PERCEPTUAL–MOTOR INTERACTION: SOME
IMPLICATIONS FOR HUMAN–COMPUTER
INTERACTION

Romeo Chua
University of British Columbia

Daniel J. Weeks and David Goodman
Simon Fraser University

Perceptual–Motor Interaction: A Behavioral Emphasis ........................................ 24
Human Information Processing and Perceptual-Motor Behavior .................................. 24
Translation, Coding, and Mapping .................................................................................. 25
Movement Dynamics and Perceptual-Motor Behavior .................................................. 25
Relative Organizational Mapping .................................................................................... 26
Perceptual–Motor Interaction: Attention and Performance ......................................... 28
Action-Centered Attention .............................................................................................. 28
Action Requirements and Attention .............................................................................. 29
Attention and Stimulus-Response Compatibility ......................................................... 29
Perceptual–Motor Interaction in Applied Tasks: A Few Examples ................................. 31
Remote and Endoscopic Surgery .................................................................................... 31
Remote Viewing ............................................................................................................ 31
Automated Compensation ............................................................................................. 32
Summary and Implications ............................................................................................ 32
References ..................................................................................................................... 33
PERCEPTUAL–MOTOR INTERACTION: A BEHAVIORAL EMPHASIS

Two of us (D.W., D.G.) can still remember purchasing our first computers to be used for research purposes. The primary attributes of these new tools were their utility in solving relatively complex mathematical problems and performing computer-based experiments. However, it was not long after that word processing brought about the demise of the typewriter, and our department secretaries no longer prepared our research manuscripts and reports (but that story is for another time). It is interesting to us that computers are not so substantively different from other tools such that we should disregard much of what the study of human factors and engineering psychology has contributed to our understanding of human behavior in simple and complex systems. Rather, it is the computer's capacity for displaying, storing, and processing information that has led us to the point at which the manner with which we interact with such systems has become a research area in itself.

In our studies of human–computer interaction (HCI) and perceptual–motor interactions in general, we have adopted two basic theoretical and analytical frameworks as part of an integrated approach. In the first framework, we view perceptual–motor interactions in the context of an information–processing model. In the second, we use analytical tools that allow detailed investigations of both static and dynamic interactions. The purpose of this chapter is to outline this approach and some of the current empirical work on perceptual–motor behavior we believe have considerable implications for those working in HCI.

Human Information Processing and Perceptual-Motor Behavior

For many scientists interested in perceptual-motor behavior, the information-processing framework has traditionally provided a major theoretical and empirical platform for the study of these interactions. The study of perceptual-motor behavior within this framework has inquired into such issues as the information capacity of the motor system (e.g., Fitts, 1954), the attentional demands of movement (e.g., Posner & Keele, 1969), motor memory (e.g., Adams & Dijkstra, 1966), and processes of motor learning (e.g., Adams, 1971). The language of information processing (e.g., Broadbent, 1958) has provided the vehicle for discussions of mental and computational operations of the cognitive and perceptual-motor system (Posner, 1982). Of interest in the study of perceptual-motor behavior is the nature of the cognitive processes that underlie perception and action.

The information-processing approach describes the human as an active processor of information, in terms that are now commonly used to describe complex computing mechanisms. An information-processing analysis describes observed behavior in terms of the encoding of perceptual information, the manner in which internal psychological subsystems utilize the encoded information, and the functional organization of these subsystems. At the heart of the human processing system are processes of information transmission, translation, reduction, collation, and storage (e.g., Fitts, 1964; Marteniuk, 1976; Stelmach, 1982; Welford, 1968).

Consistent with a general model of human information processing (e.g., Fitts & Posner, 1967), three basic processes have been distinguished historically. For our purposes, we refer to these processes as stimulus identification, response selection, and response programming. Briefly, stimulus identification is associated with processes responsible for the perception of information. Response selection pertains to the translation between stimuli and responses and the selection of a response. Response programming is associated with the organization of the final output (see Proctor & Yu, this volume, for a more detailed discussion of these processes).

A key feature of models of information processing is the emphasis on the cognitive activities that precede action (Marteniuk, 1976; Stelmach, 1982). From this perspective, action is viewed only as the end result of a complex chain of information-processing activities (Marteniuk, 1976). Thus, chronometric measures, such as reaction time and movement time, as well as other global outcome measures, are often the predominant dependent measures. However, even cursory examination of the literature indicates that time to engage a target has been a primary measure of interest. For example, a classic assessment of perceptual-motor behavior in the context of HCI and input devices was conducted by Card, English, and Burr (1978; see also English, Engelhart, & Berman, 1967). Using measures of error and speed, Card et al. (1978) had subjects complete a cursor positioning task using four different control devices (mouse, joystick, step keys, and text keys). Data revealed the now well-known advantage for the mouse. Of interest is that the speed measure was decomposed into homing time and positioning time. The former denoted the time that it took to engage the control device and initiate cursor movement, and the latter the time to complete the cursor movement. Although the mouse was actually the poorest device in terms of the homing time measure, the advantage in positioning time produced the faster overall time. That these researchers sought to glean more information from the time measure acknowledges the importance of the movement itself in perceptual-motor interactions such as these.

