From Embodiment, Ego-space, and Action
edited by R.L. Klatzky, B. Mac Whinney,
& M. Behrmann
New York: Taylor & Francis, 2008

1

Measuring Spatial Perception with Spatial Updating and Action

Jack M. Loomis and John W. Philbeck

Measurement of perceived egocentric distance, whether of visual or auditory targets, is a topic of fundamental importance that is still being actively pursued and debated. Beyond its intrinsic interest to psychologists and philosophers alike, it is important to the understanding of many other topics which involve distance perception. For example, many complex behaviors like driving, piloting of aircraft, sport activities, and dance often involve distance perception. Consequently, understanding when and why errors in distance perception occur will illuminate the reasons for error and disfluency in these behaviors. Also, the understanding of distance perception is important in the current debate about the "two visual systems," one ostensibly concerned with the conscious perception of 3-D space and the other with on-line control of action. Similarly, determining whether nonsensory factors, such as intention to act and energetic state of the observer, influence perceived distance, as has been claimed (e.g. Proffitt, Stefanucci, Banton, & Epstein, 2003; Witt, Proffitt, & Epstein, 2004, 2005) depends critically on the meaning of distance perception and how it is to be measured. Still another topic where measurement of distance perception is critical is spatial updating (the imaginal updating of a target perceived only prior to observer movement) involving observer translation. Being able to measure the accuracy of spatial updating depends upon being able to partial out errors due to misperception of the initial target distance (Böök & Gärling, 1981; Loomis, Klatzky, Philbeck, & Golledge, 1998; Loomis, Lippa, Klatzky, & Golledge, 2002; Philbeck, Loomis, & Beall, 1997). Finally, measurement of distance perception is important for the development of effective visual and auditory displays of 3-D space. Indeed, developing virtual reality systems that exhibit naturally appearing scale has proven an enormous challenge, both for visual virtual reality (Loomis & Knapp, 2003) and for auditory virtual reality (Loomis, Klatzky, & Golledge, 1999), and there has been a spate of recent research articles concerned with understanding the causes for uniform scale compression in many visual virtual environments (e.g., Creem-Regehr, Willemsen, Gooch, & Thompson, 2005; Knapp, 1999; Knapp & Loomis, 2004; Sahm, Creem-Regehr, Thompson, & Willemsen, 2005; Thompson, Willemsen, Gooch, Creem-Regehr, Loomis et al., 2004). Virtual reality systems that successfully create a realistic sense of scale will enjoy even greater aesthetic impact and user acceptance and will prove even more useful in the training of skills, such as safe road crossing behavior by blind and sighted children.

Indirectness of Perception

Naïve realism is the commonsense view that the world we encounter in everyday life is identical with the physical world that we come to know about through our schooling. Following decades of intellectual inquiry, philosophers of mind and scientists have come to an alternate view referred to as "representative realism"—that contact with the physical world is indirect and that what we experience in everyday life is a representation created by our senses and central nervous system (e.g., Brain, 1951; Koch, 2003; Lehar, 2003; Loomis, 1992; Russell, 1948; Smythies, 1994). Indeed, this representation, generally referred to as the phenomenal world, is so highly consistent and veridical that we routinely make life-depending decisions without ever suspecting that the perceptual information upon which we are relying is once removed from the physical world. The high degree of functionality of the perceptual process accounts for its being self-concealing and for the reason that most laypeople and indeed many scientists think of perception as little more than attention to aspects of the environment.

The representational nature of perceptual experience is easy to appreciate with color vision because the mapping from physical stimulation to perceptual space entails a huge loss of information, from the many dimensions of spectral lights to the three perceptual dimensions of photopic color vision. In order to appreciate the representational nature of perception more generally it is helpful to keep in mind such perceptual phenomena as diplopia, binocular stereopsis elicited by stereograms, geometric visual illusions, and motion illusions; such phenomena point to a physical world beyond the world of appearance. Although experiencing such phenomena momentarily reminds us of the representational nature of perception, we too easily lapse back into naïve realism when driving our cars, engaging in sports activity, and interacting with other people. It is quite an intellectual challenge to appreciate that the very three-dimensional world we experience in day-to-day life is an elaborate perceptual representation. Indeed, many people seem to be naïve realists when it comes to visual space perception, for they think of visual space perception largely as one of judging distance. But, visual space perception is so much more than this—it gives rise to our experience of the surrounding visual world, consisting of surfaces and objects lying in depth (e.g., Gogel, 1990; Howard & Rogers, 2002; Loomis, Da Silva, Fujita, & Fukusima, 1992; Marr, 1982; Ooi, Wu, & He, 2006; Wu, Ooi, & He, 2004). Virtual reality makes the representational nature of visual space perception obvious (Loomis, 1992), for the user experiences being immersed within environments which have no physical existence (other than being bits in computer memory). Teleoperator systems are useful for drawing the same conclusion. Consider a visual teleoperator system consisting of a head-mounted binocular display and externally mounted video cameras for driving the display. The user of such a teleoperator system experiences full presence in the physical environment while being intellectually aware that the visual stimulation comes only indirectly by way of the display. Because the added degree of mediation associated with the display pales in comparison with the degree of mediation associated with visual processing, the representational nature of perception when using a teleoperator points to the representational nature of ordinary perception.

How one conceives of perception determines how one goes about measuring perceived distance. For the researcher who accepts naïve realism, perceiving distance is simply a matter of judging distance in "physical space." Under this conception, one can simply ask the observer how far away objects are and then correct for any judgmental biases, such as reporting 1 m as 2 m. In contrast, for researchers who adhere to the representational conception, the measurement of distance perception is a major challenge, inasmuch as one is attempting to measure aspects of an internal representation. Because one starts with behavior of some kind (e.g., verbal report, action) and because there can be distortions associated with the readout from internal representation to behavior, measurement of perception depends on a theory connecting internal representation to behavior, a theory that is best developed using multiple response measures (e.g., Foley, 1977; Philbeck & Loomis, 1997).

Some Methods for Measuring Perceived Distance

Verbal report and magnitude estimation are two traditional methods for measuring perceived distance (Da Silva, 1985). Figure 1.1 gives the results of a number of studies using verbal report for target distances out to 28 m (Andre & Rogers, 2006; Foley, Ribeiro, & Da Silva, 2004; Kelly, Loomis, & Beall, 2004; Knapp & Loomis, 2004; Loomis et al., 1998; Philbeck & Loomis, 1997). The data sets are generally well fit by linear functions with 0 intercepts, but the slopes are generally less than 1.0. The mean slope is 0.80.

