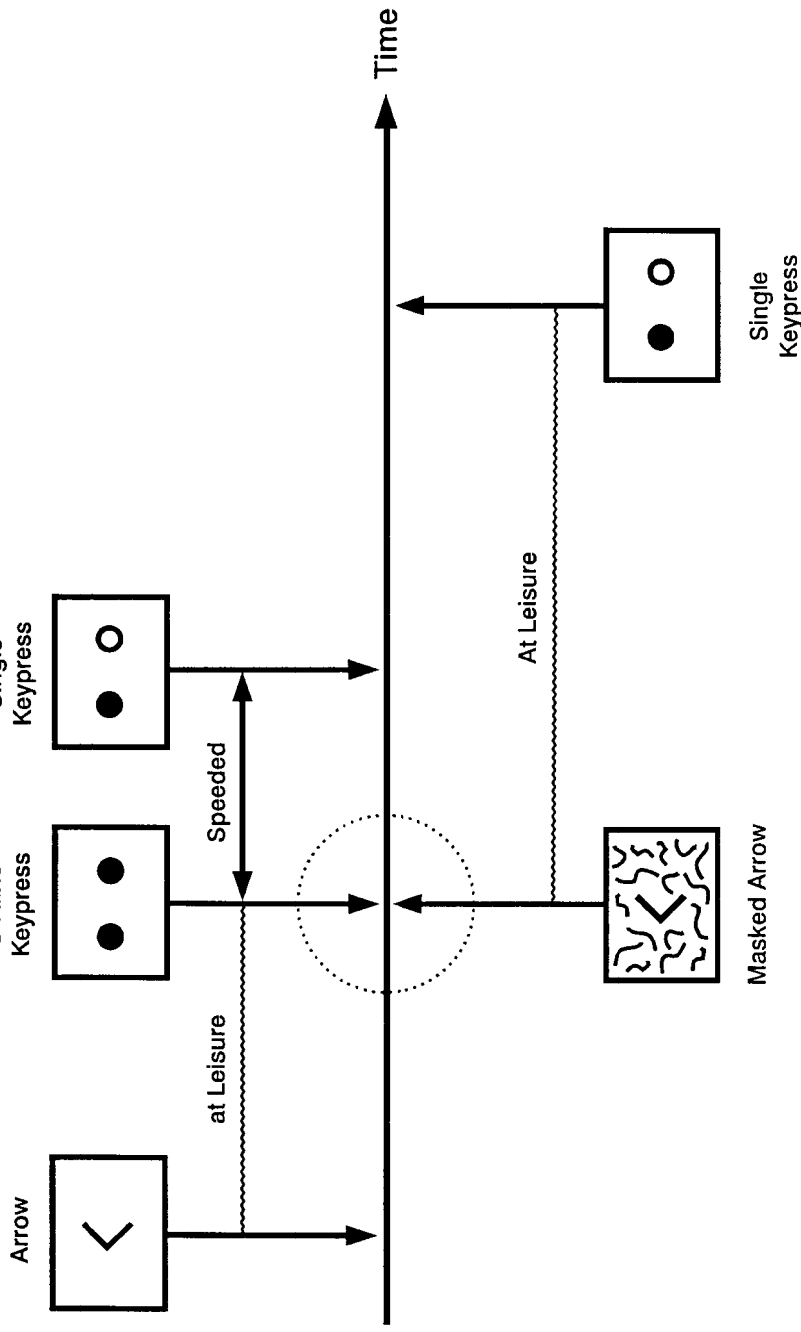


FIG 9. Interlocking between the speeded action task (upper part) and the unspeeded perceptual task (lower part) in the paradigm developed by Müsseler and Hommel (1997). See text for details.

the two responses involved. It was shown that the inverted compatibility effect survives when one takes measures to control these factors. To sum up, we may therefore (a) reiterate and reinforce our basic assumption that our mental machinery disposes of common codes for perception and action, and (b) suggest the additional assumption that these codes cannot be used for both of these functions at the same time.