The fact that various pointing devices depend on hand movement to control cursor movement has led to the emphasis that researchers in HCI have placed on Fitts' law (Fitts, 1954) as a predictive model of time to engage a target. The law predicts pointing (movement) time as a function of the distance to and width of the target. The impact of Fitts' law is most evident by its inclusion in the battery of tests to evaluate computer pointing devices in ISO 9241-9. We argue that there are number of important limitations to an exclusive reliance on Fitts' law in this context. First, although the law predicts movement time, it does so on the basis of distance and target size. Consequently, it does not allow for determining what other factors may influence movement time. Specifically, Fitts' law is often based on a movement to a single target at any given time (although it was originally developed using reciprocal movements between two targets). However, in most HCI and graphic user interface contexts, there is an array of potential targets that can be engaged by an operator.
I. Perceptual-Motor Interaction: Some Implications for Human-Computer Interaction

Discuss later in this chapter, the influence of these distractor targets on movements to the imperative target can be significant.

Second, we suggest that the emphasis on Fitts' law has diverted attention from the fact that cognitive processes involving the selection of a potential target from an array are an important, and time-consuming, information processing activity that must precede movement to that target. Indeed, the Hick-Hyman law (Hick, 1952; Hyman, 1953) predicts the decision time required to select a target response from a set of potential responses. In fact, if an operator executes the decision and movement components sequentially, then the time to complete the task will be the sum of the times predicted by the Hick-Hyman and Fitts' laws. However, an operator may opt to make a general movement first and select the final target destination concurrently. Under such conditions, Hoffman and Lim (1997) reported interference between the decision and movement component that was dependent on their respective difficulties. Finally, although Fitts' law predicts movement time given a set of movement parameters, it does not actually reveal much about the underlying movement itself. Indeed, considerable research effort has been directed toward revealing the movement processes that give rise to Fitts' law. For example, theoretical models of limb control have been forwarded that propose that Fitts' law emerges as a result of multiple submovements (e.g., Crossman & Gooden, 1988), or as a function of both initial movement impulse variability and subsequent corrective processes late in the movement (Meyer, Abrams, Kornblum, Wright, & Smith, 1988). These models highlight the importance of conducting detailed examination of movements themselves as a necessary complement to chronometric explorations only.

Translation, Coding, and Mapping. As outlined previously, the general model of human information processing (e.g., Fitts & Posner, 1967) distinguishes three basic processes: stimulus identification, response selection, and response programming. Although stimulus identification and response programming are functions of stimulus and response properties, respectively, response selection is associated with the translation between stimuli and responses (Welford, 1968).

Translation is the set of the human interface between perception and action. Moreover, the effectiveness of translation processes at this interface is influenced by a large extent by the relation between perceptual inputs (e.g., stimuli) and motor outputs (e.g., responses). Indeed, since the seminal work of Fitts and colleagues (Fitts & Deninger, 1954; Fitts & Seeger, 1953), it has been repeatedly demonstrated that errors and choice reaction times to stimuli in a spatial array are shorter when the stimuli are mapped onto responses in a spatially compatible manner. Fitts and Seeger (1953) referred to this finding as stimulus-response (S-R) compatibility and ascribed it to cognitive codes associated with the spatial locations of elements in the stimuli and response arrays. Presumably, it is the degree of coding and recoding required to map the locations of stimuli and response elements that determine the speed and accuracy of translation and thus response selection (e.g., Wallace, 1971).

The relevance of studies of S-R compatibility to the domain of human factors engineering is paramount. It is now well understood that the design of an optimal interface in which effective S-R translation facilitates fast and accurate responses is largely determined by the manner in which stimuli and response arrays are arranged and mapped onto one another (e.g., Bayer, Miller, & Lewis, 1988; Chapanis & Lindenbaum, 1959; Proctor & Van Zandt, 1994).

Movement Dynamics and Perceptual-Motor Behavior

As discussed previously, many HCI situations involve dynamic perceptual-motor interactions that may not be best indexed merely by chronometric methods (cf. Card, English, & Burr, 1978). Indeed, as HCI moves beyond the simple key press interfaces characteristic of early systems to include virtual and augmented reality, teleoperation, gestural and haptic interfaces, among others, the dynamic nature of perceptual-motor interactions are even more evident. Consequently, assessment of the actual movement required to engage such interfaces may be more revealing.

To supplement chronometric explorations of basic perceptual-motor interactions, motor behavior researchers have also advocated a movement process approach (Kelso, 1982). The argument is that to understand the nature of movement organization and control, analyses should also encompass the movement itself, and not just the activities preceding it (e.g., Kelso, 1982; Marteniuk, MacKenzie, & Leavitt, 1988). Thus, investigators have examined the kinesematics of movements in attempts to further understand the underlying organization involved (e.g., Brooks, 1974; Chua & Elliott, 1993; Elliott, Carson, Goodman, & Chua, 1991; Kelso, Southard, & Goodman, 1979; MacKenzie, Marteniuk, Dugas, Liske, & Eckmeier, 1987; Marteniuk, MacKenzie, Jeanerod, Athens, & Dugas, 1987).

This emphasis on movement dynamics is evident in the dynamical systems approach to the study of perception and action. This theoretical perspective seeks to explain perceptual-motor behavior in terms of fundamental, physical laws and principles (Jeka & Kelso, 1989; Kelso, 1995; Turvey, 1990). The dynamical systems framework is characterized by the application of the tools and principles from physical biology, synergetics (self-organization), and non-linear dynamics. In the language of dynamics, movement systems are thought of as self-organizing systems, whereby patterns emerge from the interaction of the many variables inherent in the system. Thus, the theoretical and analytical tools and principles of synergetics and nonlinear dynamics become relevant to the study of perceptual-motor systems and perceptual-motor interactions (e.g., Schön & Kelso, 1988).

