

# Action-Perception Patterns Emerge From Coupling and Adaptation

G. Schöner

*Centre de Recherche en Neurosciences Cognitives  
C. N. R. S., Marseille, France*

T. M. H. Dijkstra

*Department of Psychology  
Ohio State University*

J. J. Jeka

*Department of Kinesiology  
University of Maryland*

Organisms function in particular environments, the properties of which are reflected in the structure of their nervous systems. Therefore, sensory information may directly specify motor behaviors. We argue that such specification involves the coupling of sensory information into appropriately structured control systems that generate action. The nature of this coupling as well as the structure of the control systems reflect properties of the environment. This is most dramatically demonstrated when adaptive processes adjust the underlying control system in response to changes in the environment. Experimental and modeling work on posture in perturbed visual and haptic environments is reviewed to provide evidence for these arguments. Theoretical modeling and autonomous robotics work that goes beyond the posture example of perception-action coupling is briefly discussed, primarily to point out that the integration of multiple behavioral constraints is a nontrivial problem that has received too little attention.

Gibson's (1958/this issue) article, which we celebrate here, discusses three circles of ideas, which we briefly discuss to position our own contribution.

## STIMULATION IS CONSTRAINED ECOLOGICALLY

Patterns of sensory stimulation are not arbitrary physical signals. They are physical signals that arise from the outer world in which the organism is immersed. Optic stimulation originates typically from light reflected from surfaces, for instance. These surfaces are often at rest relative to each other and are often rigid. Light sources are often at rest. Acoustic stimulation often originates from physical events that are localized in space. Tactile stimulation typically results from touching surfaces, again often rigid, sometimes at rest. Perception is apparently sensitive to these properties of the outer world. The perception of depth, of shape, of motion, and so forth is to some extent invariant under changes of the conditions of stimulation such as changes of the distributions of light sources or changes of the aspect of the viewer, and so forth.

This hypothesis, that perception is sensitive to properties of the outer world, was expressed by Gibson in his 1958 article and in more elaborate form later (as well as by others). Modern robot vision is in large part based on this view, most explicitly in the domain of active vision (e.g., Aloimonos & Rosenfeld, 1991), although an intellectual heritage from Gibson is not typically acknowledged and historically these ideas might have arisen independently in robot vision. Perceptual science is also largely influenced by such ideas, although again they are not always traced back to Gibson (and might not always be traceable in this way). A major criticism of Gibson's view of perception is aimed at his implied denial of the need to analyze the processing of the stimulus that leads to the representation of invariant information. In the 1958 article, this denial is not actually apparent. Our work on the visual and haptic control of posture is, of course, based on this view, as is the work on the visual elicitation of action and the autonomous robotics work.

## PERCEPTION-ACTION BEHAVIORS

The fact that sensory stimulation is constrained by the way it arises from the environment makes it possible to extract specific information from the sensory array with which a number of basic locomotory behaviors of animals can be controlled. This was Gibson's second idea. His analysis implied, although this was not explicitly stated, that such information can often be obtained without first segmenting the sensory pattern. That means that the sensory array need not be segmented into foreground and background, or into separate subsets that are matched onto objects in the outer world. Information relevant to the control of locomotory behavior can be obtained from the "raw" sensory array. A recent study by Warren and Saunders (1995) provided evidence that humans perceive heading direction from an unsegmented visual array.

This idea is used quite naturally today in a number of applications in autonomous robotics. The fundamental optic flow equation, for instance, is used to extract information about the motion state of a moving observer that can then be

used to control observer motion (cf. Neven & Schöner, 1996; Sandini, Gandolfo, Grosso, & Tistarelli, 1993). Realizations of such systems have been obtained only over the last 10 years or so, in part, because the computational problems encountered when this information must be extracted from image sequences have become tractable in real time only with improved embedded computation.

In psychology, this idea is most prominently pursued within the paradigm of ecological psychology, to which this journal is dedicated. Of the many different proposals listed by Gibson, the control of posture by optic flow and the initiation of action in response to optic flow information have been most intensely studied. By contrast, obstacle avoidance based on optic flow, homing, and so forth, have received little attention, although robotic implementations demonstrate the feasibility of these proposals.

This article aims to examine these proposals. By reviewing our own experimental and theoretical work on the stabilization of posture by unsegmented visual and haptic information as well as theoretical work on the control of movement initiation by optic flow, we provide evidence in support of two hypotheses: (a) Coupling to unsegmented sensory information is the basis of simple action-perception patterns, and (b) adaptive change of the motor control system into which such information couples reveals that action-perception patterns are not the result of a feed-forward computational system in which the stimulus array drives the action but are adaptive behaviors in which the system changes so as to tune its parameters to incoming sensory information. *Adaptive* refers to motor responses that do not result simply from the reduction of an error signal. For instance, during posture in a moving optic environment, the postural control system changes its eigenfrequency to match the frequency of imposed oscillation.

The first of these two points is probably in line with Gibson's views as expressed in the 1958 article. The notion of control is implicit. It is thus useful for us to clarify to which extent control theory is an appropriate theoretical framework to formalize this aspect of Gibson's program. We comment on this issue on several occasions and devote a section on a necessary extension of the control theoretic framework. There is an aspect, however, of ideas in perception that are somewhat inspired by Gibson's program, but that may be in conflict with the importance we attribute to coupling. We refer to the notion that based on assumptions about the outer world, particular parameters may be "computed" from the sensory array, which then "specify" a particular motor act. This notion is neither sufficient (because the stabilization of a perception-action linkages is a nontrivial aspect that goes beyond specification in this sense), nor necessary (because perception-action patterns may be stabilized without explicit estimation of underlying parameter values). We comment on this problem in the Discussion section while reviewing "less simple perception action patterns."

The second point is probably less in line with Gibson's proposals, as it emphasizes the dependence of perception-action patterns on other factors than the sensory array, here, the internal state of adaptive processes.

## SPATIAL ORIENTATION

The third circle of ideas acknowledges explicitly the need for other types of processes to account for discrimination and identification. These ideas reach into cognition. It is interesting to note that Gibson arrived at the insight that such ideas are necessary, as this aspect is not often emphasized within the framework of ecological psychology. This circle of ideas is, however, the least developed, and one might wonder to what extent Gibson or his students have contributed to this domain. Gibson's stance might need to be interpreted more in relation to behaviorist ideas still topical at the time. For lack of space, we briefly allude to the relation between our dynamic ideas and this problem in the Discussion. This is, however, one point at which we depart from the program of ecological psychology by emphasizing that representations can become quite independent of the sensory array.

## THE DYNAMIC APPROACH

Our emphasis on coupling and adaptation requires a framework in which the system that sensory information couples into and that adapts can be characterized. That framework is provided by ideas from a dynamic approach to understanding coordinated movement (e.g., Schöner & Kelso, 1988). The central concept is stability, which captures the capacity of perception-action systems to maintain particular states in the face of the various types of perturbations to which the system is continuously exposed. Continuous variables characterize such stable states and excursions from those states. Such variables are chosen to accommodate different observed stable states of the system. In contrast to the theory of control, the choice of state variables is not based on an attempt to divide the system up into a plant and a controller. That latter distinction is problematic for nervous control, in particular, when the presumed plant changes in response to the experienced control signals.