Concerns about the possible intrusion of knowledge and belief into such judgments (Carlson, 1977; Gogel, 1974) have prompted the search for alternative methods. So-called indirect methods make use of other perceptual judgments thought to be less subject to intrusion by cognitive factors and then derive estimates of perceived distance by way of theory. Several of these methods rely on so-called percept-percept couplings. Space perception researchers have long known that perceptual variables often covary with one another (Epstein, 1982; Gogel, 1984; Sedgwick, 1986). In some cases these covariations may be the result of joint determination by common stimulus variables, but in other cases variation in one perceptual variable causes variation in another (Epstein, 1982; Gogel, 1990; Oyama, 1977); such causal covariations are referred to as percept-percept couplings. The best known coupling is that between perceived size and perceived egocentric distance and is referred to as size-distance invariance (Gilinsky, 1951; McCready, 1985; Sedgwick, 1986). Size-distance invariance is the relationship between perceived size (S') and perceived egocentric distance (D') for a visual stimulus of angular size α: $S' = 2D' \tan(\alpha/2)$. A special case of size-distance invariance is Emmert's Law—varying the perceived distance of a stimulus causes perceived size to vary proportionally, with angular size held constant. Another coupling of perceptual variables is that between the perceived distance of a target and its perceived motion (Gogel, 1982, 1993). Gogel demonstrated that the perceived motion of an object can be altered by mere changes in its perceived distance while keeping all other variables constant. He developed a quantitative theory for this coupling between perceived distance and perceived motion and applied it in explaining the apparent motion of a variety of stationary objects, such as depth-reversing figures and the inverted facial mask (Gogel, 1990). The existence of percept-percept couplings is methodologically important, for these couplings can be used to measure perceived distance in situations where the researcher wishes observers not to be aware that perceived distance is being measured. Judgments of perceived size and perceived motion have been used to measure perceived distance (e.g., Gogel, Loomis, Newman, & Sharkey, 1985; Loomis & Knapp, 2003) and to demonstrate the effect of an experimental manipulation on perceived distance (e.g., Hutchison & Loomis, 2006a).

Another indirect method of measuring perceived distance involves judgments of collinearity and relies on the perception of exocentric direction. A visible pointer is adjusted by the observer to be aligned with the target stimulus (Wu, Klatzky, Shelton, & Stetten, 2005); these authors used the method to measure the perceived distance of targets within arm's reach under the assumption that the pointer is perceived correctly. Application of the method to the measurement of large perceived distances seems promising, but the method will have to compensate for systematic biases in exocentric direction perception (Cuijpers, Kappers, & Koenderink, 2000; Kelly et al., 2004).

Still other indirect methods rely on judgments of perceived exocentric extent and attempt, by way of theory, to construct scales of perceived distance. The best known example is the work by Gilinsky (1951) and, more recently, Ooi and He (2007). In their experiments, observers constructed a set of equal-appearing intervals on the ground extending directly away from the observer. The more distant intervals had to be made progressively larger in order to appear of constant size. Assuming that perceived egocentric distance over the ground plane to a given point is the concatenation of the equal appearing intervals up to that point, the derived perceived distance can be associated with the corresponding cumulative physical

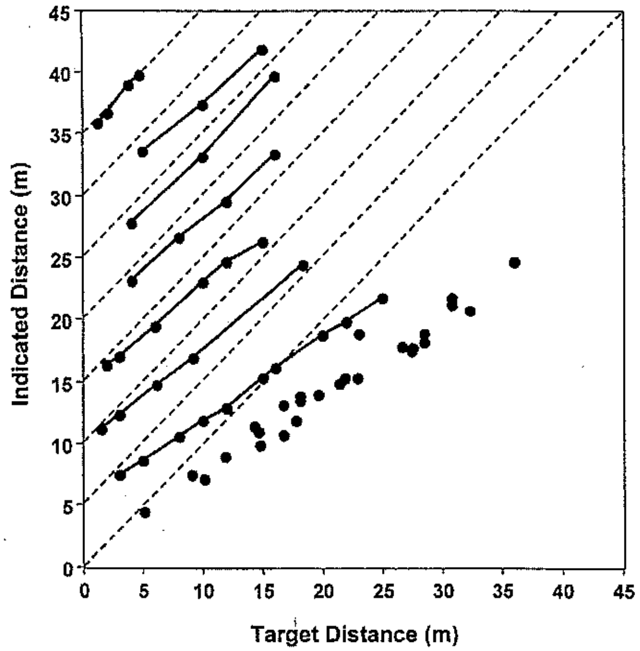

Figure 1.1 Summary of verbal reports of distance for visual targets. The data from the different studies have been displaced vertically for purposes of clarity. The dashed line in each case represents correct responding. Sources from top to bottom: Experiment 3 of Philbeck and Loomis (1997), experimental condition of Experiment 2A of Loomis et al. (1998), calibration condition of Experiment 2A of Loomis et al. (1998), results for gymnasts in Experiment 1 of Loomis et al. (1998), results of full field of view condition of Knapp & Loomis (2004), mean data from the control conditions in the 3 experiments of Andre and Rogers (2006), results of Kelly, Loomis, and Beall (2004), and egocentric distance judgments of Foley et al. (2004).

distance. The derived function of perceived egocentric distance is compressively nonlinear even within 10 m and under full-cue conditions. Because the derived function is noticeably discrepant with other functions to be discussed here and because there are process interpretations for doubting that the derived function is indeed a measure of perceived distance, we do not discuss it further.

Methods Based on Action and Spatial Updating

Given the importance of distance perception and the lack of consensus about how to measure it, researchers have occasionally proposed new measurement procedures. Here, we focus on relatively new methods for measuring perceived distance that rely on action, sometimes with the involvement of spatial updating. The typical procedure begins with the stationary observer viewing or listening to a target stimulus. After this period of "preview," further perceptual information about the target is removed by occluding vision and hearing, and the observer attempts to demonstrate knowledge of the target's location by some form of action (e.g. pointing, walking, or throwing a ball). *Visually directed pointing* was a term coined by Foley and Held (1972) to refer to blind pointing with the finger to the 3-D location of a visual target that had been previously viewed. This type of response has been used in other studies to measure the perceived locations of visual targets within arm's reach (e.g., Bingham, Bradley, Bailey, & Vinner, 2001; Foley, 1977; Loomis, Philbeck, & Zahorik, 2002). For more distant targets, ball or bean bag throwing has been used (Eby & Loomis, 1987; He, Wu, Ooi, Yarbrough, & Wu, 2004; Sahm et al., 2005; Smith & Smith, 1961). Another form of visually directed action, blind walking (sometimes called "open loop walking"), has been used to study the perception of distances of "action space" (distances beyond reaching but within the range of most action planning; Cutting & Vishton, 1995); here the observer typically views a target on the ground and attempts to walk to its location without vision. These various forms of open-loop behavior, along with others to be discussed, are referred to collectively as "perceptually directed action."

