

19. Model-Based Control of Perception/Action

Jack M. Loomis and Andrew C. Beall

Department of Psychology
University of California, Santa Barbara, CA, USA

1 INTRODUCTION

The coupling of perception and action in humans and other species is one of the major questions in cognitive science and neuroscience. The empirical domain of perception/action includes reaching and grasping, skilled use of tools, visual control of locomotion (e.g., driving and flying), dance, and individual and team sports. Understanding of perception/action in the human requires research on perceptual and cognitive representations of 3-D space, internal representation of the body, motor control of the limbs and effectors, and internalization of vehicle dynamics. Like many others in the field, we believe that research on perception/action will advance more quickly in the future by giving greater emphasis to active control tasks over simple psychophysical judgments and by studying skilled actions of moderate complexity, like driving and sport activity, instead of the stripped down tasks of the past that have permitted experimental rigor at the cost of failing to engage many of the perceptual, cognitive, and motor processes involved in complex action.

2 THREE CONCEPTIONS OF PERCEPTION/ACTION

There are three prominent conceptions of how perception controls action. These are presented in Figure 1 in the context of visual control of action (see also Philbeck et al., 1997; Warren, 1998). The first of these, the “ecological” conception, originated with Gibson (Gibson, 1950, 1958, 1966, 1979) and has since been elaborated by many others (e.g., Flach & Warren, 1995; Lee, 1993; Turvey, 1977; Turvey & Carello, 1986; Warren & Wertheim, 1990; Warren, 1998). An important assumption is that the rich and multifaceted sensory

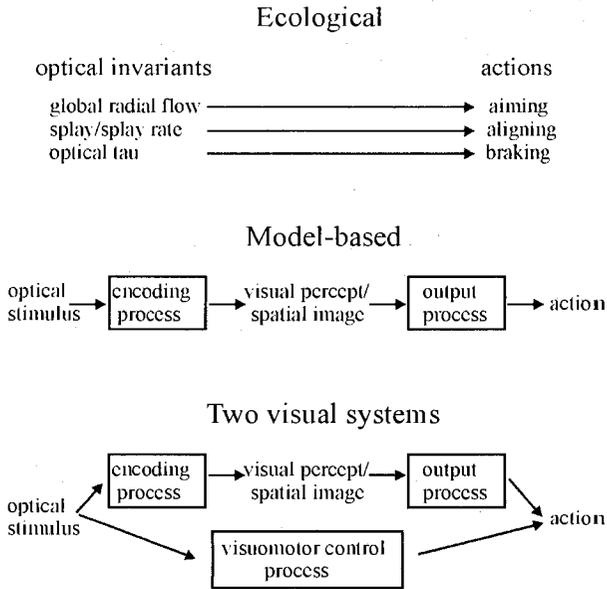


Figure 1. The three conceptions of how visual perception controls action.

stimulation of a moving actor is sufficient for understanding the control of action (i.e., additional information, like that stored in the actor's memory, need not be brought to bear). Moreover, it is assumed that very specific aspects of this changing sensory stimulation are tightly coupled to particular aspects of the desired action. Thus, in connection with the visual control of locomotion, aiming toward a point target might be controlled by global radial outflow (e.g., Warren, 1998; Warren et al., 1988), braking might be controlled by the temporal derivative of tau (Lee, 1976; Yilmaz & Warren, 1995), and aligning with a straight path might be controlled using splay (the angle between the perspective image of the path and environmental vertical) and its temporal derivative, splay rate (Beall & Loomis, 1997; Calvert, 1954; Riemersma, 1981). Specific optical variables that are sufficient for controlling particular actions are referred to as optical invariants. Strategies for utilizing these optical invariants are referred to as optic flow rules (e.g., Loomis & Beall, 1998) or laws of control (Warren, 1998). Some ecological researchers consider the mechanisms of perception and cognition and the internal representations resulting from them to be superfluous to functional explanations of action.

The second conception, referred to here as the "model-based" conception (Loomis & Beall, 1998; Warren, 1998), derives from the long tradition of

research on space perception. In this conception, spatiotemporal stimulation arising from activity within the physical world is registered by the sense organs and processed by the central nervous system to yield an internal representation of the perceiver's environment. In the case of vision, the internal representation of a single target is the visual percept, and the visual representation of surrounding 3-D space is referred to as visual space. To generalize to all of the spatial senses (vision, audition, and touch), we use the corresponding terms of percept and perceptual space. These internal representations are much simpler in informational content than the spatiotemporal stimulation to which they correspond, for much of the stimulus information is lost in the perceptual processing. A great deal of research over the years has been directed toward establishing the correspondence between physical and perceptual space and toward understanding the sensory and perceptual processes underlying this representation. It is generally recognized that sensory stimulation, when available, is a primary determinant of perceptual space. Thus, identical patterns of sensory stimulation, whether arising from real objects or virtual objects (those simulated by a virtual reality system) ought to result in nearly identical percepts (e.g., Loomis, 1992). In addition, internal assumptions (such as smoothness of optic flow, smoothness of surfaces, and rigidity of moving objects) and expectancies based on past experience also serve as important determinants of perceptual space. Thus, when stimulus determinants are greatly reduced (as in dimly lit environments), perceptual space is still well defined even when discrepant with respect to physical space (e.g., Ooi et al., 2001; Philbeck et al., 1997).

In this conception, concurrent percepts play a causal role in the control of action (for a more extensive argument, see Philbeck et al., 1997). The actor plans his/her action with respect to perceptual space and observes the consequences of the action within it. Because of the close correspondence between physical space (or action space) and perceptual space, the actor is usually successful in carrying out the desired action. However, even when perceptual input is interrupted (e.g., by closing the eyes or blocking the ears), continued action towards some goal is still often successful, indicating that more abstract spatial images mediate such actions; in this case, these spatial images are the causes of behavior in the same way that percepts are causes of behavior when sensory input is available. Besides these internal representations of the physical world (percept and spatial image), the model-based conception depicted in Figure 1 postulates an internal representation in the output process that intervenes between the percept or spatial image and the action. More will be said about this later.