The dynamic approach deals naturally with situations in which multiple stable states of a system coexist. Multistability, but also change of stability, constrain the choice of dynamical variables used to characterize the system. The different stable states must be describable by the chosen variables, and changes of stability must be captured by the dynamics of these variables. Classical control theory deals with systems that have a unique *set point* (stable state), the deviation from which is the control variable. Such a control variable cannot be defined if there are multiple simultaneous states (deviation from which set point?). We report later on some preliminary data and a speculative model to illustrate a case where the dynamic approach clearly departs from classical control theory.

## DYNAMIC MODEL OF POSTURE

During posture the body is maintained upright in the gravitational reference frame. This is one of the simplest whole-body motor behaviors and is an important element

of all locomotory behaviors. Multiple sources of sensory information are known to contribute to postural behavior including vestibular, visual, and somatosensory information, but also more indirect factors such as efference copy or central set (Horak, Diener, & Nashner, 1989). Much of the postural literature attempts to discern the role of different subsystems (e.g., vestibular system, particular muscle groups, etc.) or different biomechanical components (e.g., head vs. trunk) in overall postural control.

Modeling postural behavior within the dynamic approach does not aim at an absolute description of the postural system valid under any experimental condition. Instead, postural behavior is characterized with respect to particular behavior situations. This reflects the view that the fundamental organization of the nervous system is not in terms of anatomical or physiological subsystems, but in terms of elementary behaviors. We define elementary behaviors as classes of behaviors. As long as a behavior changes continuously and remains stable as environmental conditions are changed, it continues to belong to the same class. A new elementary behavior is encountered whenever a behavior goes through an instability. For example, the effect on postural sway of a optic array of random dots moving coherently is strongly dependent on the number of dots in the display. A small number of dots has a negligible effect on sway. The moving dots are interpreted as object motion in the environment. As the number of moving dots increases, a dramatic change occurs, as postural sway entrains to the moving visual display (Brandt, Wist, & Dichgans, 1975). This change signifies the interpretation of object (foreground) to environmental (background) motion, each of which may be considered an elementary behavior that reorganizes the response of the nervous system. From this perspective, physiological or biomechanical subsystems are not in themselves separable units, even though they are often analyzed in this manner. Multiple subsystems may be recruited to interact as single units in terms of the elementary behaviors.

We have been studying postural behavior involving small sway amplitudes for a number of years. Under such conditions, normal participants move as a rigid body, such that a single degree of freedom for anterior-posterior and a second degree of freedom for medial-lateral sway characterize sufficiently the movement state of the body. These position variables are candidate behavioral variables because the stable states obtained by postural behavior can be expressed as constant values of these variables. The dynamics of these variables is plausibly at least a second order differential equation, because oscillatory components of body sway are frequently observed. The simplest (i.e., linear) dynamical model of a stable postural state is thus

$$\ddot{x} + \alpha\dot{x} + \omega^2x = 0 \quad (1)$$

Here, only one positional degree of freedom,  $x$ , is taken into account. When perturbations are applied to either anterior-posterior or to medial-lateral sway, we assume,

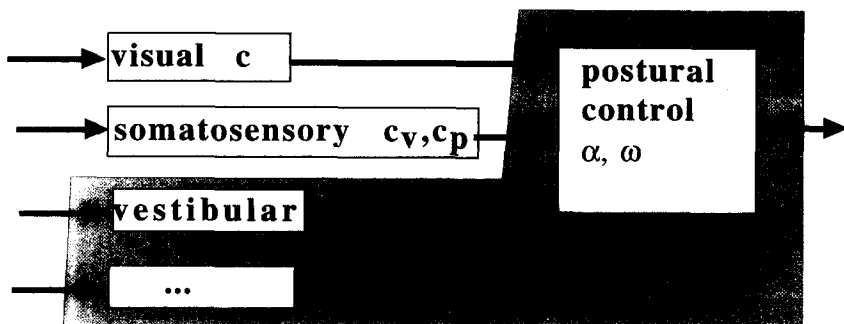


FIGURE 1 The analysis of the postural control system into its different sensory and motor components is based on the idea that when a particular type of sensory information is varied (in Figure 1, visual and somatosensory), then observed changes of postural behavior can be attributed to the properties of particular sensory-motor channels (the two boxes outside the hatched region). The ensemble of other modules (hatched region) is assumed independent of the variations induced in the chosen channels. If this is not the case, this approach loses its predictive power: Each time environmental conditions are changed, a new system is confronted. Note that such changes go beyond adaptation in the simplest control theoretic sense, which is also typically restricted to a single sensory channel (e.g., Kalman gain).

that these two components of sway can be analyzed separately.<sup>1</sup> The postural variable  $x$  refers to the matching degree of freedom. The two parameters of this linear model are the eigenfrequency,  $\omega$ , and the damping coefficient,  $\alpha$ . These names are chosen in analogy to the meaning of the corresponding parameters in the dynamic equation of physical oscillators. They do not have any deeper meaning than that.

This equation simply describes the overall effect of the various control systems contributing to postural stability, not giving any particular prominence to the mechanical contributions to that stability. The fact that, mathematically, this is the equation of a damped harmonic oscillator suggests an analogy with the dynamics of a damped pendulum. Our formulation has, however, nothing to do with the biomechanical description of posture as an inverted pendulum problem. That description does not address how the mechanically unstable inverted equilibrium solution is stabilized. Conversely, the present model does not address how the different effector degrees of freedom are coordinated to generate rigid "pendulum" motion. It should be noted that the same mathematical model has been hypothesized to underlie postural control by Johansson, Magnusson, and Åkesson (1988).

What are different constituents of this control system? The typical procedure to answer this question involves varying the conditions of one particular sensory channel and studying its effect on postural control. This leads to systematic identification of the postural control system if the postural control system remains con-

<sup>1</sup>This does not mean that the dynamics of the different degrees of freedom are independent. In the experiments summarized here, we have not looked for such a coupling across degrees of freedom, and therefore, we do not include such couplings in our models.

stant while the sensory situation is changed (Figure 1). The assumption of a constant control system turns out to be wrong, however, as we illustrate later. We have shown this for postural control while varying two sensory inputs independently: vision and somatosensation. In both cases, in the best Gibsonian manner, we have not presented arbitrary stimuli based on a complete representation of stimuli (e.g., "white noise"), but stimuli that correspond to a well-defined situation in the world: a virtual "moving wall" from which visual information is obtained and a physically moving contact surface from which somatosensory stimulation derives.

### Posture in the Visual World

The contribution of vision to postural behavior can be studied in the moving room paradigm introduced by Lee and colleagues (Lee & Aronson, 1974; Lee & Lishman, 1975). What can be expected? For anterior-posterior motion, a moving observer in a stationary environment senses a velocity field (optic flow) that is zero whenever the observer is at rest. It changes sign (from expansion to contraction) whenever the velocity of the observer in the visual frame changes sign. This is therefore a potential control signal for posture (in the sense of visual serving). Visual expansion as such does not specify the gain (how much to react to particular levels of flow). The desirable gain depends on the depth of the scene. At large depths only small flows are generated and high sensitivity is needed; at small depths, large flows are generated and only low sensitivity is need. Moreover, flow increases in amplitude with distance from the focus of expansion-contraction. Do all locations of the visual array contribute equally or appropriately weighted with eccentricity? The visual expansion rate

$$e(t) = \frac{\dot{r}(t)}{r(t)} = -\frac{\dot{d}(t)}{d(t)} \quad (2)$$

is the ratio of optic flow,  $\dot{r}$ , at eccentricity,  $r$ , with eccentricity itself.<sup>2</sup> Under the assumptions of geometrical optics, it is equal to the inverse ratio of the rate of change of visual depth,  $\dot{d}$ , and visual depth,  $d$ . This measure of visual motion is, under some conditions, homogeneous across the visual array (e.g., at constant visual depth across the array). Lee (1976) first introduced this ratio, proposing that this parameter could be used by organisms to control their approach to surfaces. Schöner (1991) hypothesized that for anterior-posterior postural sway, this ratio was the form in which visual information couples into the postural control system

$$\ddot{x} + \alpha\dot{x} + \omega^2x = -ce(t) \quad (3)$$

<sup>2</sup>Strictly speaking,  $\dot{r}$ , by itself is the visual expansion rate, while  $e(t)$  is the ratio of visual expansion rate and eccentricity. This ratio is independent of the size of the patch, the expansion of which is considered. For simplicity we use the term *visual expansion rate* to refer to this ratio.