Many studies have used blind walking to assess the accuracy of perceiving the distances of targets viewed on the ground under full-cue conditions, for distances up to 28 m (Andre & Rogers, 2006; Corlett, Byblow, & Taylor, 1990; Corlett & Patla, 1987; Creem-Regehr et al., 2005; Elliott, 1987; Elliott, Jones, & Gray, 1990; Knapp & Loomis, 2004; Loomis et al., 1992; Loomis et al., 1998; Messing & Durgin, 2005; Rieser, Ashmead, Talor, & Youngquist, 1990; Steenhuis & Goodale, 1988; Thomson, 1983; Wu et al., 2004). Figure 1.2 shows many of the results, with the data sets shifted vertically for purposes of clarity. Except for two data sets, perceived distance is proportional to physical distance with no evidence of systematic error (slopes of the best fitting linear functions are generally close to 1, and intercepts are near zero). In contrast, when the same task, modified for audition, is used to study distance perception of sound-emitting sources heard out-of-doors, systematic errors are observed over the

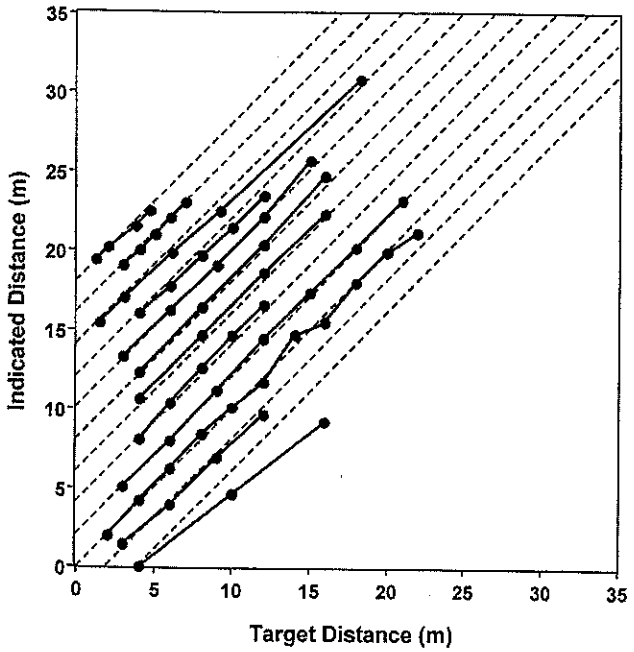

Figure 1.2 Summary of blind walking results for vision. The data from the different studies have been displaced vertically for purposes of clarity. The dashed line in each case represents correct responding. Sources from top to bottom: Experiment 3 of Philbeck and Loomis (1997), Experiment 1 of Wu, Ooi, and He (2004), mean data from the control conditions in the 3 experiments of Andre and Rogers (2006), results of Elliott (1987), average of two groups of observers from Experiment 1 of Loomis et al. (1998), Experiment 2b of Loomis et al. (1998), Experiment 1 of Loomis et al. (1992), Thomson (1983), Rieser et al. (1990), Steenhuis and Goodale (1988), and Experiment 2a of Loomis et al. (1998).

same range of distances (Ashmead, DeFord, & Northington, 1995; Loomis et al., 1998; Speigle & Loomis, 1993). Figure 1.3 shows representative results; this time, the data sets have not been shifted vertically. The best linear functions have slopes close to 0.5, indicating response compression relative to the stimulus range, and there is considerable variability in the intercepts. We should mention, however, that in a recent review of these and other results obtained using other response measures including verbal report, Zahorik, Brungart, and Bronkhorst (2005) fit power functions to the data and generally found exponents less than 1.0, the interpretation being that perceived auditory distance is a compressively nonlinear function of source distance. Still, for the range of distances in Figure 1.3, the conclusion that they are linear functions with roughly constant slope but varying intercept is justified. The source of the variation in intercept is a mystery.

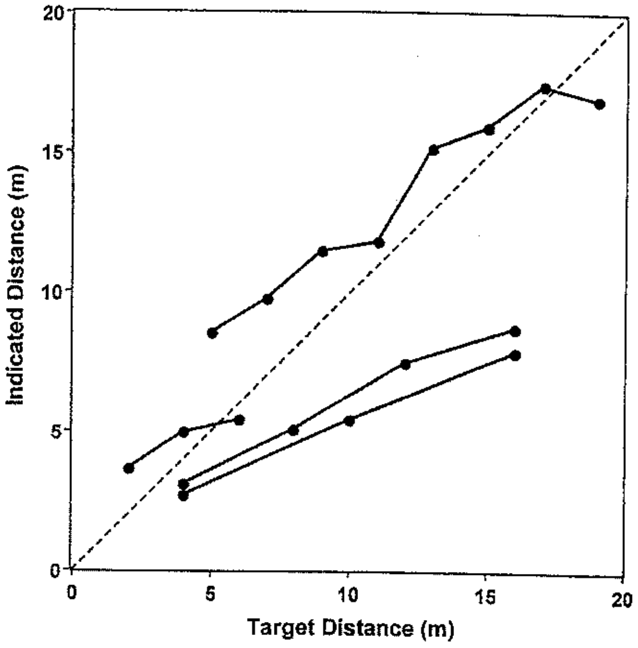

Figure 1.3 Summary of blind walking results for audition. These are the actual data and have not been displaced vertically for purposes of clarity. The dashed line represents correct responding. Sources from top to bottom: Ashmead et al. (1995), Speigle and Loomis (1993), Experiment 1 of Loomis et al. (1998), and Experiment 2a of Loomis et al. (1998)

With vision, systematic errors do arise. Sinai, Ooi, and He (1998) and He et al. (2004) have found that when the ground surface is interrupted by a gap, visual targets resting on the ground are mislocalized even with full cue viewing. Larger systematic errors occur when visual cues to distance are minimal. Figure 1.4 gives the results of a study by Philbeck and Loomis (1997) in which blind walking responses and verbal report were obtained under two conditions: reduced cues (luminous targets of constant angular size at eye level in the dark) and full-cues (the same targets placed on the floor with normal room lighting). When cues were minimal, both types of judgment showed large systematic errors, and when cues were abundant, both types of judgment showed small systematic errors. This study also showed that when the verbal responses were plotted against the walking responses in these and two other conditions, the data were

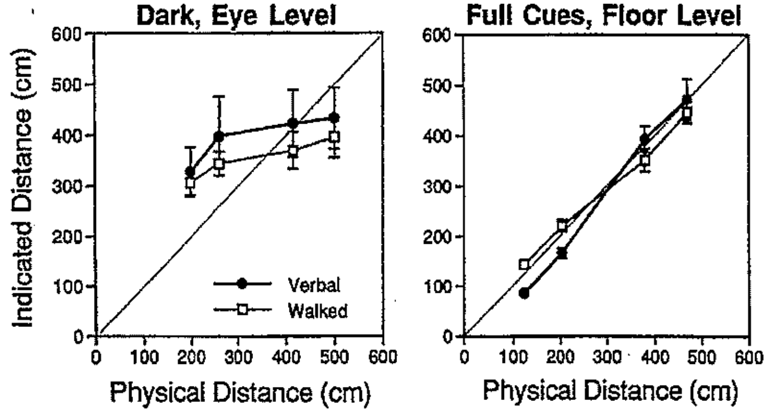

Figure 1.4 Results of an experiment using both verbal report and visually-directed blind walking in reduced-cue and full-cue conditions. Adaptation of Figure 5 from Philbeck, J. W. & Loomis, J. M. (1997). Comparison of two indicators of visually perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 72–85.

well fit by a single linear function, suggesting that variations in the two response measures are controlled by the same internal variable, visually perceived distance.

A related experiment was concerned with measuring the perceptual errors in visual virtual reality (Sahm et al., 2005). Observers performed blind walking and bean bag throwing to targets in both a real environment and a virtual environment modeled on the real environment. Prior to testing, observers were given feedback only about their throwing performance in the real environment. The results are given in Figure 1.5. The fact that the transition from the real to virtual environment produces the same errors in walking and throwing supports the claim that the two actions, one that involves locomotion and the other that does not, are controlled by the same internal variable, visually perceived distance. In addition, the results provide further support for the growing consensus that current virtual reality systems produce under perception of distance (Knapp, 1999; Thompson et al., 2004).