Furthermore, and of greatest utility in our work, is that with the application of these tools, the set of measurable, dependent, variables are extended to capture the richness of movement dynamics more adequately.

A key element of the dynamical systems approach is the identification of patterns (e.g., perceptual-motor patterns or interactions) relevant to the system under study. The primary strategy for identifying these patterns is to find transitions, situations in which one observes qualitative changes in the system's behavior. The transition demarcates one pattern from another, and the qualitative change allows one not only to distinguish
between the patterns, but also to identify the relevant dimension of the pattern (Jeka & Kelso, 1989; Kelso, 1995). It is also the range about the transition that helps to identify the relevant variable that characterizes the pattern itself. A second important element is the study of stability and loss of stability of the patterns. The study of the system's stability or instability allows a determination of the system's dynamics. It is the stability of a given pattern that distinguishes it from others, characterizing the state in which the system resides. Moreover, the loss of stability is hypothesized to be a mechanism that effects a change in pattern (Jeka & Kelso, 1989; Kelso, 1995). It is the loss of stability of a pattern that may lead to a transition to a new pattern, one distinguished by its greater stability. In the section that follows, we describe our efforts to integrate compatibility phenomena and dynamic perceptual-motor interactions under one theoretical framework. This is followed by some recent data from our lab that has attempted to apply the framework.

Relative Organizational Mapping

Our studies of perceptual–motor interactions, and compatibility phenomena more generally, has led us to propose a relative organizational mapping model (Chua & Weeks, 1997). Consistent with other conceptual divisions forwarded by Fitts and colleagues, among others (e.g., Fitts & Deninger, 1954; Fitts & Posner, 1967; Kantowitz, Triggs, & Barnes, 1990; Kornblum, Hashrouqi, & Osman, 1990), the model proposes three levels of organization: global relation, configuration, and mapping (see p. 1.1). The mapping relation refers to how stimulus events are mapped onto response events. The configuration relation refers to the orientation of the stimulus and response arrays with respect to the other. The global relation refers to the overall relation in space between the stimulus display and response array. For example, in a prototypical two-choice reaction time task, the assignment of a given stimulus event to a response event would be categorized as a mapping relation. Whether or not the stimulus display and response array are arranged along the same spatial dimension (e.g., parallel or orthogonal) would be a configuration relation. Lastly, where the response array is physically located with respect to the stimulus display would be subsumed under global relation.

The three levels of organization can be considered as a nested hierarchy. Mapping bears directly on the response (i.e., stimuli are assigned to responses on the basis of the mapping). Mapping is nested within a configuration, which may affect performance indirectly through its influence on the relation between the mapping and the response. Both mapping and configuration are, in turn, subsumed within a global spatial relation. The global relation may affect performance indirectly through its constraint on the relation between the configuration, mapping, and the response action. The hierarchy is not necessarily meant to suggest that one level has precedence over the other. Rather, it provides a means to conceptualize potential constraints on perceptual–motor interactions and to provide a framework for our experimental manipulations. These levels of organization reveal the complex factors that a human operator must contend with to translate between perceptual and motor workspaces. Moreover, not only does the model provide a useful taxonomy for standard compatibility phenomena involving discrete stimuli and response events (as do a number of models and theories—Kantowitz et al., 1990; Kornblum et al., 1990), together with the analytical tools used to study movement dynamics, the model also provides a platform for examining the manner in which response selection and response execution become connected in dynamic perceptual–motor interactions.

In a simple and elegant study that acted as a catalyst for our work on relative organizational mapping, Worringham and Beringer (1989) examined the influence of operator orientation on visual-motor performance, thereby providing us with the requisite experimental conditions for the application of our distinction between levels of spatial organization. In their study, participants used a joystick to move a cursor on a monitor from a central location to targets positioned radially about a central location. The authors noted three types of compatibility relations that could exist between movement of the joystick and cursor, and the different orientations of the operator with respect to the display and control. Control-display compatibility referred to a relation in which the direction of control motion corresponded spatially to motion of the cursor on the display, independent of operator orientation. Visual-motor compatibility referred to the relation in which the directions of control motion and cursor motion were referenced to the visual field axis of the operator, if the operator was looking at either control or display. Visual-trunk compatibility referred to the relation in which the cursor motion was referenced to the visual field axis, and control motion was referenced to the body midline. Depending on the orientation of the operator, and the spatial position of the display with respect to the control, different levels of either one, two, or all three of these compatibility relations could exist (see Worringham & Beringer, 1989). Worringham and Beringer found that participants' visual-motor task performance was influenced by the spatial relations between the display, control, and operator. Moreover, they also found that situations that yielded compatible relations according to a visual-motor reference (visual-motor compatibility) led to superior performance compared with other compatibility relations (display-control and visual-trunk compatibility).

![Figure 1.1: Relative organizational mapping: levels of organization in perceptual-motor interaction.](image-url)
Using Worringham and Beringer's (1989) study as an empirical starting point, we (Chua, Weeks, Ricker, & Poon, 2001) imported a subset of their task conditions and examined these within the context of examining interactions between the organizational levels of mapping, configuration, and global relation. Specifically, in our set-up, participants sat with the stimulus display (a rectangular panel with a light-emitting diode toward each end) located either directly in front of them, aligned with their midline, or directly to their left. The display panel could be oriented either vertically or horizontally. A control lever (allowing lefward and rightward rotational movements of the forearm) was located either in front, in right ipsilateral space, or directly to the participant's right side. In the first case, participants could reach forward to grasp the control, whereas in the second, the participants reached to their right. We assigned the relation between display-stimulus events and control-response events to the mapping level of organization, the orientation of the display with respect to the control to the configuration level, and the spatial relation between the display, control, and operator to the global relation level. Thus, with the set-up shown in Fig. 1.2, we manipulated the orientation of the participant (operator) with respect to the locations of the display and control to examine the influence of global spatial relations. We varied display orientation to examine the influence of spatial configuration. Finally, we varied mapping rules to examine the effects of spatial mapping.