The third conception, termed "two visual systems", has attained prominence in recent years because of the confluence of two trends: the increasing recognition of the functional modularity of mind (e.g., Fodor, 1983) and the deepening understanding of the functional architecture of the

brain, especially those pathways dealing with vision (Milner & Goodale, 1995). Turvey (1977) was one of the first to express the view that visually-based action might be controlled by levels of the nervous system that are inaccessible to conscious awareness. More recently, evidence has been accumulating that conscious visual perception is sometimes dissociated from those processes controlling visually-based action. "Blindsight", the visually-based capacity for performing certain actions like pointing to and reaching for targets in the absence of visual experience but with visual input available, is perhaps the strongest evidence of dissociation (Weiskrantz, 1986). Milner and Goodale (1995) summarize much of the evidence from patients with brain lesions, including their own work with a brain damaged but sighted patient who can readily reach for visual objects that she cannot otherwise discriminate. Their research and other research like it (e.g., Bhalla & Proffitt, 1999) point to the existence of two distinct visual systems. The visuomotor control system acts on concurrent perceptual input and need not involve conscious awareness. The other system is the same as that discussed in connection with the model-based approach.

For our present purposes, we group visuomotor control with the ecological conception, for both operate solely on concurrent sensory input. Here, we will be most interested in the contrast between the idea of control of action by concurrent optic flow information and that of control by a 3-D representation of surrounding space that may be concurrent with sensory input (i.e., a percept) or not (i.e., a spatial image).

3 SIGNIFICANCE OF CHANGE BLINDNESS

Recent research on change blindness has shown that the impression of a perceptual space that is coherent over extended epochs and regions of surrounding space is illusory (Rensink, 2002); instead, the perceptual representation that is constructed concurrently with sensory input exhibits coherence only over very limited regions of space and time. However, to the extent that perceptual representations of surrounding space are involved in controlling action, change blindness implies that these representations must be limited in spatiotemporal scope. Even so, these perceptual representations must be sufficient, given the success of most perception/action. Thus, for an actor running over terrain, the actor's brain creates a representation of the ground surface "just in time", a representation that is usually sufficient to permit safe travel. The recent work of Land and Hayhoe (2001) on the patterns of eye fixations while performing mundane tasks is consonant with the idea that visual perception is highly selective in creating representations that are "just in time" and sufficient for the action being controlled.

4 ELABORATION OF MODEL-BASED CONTROL

The brief discussion above of model-based control focused on perception, the input component of the perception/action cycle. Here we elaborate on this input component as well as extend the idea of model-based control to the output side as well.

4.1 Input Side

4.1.1 Percept

Our starting point is the recognition that percepts (perceptual representations) of the surrounding environment mediate most interactions between the human actor and his/her environment. One of the profound ideas to come out of philosophical and scientific inquiry is the realization that the world we see, hear, and touch in everyday life is not the physical world itself but an amazingly functional representation of the physical world created by our senses and central nervous system (e.g., Lehar, 2002; Loomis, 1992; Russell, 1948; Smythies, 1994).

One convincing argument comes from the study of color vision. Research on color vision has made it abundantly clear that the tri-dimensionality of color is a consequence of visual processing rather than a property of spectrally varying light. The fact that color is part of the fabric of the surfaces in visual space is a strong clue that these surfaces are themselves representations of the corresponding physical surfaces. An interesting phenomenon that also makes this point is "apparent concomitant motion", the experience of seeing compelling motion in depth-reversed objects as the person moves his/her head side-to-side (see Gogel, 1990 for an explanation); the apparent movement of a concave mask of a human face is a familiar example but much more compelling are the 3-D paintings by the British artist Patrick Hughes that are extremely rich in visual perspective.

The idea that visual space is an internal representation can be extended to the other senses, as well as to the cognitive, social, and emotional realms. Accordingly, philosophers have referred to the world of everyday experience as the phenomenal world (e.g., Smythies, 1994). It includes all of perceptually external space, the phenomenal body (body image), and the province of subjective experience (Brain, 1951; Lehar, 2002; Loomis, 1992; Ramachandran & Blakeslee, 1998; Russell, 1948; Simmel, 1966; Smythies, 1994). The evidence for this view provided by perceptual science, neurology, and cognitive neuroscience is overwhelming. Because the phenomenal world

is such a functional representation of the physical world, most laypeople and many scientists, even some perceptual scientists, fail to appreciate that contact with the physical world is indirect. Given that the phenomenal world can be so functional as to be self-concealing of its very nature, it is difficult to imagine that the phenomenal world is merely epiphenomenal. In a later section, we will present evidence that visual and auditory percepts in 3-D space are indeed causes of action.

4.1.2 Spatial Image

Even when visual, auditory, and haptic stimulation about the surrounding environment is temporally blocked, people continue to maintain some sort of internal 3-D representation of the environment that is of sufficient fidelity to support continued action for a short period of time (see section below on perceptually directed action). We refer to this representation as a spatial image. As the person moves through space, this representation is updated in much the way that the percept would be updated were sensory stimulation still available.

4.2 Output Side

The idea of model-based control of perception/action is not limited to the idea of an internal representation of the surrounding environment but applies to the output side as well. We begin by recognizing the complexity of motor control. Besides moving ourselves about under our own propulsion, we also perform many activities in which we maneuver self-propelled vehicles through space or in which we manipulate tools and other implements to achieve our goals. In many cases, feedback control is sufficient to perform the action (e.g., tracing a line with a pencil, steering a motorboat so as remain directly behind a leader moving with constant speed) but in many other cases, feedforward control is either essential (as when sensory feedback is temporarily lost) or is used to reduce tracking error (as when steering along a curving road) to a manageable level that allows for stable feedback control (McRuer et al., 1977). Fluent and effective control of complex behavior requires that the human has internalized the dynamics of the body, vehicle, and/or tool (Godthelp, 1986; Godthelp & K appler, 1988; Loomis, 1992; Loomis & Beall, 1998; McRuer et al., 1977; Miall et al., 1993; Wolpert & Ghahramani, 2000; Wolpert et al., 1995). It is obvious that the skill of performing gymnastic maneuvers, of manipulating a tool (e.g., using a scalpel in surgery, wielding a tennis racket in tennis play) or of controlling a vehicle

(e.g., automobile, airplane, helicopter, vessel) comes only with extensive practice. In such cases, the skilled actor has created an internal model of the dynamics of the plant (body, vehicle, or tool). Once acquired, the flexible internal model can be modified by parametric adjustment to allow for transfer of the skill to other plants that are functionally similar (e.g., when transitioning from a light and maneuverable aircraft to a heavy and sluggish one).