Triangulation Methods

The similarity of the walking measures and verbal reports above might be taken as evidence of a simple strategy for performing blind

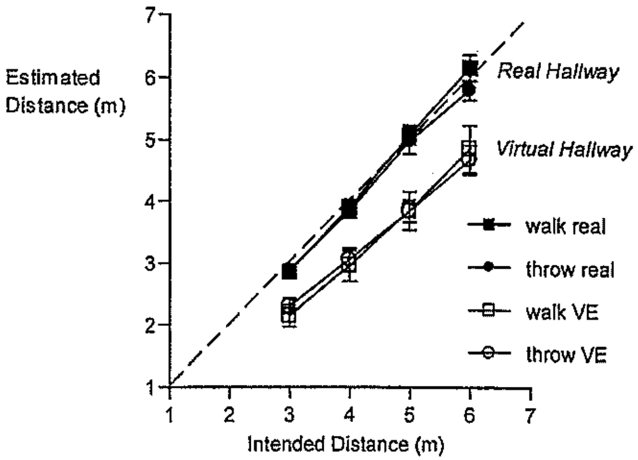

Figure 1.5 Results of an experiment using both visually-directed blind walking and visually directed throwing in real and virtual environments. Reprinting of Figure 3 of Sahm, C. S., Creem-Regehr, S. H., Thompson, W. B., & Willemsen, P. (2005). Throwing versus walking as indicators of distance perception in similar real and virtual environments. *ACM Transactions on Applied Perception, 2*, 35–45.

walking—while perceiving the target, estimate its distance in feet or meters, and then, with vision and hearing occluded, walk a distance equal to the estimate.

Whereas the blind walking task might well be performed using this simple strategy, there are other closely related tasks that cannot. Foremost are triangulation tasks that require the observer to constantly update the estimated location of the target while moving about in the absence of further perceptual input specifying its location. Figure 1.6 depicts three triangulation tasks that have been used. In "triangulation by pointing," the observer views (or listens

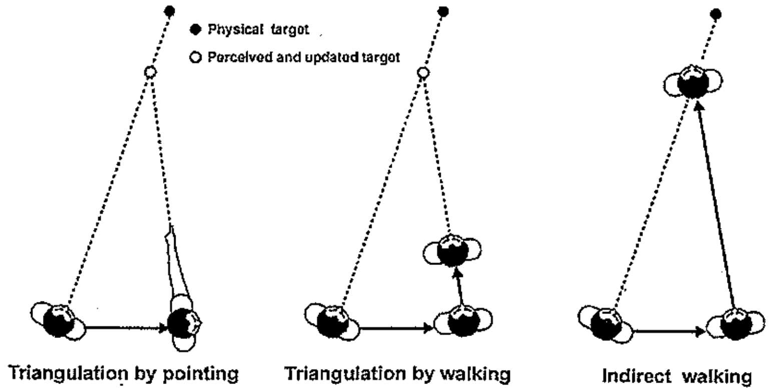

Figure 1.6 Three triangulation methods (see text for explanation).

to) a target and, then without further perceptual information about its location, walks along a straight path to a new location (specified by auditory or haptic signal) and then points toward the target. The pointing direction is used to triangulate the initially perceived and spatially updated target location. In one variant, the arm orientation was monitored continuously as the observer walked along a straight path (Loomis et al., 1992) after viewing a target on the ground up to 5.7 m away; the average pointing responses were highly accurate. "Triangulation by walking" (also, "triangulated walking") is similar to triangulation by pointing except that after the initial straight path, the observer turns and walks a short distance toward the target. The walking direction after the turn is used to triangulate the perceived and updated target location. Finally, in the "indirect walking" version of triangulation, the observer walks to a turn point (specified by auditory or haptic signal), and then attempts to walk the rest of the way to the updated target location. Figure 1.7 gives the results of a number of experiments using the different triangulation methods to measure the perceived distance of targets viewed under full-cue conditions (Fukusima, Loomis, & Da Silva, 1997; Knapp, 1999; Loomis et al., 1998; Philbeck et al., 1997; Thompson et al., 2004); as in Figure 1.2, the data sets have been vertically shifted for clarity. Although the data are more variable than with blind walking (Figure 1.2), they indicate overall that perceived distance is proportional to target distance with little systematic error.

A Model of Perceptually Directed Action

Because blind walking and the triangulation methods just mentioned rely on actions that occur after the percept has disappeared, it might be argued that these methods cannot be used to measure perception because the action depends upon postperceptual processes (e.g., Proffitt et al., 2006). However, we maintain that a valid measurement method is one for which the variations in the indicated values (those resulting from the measurement process) are coupled to variations in the variable being measured and for which a calibration between the two has been established (Hutchison & Loomis, 2006b). As with any measurement device (e.g., a thermometer with an electronic display), the indirectness of the mechanism between the variable being measured and the indicator has no bearing on whether the indicated values are proper measures of the variable of interest, here perceived distance. In the case of perceptually directed action, what is required is a theory linking the indicated value to perceived distance. Action can be used to measure perception provided that the postperceptual processes introduce no systematic biases or, if they do, that the biases can be corrected for by way of calibration. Of course, as with any measurement device or method, the precision of measurement will ultimately be limited by random noise associated with each of the subsequent processes, even if systematic biases can be eliminated by calibration.

Here, we present a model of perceptually directed action that links the perceptual representation to the observed behavior. The model

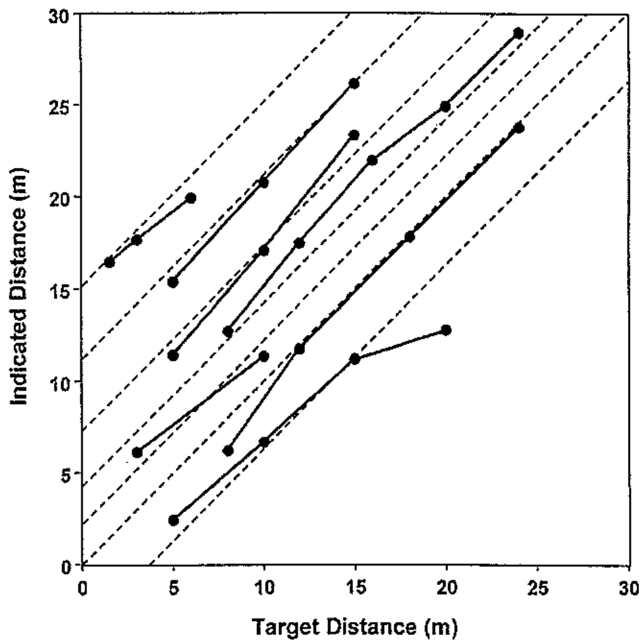

Figure 1.7 Summary of triangulation results for vision using triangulation by pointing, triangulation by walking, and indirect walking, all obtained under full-cue conditions. The data from the different studies have been displaced vertically for purposes of clarity. The dashed line in each case represents correct responding. Sources from top to bottom: results of indirect walking by Philbeck et al. (1997), results of triangulation by walking in real environment (Thompson et al., 2004), outdoor results of Knapp (1999), average of two conditions from Experiment 3 (triangulation by walking) of Fukusima et al. (1997), Experiment 3 (direct and indirect walking) of Loomis et al. (1998), average of two conditions from Experiment 4 (triangulation by walking) of Fukusima et al. (1997), and average of two conditions from Experiment 2 (triangulation by pointing) of Fukusima et al. (1997).