In an initial experiment (Chua et al., 2001, Experiment 1), we used a choice reaction time paradigm in which participants had to make rapid, discrete, rotational movements in response to the brief onset of a visual stimulus. In a second experiment (Chua et al., 2001, Experiment 2), we used a coordination paradigm in which participants had to synchronize continuous rhythmic movements with an oscillating visual display. Thus, we examined independently the impact of the three levels of spatial organization on both discrete actions that emphasized selection of a response, and continuous actions that emphasized dynamic perceptual-motor coordination. Our results showed that, for discrete responses, the spatial mapping that yielded faster responding was dependent on the display-control array configuration and the global relation. When the display and control dimensions were parallel to one another (e.g., movement in horizontal plane), participants appeared to code the spatial aspects of the display and control in a manner that was unaffected by the global spatial relation. Thus, one mapping rule yielded an advantage that was not affected by the global relation. When the display and control dimensions were orthogonal to one another (e.g., vertical display, horizontal control movement), there was a tendency for the direction of mapping effects to be influenced by the global relation. The results for the coordination task showed that spatial configuration had an impact on whether or not performance differences between spatial mapping rules emerged. Differences in coordination performance between mapping conditions was evident only under a parallel configuration. At the global relation level, there was no influence on accuracy and stability of perceptual-motor coordination under different configuration or mapping conditions. In both the discrete task (Experiment 1) and the coordination task (Experiment 2), our findings pointed to the consistent adoption of a visual-motor reference (e.g., see Worringham & Beringer, 1989), even if the end result was that the direction of motion of the stimulus display and the motion of the response were in actuality spatially incompatible (see Chua et al., 2001). The results thus support the proposal that a visual-motor frame of reference may dominate over others (Worringham & Beringer, 1989). More importantly, these findings suggested that, in determining the degree of compatibility between a stimulus display and a control array, the determination must include consideration of the fact that a human operator is physically oriented, and must translate between stimulus and response events.

In our initial proposal of the relative organizational mapping framework (Chua & Weeks, 1997; Chua, Weeks, & Goodman, 1996), we presented a series of experiments that examined spatial compatibility effects in coordinative actions. Consistent with our most recent work described previously (Chua et al., 2001), we showed that spatial compatibility effects that were observed in discrete tasks could also be demonstrated during performance of continuous coordination tasks. These spatial compatibility effects were sufficiently robust so as to influence performance even when the spatial stimulus dimension was irrelevant for the task. In one study (Chua, 1995), participants coordinated rhythmic leftward and rightward movements of a control lever with a visual stimulus that oscillated between two spatial locations (cf. Chua et al., 2001). The stimulus changed colors such that the two oscillation points were distinguished by a particular color. The position of the two oscillation points was also varied, such that the oscillation would begin along a horizontal plane, then rotate such that the two oscillation
points would exchange positions (e.g., a counter-clockwise rotation of 180 degrees). The changing orientation of the stimulus resulted in a change in the stimulus display-control configuration. The endpoints of the leftward and rightward movements were mapped onto the stimulus colors—the actual position of the stimulus was irrelevant. Nevertheless, the results showed that coordination performance (in both accuracy and variability) was influenced by the orientation of the stimulus display relative to the spatial dimension of the control action. When both motion of the display and control happened to correspond, coordination performance was improved.

This impact of spatial stimulus information, even when it is not directly relevant to the task, has led us to speculate about the influence that symbolic or implied spatial information might have on the execution of continuous tasks. Symbolic spatial information has been shown to exert effects on reaction time in prototypical compatibility tasks (e.g., Weeks & Proctor, 1990). Presently, we have turned our focus toward symbolic cues to investigate whether or not similar effects can be found in tasks that require continuous perceptual–motor interactions. In addition, the effects described previously reminded us of the Simon effect in choice reaction time tasks (see later in this chapter). Briefly, the Simon effect (Simon, 1968) occurs when an irrelevant spatial stimulus attribute interferes with response selection. As described later in this chapter, the Simon effect has been linked to the operation of attention-based mechanisms. Indeed, attention-based accounts of the Simon effect has become an important point of contact in considerations of the role of attention in perceptual–motor interactions.

PERCEPTUAL–MOTOR INTERACTION: ATTENTION AND PERFORMANCE

The vast literature on selective attention and its role in the filtering and selection of information (e.g., Cherry, 1953; Deutsch & Deutsch, 1963; Treisman, 1966a, 1966b, 1986; Treisman & Gelade, 1980) has no doubt been informative in the resolution of issues in HCI pertaining to stimulus displays and inputs (e.g., the use of color and sound). However, attention can be thought of as not a unitary function, but rather as a set of information-processing activities that are important for perceptual, cognitive, and motor skills. Indeed, the evolution of HCI into the realm of augmented reality, teleoperation, gestural interfaces, and other areas that highlight the importance of dynamic perceptual–motor interactions necessitates a greater consideration of the role of attention in the selection and execution of action. Recent developments in the study of selective attention mediate perception and action, and more importantly, how action in turn influences attentional processes, are poised to make such a contribution to HCI. We will turn to a brief review of these developments and some thoughts on their potential relevance to HCI.