5 LEVELS OF CONTROL

It is widely recognized that the control of locomotion, like the control of action more generally, is a hierarchy of control problems (Dickmanns, 1992; Lee & Lishman, 1977; McRuer et al., 1977); for each problem, a goal is set and actions are then taken to achieve that goal. With respect to controlling locomotion, McRuer et al. (1977), identified three levels of control: "navigation", "guidance", and "control". The uppermost level, "navigation", refers to the overall selection of a route within geographical space and will not be discussed further. The next lower level, "guidance" (McRuer et al., 1977) involves planning an optimal path to the goal based on current perceptual information and then attempting to follow that path. For steering a car along a curving road, McRuer et al. (1977) assume that guidance depends on perception of the road ahead and that control inputs to the vehicle are largely feedforward (see also Donges, 1978; Godthelp, 1986; Godthelp & K ppler, 1988; Land, 1998). Gibson and Crooks (1938) presented an intriguing analysis of guidance in connection with car driving employing the notion of the "field of safe travel", an internal representation of the open space in front of the driver that is momentarily suitable for forward progress (for a similar idea in the context of autonomous vehicles see Dickmanns, 1992). The lower level, referred to as "control" (McRuer et al., 1977) or "stabilization" (Donges, 1978), involves feedback control of the vehicle's motion relative to edges of the road. There are two obvious possibilities for this level of control corresponding to the model-based and ecological conceptions – either the actor is perceiving 3-D layout of the road and judging the location of the vehicle in relation to this 3-D representation or the actor is using optical invariants.

Although the two lower levels, guidance and control, were identified in connection with the control of locomotion, they obviously apply to many other forms of action (e.g. Patla, 1998). A basketball player intending to shoot a goal obviously uses feedforward control in planning how to approach the basket while using concurrent visual feedback to judge his/her movement with respect to the basket and with respect to other players. A dancer on stage

has a plan of how and where to maneuver in space but uses visual information to monitor how well the executed action compares with what is intended. Finally, a person using heavy machinery to grab a large object and place it on top of another first plans the action and then uses visual feedback to monitor its execution.

6 PERCEPTUALLY-DIRECTED ACTION

In many circumstances, perceptual information is continuously available to provide closed-loop control of action. Contrasting with this is action that is carried out after the feedback loop with the surrounding environment has been opened (e.g., by occluding vision). An example is perceiving a target's location from a fixed vantage point (visual or auditory), and then attempting to walk to the target without further perceptual information about its location during the traverse. When vision provides the initial information, this type of open-loop behavior has been referred to as visually-directed action (Foley & Held, 1972). Perceptually directed action is the general term involving any of the spatial senses (Loomis et al., 1998).

Thomson (1980, 1983) was the first to do systematic research on visually directed action. He showed that, on average, adults walk blindly with little systematic error to targets up to 20m away. Since then, a number of studies have confirmed this basic result for vision using blind walking and related tasks when visual cues to distance are plentiful (e.g. Amorim et al., 1997; Corlett, 1986; Elliot, 1987; Elliot et al., 1990; Farrell & Thomson, 1999; Fukusima et al., 1997; Loomis et al., 1992, 1998; Rieser et al., 1990; Steenhuis & Goodale, 1988). In addition, several studies have shown that when distance information is inadequate for veridical perception, actors performing the same tasks exhibit systematic error consistent with visual distance errors measured in other ways (Ooi et al., 2001; Philbeck & Loomis, 1997; Philbeck et al., 1997). Other research has demonstrated perceptual directed action with auditory targets (e.g., Ashmead et al., 1995; Loomis et al., 1998) and with haptic targets (e.g., Hollins & Kelley, 1988).

A notable feature of perceptually directed action is its flexibility. It might be thought that a person performing perceptually directed action preprograms the motor response prior to moving and, once underway, simply executes the preprogrammed response ballistically; this is what Miall et al. (1993) refer to as control by an inverse dynamic model. Thomson (1983) showed ballistic execution of a motor program could not explain some types of perceptually directed action. Actors were visually presented with a target on the ground some distance ahead. While walking without vision toward the target, actors were cued at some point to stop and then throw a beanbag the rest of the way.

Accuracy was high even though they were unaware at which point they would be cued to throw.

Fukushima et al. (1997) provided similar evidence. On each trial of the experiment, the actor viewed a target and then walked without vision along a straight path in a direction oblique to the target. On command from an experimenter, the actor then turned and walked in the direction of the target. Because the actor did not know when to turn, the accurate terminal walking direction indicates that the actor could not have been preprogramming the response. Still other evidence comes from an experiment by Philbeck et al. (1997). Here actors viewed a target, after which they walked blindly to the target along one of three paths. Because they were not cued as to which path to take until after vision was blocked, the excellent updating performance indicates that actors were not preprogramming the response (Figure 2). As Philbeck et al. (1997) noted, the mode of locomotion toward a goal is completely flexible; once the goal has been established a person can switch from walking to sidestepping to crawling on the knees. In a similar vein, the first author has informally demonstrated the ability to walk blindly to a previewed target while another person pushes against him to impede direct access to the target. These perturbations result in surprisingly little error. This is further evidence of the flexible nature of perceptually directed action.