The visual expansion rate is coupled negatively into the postural dynamics so that an expanding flow ( $e > 0$ ) accelerates the body backwards; a contracting flow accelerates the body forwards. This models a control behavior in which body sway tends to reduce optic flow.

The visual expansion rate has a contribution that is monotonically related to the velocity,  $\dot{x}$ , of the body in the outer frame. In a static optic environment, an outflow is generated ( $e > 0$ ) when the body moves forward ( $\dot{x} > 0$ ); an inflow is generated ( $e < 0$ ) when the body moves backward ( $\dot{x} < 0$ ). Coupling the expansion rate into the postural dynamics thus accounts for the stabilizing effect of visual information on posture. In fact, for large optic distances ( $d > x$ ), the expansion rate in a static optic environment can be approximated as

$$e(t) = \frac{\dot{x}}{x+d} \approx \frac{\dot{x}}{d} \quad (4)$$

In this case, the visual contribution has the form of a linear damping function, which increases the effective damping to  $\tilde{\alpha} = \alpha + c/d$ :

$$\ddot{x} + \tilde{\alpha}\dot{x} + \omega^2 x = 0 \quad (5)$$

One way to probe this coupling structure is to induce perturbations into the optic environment. In the simplest case, the optic environment may be moved sinusoidally in depth:  $d(t) = d_0 + \Delta d \sin(\omega_d t)$ . Under such conditions, there are two contributions to optic flow—the flow created by movement of the body and the flow created by the moving optic scene:

$$e(t) \approx \frac{\dot{x}}{d_0} - \frac{\omega_d \Delta d \cos(\omega_d t)}{d_0} \quad (6)$$

Again we assumed that the average visual depth of the scene ( $d_0$ ) is much larger than the sway amplitude and the amplitude of visual motion,  $\Delta d$ . The dynamic model of postural control is then simply a driven damped harmonic oscillator:

$$\ddot{x} + \tilde{\alpha}\dot{x} + \omega^2 x = \tilde{c} \cos(\omega_d t) \quad (7)$$

with effective coupling strength  $\tilde{c} = c \omega_d \Delta d / d_0$ .

Qualitatively, this is consistent with the early experiments in which participants were exposed to an oscillatory optic environment and exhibited typical driven oscillator behavior (e.g., van Asten, Gielen, & Denier van der Gon, 1988). We can be more precise, however. First, the assumed coupling to expansion rate predicts more specifically that the strength of the driving force generated by a moving optic scene decreases with increasing visual depth (because  $d_0$  appears in the denomina-



tor of the driving term). Second, the stabilizing effect of vision is likewise predicted to decrease with increasing visual depth (same argument of the effective damping term).

These predictions were tested in the experiment of Dijkstra, Schöner, and Gielen (1994). Quietly standing participants were exposed to a optic scene (random dots lying on a virtual "moving wall") that moved in depth with moderate amplitude (4 cm) around different mean depths (25, 50, 100 and 200 cm) at a frequency (0.2 cycles/sec) that approximately matches typical postural frequencies. Depth was defined both by stereoscopic cues as well as by motion parallax generated in a virtual reality setup, in which measured head position served to update the display consistent with a particular visual depth (see Figure 2).

What happens in this experimental setting is that participants sway rhythmically, approximately in-phase with the oscillatory optic environment. The first observation to make is that the spatiotemporal coupling between visual motion and induced postural sway is remarkably tight. Dijkstra, Schöner, and Gielen (1994) estimated the stability of the coupling of postural sway to visual motion both by computing the variability of the relative phase between visual motion and sway and by perturbing the visual motion and measuring relaxation time to the stable phase

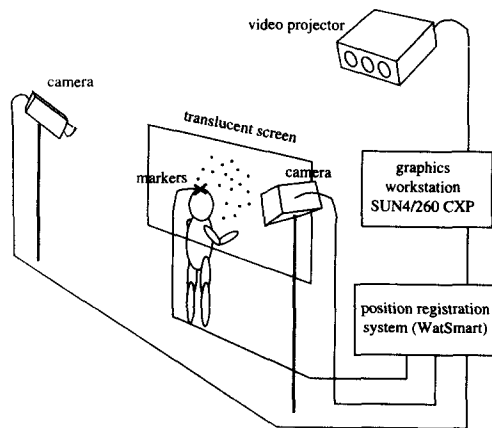


FIGURE 2 The setup for studying visual control of posture. Participants viewed a translucent screen through red/green stereo goggles. A video projector projected a stereogram onto the screen. Participants wore a helmet on which infrared markers were mounted. From these markers the positions of the eyes in space was calculated in real-time. These positions were used by a graphics workstation to update the visual scene so as to be consistent with a wall of random dots at a specific distance from the participant.

locking pattern. Both measures are expected to reveal the effective damping of the postural control system. As predicted, both measures varied with distance in a manner consistent with decreased stability for increased visual depth. Moreover, the gain of the visually induced sway decreased with increasing visual distance as predicted. That decrease was quantitatively weaker than expected; that is, gain dropped off far less than expected as visual distance was increased. This raised the issue of whether the parameters of the postural control system (i.e., damping, eigenfrequency, coupling strength to vision) are constant and independent of the optic environment.

That issue was addressed in a second experiment (Dijkstra, Schöner, Giese, & Gielen, 1994; Giese, Dijkstra, Schöner, & Gielen, 1996), in which the oscillation frequency of the optic display was varied (0.05, 0.1, 0.2, 0.3, 0.4, and 0.5 cycles/sec), while keeping peak visual velocity constant. Although the dependence of sway amplitude and relative phase between sway and visual motion was qualitatively consistent with a linear driven oscillator model, the observed resonance structure was too weakly developed to be quantitatively consistent with the model. Amplitude did not peak sharply when the system was driven at its preferred frequency of sway, nor did the relative phase between sway and visual motion change abruptly near that frequency. By estimating the model parameters trial by trial it was possible to show that these parameters do vary with the frequency of the oscillatory optic scene (although the fits confirmed, at the same

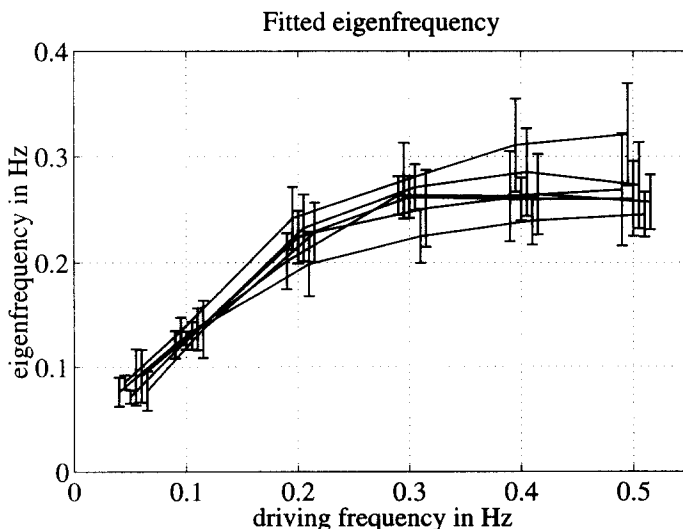


FIGURE 3 The dependence of eigenfrequency of the postural control system on modulation frequency of the optical surround.

time, the linear model structure)! The eigenfrequency of the postural control system ( $f = 2\pi\omega$ ) matched the optic driving frequency up to about 0.3 cycles/sec, where it leveled off (see Figure 3).