Action-Centered Attention

It has been suggested (Allport, 1987) that one role for human selective attention is to provide the motor system with the relevant stimulus characteristics necessary for selecting and executing an appropriate action (e.g., reaching and grasping a particular selected object). Tipper and colleagues (e.g., Tipper, Lortie, & Baylis, 1992) have suggested that, for actions such as selective reaching to an item in a cluttered multi-item environment, attention is mediated by an action-centered cognitive representation. The architecture of this representation is such that distractor items are included in the response selection computations associated with programming the motor output (e.g., the reach) for a target stimulus. In their now oft-cited study, Tipper and colleagues (1992) used a reaching paradigm in which participants reached and pointed to a target located within a 3 x 3 matrix of push buttons that denoted potential target locations. On any given trial, light emitting diodes alongside the buttons were used to cue participants to the position of the target only, or the target and co-occurring distractor. Tipper and colleagues found that responses were slowed in the presence of distractors, with the effect even more pronounced when distractors were located along the path of movement (Tipper et al., 1992).

Several investigators have both corroborated and qualified the work of Tipper and colleagues (1992). Pratt and Abrams (1994) confirmed that responses are slower in the presence of distractors, particularly when distractors are positioned along the path of the movement. Moreover, the same authors demonstrated that the interference effect of distractors were found during both response preparation (indexed by reaction time) and response execution (indexed by movement time). In addition, interference effects during the execution of the response had its locus in the terminal, corrective phase, of the movement.

The cluttered environment of response buttons used by Tipper and colleagues (1992) struck us as being analogous to the array of icons present in a typical graphical user interface. In a recent study, Lyons, Elliott, Ricker, Weeks, and Chua (1999) sought to determine whether the paradigm developed by Tipper and colleagues (1992) could be imported into a virtual environment and ultimately serve as a test bed for investigations of perceptual–motor interactions in an HCI context. The task space utilized a 3 x 3 matrix similar to that used by Tipper et al. (1992). The matrix, made up of nine blue circles, was displayed on a monitor placed vertically in front of the participant. Consistent with Tipper and colleagues' paradigm, the target was presented either in isolation or in the presence of a distractor. Consequently, at trial onset, a target circle would turn red in color; and for distractor trials a yellow circle would appear simultaneously in one of the other eight locations. The participants were required to move the mouse on the graphics tablet, which would in turn move a cursor on the monitor in the desired direction toward the target circle. The participants were unable to view their hand; the only visual feedback of their progress was from the monitor. The graphics tablet allowed the researchers to record displacement and time data of the mouse throughout the trial. In contrast to previous experiments (e.g., Meegan & Tipper, 1998; Tipper et al., 1992), the presence of a distractor had relatively little influence on performance. Lyons et al. postulated that, in a task environment in which perceptual–motor interaction is less direct (e.g., using a mouse to move a cursor on a remote display), perceptual and motor workspaces are misaligned, and the increased translation processing owing to the misalignment serves to limit the impact of distractor items.
To test this idea, Lyons et al. (1999) modified the task environment so as to align the perceptual and motor workspaces. The monitor was turned on and held screen down inside a support frame. The same 3 x 3 matrix was displayed on the monitor and reflected into a half-silvered mirror positioned above the graphics tablet allowing for sufficient space for the participant to manipulate the mouse and move the cursor to the target without vision of the hand. With this configuration, the stimulus display was presented and superimposed on the same plane as the motor workspace (i.e., the graphics tablet). Under this setup, distractor effects became evident and were consistent with an action-centered framework of attention. Taken together, these findings underscore the influence of translation requirements demanded by relative alignment of perceptual and motor workspaces. More importantly, these findings suggest that even relatively innocuous changes to the layout of the task environment may have significant impact on processes associated with selective attention in the mediation of action in an HCI context.

The behavioral consequences of selecting and executing target-directed actions in the presence of potential distractors is not limited simply to the time taken to prepare and execute the movement. Indeed, further recent investigations have provided evidence showing that increases in movement time due to the presence of distractors may be a result of deviations in the movement trajectory itself. Tipper and colleagues (Howard & Tipper, 1997; Tipper, Howard, & Jackson, 1997) have reported that the trajectory of the movement toward the target deviates away from the rest of the competing stimulus or inhibit a competing response to a distractor object that is located in proximity to the hand. Recently, Welsh and Elliott (2001) examined cursor movement trajectories using a virtual environment setup similar to Lyons et al. (1999). Interestingly, the results revealed that participants were actually attracted to the competing stimulus, Welsh and Elliott (2001) suggested that subtle methodological differences may account for the discrepancy in results. Specifically, in the task used by Howard and Tipper (1997), the location of the nontarget stimulus was known before the location of the target. This may have facilitated early inhibition of the response to the nontarget location. In contrast, in the virtual environment used by Welsh and Elliott (2001), presentation of the target and distractor was simultaneous. This may have resulted in the preparation of an initial averaged response that was subsequently corrected online during execution. Regardless, these results again reveal the impact that even subtle changes to the stimulus array can have on task performance and the need for a detailed examination of the execution of the response itself.