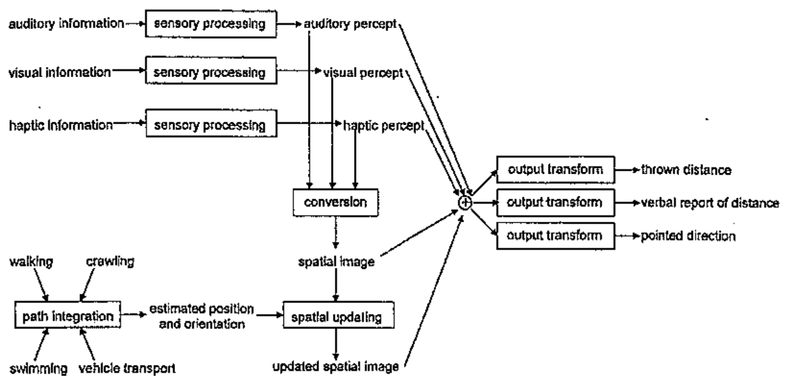

Figure 1.8 A block diagram of perceptually directed action (see text for explanation).

involves a number of processing stages (Figure 1.8). For similar models, see Böök and Gärling (1981), Loomis et al. (1992), Medendorp, Van Asselt, and Gielen (1999), and Rieser (1989). First, the visual, auditory, or haptic stimulus gives rise to the percept, which may or may not be coincident with the target location (Figure 1.9). Accompanying the percept is a more abstract and probably more diffuse "spatial image," which continues to exist in representational space even after the percept ceases. There is evidence that the spatial images from different modalities are functionally equivalent, perhaps even amodal in nature (Avraamides, Loomis, Klatzky, & Golledge, 2004; Klatzky, Lippa, Loomis, & Golledge, 2003; Loomis, Lippa et al., 2002). We assume that the spatial image is coincident with the percept, but future research may challenge this assumption; for now, it appears that 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 (Böök & Gärling, 1981; Loarer & Savoyant, 1991; Loomis et al., 1992; Loomis, Klatzky, Golledge, & Philbeck, 1999; Rieser, 1989; Thomson, 1983). At any point in the traverse, as depicted in Figures 1.8 and 1.9, the observer may be asked to make some nonlocomotion response, such as pointing at or throwing to the target or verbally reporting the remaining distance. The response processes clearly are different for different types of response. An important assumption, to be discussed later, is that the response is computed in precisely the same fashion whether based on the concurrent percept of the target or on the spatial image of the target (whenever the percept is absent). This assumption is depicted in Figure 1.8 by the convergence of percepts, initial spatial images, and updated spatial images onto the output transforms for different types of responses.

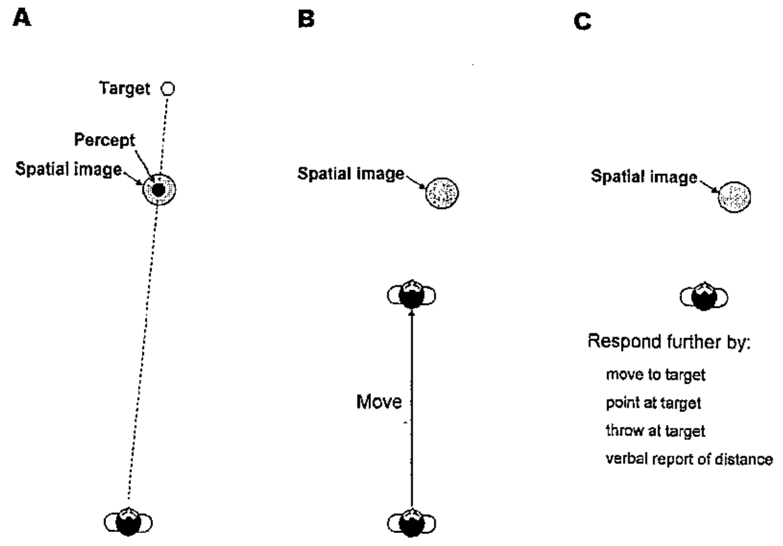

Figure 1.9 From left to right, depiction of 3 successive moments during perceptually directed action. A. The observer perceives a target closer than its physical distance. Accompanying the perceived target is a more abstract and spatially diffuse spatial image. B. With the stimulus and its percept no longer present, the observer moves through space, updating the egocentric distance and direction of the target. If path integration is accurate (as depicted here), the spatial image remains stationary with respect to the physical environment. C. After moving, the observer can make another response to the updated spatial image, by continuing to move toward it, by pointing at it, by throwing at it, or by making a verbal report of the distance remaining.

Not depicted in Figure 1.8 is a nonperceptual input to the creation of a spatial image. Loomis and his colleagues (Avraamides et al., 2004; Klatzky et al., 2004; Loomis, Lippa et al., 2002) have shown that once a person forms a spatial image, whether based on a spatial percept or based on spatial language, subsequent behaviors (like spatial updating and exocentric direction judgments) appear to be indifferent to the source of the input, suggesting that spatial images based on different inputs might be amodal. The implication is that the spatial image produced by vision, hearing, or touch, can, in principal, be modified by higher-level cognition so as not to be spatially coincident with the percept. Whether this dissociation between percept and spatial image ever occurs remains to be determined, but the evidence to be reviewed is consistent with the assumption of coincidence.

Accurate path integration and consequent accurate updating mean that the spatial image remains fixed with reference to the physical environment. If path integration is in error, multiple updated targets move with respect to the physical environment, but they move together rigidly. If only sensed speed but not heading during path integration is off by a constant factor, blind walking to a target over very different paths will cause the terminal points to coincide even though the convergence point will not coincide with the initially perceived location. Comparisons of indirect walking responses with direct walking responses to the same visual targets indicate that walking results in accurate path integration and consequent accurate spatial updating (Loomis et al., 1998, Philbeck et al., 1997). Figure 1.10 gives the average terminal locations for direct and indirect walking responses to visual and auditory targets in one of these studies (Loomis et al., 1998). The congruence of the direct and indirect terminal points and their near coincidence with the visual targets demonstrate the accuracy of updating. The fact that the terminal points for audition are further from the auditory targets (along the initial target directions) than the vision signifies the poorer accuracy of auditory distance perception compared to visual distance perception. Whereas actively controlled walking to targets at normal walking speeds produces accurate path integration and updating for short paths (e.g., Philbeck, Klatzky, Behrmann, Loomis, & Goodridge, 2001), walking at unusual speeds or passive transport by wheelchair or other conveyance generally results in degraded updating performance (Israël, Grasso, Georges-François, Tsuzuku, & Bathos, 1997; Juurmaa and Suonio, 1975; Marlinsky, 1999a, 1999b; Mittelstaedt & Glasauer, 1991; Mittelstaedt & Mittelstaedt, 2001; Sholl, 1989), especially for older adults (Allen, Kirasic, Rashotte, & Haun, 2004).