An explanation of perceptually directed action consistent with the above results relies on several internal representations. First, visual, auditory, or haptic input specifies the perceived target location within the current perceptual model. Accompanying this percept is the spatial image, mentioned earlier. This spatial image of the target location is initially coincident with the percept. It continues to exist even after cessation of perceptual input causes the percept to disappear (as when closing the eyes). Depending on the circumstances, the percept may or may not be veridical (see the results for the Light on and Lights off conditions in Figure 3). Because the spatial image is coincident with the percept, error in the percept is carried over to the spatial image. When the actor begins moving, sensed changes in position and orientation (path integration) result in spatial updating of the spatial image. Clear evidence for continuous updating of the spatial image was provided by an experiment by Loomis et al. (1992) in which actors walked without vision and attempted to continually point toward a target that had been viewed prior to walking; pointing direction was accurate during the entire traverse. The fact that errors in perceiving target locations, as measured using verbal report and other non-motor perceptual judgments, are also reflected in perceptually directed action (vision-based reaching: Foley, 1985; vision-based locomotion: Loomis & Knapp, in press; Philbeck & Loomis, 1997; audition-based locomotion: Loomis et al., 1998) is evidence that perceptual space (or more generally, the phenomenal world) is not an epiphenomenon but indeed a cause of action. Incidentally, performance of perceptually directed action is not

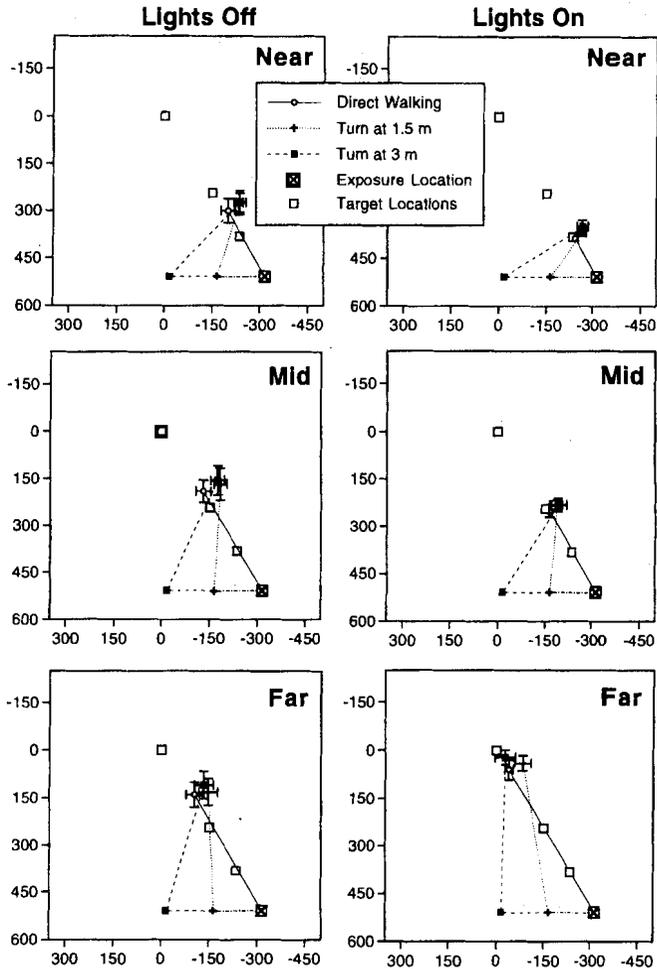


Figure 2. Results of an experiment on perceptually directed action. Glowing rectangular targets of constant angular size were viewed in darkness or with the room lights on. After the observer viewed the target on a given trial, he/she closed the eyes and was then informed whether to walk directly to the target or along one of two indirect paths. For the latter, the initial segment was aligned with a wall, which the observer could feel while walking. On each trial, the target was placed 1.5, 3.1, or 6.0 m from the origin (lower-right corner of each panel). The data are the centroids of the terminal locations of six walking trajectories of each of eight observers in response to each target and condition along each of the three paths. The convergence of the walking paths in each panel indicates that observers were directing their walking behavior to the same points in space without knowing in advance which path they would be asked to walk. Perception was more accurate with lights on, as indicated by the fact that the centroids in this condition were closer to the targets. [Fig. 5 from Philbeck, Loomis, & Beall, 1997; used by permission]

easily interpretable within the ecological conception. In vision-based tasks, for example, once perceptual input about the target's location is removed, there is no stimulus support for the ensuing action.

Because we rarely close our eyes or block our hearing while acting in space, perceptually directed action as it has been typically studied may seem largely irrelevant to the understanding of perception/action. We note, however, that perceptually directed action is more common than it first appears. When one is moving toward a visually specified goal, its temporary occlusion by intervening objects poses no problem, for the goal is spatially updated in the meantime. Similarly, team sports and dancing with a partner involve spatial updating of objects and other people (who are frequently in motion!) when they pass out of the field of view. Although certainly keeping targets of interest in sight ought to result in more accurate interaction with them than spatial updating allows, the importance of spatial updating is unmistakable in providing a robustness and flexibility to action that would otherwise not exist.

Besides these examples of perceptually directed action in which the actors move through space under their own propulsion, there have been a few studies on car steering in which the visual feedback loop has been opened, forcing the actors to continue steering without visual feedback (Hildreth et al., 2000; Godthelp, 1986; Godthelp et al., 1984; Land, 1998; Wallis et al., 2002; see also Larish & Andersen, 1995, in connection with flying). The research shows that drivers can continue steering for very brief periods following the cutoff of vision, but performance quickly degrades. In open-loop lane changing in particular, actors make the correct initial turn toward the next lane but after arriving in the lane, fail to make the opposite turn to align the vehicle with the turn (Wallis et al., 2002). This is a clear case in which perceptually directed action fails rather drastically. However, lane changing at highway speeds, which involves small control inputs and subtle inertial cues about the car's motion, may not be a particularly good candidate for open-loop control, and there may be better candidates. For example, it is possible that an expert driver who has formed a very good internal model of the dynamics of his/her car and who has strong inertial cues available, might be able to view a layout of several markings in a parking lot and, from a standing start, travel over them without vision. Performance would surely fall short of that possible with continuous vision, but if the person could do the task at all, it would support the idea of model-based control in connection with such a complex behavior.

The examples of perceptually directed action discussed above indicate that an actor, after receiving preview information about the environment, can temporarily control his/her action with respect to the environment without concurrent information specifying errors in position and orientation relative to the environment. Without this concurrent information, the actor has no feedback from the environment. To the extent that the action is still reliably

related to the environment, the actor is exercising feedforward control with respect to the environment. Is there any sense in which the actor can be said to be exercising feedback control? The answer is yes, for the person is updating an estimate of his/her position and orientation with respect to a spatial image of the environment. The person modifies commands to the musculature in order for the internally represented trajectory to match the desired trajectory through the represented environment.