## CONCLUSIONS

Two general conclusions and two specific suggestions can be derived from this work. The first conclusion is that under conditions where stimuli share some features with planned actions, these stimuli tend, by virtue of similarity, either to induce those actions or interfere with them, depending on the structure of the task at hand. This implies that there are certain products of perception on the one hand and certain antecedents of action on the other that share a common representational domain. This is the common-coding principle. The second conclusion is that actions are planned and controlled in terms of their effects; that is, that representations of action effects play an important role in the planning and the control of those actions. This is the action-effect principle.

The first suggestion is that common coding may apply to highly processed or cognitive levels of coding; that is, to rather late products of perception and rather early antecedents of action. The second suggestion is that there may be a mechanism that protects ongoing action planning from the processing of incoming stimulus information, and vice versa, if they both address identical structures in the common-coding domain.

Taken together, we believe that our results encourage us to pursue this approach further as a theoretical guideline for the study of relationships between perception and action.

Manuscript received April 1996

Revised manuscript received August 1996

## REFERENCES

- Aschersleben, G. (1994). *Afferente Informationen und die Synchronisation von Ereignissen*. Frankfurt/Main: Lang.
- Aschersleben, G., & Prinz, W. (1995). Synchronizing actions with events: The role of sensory information. *Perception and Psychophysics*, *57*, 305–317.
- Aschersleben, G., & Prinz, W. (1997). Delayed auditory feedback in synchronization. *Journal of Motor Behavior*, *29*, 35–46.
- Bauer, D.W., & Miller, J. (1982). S-R compatibility and the motor system. *Quarterly Journal of Experimental Psychology*, *34*, 367–380.
- Beric, A. (1992). Cortical somatosensory evoked potentials in spinal cord injury patients. *Journal of Neurological Sciences*, *107*, 50–59.



- Descartes, R. (1664). *Traité de l'Homme*. Paris: Girard.
- Dunlap, K. (1910). Reactions on rhythmic stimuli, with attempt to synchronize. *Psychological Review*, 17, 399–416.
- Fraisse, P. (1980). Les synchronisations sensori-motrices aux rythmes. In J. Requin (Ed.), *Anticipation et comportement*, pp. 233–257. Paris: Centre National.
- Franzén, O., & Offenloch, K. (1969). Evoked responses: Correlates of psychophysical magnitude estimates for tactile stimulation in man. *Experimental Brain Research*, 8, 1–18.
- Gehrke, J. (1996). *Kinematische Analyse repetitiver Bewegungen*. Doctoral dissertation, University of Munich.
- Gehrke, J., Aschersleben, G., & Prinz, W. (in prep.). Processing of afferent feedback and the timing of actions.
- Hashimoto, I., Gatayama, T., Yoshikawa, K., & Sasaki, M. (1992a). Somatosensory evoked potentials: Correlates of psychophysical magnitude estimations for air-puff stimulation of the face in man. *Experimental Brain Research*, 88, 639–644.
- Hashimoto, I., Gatayama, T., Yoshikawa, K., & Sasaki, M. (1992b). Somatosensory evoked potentials: Correlates of psychophysical magnitude estimations for air-puff stimulation of the foot in man. *Experimental Brain Research*, 92, 318–325.
- Hommel, B. (1990). *Quellen der Interferenz beim Simon-Effekt*. Doctoral dissertation, University of Bielefeld.
- Hommel, B. (1993). Inverting the Simon effect by intention. *Psychological Research*, 55, 270–279.
- Johansson, R.S. (1979). Tactile afferent units with small and well demarcated receptive fields in the glabrous skin area of the human hand. In D.R. Kenshalo (Ed.), *Sensory function of the skin on humans*, pp. 129–152. New York: Plenum Press.
- Johnson, D., Jürgens, R., Kongehl, G., & Kornhuber, H.H. (1975). Somatosensory evoked potentials and magnitude of perception. *Experimental Brain Research*, 22, 331–334.
- Koffka, K. (1925). *Die Grundlagen der psychischen Entwicklung. Eine Einführung in die Kinderpsychologie*. Osterwieck: Zickfeldt.
- Kolers, P.A., & Brewster, J.M. (1985). Rhythms and responses. *Journal of Experimental Psychology*, 11, 150–167.
- Kornblum, S. (1992). Dimensional overlap and dimensional relevance in stimulus–response and stimulus–stimulus compatibility. In G.E. Stelmach & J. Requin (Eds.), *Tutorial in motor behavior II*, pp. 743–777. Amsterdam: North-Holland.
- Kornblum, S. (1994). The way irrelevant dimensions are processed depends on what they overlap with: The case of Stroop- and Simon-like stimuli. *Psychological Research/Psychologische Forschung*, 56, 130–135.
- 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 and Performance*, 21, 855–875.
- Kornblum, S., Hasbroucq, T., & Osman, A. (1990). Dimensional overlap: Cognitive basis for stimulus–response compatibility—a model and taxonomy. *Psychological Review*, 97, 253–270.
- Lesser, R.P., Köhle, R., & Lüders, H. (1979). Effect of stimulus intensity on short-latency somatosensory evoked potentials. *Electroencephalography and Clinical Neurophysiology*, 47, 379–382.
- Liberman, A.M., Cooper, F.S., Shankweiler, D.P., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.
- Lippa, Y. (1996a). *Über die Vermittlung von Wahrnehmung und Handlung: Kompatibilitätseffekte unter Verwendung orthogonaler Reiz- und Handlungsdimensionen*. Doctoral dissertation, University of Munich.