In other words, even when only visual information was manipulated, the eigenfrequency of the postural control system depended on that visual information. This may best be understood as a result of an adaptive process, in which the sway generated by oscillatory optic stimulation leads to adjustments in the nonvisual part of the postural control system that make the postural control system more responsive to visual stimulation (by adjusting eigenfrequency). This is, apparently, how the postural control system achieves better than predicted coupling to the moving optic environment (e.g., larger gain at high frequencies). The adaptation of eigenfrequency cannot be modeled as a minimization of the error between driving frequency and eigenfrequency, as would be standard in control theory. First, such an error-minimizing model would be specific to periodic motion and thus would not work for more general movements of the environment. Second, it would not capture the observed leveling off of adaptation beyond driving frequencies of 0.3 Hz. The adaptation calls into question the strategy of analysis, in which each sensory channel is studied in isolation. The control system for posture cannot be separated into sensory and motor components.

### Posture in the Haptic World

Does this result imply that visual input to the postural control system is dominant, so that posture in the visual frame overrides posture in the inertial frame? The answer is probably in the negative, supported by some experimental work on somatosensory driving of postural sway.

Lightly touching a stationary surface with a fingertip stabilizes posture in the experimental setup shown in Figure 4. (Jeka & Lackner, 1994, 1995). Jeka and colleagues exploited this fact to study the dynamic coupling of somatosensory information from touch into the postural control system. An oscillatory "tactile" environment was created by moving the surface (touch bar), on which the index finger rested, sinusoidally at very small amplitudes (3–32 mm). The frequency of this tactile motion was varied from 0.1 to 0.5 cycles/sec at constant amplitude in one experiment (Jeka, Schöner, Dijkstra, Ribeiro, & Lackner, 1997) and from 0.1 to 0.8 cycles/sec at constant peak velocity in another experiment (Jeka, Oie, Schöner, Dijkstra, & Henson, 1998). In all cases, touch was light: Applying more than one Newton set off an alarm that participants successfully suppressed.

The results of these experiments were likewise analyzed from the perspective of a (linear) dynamic control model. In which way might haptic information couple into postural control? The requirement to minimize the transmitted force at the fingertip may lead to the acquisition of a sensory-motor mapping that permits participants to maintain the fingertip immobile in the outer frame even as the

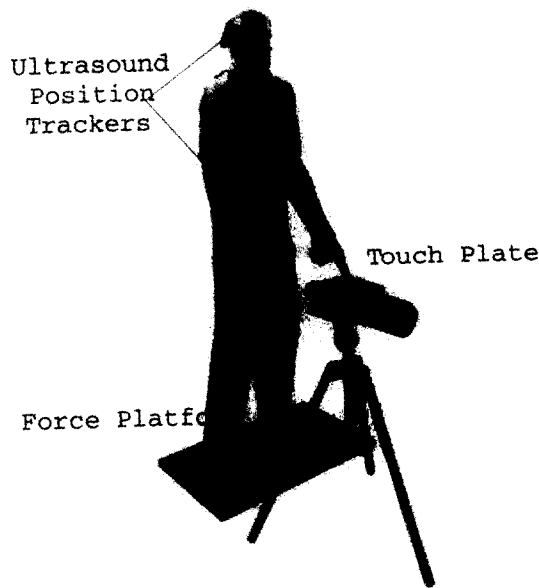


FIGURE 4 A participant depicted on the force platform with his right index fingertip on the touch plate positioned at waist level. The plate moved sinusoidally in the medial-lateral plane at different frequencies (0.1–0.8 Hz) in the moving plate conditions. For the sake of illustration, the alarm is sounding due to an applied force on the touch plate of greater than 1N (about 100 grams). This occurred on less than 5% of all experimental trials.

body sways. Such a map would provide information about the position of the body in space given the current sensory signal at the fingertip. Its contribution to postural control could be modeled by a coupling term that is linear in body position. On the other hand, a velocity might be directly sensed at the fingertip. This velocity is the relative velocity between the body and the touch surface. If, in a first approximation, the finger moves with the body, this velocity is thus the difference between the velocity of the body in the outer frame,  $\dot{x}$ , and the velocity,  $\dot{d}$ , of the touch surface in that frame. In summary, the simplest model has two coupling terms

$$\ddot{x} + \alpha \dot{x} + \omega^2 x = -c_p x - c_v (\dot{x} - \dot{d}) \quad (8)$$

The signs are chosen so as to express that touch tends to counteract movement relative to the touch surface.

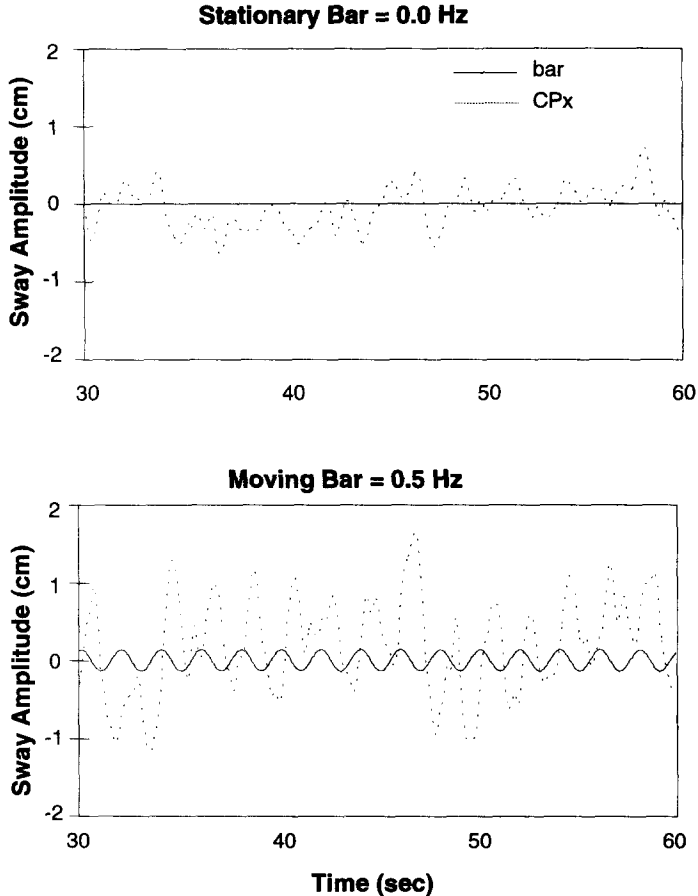


FIGURE 5 Overlaid time series of center of pressure (CPX), head (HX), and touch bar (TBX) displacement when the touch plate was (a) stationary and (b) moving at 0.5 Hz. Note how head and body sway entrain to the touch plate movement in (b) while in (a) body sway has no characteristic frequency.