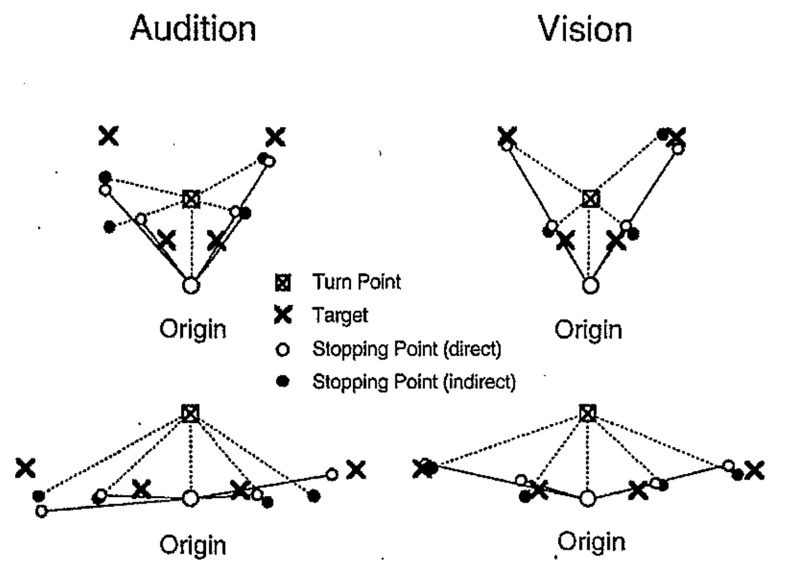

Figure 1.10 Stimulus layout and results of an experiment on spatial updating of visual and auditory targets by Loomis et al. (1998). The observer stood at the origin and saw or heard a target (X) located either 3 or 10 m distant at an azimuth of -80°, -30°, 30°, or 80°. Without further perceptual information about the target, the observer attempted to walk to its location either directly or indirectly. In the latter case, the observer was guided forward 5 m to the turn point and then attempted to walk the rest of the way to the target. The small open circles are the centroids of the direct path stopping points for the 7 observers, and the small closed circles are the centroids for the indirect path stopping points. Reproduction of Figure 5.6 from 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.

A more recent variant of perceptually directed action provides a means of measuring the 3-D perceived location of a visual stimulus perceived to lie above the ground plane (Ooi, Wu, & He, 2001, 2006). Figure 1.11 depicts the procedure and typical pattern of results for luminous visual targets placed on the ground in an otherwise dark room. At the left, the observer views the target. Because of insufficient distance cues, the target is perceived to be closer than it is, resulting in the percept being elevated. The observer, wearing a blindfold, then walks out to the perceived and updated target and gestures by placing the hand at its location. Despite errors in distance (which are like those

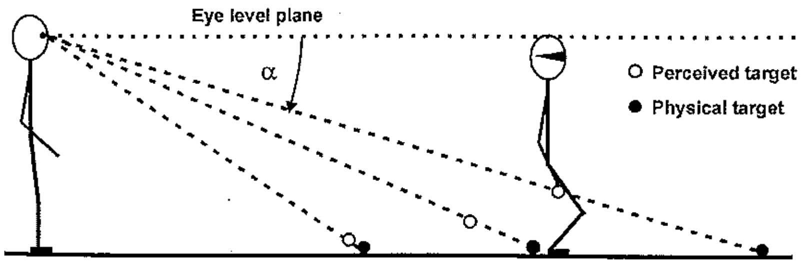

Figure 1.11 Procedure and typical results for the experiments of Ooi, Wu, and He (2001, 2005). A glowing target was placed on the ground in an otherwise dark room. At the left, the observer views the target. Because of insufficient distance cues, the target is perceived to be closer than it is, resulting in the percept being elevated. The observer, wearing a blindfold, then walks out to the perceived and updated target and crouches to place the hand at its location. The angle α is "angular declination" (or "height in the field"), which is an important cue to egocentric distance.

reported by Philbeck and Loomis [1997]; see Figure 1.4), the indicated locations lay in very nearly the same direction as the targets as viewed from the origin. Given the complexity of this "blind-walking–gesturing" response, this is a remarkable result. It is difficult to imagine any interpretation of this result other than one of systematic error in perceiving the target's distance followed by accurate path integration and spatial updating. Similar evidence supporting this interpretation comes from the aforementioned studies involving both direct and indirect walking (Loomis, Klatzky et al., 1998; Loomis, Lippa et al., 2002; Philbeck et al. 1997). When people walked along indirect paths while updating, they traveled to very nearly the same locations as when traveling along direct paths. Importantly, when the terminal points clearly deviated from the targets in terms of distance, indicating distance errors, the directions were nonetheless quite accurate (Loomis, Klatzky et al., 1998; Loomis, Lippa et al., 2002; Philbeck et al. 1997). The conclusion is strong that observers were traveling to the perceived and updated target locations.

Still further evidence that perceptually directed action can be used to measure perceived location and, thus, perceived distance, comes from recent work by Ooi et al. (2006). They performed two experiments, one involving the blind-walking—gesturing response to luminous point targets in the dark—and the other involving judgments of the shapes of luminous figures, also viewed in the dark. The indicated locations and judged slants were consistent with the targets in the two tasks being located on an implicit surface, extending from near the observer's feet and moving outward while curving upward; the authors hypothesize that the implicit surface reflects intrinsic biases of the visual system, like the specific distance tendency (Gogel & Tietz, 1979). It is significant that two such very different responses, one involving action and other involving judgments of shape, can be understood in terms of a unitary perceptual process.

The flexibility of perceptually directed action is indicated by studies demonstrating on-line modification of the response. In a highly influential paper, Thomson (1983) reported an experiment that prevented the observer from executing a preplanned response. On a given trial, the observer viewed a target on the ground some distance ahead. After the observer began blind walking toward the target, the experimenter gave a signal to stop and throw a beanbag the remaining distance. Accuracy was high even though observers did not know at which point they would be cued to throw, demonstrating a flexible response combining two forms of action. Another example of on-line modification comes from one of the experiments on direct and indirect walking to targets (Philbeck et al., 1997). Here observers viewed a target, after which their vision and hearing were occluded. On cue from the experimenter, the observers walked to the target along one of three paths. Because they were not cued as to which path to take until after vision was occluded, the excellent updating performance indicates that observers could not have been preprogramming the response. Given these results, it appears that action directed toward a goal is extremely flexible. Presumably, once a goal has been established, any combination of actions, including walking, sidestepping, crawling, and throwing can be assembled "on the fly" to indicate the location of an initially perceived target. For other evidence of on-line adjustment of perceptually directed action, see the paper by Farrell and Thomson (1999).