**Action Requirements and Attention.** The first round of basic research into the action-centered model of attention has been focused primarily on the examination of the spatial locations of distractors with respect to the target. In that context, an action-centered framework could offer a useful perspective for the spatial organization of perceptual information presented in an HCI context. However, often the reason for engaging a target in an HCI task is because the target symbolically represents an outcome or operation to be achieved. Indeed, this is what defines a target as an icon—target features symbolically carry a meaning that defines it as the appropriate target. An emerging line of research investigations also under the umbrella of the action-centered framework have turned to issues pertaining to the impact of intrinsic features of targets and distractors. For example, Jervis, Bennett, Thomas, Lim, and Castiello (1999) have examined the relation between semantic categories to which targets and distractors belong (fruits vs. three-dimensional shapes). They have shown distractor interference effects on grasping the target object when the target and distractor belong to different semantic categories. When distractors and targets come from similar semantic categories (e.g., both fruits), distractors may not be of behavioral importance to the task and therefore can be inhibited (Castiello, 1996; Jervis et al., 1999).

Although using objects from different semantic categories may effectively distinguish the visual similarity between targets and distractors, the physical features of these objects may still evoke similar response requirements (i.e., a similar reach and grasp). An interest in the application of the action-centered model to human factors and HCI led Weir et al. (2002) to consider whether or not distractor effects could be elicited based on the specific actions required to engage a target and distractor object. The question was whether the engagement properties of target and distractor objects (i.e., turn or pull) in a control array would mediate the influence of the distractor on the control of movement. In that study, participants executed their movements on a control panel that was located directly in front of them. On some trials, the control panel consisted of a single pull-knob or right-turn dial located at the midline either near or far from a starting position located proximal to the participant. On other trials, a second control device (pull-knob or dial) was placed into the other position on the display. If this second device was present, it served as a distractor object and was to be ignored. The findings suggested that, when moving in an environment with distracting stimuli or objects, competing responses may be programmed in parallel. When the distractor is different from the target object, greater interference is present when the competing responses are being programmed, resulting in an increased response time. The implication is that the terminal action required to engage a target object can also be important to movement planning and execution.

One potential area of relevance to HCI is whether terminal action requirements will have a similar impact on perceptual-motor interactions within a virtual environment. Current work in our labs (Robertson, Chua, & Weeks, in preparation) is focusing on the influence the action and engagement properties of target and distractor objects on response competition and selection for action in the same task environment we have used previously (cf. Lyons et al., 1999). Thus, in our studies of the action-centered model of attention and its relevance to HCI, we have progressed to considering the response execution requirements of the task environment, in addition to the spatial layout of the environment.

**Attention and Stimulus-Response Compatibility**

To this point, we have separated the discussion of attention from the issues related to translation and perceptual-motor interaction. However, the action-centered model of selective attention
clearly advocates the view that attention and action are intimately linked. The fundamental premise is that attention mediates perceptual-motor interactions, and these, in turn, influence attention. In line with this perspective, the role of attention in the translation between perceptual inputs and motor outputs has also received considerable interest over the past decade. As discussed previously, a key element in the selection of an action is the translation between stimuli and responses, the effectiveness of which is influenced to a large extent by the spatial relation between the stimulus and responses. The degree of coding and recording required to map the locations of stimuli and response elements has been proposed to be a primary determinant of the speed and accuracy of translation (e.g., Wallace, 1971). Attentional processes have been implicated in the issue of how relative spatial stimulus information is coded. Specifically, the orienting of attention to the location of a stimulus has been proposed to result in the generation of the spatial stimulus code.

Initial interest in the link between attention orienting and spatial translation has emerged as a result of attempts to explain the Simon effect. The Simon effect (Simon, 1968; Simon & Rudell, 1969), often considered a variant of spatial S-R compatibility, occurs in a situation in which a nonspatial stimulus attribute indicates the correct response and the spatial attribute is irrelevant to the task. Thus, the spatial dimension of the stimulus is an irrelevant attribute, and a symbolic stimulus feature constitutes the relevant attribute. Although the spatial stimulus attribute is irrelevant to the task, faster responding is found when the position of the stimulus and the position of the response happen to correspond. A number of researchers (e.g., Umiltà & Nicolotti, 1992) have suggested that attentional processes may be a unifying link between the Simon effect and the spatial compatibility effect proper. Specifically, the link between attention and action in these cases is that a shift in attention is postulated to be the mechanism that underlies the generation of the spatial stimulus code (e.g., Nicolotti & Umiltà, 1989, 1994; Proctor & Lu, 1994; Rubichi, Nicolotti, Iani, & Umiltà, 1997; Stoffier, 1991; Stoffier & Umiltà, 1997, Umiltà & Nicolotti, 1992). According to an attention-shift account, when a stimulus is presented to the left or right of the current focus of attention, a reorienting of attention occurs toward the location of the stimulus. This attention shift is associated with the generation of a spatial code that specifies the position of the stimulus with respect to the last attended location. If this spatial stimulus code is congruent with the spatial code of the response, then S-R translation, and, hence response selection, is facilitated. If the two codes are incongruent, response selection is hindered.