The simplest, but perhaps most remarkable result is shown in Figure 5. Lightly touching an oscillating surface induces periodic sway, which is coupled remarkably tightly to the surface motion. This can be established by observing the coherence between sway and touch bar motion, which is high ( $> 0.9$  at all frequencies), similar to the visual case although the touch bar amplitude and hence the sway amplitude is smaller here. We have argued that relatively small driving amplitudes lead to high coherence with postural sway because the amplitude is compatible with sway amplitudes typically produced by the nervous system. Strong coupling is also

evidenced by the low variability of the relative phase between touch bar position and postural position.

More specifically, the relative phase between sway and touch bar motion as well as the *gain* (i.e., the ratio between the amount of sway motion and touch bar motion at the driving frequency) both display qualitatively the typical pattern of a linear driven oscillator. Figure 6 shows how relative phase decreases with increasing frequency, while gain has a maximum at intermediate frequencies, matching the typical frequencies of sway in stationary environments.

A closer, quantitative look reveals some interesting details and problems, however. Assuming that the model parameters damping, eigenfrequency, and the coupling constants do not depend on the frequency of the touch bar, these parameters

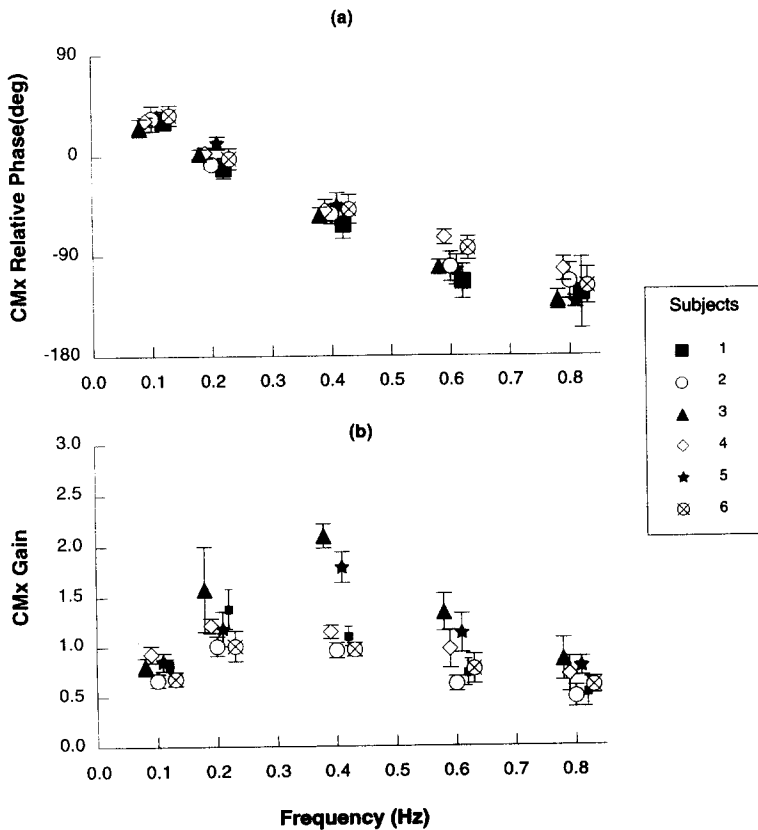


FIGURE 6 (a) Mean relative phase of the center of mass (CMX) versus touch plate displacement for all participants at each frequency of touch plate movement. (b) Mean CMX gain for all participants at each frequency of touch plate movement.

can be estimated by fitting the predicted relative phase and gain to the experimentally obtained phase and gain functions. The quality of the fits and the systematic errors (deviations that are clear functions of touch bar frequency) can then be interpreted.

First, such fitting enabled Jeka and colleagues (1998) to determine that both coupling to touch bar position and to touch bar velocity was necessary. If one of the two coupling coefficients was set to zero, the best fit remained bad. Essentially, the fact that phase lags beyond  $90^\circ$  implies that coupling to position must be taken into account. Similarly, the fact that positive relative phases occur (so that sway motion "anticipates" touch bar motion) implies that velocity coupling must be taken into account. This contrasts with the situation in which sway was driven by visual motion in depth. In that case, coupling to position did not emerge as a necessary component of the model. It is quite natural to assume that visual motion in depth does not clearly define a positional spatial reference frame, whereas somatosensory motion does. Constructing a reliable depth estimate from visual information is clearly a much more difficult problem than computing positional information from touch sensors. In the first case, a number of fundamental problems in vision must be resolved, whereas the latter depends merely on a model of the touch sensor, which the nervous system may have learned.

A second observation from these fits is that adaptation might likewise take place in this haptic scenario as in the visual case. Evidence for this arises somewhat more indirectly as fitting the parameters of the linear model trial by trial is not yet complete. The systematic deviations from the constant parameter linear model in phase and gain make it possible, however, to obtain some intuition about these adaptive changes. Thus, for instance, is the gain typically underestimated in the model at high frequencies. This is consistent with an adaptation of eigenfrequency similar to the one observed more directly in the visual case. Moreover, the relative phase is estimated as a less pronounced function of frequency than that observed (Jeka et al., 1998). This is also consistent with an adaptation of eigenfrequency. Thus, the assumption that the parameter values of the linear postural model remain constant as the touch bar frequency is varied is clearly violated. We confirmed this by fitting higher order models to the data, which did not produce better results.

Overall, the weak resonance structure (i.e., no clear peak of gain at any frequency) and high level of the gain at all frequencies as well as the strong coupling (i.e., high coherence) at all frequencies indicates that adaptive changes in response to a moving touch surface make the system more sensitive to somatosensory information. This result calls into question the strategy of analyzing the different sensory channels separately. As the haptic channel is analyzed, the control system changes central parameters such as eigenfrequency and damping, which presumably are shared by other sensory channels. Moreover, this outcome shows that these adaptive changes are not reflective of some inherent dominance of particular sensory input channels. Although the cited results were

obtained when participants held eyes closed, recent unpublished data of Jeka's group have revealed the same pattern for phase and gain when eyes were open in a stationary environment.

### BEYOND CONTROL THEORY: SWITCHING COUPLING MODE IN RESPONSE TO PERCEIVED MOTION

We have touched on differences between the control-theoretic viewpoint and the dynamic approach. The fact that postural behavior was modelled as having a single stable state (upright posture mapped onto  $x = 0$  by choice of coordinates) has not made those differences apparent at the level of the mathematical modeling. Here we describe an attempt to model a phenomenon that lends itself to analysis within the dynamic approach, but less easily from a control-theoretic viewpoint. The phenomenon consists of a change in control mode as a moving stimulus is either perceived as background or as foreground motion. This change may be controlled, for instance, by the amplitude of visual motion. Because we only have very preliminary data on the problem, the modeling is more speculative in nature. The goal is to illustrate the principles, rather than to provide an accurate account.