- Lippa, Y. (1996b). A referential-coding explanation for compatibility effects of physically orthogonal stimulus and response dimensions. *Quarterly Journal of Experimental Psychology*, 49A, 950–971.
- Mach, E. (1922). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen* (9. Aufl.). Jena: Gustav Fischer.
- MacKay, D.G. (1987). *The organization of perception and action: A theory for language and other cognitive skills*. New York: Springer-Verlag.
- Massaro, D.W. (1990). An information-processing analysis of perception and action. In O. Neumann & W. Prinz (Eds), *Relationships between perception and action: Current approaches*, pp. 133–166. Berlin: Springer-Verlag.
- Meltzoff, A.N., Kuhl, P.K., & Moore, M.K. (1991). Perception, representation, and the control of action in newborns and young infants: Toward a new synthesis. In M.J. Weiss & P. Zelazo (Eds), *Newborn attention: Biological constraints and the influence of experience*, pp. 377–411. Norwood, NJ: Ablex.
- Morton, J., Marcus, S.M., & Frankish, C. (1976). Perceptual centers (P-centers). *Psychological Review*, 83, 405–408.
- Müsseler, J. (1995). *Wahrnehmung und Handlungssteuerung. Effekte kompatibler und inkompatibler Reize bei der Initiierung und Ausführung von Reaktionssequenzen*. Aachen: Shaker.
- Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance*, 23.
- Paillard, J. (1949). Quelques données psychophysiologiques relatives au déclenchement de la commande motrice. *L'Année Psychologique*, 48, 28–47.
- Rappaport, M., Portillo, S.R., & Leonard, J. (1992). Effects of stimulus intensity on posterior tibial nerve somatosensory evoked potentials. *Clinical Electroencephalography*, 23, 89–94.
- Riggio, L., Gawryszewski, L., & Umiltà, C. (1986). What is crossed in crossed-hands effects? *Acta Psychologica*, 62, 89–100.
- Sanders, A.F. (1980). Stage analysis of reaction processes. In G.E. Stelmach & J. Requin (Eds), *Tutorials in motor behavior*, pp. 331–354. Amsterdam North-Holland.
- Simon, J.R., Hinrichs, J.V., & Craft, J.L. (1970). Auditory S–R compatibility: Reaction time as a function of ear–hand correspondence and ear–response–location correspondence. *Journal of Experimental Psychology*, 86, 97–102.
- Viviani, P. (1990). Motor–perceptual interactions: The evolution of an idea. In M. Plattelli-Palmarini (Ed.), *Cognitive science in Europe: Issues and trends*, pp. 11–39. Monograph Series No. 1. Boulder, CO: Golem.
- Vos, P.G., Mates, J., & van Kruysbergen, N.W. (1995). The perceptual centre of a stimulus as the cue for synchronization to a metronome: Evidence from asynchronies. *Quarterly Journal of Experimental Psychology*, 48A, 1024–1040.
- Wallace, R.A. (1971). S–R compatibility and the idea of a response code. *Journal of Experimental Psychology*, 88, 354–360.
- Welford, A.T. (1968). *Fundamentals of skill*. London: Methuen.