Systematic error in the person's observed trajectory can come about as the result of systematic error in the percept and in the updating of the spatial image (Loomis et al., 2002). We have already seen the effects of systematic error in perception (Figure 3); actors consistently traveled to the wrong location in the Lights off condition in the study by Philbeck et al. (1997), regardless of the path traveled. One way that systematic error in spatial updating might come about is through error in updating the actor's estimate of position and orientation. In the absence of direct sensory information about position and orientation (e.g., a position fix), the actor must use path integration, based on information about velocity and/or acceleration of one's displacements and turns. Lacking allothetic information from the environment about velocity and acceleration, the actor needs to rely on idiothetic information (e.g., inertial cues, proprioceptive cues, and efference copy of the commands sent to the musculature). The accuracy with which people perform perceptually directed action using walking indicates that path integration is quite accurate with walking, at least over the short term (Loomis et al., 1999).

Would it be possible to perform perceptually directed action while driving a car without vision, as in the above thought experiment involving targets on the ground? Here, proprioceptive cues would be non-existent, leaving only inertial cues and efference copy. Taking this even further, what if the person had to perform the task while driving a car simulator without a motion base? Here, after initial preview of the scene, the actor would have only efference copy of the control inputs to the simulator as the basis for keeping track of his/her trajectory. Although control might at first seem impossible in this task, an actor who has internalized the vehicle dynamics through extensive practice might nevertheless be able to perform moderately well. In this case, we would say that the actor is using feedforward control with respect to the external world and feedback control with respect to the internal model (updated spatial image).

7 A CLOSER LOOK AT OPTIC FLOW

We now consider the case of feedback control of perception/action, with a focus on optic flow. The prototypical optic flow field of a translating observer

is that depicted in the well known image of a flying bird (Gibson, 1966). For pure translation, the optic flow diverges from the point in the optic array (focus of expansion), which corresponds to the direction of travel and converges on the opposite pole (e.g., Gibson et al., 1955; Nakayama & Loomis, 1974). This prototypical flow field, corresponding to point features, is the stimulus used in the many studies of heading perception (e.g., Crowell & Banks, 1993; Warren et al., 1988). Aside from the important issue of whether heading perception is actually involved in directional control (aiming toward a point, moving into and maintaining alignment with a straight path, and steering along a curving path), (Beall & Loomis, 1996; Rushton et al., 1998; Wann & Land, 2000; Warren et al., 2001), a more basic issue is whether the optic flow used in directional control need be at all like the prototypical flow field.

Beall and Loomis (1996) observed that in the case of driving along a straight road with only solid lane markers visible, maintaining alignment with the road in the presence of crosswinds does not require sensing of the longitudinal component of motion (parallel to the road); the splay of the lane markers is sufficient for lateral control and information about the longitudinal component (by providing ground texture) does not improve steering performance. Gordon (1966; see also Biggs, 1966; Donges, 1978; and Wann & Land, 2000) made a similar point in connection with steering a curving path of constant curvature in which solid lane markers are present (Figure 3a). Once one is moving about such a curve, it is sufficient to provide steering inputs to keep the perspective view of the road constant. If the perspective view of the road begins to move to the left and to distort in shape, one can compensate with a steering input that causes the perspective view to once again remain constant. As with the case of steering a straight path, it is quite possible that adding information about the longitudinal component of vehicle motion (e.g., by adding ground texture) would fail to improve steering performance (Figure 3b). Although these cases of steering are special cases where sensing one's longitudinal speed is irrelevant to the task and other cases of steering surely do depend upon sensing the full motion state of the vehicle from optic flow, they remind us that no single optic flow analysis, based on the prototypical flow field, is suited for understanding the control of locomotion and other forms of perception/action.

A separate issue concerning closed-loop control of action using vision is whether continuous optic flow or intermittent sampling of the optic array (the perspective view of visual space) is important. As an ideal, optic flow corresponds to continuous change in the angular positions of visible features (Gibson et al., 1955; Nakayama & Loomis, 1994). In practice, however, experimenters use a rapid succession of static images (e.g. on a CRT) that is visually indiscriminable from continuous flow or very nearly so. An important question is whether perceptually smooth optic flow is required in most tasks

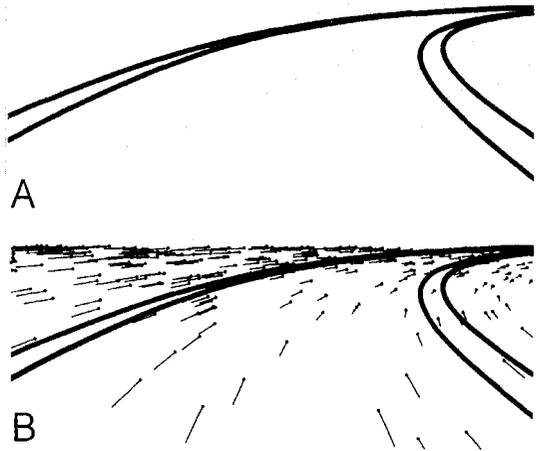


Figure 3. a) A driver's view of a curving road with constant curvature at slightly different times while the vehicle is moving laterally within the lane. The change in perspective, even without flow signaling longitudinal motion of the vehicle, is sufficient to steer. b) Adding ground texture to the information of panel (a) adds information about the longitudinal motion of the vehicle but quite possibly does not contribute to steering performance.

of visually controlled perception/action. The answer is probably yes in some cases (e.g., postural control; see Bardy et al., 1996) but decidedly no in many others. For directional control of a vehicle (aiming toward a point, maintaining alignment with a straight path, steering a curving path), it is clear that intermittent view of the environment ahead will often be as good as visually smooth flow. However, the intermittency required for accurate control of steering is surely related to the variation in road curvature per unit time, the vehicle dynamics, and other factors (Godthelp et al., 1984). Thus, we conclude that "optic change" in the sampled images of the optic array, rather than continuous optic flow, is often sufficient for precise visual control of action. To keep the terminology simple, however, we will generally use optic flow to refer to both continuous flow and optic change.

8 OPTIC FLOW AND PERCEIVED FLOW

In many situations of visually controlled action, the optic flow produced by locomotion results in a visual streaming or perceived flow of the environment. Indeed, this perceived flow is surely considered by some researchers to be identical with the optic flow that causes it. Yet, it is important to keep the two separate, for one is the stimulus and the other is perceptual. Indeed, Gogel (1990) has provided abundant evidence that

perceived depth is a fundamentally important determinant of how optic flow is perceived.