The stimuli used in the Dijkstra, Schöner, Giese, et al. (1994) experiments had rather small amplitudes of motion (a few centimeters in the visual case, a few millimeters in the haptic case) whereas most other studies used larger amplitudes (e.g., 50 cm or more in van Asten et al., 1988). Whereas we found near constant gain of about one (independently of frequency), other authors found much smaller gains, which decreased with increasing frequency. We hypothesize that the postural control system may switch control mode when external (foreground) motion is detected, which is not caused by self-motion. Such detection changes the interpretation of the sensed motion from self- to object motion. In one mode, the system responds to the expansion rate,  $e(t)$ , in the visual case, or to relative motion of the center of mass and the touchbar surface in the haptic case (this is consistent with the model by Schöner, 1991). In the other mode, the system couples to the perceived external motion,  $D(t)$ , which does not depend on the postural state,  $x$  and  $\dot{x}$ . Each of the two coupling terms has a coupling strength ( $c_s$  for the self-motion strategy and  $c_o$  for the object motion strategy) and a degree of activation ( $w_s$  for self-motion,  $w_o$  for object motion):

$$\ddot{x} + \alpha\dot{x} + \omega^2x = -c_s w_s e(t) - c_o w_o D(t) \quad (9)$$

(notation for the visual case). The activation variables range from 0 (corresponding mode is "off") to 1 (corresponding mode is "on"). The activation variables are controlled through a dynamical system that expresses the following constraints. For large input amplitude, an attractor at 1 is stabilized, for small input amplitude an at-



tractor at 0 is stabilized, and the two activation variables compete with each other. A typical dynamics of this type is (cf. Schöner, 1989)

$$\dot{w}_i = \alpha_i w_i (1 - w_i^2) - \gamma w_i w_o^2 \quad (10)$$

$$\dot{w}_o = \alpha_o w_o (1 - w_o^2) - \gamma w_o w_i^2 \quad (11)$$

where  $\alpha_i$  and  $\alpha_o$  determine the competitive advantage of each mode, and  $\gamma$  represents the strength competition. How the stimulus amplitude controls the factors  $\alpha_i$  and  $\alpha_o$  is another issue, which we gloss over here (a simplest idea is that perceiving foreground motion leads to large values of  $\alpha_o$ , not perceiving such motion leads to small values of  $\alpha_o$ ). Such dynamics typically possesses a bistable regime, in which either of the two activation variables may win the competition. There are, of course, many other mathematical formulations of such activation dynamics, including "neural" ones (our formulation is the simplest in terms of bifurcation theory; cf. Schöner & Dose, 1992).

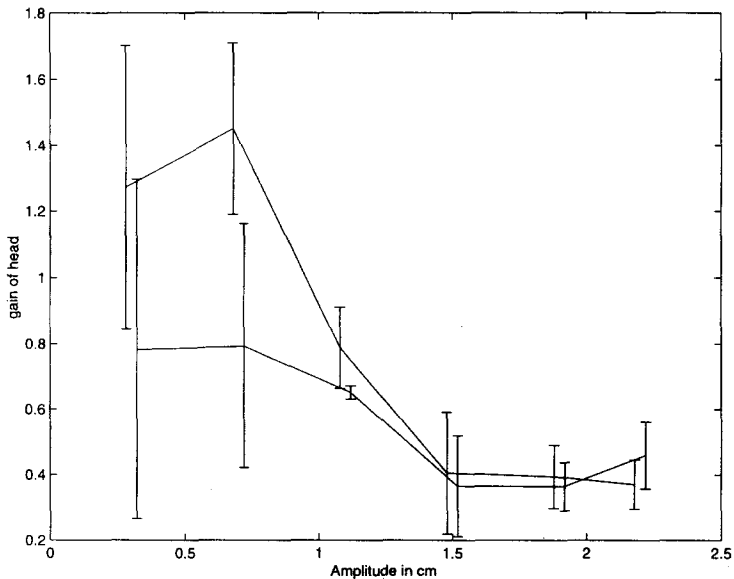


FIGURE 7 Gain of the head versus amplitude of touchbar for two participants. Errorbars denote standard deviation. These are results of a preliminary experiment to illustrate the transition from sensed motion interpreted as self-motion to sensed motion interpreted as object motion.

It is hard to conceive of a control-theoretic model of this state of affairs. The system is not minimizing any error in the object motion case. Also, adaptive control does not provide the right framework because the system is not tuning a parameter of the postural controller. Instead the activation variables describe the switching between control modes in stable and smooth fashion (e.g., including hysteresis).

We have performed a preliminary experiment that illustrates that the aforementioned modeling is on the right track. In the light-touch setup (Figure 4) we had participants touching the moving touchbar. We kept frequency constant at 0.2 Hz but varied amplitude from 3 mm up to 22 mm. The amplitude of 3 mm is at the 2-point threshold whereas in the 22 mm condition the touchbar can clearly be sensed as moving independently. The gain of the head movement is illustrated in Figure 7. For the amplitudes of 3 and 7 mm the gain is around 1 and for larger amplitude it falls off. For amplitudes below 10 mm, the touchbar motion is interpreted as self-motion and above this threshold it is interpreted as object motion.

## DISCUSSION

By studying the stabilization of posture by visual and other sensory information, we have illustrated that an appropriate analysis of this behavior is in terms of coupling of sensory information into a dynamical system governing postural control. Coupling involves additional parameter (coupling strength), and the effects of coupling depend on the temporal characteristics of both sensory information and control system. It is therefore not sufficient to talk about the specification of information relevant to posture by the sensory array as such specification refers purely to the computation of the desired state of the postural system. Second, we found that the perception-action loop is not only an online control system, in which the current sensory information affects the effector system. The system as a whole adjusts to particular conditions of stimulation and behavior. Thus, exposing a standing participant to a moving optic environment leads to a postural control system specific to that situation. Characterizing the system under those conditions does not yet predict how the system will behave in another sensory environment. In this respect, the classical strategy of separately analyzing different sensory channels by perturbing them separately cannot be expected to lead to a complete analysis of postural control.

These two points are more dramatically illustrated if we look into "higher" perception action behaviors, such as the next few ones listed by Gibson.

### The Initiation of Movement

A more complex class of perception-action behaviors involves the initiation of movement, the activation of a motor behavior, in response to sensory events of varying complexity. Catching a ball, anticipating impact, soft collision such as

when stopping at a mark, hitting a mark on the ground while running, and so forth, are examples of perception-action behaviors in which a movement event is related to a sensory event (for review, see Lee & Young, 1986; Warren, Young, & Lee, 1986). Lee and colleagues looked into a class of such behaviors with a view to identify parameters that can be obtained from the sensory array and that specify the timing of the action to be initiated. The visual expansion rate divided by eccentricity, or the reciprocal of this ratio, the tau-parameter, was identified (Lee, 1976) as a potential parameter of this nature in a wealth of different situations including gannets that fold their wings in time before diving into water (Lee & Reddish, 1981), humans anticipating fall by cocontracting appropriate muscles (Dietz & Noth, 1978), or flies extending their legs to anticipate the impact on landing (Borst, 1986; Wagner, 1982). At constant approximation speed to a surface, this parameter identifies the time until contact. Empirical evidence in support of this parameter was obtained when the times at which action was initiated were constant if scaled to the tau parameter as initial distance or speed varied. This scaling holds to good precision in cases in which approach velocity is constant, less so when approach is accelerated.