Despite the involvement of cognitive and locomotor processes in perceptually directed action, the results of a number of experiments demonstrate that this method provides a pure, albeit indirect, measure of perceived distance (and direction). They do so by demonstrating that cognitive and motor processes contribute little to the systematic error of task performance. Especially compelling are the results demonstrating the absence of systematic error in path integration, spatial updating, and response execution through the congruence of direct and indirect walking paths (Loomis, Klatzky et al., 1998; Loomis, Lippa et al., 2002; Philbeck et al., 1997; see Figure 1.10) and the result of Ooi et al. (2001, 2006) showing that terminal location of a complex spatial response is consistent with the initial target direction. In addition, the close coupling of action and verbal responses in both reduced-cue and full-cue conditions (Philbeck & Loomis, 1997), the close coupling of blind walking and throwing responses in both real and virtual environments (Sahm et al., 2005), and the close coupling of blind walking/gesturing and shape judgments (Ooi et al., 2006; Wu et al., 2004) is further support that the action-based method provides a measure of distance perception. A later section showing that spatial updating can be used to correct for biases in verbal report provides still further evidence.

Role of Calibration in Perceptually Directed Action

Although we believe that cognitive and motor processes do not contribute appreciably to systematic errors in perceptually directed action, such processes are clearly required to control and execute the response. In light of this, and given the very high accuracy of visually directed action under full-cue conditions, there is good reason to believe that perceptually directed action depends on some adaptation process that keeps the action system in calibration. Indeed, several forms of adaptation of perceptually directed action have been demonstrated (Durgin, Fox, & Kim, 2003; Durgin, Gigone, & Scott, 2005; Durgin, Pelah, Fox, Lewis, Kane et al., 2005; Ellard & Shaughnessy, 2003; Mohler, Creem-Regehr, & Thompson, 2006; Ooi et al., 2001; Philbeck, O'Leary, & Lew, 2004; Richardson & Waller, 2005; Rieser, Pick, Ashmead, & Garing, 1995; Witt et al., 2004). Does adaptation call into question the claim that perceptually directed action can be used to measure perception?

Because perception, path integration/spatial updating, and response execution all contribute to perceptually directed action, adaptation of any of these processes or of the couplings between them can be expected to alter task performance. Adaptation that alters perceived distance will influence all measures of distance perception (verbal report, size-based measures, and all action-based measures, including walking and throwing) for whatever sensory modality has been adapted. In the aforementioned study, Ooi et al. (2001) used prism adaptation during a prior period of walking with vision to alter the effective "angular declination," one of the cues in sensing target distance (represented by α in Figure 1.11). The results showed that perceived distance was indeed altered because walking and throwing responses were similarly affected, even though only walking was used during adaptation.

Adaptation of the path integration process alters the gain of sensed self-motion relative to actual physical motion and is expected to have a uniform influence on walked distance. That is, if the gain is halved, observers ought to walk twice as far in order to arrive at the updated targets, and this effect should apply to all targets regardless of initial distance and the path taken. If the adaptation of walking speed is specific to walking direction, then the effect of adaptation on path integration will depend upon the walking direction. Rieser et al. (1995) had observers walk on a treadmill while being pulled by a tractor and altered the normal relationship between vision and the proprioceptive cues of walking. Adaptation to this altered relationship produced reliable changes in the walked distance to previewed targets but did not affect throwing to targets, a result that rules out perceptual adaptation. The magnitude of recalibration depended upon walking direction. A likely interpretation is that sensed self-motion was altered by the adaptation process. Durgin and his colleagues have additional evidence of recalibration of sensed self-motion (Durgin, Fox et al., 2003; Durgin, Gigone et al., 2005; Durgin, Pelah et al., 2005), although their results and interpretation of the cause of recalibration differ somewhat from those of Rieser et al. (1995). It might be thought that recalibration involves a comparison of sensed self-motion signaled by idiothetic (proprioceptive and inertial) cues and the overall pattern of optic flow, but recent results by Thompson, Mohler, and Creem-Regehr (2005) show that the recalibration depends upon the perceived scale of the environment, with optic flow held constant. This means that recalibration is determined by the comparison between sensed self-motion signaled by idiothetic cues and the sensed self-motion based on visual perception of the environment, which depends upon distance perception.

The accuracy of visually directed action under full cues (Figures 1.2 and 1.7) clearly relies on calibration of the gain of walking relative to visual perception. This calibration is likely to be induced by sensing of just the near visible environment (e.g., Thompson et al., 2005). An interesting implication is that adaptation of visually sensed self-motion ought to affect any form of action based on spatial updating regardless of the modality with which the target is perceived; thus, the indicated distance to a target based on perceptually directed action will be altered the same amount by visual recalibration whether the targets are visual, auditory, or haptic, provided that the initially perceived distances are the same for the different modalities.

An experiment comparing the effects of feedback on blind walking and verbal reports (Mohler et al., 2006) showed evidence of a form of recalibration not confined to the action system. To induce recalibration, the authors took advantage of the systematic underperception of distance in virtual reality. Observers gave verbal reports and performed blind walking to targets seen in the virtual environment before and after getting feedback about their errors. During the feedback phase, observers blind walked to the estimated locations of the targets and were given feedback about their errors. Both walking responses and verbal reports showed considerable improvement with the feedback. Because verbal reports were affected, the recalibration cannot be confined to a change in sensed self-motion. Although the common recalibration is consistent with a modification of perceived distance, the authors conclude that it is more likely a result of a cognitive rule that influences both types of responses. If true, this might be the result of a cognitive alternation in the spatial image so that it does not coincide with the perceived target. If so, triangulation responses ought to be similarly affected.

Still another form of adaptation has been demonstrated by Ellard and Shaughnessy (2003). In their experiment, observers viewed targets at varying distances on different trials and blind walked to them. For two of the targets, observers were given false feedback about the accuracy of their responses. Telling observers that they had undershot the target resulted in overshooting on subsequent trials. This form of adaptation was specific to the targets for which false feedback was given.

The result by Ellard and Shaughnessy (2003) raises the possibility of a form of adaptation that might undermine the claim that perceptually directed action measures perception. In particular, this type of adaptation could potentially explain the linearity (proportionality) between responded distance and visual target distance under full cues even if perceived distance should in fact be a compressively nonlinear function of target distance, as claimed, for example, by Gilinsky (1951). It would have to be a type of adaptation that modifies neither perception (which affects all responses, whether action-based or not) nor path integration (which affects all walked distances by the same scale factor). In addition, because of the aforementioned triangulation results, it would have to affect the coupling between perceived distance and sensed displacements from the origin, regardless of path taken, and do so in a way that varies nonlinearly with distance from the origin (so as to compensate for the putative nonlinearity between target and perceived distances).

There are at least three lines of evidence against the hypothesis that this type of distance specific adaptation undermines the measurement of perceived distance. The first is that people rarely view distant targets and then walk to them without vision. Error feedback following such blind walking would be needed to "calibrate" perceptually directed action based on the putative nonlinearity between target distance and perceived distance.