Because optic flow and its discrete version, optic change, often have perceptual concomitants, the question arises as to whether perceived flow and changes in the perceived visual scene are instead the actual basis for visually controlled perception/action. This is a broad question about which little is known. However, one study addressing this issue (Beusmans, 1998) showed that changes in how a visual scene was perceived in terms of depth did influence the perception of heading. This is a hint that optic flow, when it is the basis for heading perception, might act through perceived flow.

We have recently discovered that a variety of active control tasks of various spatial behaviors can be effectively accomplished without any optic flow whatsoever. For this research, we have exploited the random dot cinematogram (RDCs) with single-frame lifetimes (Julesz, 1971); this is a sequence of random-dot stereograms portraying a changing 3-D stimulus in which the base pair of random dots changes every frame. Julesz (1971) first demonstrated that observers are readily able to perceive a stereo-defined square moving laterally across a visual display despite the very strong perception of scintillation. More recently, Gray and Regan (1996, 1998) have shown that observers are able to judge time to contact and the direction of motion of single moving targets presented in these types of displays.

In our work, the second author has developed a technique for combining RDCs with single-frame lifetimes with our immersive virtual reality system that we have been using to study complex behavior (Loomis et al., 1999). With this technique, the second author can switch from normal binocular display of a richly textured virtual environment to a scintillating RDC display of the same environment (without textures). At a 60 Hz graphics update rate in each eye, he can cast up to 1000 random dots onto the surfaces of the virtual environment within the observer's field of view. These dots are spread over the field of view with constant angular density. With the single-frame lifetimes of each video frame, each eye simply is exposed to a scintillating field of uniform dot density. Thus, there is no optic flow signal correlated with the spatial structure of the virtual environment. Amazingly, when viewing with both eyes, observers are able to perceive complex environments almost immediately. Even more amazing is the ease with which they can perform complex spatial tasks, such as walking through a cluttered field of obstacles, walking to intercept a rapidly moving object, and steering a simulated vehicle along a curving road. Figure 4a is a stereogram (for cross-fusion) depicting a path represented by a raised centerline. Figure 4b is the same path rendered as a random-dot stereogram.

Research underway is exploring the extent to which performance of visually controlled actions carried out with the RDC approaches that of the same actions using the traditional optic flow stimulus. Because visual

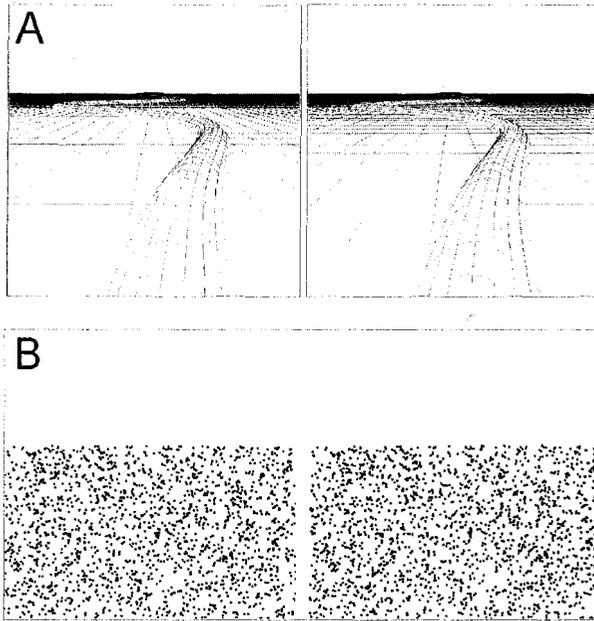


Figure 4. a) Cross-fusion stereogram of a path represented by a raised centerline. b) Random-dot stereogram of the same view. A scintillating RDC consists of many frames like this, each with a different base pair of random dots (Julesz, 1971). Surprisingly, people can perform a wide variety of visually controlled actions with scintillating RDCs even though there is no optic flow relating to the task.

perception of RDC stimuli is impaired relative to vision of normal optic flow stimuli (Gray & Regan, 1998), performance of complex actions is somewhat poorer than it is with normal stimulation. For example, in a steering task using the stimuli of Figure 4, the root-mean-squared error showing how closely the simulated vehicle tracks the path is approximately 75% greater with the scintillating RDC (Loomis & Beall, 2003). The important question is whether, when monocular and stereoscopic vision are matched for these lower level processing differences, performance of these complex actions is also matched (see Gray & Regan, 1998 for a treatment of the same issue in connection with the processing of single moving targets).

Our work so far indicates that when adjusting for these processing differences, performance of complex action is even more closely matched (Loomis & Beall, 2003). If further research shows a near functional equivalence between optic flow and this binocularly-mediated "flow", for example, with respect to eye fixations and pursuit eye movements during the performance of complex action, this would support the hypothesis of a final common pathway of perceived flow through which both forms of flow exert

their influence in controlling action. This would implicate internal representation in the control of perception/action, even in this case of feedback control involving optic flow. However, we note that conventional optic flow, even if acting through perceived flow, surely will allow for better performance of certain complex actions (e.g., catching a baseball) than binocularly-mediated flow when those actions require higher spatial and temporal bandwidths of visual processing than is afforded by pure binocular vision.

ACKNOWLEDGEMENTS

The authors' ideas and empirical results presented in this chapter are based on research supported by Office of Naval Research grant N00014-95-1-0573 and Air Force Office of Scientific Research grant F49620-02-1-0145. The authors thank Willard Larkin for comments on an earlier version of the manuscript.