In line with Gibson's discourse, this ensemble of results has been theoretically summarized by saying that action is initiated in response to particular threshold values of parameters such as  $\tau$  (or  $\dot{\tau}$ ). These parameters are directly specified from the sensory array. This manner of initiating action in response to changes in the sensory array has been referred to as the "gearing of action to perception" (Bootsma & van Wieringen, 1990; Lee, Young, Reddish, Lough, & Clayton, 1993), implying an aspect of control. Schöner (1994) showed how the initiation of an action in response to changes in the sensory array can be achieved while the underlying control system is continuously coupled to sensory parameters. In his account, a dynamical system controlling action has at least two distinct states, one corresponding to the postural state before action is initiated, the other to the postural state after action is initiated. The sensory parameter controls an instability in which the first state is rendered unstable and the system switches to the second state. The structure of the dynamical system that generates such nonlinear behavior then leads to coordination of action with the sensory influx. This account predicts that continuous corrections of relative timing between the action and the sensory event occur during the approach. Experimental evidence of this sort had earlier been obtained by Lee and colleagues (1983) in jumping to hit a ball and by Bootsma and van Wieringen (1990) in the table tennis stroke (and it was that evidence that motivated the notion of gearing action to perception).

This proposal deals quite naturally also with accelerated approach situations. The perception-action coupling needs to be of appropriate strength and functional form to assure stable relative timing of action to the sensory event. This account thus leads quite naturally to the hypothesis that processes of adaptive adjustment exist that tune this coupling to the particular sensory situation based on feedback about the success of the perception-action loop.

### **Less Simple Perception Action Patterns: Obstacle Avoidance, Homing, Spatial Orientation**

Very little work has been done on perception–action systems that go beyond these simplest two cases. If we think of obstacle avoidance as a perception–action pattern, then it contains a new element compared to posture and the initiation of action that might best be characterized as a form of decision making. An individual obstacle, for instance, may be circumnavigated to the left or to the right (in two dimensions). In the presence of a number of obstacles, movement paths might emerge that pass in between two obstacles or such paths might be avoided if the obstacles are sufficiently close to each other.

The idea of coupling parameters that are derived directly from the sensory array into nonlinear dynamical systems that generate action has been demonstrated in synthesis. Autonomous robot vehicles have been built that avoid obstacles, find a home position, stop softly, and so forth (Bicho & Schöner, 1997; Neven & Schöner, 1996; Schöner, Dose, & Engels, 1995; Steinhage & Schöner, 1997).

The work abounds in Gibsonian-style exploitation of the ecology of vision. For instance, the optic flow can be used to avoid obstacles simply by detecting regions in which flow increases with eccentricity more than average, suggesting smaller visual depth. Homing can be based on a computed optic flow between a memorized image taken at the home position and the current image (Neven & Schöner, 1996). Behavior is generated by enacting the stable states of a dynamical system of heading direction, into which the sensory information is coupled either attracting to targets or repelling from obstacles. The approach works quite well even if sensory information at a very low level is used. For instance, Bicho and Schöner (1997) implemented obstacle avoidance and photo taxis (moving toward light sources) with only five infrared detectors and two photoresistors. In such cases, one cannot speak of obstacles as objects detected in the outer world. Instead, every sensor is assumed to specify an obstacle, from which heading direction is repelled. The strength of this repulsion is tuned, however, with the detected distance so that at sufficiently large distance the repulsive force is zero.

Decision making emerges from such coupled perception action systems through instabilities and multistability. For instance, an obstacle detector generating a single repulsive contribution to the heading direction dynamics already specifies two attractors—one to the left and one to the right of the zone of repulsion. Two detectors may lead to a single repeller if the repulsion zones are sufficiently broad. But they may also generate two separate repellers with an attractor in between when the repulsion zones are sufficiently narrow. In the first case, the vehicle avoids the obstacle zone by going around it. In the second case, it may steer a course that cuts through the zone.

Another conceptually interesting issue is related to specification. In homing, for instance, the optic flow between memorized image and current image specifies the direction into which the robot needs to move to minimize that flow and thus to re-

turn to the home position. This information does not specify the home position (e.g., in terms of distance to the home position). In a closed-loop control system, however, homing can still be achieved as a form of visual servoing, in which action (moving the robot into a particular heading direction with a particular velocity) is based on the current visual information. A stable stationary state is reached when the optic flow between the memorized image taken at the home position and the current image is zero.

For certain behaviors this form of achieving a behavior in the absence of precise specification is not sufficient. For instance, when different sources of navigational information such as dead-reckoning and optic landmark information must be fused or reconciled, a common data format must be found. This leads to the necessity of representing also the presence or absence of sensory information. For example, although no landmark has been recognized, visual navigational information is not available (Steinhage & Schöner, 1997). The need for additional concepts in such cases was recognized by Gibson (see second paragraph of his section IV). Much of cognitive science is addressing such issues. A task that is not yet satisfactorily solved consists of understanding how representational concepts are reconciled with dynamic concepts. First steps in such a direction have been made through dynamic field theory (Schöner, Kopeck, & Erlhagen, 1997).

## CONCLUSION

Gibson provided in his 1958 article a useful behavioral analysis of perception action systems. Two aspects that we emphasized here were, however, less clearly articulated: (a) The processes of steering action through sensation is based on coupling, and (b) the notion of specification, which is sometimes used to describe how desired behavioral states are determined by the sensory array, is but one aspect of coupling, that is neither necessary (as coupling can stabilize states that were never explicitly computed from the sensory array) nor sufficient (as the computation of the desired state of a control system leaves open the nature and strength of the stabilization processes). Adaptive change of the entire control system reflects that such perception-action coupling is an active processes involving internal representation of parameters that reflect the current environment.

We have addressed these issues by employing a set of concepts based on the theory of dynamical systems as a language in which perception-action systems can be analyzed. Employing this formal language offers advantages over the plain English used by Gibson (incidentally, although Gibson refers to his analysis of locomotory behaviors as *formulae*, he abstains from translating them into mathematical equations. Those of us who try to do just that know that it is far from trivial.) First, the formalized account leads to quantitative prediction. In particular, quantitative fit of the parameter values of a dynamical model contributes both to a precise description of the system (useful, for instance, for clinical evaluation), but also uncovers

weak spots of the model and thus of the account. In our examples, the fact that certain parameters depended on task and stimulus forced us to interpret the model differently and led to the hypothesis that adaptive processes are at work. For instance, the eigenfrequency of the postural control system was apparently altered by exposure to a moving optic surround. Second, dynamical systems theory provides a concept for what is the simplest system compatible with a set of observed behaviors. Such normal forms of dynamical models are useful, because they express the necessary properties that any other (e.g., more mechanistic) model must possess. This method of normal form modeling was used by Schöner (implicit in the 1991 article, explicit in the 1994 article), where more details can be found.

Some important problems of organizing perception–action behaviors have been largely overlooked by Gibson and apparently continue to be ignored in large part by the research community. First, the integration of different behaviors into the same system is a difficult problem. This is well known in autonomous robotics where the problem arises from the start. In nervous systems this relates to how elementary behaviors are activated or deactivated, and how they interact with each other and how they reconcile contradictory demands. From the theoretical perspective that we have reviewed, integration can be achieved through nonlinear dynamical systems that reside at sufficiently abstract levels at which behaviors are generated. Although this program has been largely realized in synthesis now, its role in the analysis of real nervous systems must still be clarified.