Recent work in our lab has also implicated a role for attention shifts in compatibility effects and object recognition. In these studies, Lyons, Weeks, and Chua (2000a, 2000b) sought to examine the influence of spatial orientation on the speed of object identification. Participants were presented with video images of common objects that possessed a graspable surface (e.g., a tea cup, frying pan) and were instructed to make a left or right key press under two distinct mapping rules, depending on whether the object was in an upright or inverted vertical orientation. The first mapping rule required participants to respond with a left key press when the object was inverted and a right key press when the object was upright. The opposite was true for the second mapping rule. The orientation of the object's graspable surface was irrelevant to the task. The results showed that identification of object orientation was facilitated when the graspable surface of the object was also oriented to the same side of space as the response (see also Tucker & Ellis, 1998). In contrast, when participants were presented with objects that possessed symmetrical graspable surfaces on both sides (e.g., a sugar bowl with two handles), identification of object orientation was not facilitated. Lyons et al. (2000a) also showed that response facilitation was evident when the stimuli consisted simply of objects that, though may not inherently be graspable, possessed a left-right asymmetry. Taken together, these results were interpreted in terms of an attentional mechanism. Specifically, Lyons et al. (2000a, 2000b) proposed that a left-right object asymmetry (e.g., a protruding handle) may serve to capture spatial attention (cf. Tucker & Ellis, 1998). If attention is thus oriented toward the same side of space as the ensuing action, the spatial code associated with the attention shift (e.g., see previous discussion) would lead to facilitation of the response. In situations in which no such object asymmetry exists, attentional capture and orienting may be hindered, and, as a result, there is no facilitation of the response.

Taken into the realm of HCI, it is our position that the interplay between shifts of attention, spatial compatibility, and object recognition will be a central human performance factor as technological developments continue to enhance the directness of direct-manipulation systems (cf. Shneiderman, 1983, 1992). Specifically, as interactive environments become better abstractions of reality with greater transparency (Rutkowski, 1982), the potential influence of these features of human information-processing will likely increase. Thus, it is somewhat ironic that the view toward virtual reality, as the solution to the problem of creating the optimal display representation, may bring with it an unintended consequence (Tenner, 1996). Indeed, the operator in such an HCI environment will be subject to the same constraints that are present in everyday life.

A goal of human factors research is to guide technological design to optimize perceptual-motor interactions between human operators and the systems they use. Thus, the design of machines, tools, interfaces, and other sorts of devices utilizes knowledge about the characteristics, capabilities, as well as limitations, of the human perceptual-motor system. In computing, the development of input devices such as the mouse and graphical user interfaces was intended to improve HCI. As technology has continued to advance, the relatively simple mouse and graphical displays have begun to give way to exploration of complex gestural interfaces and virtual environments. This development may perhaps, in part, be a desire to move beyond the artificial nature of such devices as the mouse to ones that provide a better mimic of reality. Why move an arrow on a monitor using a hand-held device to point to a displayed object, when instead, you can reach and interact with the object. Perhaps such an interface would provide a closer reflection of real-world interactions, and the seeming ease with which we interact with our environments, but also subject to the constraints of the human system.
PERCEPTUAL–MOTOR INTERACTION
IN APPLIED TASKS: A FEW EXAMPLES

As we described at the outset of this chapter, the evolution of computers and computer-related technology has brought us to the point at which the manner in which we interact with such systems has become a research area in itself. Current research in motor behavior and experimental psychology pertaining to attention, perception, action, and spatial cognition is poised to make significant contributions to the area of HCI. In addition to the continued development of a knowledge base of fundamental information pertaining to the perceptual motor capabilities of the human user, these contributions will include new theoretical and analytical frameworks that can guide the study of HCI in various settings. In this final section, we highlight just a few specific examples of HCI situations that offer a potential arena for the application of the basic research that we have outlined in this chapter.

Remote and Endoscopic Surgery

Recent work by Hanna and colleagues (Hanna, Shimi, & Cuschieri, 1998) examined task performance of surgeons as a function of the location of the image display used during endoscopic surgical procedures. In their study, the display was located either in front, to the left, or to the right of the surgeon. In addition, the display was placed either at eye level or at the level of the surgeon’s hands. The surgeons’ task performance was observed with the image display positioned at each of these locations. Hanna et al. (1998) showed that the surgeons’ performance was affected by the location of the display. Performance was facilitated when the surgeons were allowed to view their actions with the monitor positioned in front and at the level of the immediate workspace (the hands). In terms of our relative organizational mapping framework (see earlier in this chapter), the position of the display relative to the surgeon and the workspace is a global relation. We would suggest that the optimal display location (or global relation) in this task environment placed less translation demands on the surgeon during task performance. Given that surgeons work with a team of support personnel each with a different vantage point relative to the operating field, an interesting empirical question will be the manner in which perceptual–motor workspace can be effectively optimized for each team member.

Similar findings have also been demonstrated by Mandryk and MacKenzie (1999). These investigators also examined the impact of display location on endoscopic surgical performance. In addition to the frontal image display location used by Hanna et al. (1998), Mandryk and MacKenzie also investigated the benefits of projecting and superimposing the image from the endoscopic camera directly over the workspace. Their results showed that performance was superior when participants were initially exposed to the superimposed viewing condition. This finding was attributed to the superimposed view allowing the participants to better calibrate the display space with the workspace. These findings are consistent with our own investigations of action-centered attention in virtual environments (Lyons et al., 1999). We would suggest that the alignment of perceptual and motor workspaces in the superimposed viewing condition facilitated performance due to the decreased translation requirements demanded by such a situation. However, the findings of Lyons et al. (1999) would also lead us to suspect that this alignment may have additional implications with respect to processes associated with selective attention in the mediation of action. Although the demands on perceptual–motor translation may be reduced, the potential intrusion of processes related to selective attention and action selection may now surface.