REFERENCES

- Amorim, M. A., Glasauer, S., Corpinot, K., & Berthoz, A. (1997). Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes, *Percept. Psychophys.*, *59*, 404-418.
- Ashmead, D. H., DeFord, L. D., & Northington, A. (1995). Contribution of listeners' approaching motion to auditory distance perception, *J. Exp. Psychol. Hum. Percept. Perform.*, *21*, 239-256.
- Bardy, B. G., Warren, W. H., & Kay, B. A. (1996). Motion parallax is used to control postural sway during walking, *Exp. Brain Res.*, *111*, 271-282.
- Beall, A. C. & Loomis, J. M. (1996). Visual control of steering without course information, *Perception*, *25*, 481-494.
- Beall, A. C. & Loomis, J. M. (1997). Optic flow and visual analysis of the base-to-final turn, *Int. J. Aviat. Psychol.*, *7*, 201-223.
- Beusmans, J. M. H. (1998). Perceived object shape affects the perceived direction of self-movement, *Perception*, *27*, 1079-1085.
- Bhalla, M. & Proffitt, D. (1999). Visual-motor recalibration in geographical slant perception, *J. Exp. Psychol. Hum. Percept. Perform.*, *25*, 1076-1096.
- Biggs, N. L. (1966). Directional guidance of motor vehicles: A preliminary survey and analysis, *Ergonomics*, *9*, 193-202.
- Brain, W. R. (1951). *Mind, Perception, and Science*. London: Blackwell.
- Calvert, E. S. (1954). Visual judgments in motion, *J. Inst. Nav.*, *7*, 233-251.

- Corlett, J. T. (1986). The effect of environmental cues on locomotor distance estimation by children and adults, *Hum. Movement Sci.*, 5, 235-248.
- Crowell, J. A. & Banks, M. S. (1993). Perceiving heading with different retinal regions and types of optic flow, *Percept. Psychophys.*, 53, 325-337.
- Dickmanns, E. D. (1992). A general dynamic vision architecture for UGV and UAV. *J. App. Intell.*, 2, 251-270.
- Donges, E. (1978). A two-level model of driver steering behavior, *Hum. Factors*, 20, 691-707.
- Elliott, D. (1987). The influence of walking speed and prior practice on locomotor distance estimation, *J. Mot. Behav.*, 19, 476-485.
- Elliott, D., Jones, R., & Gray, S. (1990). Short-term memory for spatial location in goal-directed locomotion, *Bull. Psychon. Soc.*, 8, 158-160.
- Farrell, M. J. & Thomson, J. A. (1999). On-line updating of spatial information during locomotion without vision, *J. Mot. Behav.*, 3, 39-53.
- Flach, J. M., & Warren, R. (1995). Active Psychophysics: The Relation Between Mind and What Matters. In: J. M. Flach, P. A. Hancock, J. Caird, & K. J. Vicente (Eds.) *Global Perspectives on the Ecology of Human-Machine Systems, Vol. 1. Resources for Ecological Psychology* (pp. 189-209). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Fodor, J. A. (1983). *The Modularity of Mind: An Essay on Faculty Psychology*. Cambridge, MA: MIT Press.
- Foley, J. M. (1985). Binocular distance perception: egocentric distance tasks, *J. Exp. Psychol. Hum. Percept. Perform.*, 11, 133-149.
- Foley, J. M. & Held, R. (1972). Visually directed pointing as a function of target distance, direction, and available cues, *Percept. Psychophys.*, 12, 263-268.
- Fukushima, S. S., Loomis, J. M., & Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation, *J. Exp. Psychol. Hum. Percept. Perform.*, 23, 86-100.
- Gibson, J. J. (1950). *The Perception of the Visual World*. Boston: Houghton-Mifflin.
- Gibson, J. J. (1958). Visually controlled locomotion and visual orientation in animals, *Br. J. Psychol.*, 49, 182-194.
- Gibson, J. J. (1966). *The Senses Considered as Perceptual Systems*. Boston: Houghton-Mifflin.
- Gibson, J. J. (1979). *The Ecological Approach to Visual Perception*. Boston: Houghton-Mifflin.
- Gibson, J. J. & Crooks, L. E. (1938). A theoretical field-analysis of automobile driving, *Am. J. Psychol.*, 51, 453-471.
- Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings, *Am. J. Psychol.*, 68, 372-385.
- Godthelp, H. (1986). Vehicle control during curve driving, *Hum. Factors*, 28, 211-221.
- Godthelp, H. & K appler, W. D. (1988). Effects of vehicle handling characteristics on driving strategy, *Hum. Factors*, 30, 219-229.

- Godthelp, H., Milgram, P., & Blaauw, G. J. (1984). The development of a time-based measure to describe driving strategy, *Hum. Factors*, *26*, 257-268.
- Gogel, W. C. (1990). A theory of phenomenal geometry and its applications, *Percept. Psychophys.*, *48*, 105-123.
- Gordon, D. A. (1966). Perceptual basis of vehicular guidance, *Public Roads*, *34*, 53-68.
- Gray, R. & Regan, D. (1996). Cyclopean motion perception produced by oscillations of size, disparity and location, *Vision Res.*, *36*, 655-665.
- Gray, R. & Regan, D. (1998). Accuracy of estimating time to collision using binocular and monocular information, *Vision Res.*, *38*, 499-512.
- Hildreth, E. C., Beusmans, J. M. H., Boer, E. R., & Royden, C. S. (2000). From vision to action: Experiments and models of steering control during driving, *J. Exp. Psychol. Hum. Percept. Perform.*, *26*, 1106-1132.
- Hollins, M. & Kelley, E. K. (1988). Spatial updating in blind and sighted people, *Percept. Psychophys.*, *43*, 380-388.
- Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago: University of Chicago Press.
- Land, M. F. (1998). The Visual Control of Steering. In: L. R. Harris & M. Jenkin (Eds.), *Vision and Action* (pp. 163-180). Cambridge: Cambridge University Press.
- Land, M. F., & Hayhoe, M. (2001). In what ways do eye movements contribute to everyday activities? *Vision Res.*, *41*, 3559-3565.
- Larish, J. F. & Andersen, G. J. (1995). Active control in interrupted dynamic spatial orientation: the detection of orientation change. *Percept. Psychophys.*, *57*, 533-545.
- 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. (1993). Body-Environment Coupling. In: U. Neisser (Ed.), *The Perceived Self: Ecological and Interpersonal Sources of Knowledge* (pp. 43-67). New York: Cambridge University Press.
- Lee, D. N. & Lishman, R. (1977). Visual control of locomotion, *Scand. J. Psychol.*, *18*, 224-230.
- Lehar, S. (2002). *The World in Your Head: A Gestalt View of the Mechanism of Conscious Experience*. Mahwah, N.J.: Lawrence Erlbaum Associates.
- Loomis, J. M. (1992). Distal attribution and presence. *Presence*, *1*, 113-119.
- Loomis, J. M. & Beall, A. C. (1998). Visually-controlled locomotion: Its dependence on optic flow, 3-D space perception, and cognition, *Ecol. Psychol.*, *10*, 271-285.
- Loomis, J. M. & Beall, A. C. (2003). Visual control of locomotion without optic flow. *Third annual meeting of Vision Sciences Society, Sarasota, FL, May 9-14, 2003*.
- Loomis, J. M., Blascovich, J.J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology, *Behav. Res. Meth. Instrum. Comput.*, *31*, 557-564.
- Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukusima, S. S. (1992). Visual space perception and visually directed action, *J. Exp. Psychol. Hum. Percept. Perform.*, *18*, 906-921.

- Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human Navigation by Path Integration. In: R. G. Golledge (Ed.), *Wayfinding: Cognitive Mapping and Other Spatial Processes* (pp. 125-151). Baltimore: Johns Hopkins.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action, *Percept. Psychophys.*, *60*, 966-980.
- Loomis, J. M. & Knapp, J. M. (in press). Visual Perception of Egocentric Distance in Real and Virtual Environments. In: L. J. Hettinger and M. W. Haas (Eds.), *Virtual and Adaptive Environments*. Mahwah N.J.: Erlbaum.
- Loomis, J. M., Lippa, Y., Klatzky, R. L., & Golledge, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language, *J. Exp. Psychol. Learn. Mem. Cogn.*, *28*, 335-345.
- McRuer, D. T., Allan, R. W., Weir, D. H., & Klein, R. H. (1977). New results in driver steering control models, *Hum. Factors*, *19*, 381-397.
- Miall, R. C., Weir, D. J., Wolpert, D. M., & Stein, J. F. (1993). Is the cerebellum a Smith Predictor? *J. Mot. Behav.*, *25*, 203-216.
- Milner, A. D. & Goodale, M. A. (1995). *The Visual Brain in Action*. New York: Oxford University Press.
- Nakayama, K. & Loomis, J. M. (1974). Optical velocity patterns, velocity sensitive neurons, and space perception: a hypothesis, *Percept.*, *3*, 63-80.
- Ooi, T.L., Wu, B., & He, Z. J. (2001). Distance determined by the angular declination below the horizon, *Nature*, *414*, 197-200.
- Patla, A. E. (1998). How is human gait controlled by vision? *Ecol. Psychol.*, *10*, 287-302.
- Philbeck, J. W. & Loomis, J. M. (1997). Comparison of two indicators of visually perceived egocentric distance under full-cue and reduced-cue conditions, *J. Exp. Psychol. Hum. Percept. Perform.*, *23*, 72-85.
- Philbeck, J. W., Loomis, J. M., & Beall, A. C. (1997). Visually perceived location is an invariant in the control of action, *Percept. Psychophys.*, *59*, 601-612.
- Ramachandran, V. S. & Blakeslee, S. (1998). *Phantoms in the Brain: Probing the Mysteries of the Human Mind*. New York: William Morrow.
- Rensink, R. A. (2002). Change detection, *Ann. Rev. Psychol.*, *53*, 245-277.
- Riemersma, J. B. (1981). Visual control during straight road driving, *Acta Psychologica*, *48*, 215-225.
- Rieser, J. J., Ashmead, D. H., Talor, C. R., & Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets, *Perception*, *19*, 675-689.
- Rushton, S. K., Harris, J. M., Lloyd, M. R. & Wann, J. P. (1998). Guidance locomotion on foot uses perceived target location rather than optic flow, *Curr. Biol.*, *21*, 1191-1194.
- Russell, B. (1948). *Human Knowledge: Its Scope and Limits*. New York: Simon and Schuster.
- Simmel, M. L. (1966). Developmental aspects of the body scheme, *Child Develop.*, *37*, 83-96.
- Smythies, J. R. (1994). *The Walls of Plato's Cave*. Aldershot, England: Avebury.

- Steenhuis, R. E. & Goodale, M. A. (1988). The effects of time and distance on accuracy of target-directed locomotion: does an accurate short-term memory for spatial location exist? *J. Mot. Behav.*, *20*, 399-415.
- Thomson, J. A. (1980). How do we use visual information to control locomotion? *Trends Neurosci.*, *3*, 247-250.
- Thomson, J. A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *J. Exp. Psychol. Hum. Percept. Perform.*, *9*, 427-443.
- Turvey, M. T. (1977). Preliminaries to a Theory of Action with Reference to Vision. In: R. J. Shaw & J. Bransford (Eds.), *Perceiving, Acting, and Knowing: Toward an Ecological Psychology*. Hillsdale, N.J.: Erlbaum.
- Turvey and Carello, C. (1986). The ecological approach to perceiving-acting: a pictorial essay, *Acta Psychologica*, *63*, 133-155.
- Wallis, G.M., Chatziastros, A., & Bühlhoff, H. H. (2002). An unexpected role for visual feedback in vehicle steering control, *Curr. Biol.*, *12*, 295-299.
- Wann, J. & Land, M. (2000). Steering with or without the flow: is the retrieval of heading necessary? *Trends Cogn. Sci.*, *4*, 319-324.
- Warren, R. & Wertheim, A. H. (1990). *Perception & Control of Self-Motion*. Hillsdale, NJ: Lawrence Erlbaum.
- Warren, W. H. (1998). Visually controlled locomotion: 40 years later, *Ecol. Psychol.*, *10*, 177-219.
- Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking, *Nat. Neurosci.*, *4*, 213-216.
- Warren, W. H., Morris, M. W., & Kalish, M. (1988). Perception of translational heading from optical flow, *J. Exp. Psychol. Hum. Percept. Perform.*, *17*, 28-43.
- Weiskrantz, L. (1986). *Blindsight: A Case Study and Implications*. Oxford: Oxford University Press.
- Wolpert, D. M. & Ghahramani, Z. (2000). Computational principles of movement neuroscience, *Nat. Neurosci. Supp.*, *3*, 1212-1217.
- Wolpert, D. M., Ghahramani, Z., & Jordan, M. I. (1995). An internal model for sensorimotor integration, *Science*, *269*, 1880-1882.
- Yilmaz, E. H. & Warren, W. H. (1995). Visual control of braking: a test of the tau-dot hypothesis, *J. Exp. Psychol. Hum. Percept. Perform.*, *21*, 996-1014.