A second issue relates to what in autonomous robotics is sometimes called “natural environments.” Very little is actually known about the limits of the behavioral competencies obtained through perception–action couplings. Under which environmental conditions do these behaviors break down, and to which conditions can these behaviors adapt? In light of our findings about the broad adaptation of the postural control system across different modalities and without a clear hierarchy of dominance, this question becomes particularly interesting. Why is the postural system able to adapt to oscillating optic and tactile environments? Would it likewise adapt to an irregularly moving environment? Artificially opening the visual feedback loop often leads to erratic and unstable behavior (Heisenberg & Wolf, 1988). Is this informative about the limits of the adaptive capacity of the nervous system, so that “unnatural” perturbations would not specify adaptive change? The challenge is now to study behaviors beyond the simplest perception–action patterns to reveal both the insights and limitations of Gibson’s ideas.

### ACKNOWLEDGMENTS

John Jeka was supported by NIH R29 NS35070–02 and NSF SBR–9709361 and Tjeerd Dijkstra was supported by NSF SBR–9809447.

We thank Peter Beek, Bruce Kay, and Geoffrey Bingham for their detailed comments.

## REFERENCES

- Aloimonos, Y., & Rosenfeld, A. (1991). Computer vision. *Science*, 253, 1249–1254.
- Bicho, E., & Schöner, G. (1997). The dynamic approach to autonomous robotics demonstrated on a low-level vehicle platform. *Robotics and Autonomous Systems*, 21, 23–35.
- Bootsma, R. J., & van Wieringen, P. C. W. (1990). Timing an attacking forehand drive in table tennis. *Journal of Experimental Psychology: Human Perception and Performance*, 50, 197–209.
- Borst, A. (1986). Time course of the houseflies' landing response. *Biological Cybernetics*, 54, 379–383.
- Brandt, T., Wist, E. R., & Dichgans, J. (1975). Foreground and background in dynamic spatial orientation. *Perception & Psychophysics*, 17, 497–503.
- Dietz, V., & Noth, J. (1978). Pre-innervation and stretch responses in triceps brachii in man falling with and without visual control. *Brain Research*, 142, 576–579.
- Dijkstra, T. M. H., Schöner, G., & Gielen, C. C. A. M. (1994). Temporal stability of the action–perception cycle for postural control in a moving visual environment. *Experimental Brain Research*, 97, 477–486.
- Dijkstra, T. M. H., Schöner, G., Giese, M., & Gielen, C. C. A. M. (1994). Frequency dependence of the action–perception cycle for postural control in a moving visual environment: Relative phase dynamics. *Biological Cybernetics*, 71, 489–501.
- Gibson, J. J. (this issue). Visually controlled locomotion and visual orientation in animals. *British Journal of Psychology*, 49, 182–194. (Original work published 1958)
- Giese, M., Dijkstra, T. M. H., Schöner, G., & Gielen, C. C. A. M. (1996). Identification of the nonlinear state space dynamics of the action–perception cycle for visually induced postural sway. *Biological Cybernetics*, 74, 427–437.
- Heisenberg, M., & Wolf, R. (1988). Reafferent control of optomotor yaw torque in *Drosophila melanogaster*. *Journal of Comparative Physiology*, 163, 373–388.
- Horak, F. B., Diener, H. C., & Nashner, L. M. (1989). Influence of central set on human postural responses. *Journal of Neurophysiology*, 62, 841–853.
- Jeka, J. J., & Lackner, J. R. (1994). Fingertip contact influences human postural control. *Experimental Brain Research*, 100, 495–502.
- Jeka, J. J., & Lackner, J. R. (1995). The role of haptic cues from rough and slippery surfaces in human postural control. *Experimental Brain Research*, 103, 267–276.
- Jeka, J., Oie, K., Schöner, G., Dijkstra, T. M. H., & Henson, E. (1998). Adaptive velocity and position coupling of postural sway to somatosensory drive. *Journal of Neurophysiology*, 79, 1661–1674.
- Jeka, J. J., Schöner, G., Dijkstra, T. M. H., Ribeiro, P., & Lackner, J. R. (1997). Coupling of fingertip somatosensory information to body sway during stance. *Experimental Brain Research*, 113, 475–483.
- Johansson, R., Magnusson, M., & Åkesson, M. (1988). Identification of human postural dynamics. *IEEE Transactions on Biomedical Engineering*, 35, 858–869.
- Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception*, 5, 437–459.
- Lee, D. N., & Aronson, E. (1974). Visual proprioceptive control of standing in human infants. *Perception & Psychophysics*, 15, 529–532.
- Lee, D., & Lishman, J. R. (1975). Visual proprioceptive control of stance. *Journal of Human Movement Studies*, 1, 87–95.
- Lee, D. N., & Reddish, P. E. (1981). Plummeting gannets: A paradigm of ecological optics. *Nature*, 293, 291–294.
- Lee, D., & Young, D. S., (1986). Gearing action to the environment. In H. Heuer & C. Fromm (Eds.), *Generation and modulation of action patterns* (pp. 217–230). Berlin: Springer-Verlag.
- Lee, D. N., Young, D. S., Reddish, P. E., Lough, S., & Clayton, T. M. H. (1983). Visual timing in hitting an accelerating ball. *Quarterly Journal of Experimental Psychology*, 35A, 333–346.
- Neven, H., & Schöner, G. (1996). Neural dynamics parametrically controlled by image correlations organize robot navigation. *Biological Cybernetics*, 75, 293–307.

- Sandini, G., Gandolfo, F., Grosso, E., & Tistarelli, M. (1993). Vision during action. In Y. Aloimonos (Ed.), *Active perception* (pp. 151–190). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Schöner, G. (1989). Learning and recall in a dynamic theory of coordination patterns, *Biological Cybernetics*, 62, 39–54.
- Schöner, G. (1991). Dynamic theory of action–perception patterns: The “moving room” paradigm. *Biological Cybernetics*, 64, 455–462.
- Schöner, G. (1994). Dynamic theory of action–perception patterns: The time-before-contact paradigm. *Human Movement Science*, 3, 415–439.
- Schöner, G., & Dose, M. (1992). A dynamical systems approach to task-level system integration used to plan and control autonomous vehicle motion. *Robotics and Autonomous Systems*, 10, 253–267.
- Schöner, G., Dose, M., & Engels, C. (1995). Dynamics of behavior: Theory and applications for autonomous robot architectures. *Robotics and Autonomous Systems*, 16, 213–245.
- Schöner, G., & Kelso, J. A. S. (1988). Dynamic pattern generation in behavioral and neural systems. *Science*, 239, 1513–1520.
- Schöner, G., Kopecz, K., & Erhagen, W. (1997). The dynamic neural field theory of motor programming: Arm and eye movements. In P. G. Morasso & V. Sanguineti (Eds.), *Self-organization, computational maps and motor control* (Psychology Series, Vol. 119, pp. 271–310). Amsterdam: Elsevier–North-Holland.
- Steinhage, A., & Schöner, G. (1997). Self-calibration based on invariant view recognition: Dynamic approach to navigation. *Robotics and Autonomous Systems*, 20, 133–156.
- van Asten, W. N. J. C., Gielen, C. C. A. M., & Denier van der Gon, J. J. (1988). Postural adjustments induced by simulated motion of differently structured environments. *Experimental Brain Research*, 73, 371–383.
- Wagner, H. (1982). Flow-field variables trigger landing in flies. *Nature*, 297, 147–148.
- Warren, W. H., & Saunders, J. A. (1995). Perceiving heading in the presence of moving objects. *Perception*, 24, 315–331.
- Warren, W. H., Young, D. S., & Lee, D. N. (1986). Visual control of step length during running over irregular terrain. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 259–266.