Remote Viewing

Hooper and Coury (1994) have described an intriguing and important HCI issue in the operation of periscopes in modern submarines. Traditionally, operators of standard submarine periscopes were required to physically rotate the periscope to view in different directions and, as a consequence, the operator adjusted his/her own orientation with respect to the submarine. Thus, in combination with visual cues from within the internal environment of the submarine, information pertaining to the spatial direction of the periscope was available to the operator. Hooper and Coury (1994) noted that the development of newer periscopes with advanced video systems and graphics workstations has led to a significant perceptual–motor translation issue. Specifically, the spatial information that was inherent in the physical operation of the periscope is no longer readily available in the computer displays that accompany new systems. As such, information such as periscope direction relative to the submarine and the direction of view now have to be provided within the graphic display (see Hooper & Coury, 1994, for examples). Hooper and Coury examined the best way to present this type of information. The results of their experiments indicated that, in some display situations, operators would adopt a strategy of mental rotation to determine orientation. This translation process would represent an additional cognitive demand on the operator.

An interesting aspect of this HCI scenario as described by Hooper and Coury (1994) is that it seems to be a case in which information pertaining to the global relation between perceptual and motor workspaces is incorporated into some form of image display and must therefore be extracted from that display. This rather indirect means of obtaining this information is then shown to have an impact on information processing demands on the operator. To date, our investigations involving the use of our relative organizational mapping model has been limited to the study of how different levels of spatial organization might interact to constrain perceptual–motor performance. The work by Hooper and Coury (1994) provides a new interesting avenue for us in which we might consider how information such as global relations can be used in perceptual–motor interactions, and the impact that removing or modifying this type of information might have on performance.
Automated Compensation

In the previous examples, the performance impact of the spatial relation between the perceptual and motor workspaces of an operator is readily apparent. Our own research efforts using the relative organizational mapping framework have been directed at the specific levels of spatial relations that have a bearing on perceptual-motor interactions. Although we have focused primarily on the interaction between levels of spatial organization and their influence on perceptual-motor translation, recent work by Macedo and colleagues (Macedo, Kaber, Endsley, Powanasorn, & Myung, 1998) has examined how a system might readily compensate for spatial misalignments between perceptual and motor workspaces. Specifically, Macedo and colleagues tested methods to automatically compensate for misalignments between display and operator orientation on a visual tracking task. Participants tracked an irregularly moving target on a display using a joystick-controlled cursor. Spatial alignment between the operator and the display was varied by orienting the display, operator, or both, at fixed angles of rotation. By monitoring joystick orientation and the operator's head rotation, Macedo et al. implemented a computer algorithm that served to rotate and align the axes of the joystick and the display, and automatically compensate for the misalignment. Their findings showed that tracking performance of the operator in situations of operator-display misalignment could be improved through automated compensation. These findings suggest that methods can be developed that decrease the burden of perceptual-motor translation on an operator in situations in which SR incongruencies might otherwise have a significant impact on performance. Thus, an engineering approach to facilitating the effective mapping of perceptual and motor workspaces would seem to hold promise. Whether algorithms can be developed to compensate for the impact of mapping at multiple levels of constraint will be an interesting empirical issue for the future.

Technological advances have undoubtedly served to improve the HCI experience. For example, we have progressed beyond the use of computer punch cards and command-line interfaces to more complex tools, such as graphical user interfaces and speech recognition. As HCI has become not only more effective, but also by the same token more elaborate, the importance of the interaction between the various perceptual, cognitive, and motor constraints of the human system has also come to the forefront. In this chapter, we have presented an overview of some current topics of research in perceptual-motor interactions that we believe have relevance to HCI. Clearly, considerable research will be necessary to evaluate the applicability of these potentially relevant data to specific HCI design problems. Nevertheless, the experimental work to date, along with rational considerations, leads us to conclude this chapter with a few implications for early infusion to HCI.

First, we have outlined the benefits of integrating the frameworks and tools offered by the information-processing and movement dynamics approaches to capture the richness of perceptual-motor interactions, from processing and movement preparation through movement execution. For example, one area that we are pursuing in this regard involves a more detailed evaluation of principles of control-display movement. Current design principles (e.g., Warrick's principle, scale-side principle, clockwise-for-increase principle, etc.) are generally based on the expectations and preferences of the population of users. However, faced with a unique or novel display-control arrangement, such an approach can place principles in conflict, making it difficult to determine the optimal mapping (e.g., Brebner & Sandow, 1976; Petropoulos & Brebner, 1981). We hold that the integration of information processing and movement dynamics approaches could yield dependent measures and new metrics that would be effective in evaluating dynamic human performance.

Second, we have discussed the implications of recent research on attention and SR compatibility with respect to their impact on action selection and execution. Here, the implications are more straightforward. That the allocation of attention also carries an action-centered component means that an effective interface must also be sensitive to the specific action associated with a particular response location, the action relationship between that response and those around it, and the degree of translation required to map the perceptual-motor workspaces.

Finally, we presented our relative organizational mapping model to conceptualize and address empirically, potential multilevel constraints on perceptual-motor interactions. We hold that the examination of perceptual-motor interaction as a problem involving multiple levels of constraint acknowledges a range of potential interactions that can occur in a given HCI context. Indeed, this range extends from the specific arrangement of display and response aspects of an interface, to include the immediate workspace, the workspace within a room, that room within a building, and so on. As theoretical and empirical work continues, we believe the important implication of current research for HCI and human factors situations remains, to quote Norman (1988), "getting the mappings right."
References