The second line of evidence is concerned with whether adaptation following visual feedback about open-loop walking errors generalizes to other forms of perceptually directed action. Richardson and Waller (2005; Experiment 2) had observers perform blind walking to targets in virtual reality. At the outset, their observers walked to locations only about half of the simulated target distances, along both direct and indirect paths, indicating the underperception of distance in virtual reality found by others using a variety of methods (Knapp, 1999; Sahm et al., 2005; Thompson et al., 2004). Observers were then given a period of training involving open-loop walking to the targets along direct paths; after arriving at the estimated position of the target on each trial, the observer was given explicit feedback about the undershoot error. After training, observers were once again tested using direct and indirect walking. The training eliminated 79% of the undershoot error for direct walking but only 27% of the undershoot error for indirect walking. The large difference in amounts of recalibration argues that if people get explicit feedback about their blind walking in the real world so as to allow for correct walking to compensate for putative nonlinearity in perceived distances, this type of recalibration still would not explain the high accuracy with which people perform triangulation tasks.

In a follow-up study, Richardson and Waller (2007) found that observers who were allowed to walk around in immersive virtual environments while continuously interacting with visual targets exhibited a more general form of recalibration. Contrasting with the results obtained with explicit feedback in their earlier study, the results of this study showed that the implicit feedback accompanying interaction with the environment during the training phase did allow for an equal amount of recalibration when walking open-loop along direct and indirect paths during the testing phase. Prima facie, this result appears to support the hypothesis that recalibration accounts for accurate visually directed action despite nonlinear functions for egocentric distance. However, their result is also consistent with two other hypotheses: recalibration of visual perception and recalibration of sensed self motion. Further experiments not relying on updating (e.g., verbal report and ball throwing) are needed to distinguish between the three alternative hypotheses.

The third line of evidence against the hypothesis is made possible by comparing the accurate responses to visual targets with the systematically compressed responses to auditory targets. Figure 1.12 gives the results of two experiments (Loomis et al., 1998) from 12 observers who made both verbal and blind walking responses to visual targets and 12 observers who made both types of responses to auditory targets; some observers were given targets at 4, 10, and

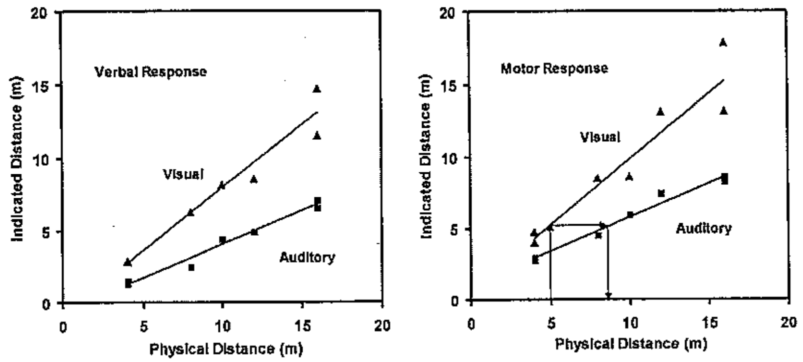

Figure 1.12 Results of two experiments on visual and auditory distance perception (Experiments 1 and 2a from Loomis et al. [1998]). The same observers made verbal and blind walking responses to both visual and auditory targets in a large open field. Seven responded to targets at 4, 10, and 16 m, and 5 observers responded to targets at 4, 8, 12, and 16 m. The best fitting linear functions are plotted as well. In the right panel, arrows indicate how, for a given visual target distance, the corresponding "equivalent" auditory distance was determined, this being the distance of the auditory target which produced the same walked distance.

16 m, and others were given targets at 4, 8, 12, and 16 m. The best fitting linear functions (with 0 intercepts) for the combined data sets are plotted as well.

If the accurate responses to visual targets reflect some sort of calibration process acting on the visually-based action process, presumably the same calibration does not apply to auditorially based action, given the very large systematic errors. Thus, the two action processes must involve different calibration functions. This means that if a visual target distance and an auditory target distance produce the same value of blind walking, the corresponding values of visually perceived and auditorially perceived distance must be different. Thus, we would expect that since the process of making a verbal report is common to both modalities, the verbal reports ought to be different for visual and auditory target distances that produce the same action response, assuming that the visually-based action responses have been calibrated through experience.

To test this idea, we have used the best fitting linear functions to the blind walking (motor) responses (Figure 1.12) to find, for each visual target distance, the corresponding auditory target distance that produced the same walking responses (see arrows in the right panel of Figure 1.12). The visual target distance and "equivalent" auditory target distances were then used to compute the corresponding verbal reports from the best-fitting linear functions. The function relating the verbal reports for vision and the verbal reports for audition is very nearly the identity function (linear function with slope of 1.04 and intercept of -0.22 m). This means that the visually-based and auditorially-based action processes are essentially identical. At least for these data, there is no evidence of a calibration of blind walking to compensate for a putative compressively nonlinear perceptual function.

On the basis of the three lines of evidence, we conclude that perceptually directed action does provide a comparatively pure measure of perception when action is properly calibrated to near surrounding space. Also based on both the perceptually directed action results of Figures 1.2 and 1.7 and the verbal report results of Figure 1.1, we conclude that visual distance perception is a linear function of target distance out to at least 25 m on the ground plane when distance cues are abundant.

Using Spatial Updating to Correct for Bias in Verbal Reports of Egocentric Distance

Figures 1.1, 1.2, and 1.7 show that visually perceived egocentric distance is a linear function of physical distance out to 25 m and that verbal reports are generally about 80% of the distances indicated by action. This systematic difference in response values does not, by itself, indicate that different internal representations of distance control the two types of responses, for Foley (1977) and later Philbeck and Loomis (1997) presented a model in which the same internal representation of distance, ostensibly perceived distance, acts through different output transforms to determine the indicated responses (see right part of Figure 1.8). Philbeck and Loomis (1997) showed that verbal reports and blind walking to targets, while systematically different, were related to each other by a fixed mapping when switching from reduced-cue to full-cue viewing (also see the above analysis in connection with Figure 1.12). This is consistent with there being just a single internal representation of distance acting through different output transforms. However, it is possible, even likely, that the output transforms for action and verbal report are sometimes affected differently by experimental manipulations, such that there is no fixed mapping between the two types of responses. For example, an observer can view a photograph and make reliable judgments of distance of depicted objects, but asking observers to perform open-loop walking to the same depicted objects is likely to be met with reluctance followed by very noisy performance. Andre and Rogers (2006) have found that experimental manipulations, including viewing targets through base up and base down prisms, can differentially affect blind walking and verbal report. They interpret their findings in terms of different internal representations of distance, but it is possible that the differential effects are produced at the level of response production and execution.

In connection with the hypothesis that action and verbal report involve the same internal representation but different output transforms, a special case can be identified in which verbal reports are subject to a systematic under reporting bias (b) that is a constant proportion of perceived egocentric distance (i.e., b<1.0). Figure 1.13 depicts the consequences of such a bias in an updating task, based on four assumptions. First, the output transform for verbal report of distance is assumed to be the same whether the report is based on a concurrent percept, a spatial image prior to updating, or an updated spatial image (right part of Figure 1.8). The same assumption is made for verbal report of direction. Third, any systematic biases in verbally reporting perceived direction are assumed to be small for both vision and hearing, an assumption that is well supported (Loomis et al., 1998; Experiment 2A). Fourth, path integration is assumed to be properly calibrated (per the previous section), with accurate spatial updating a consequence. Figure 1.13 depicts the situation where the verbal report scaling factor (b) is 0.75 times the perceived distance from the origin or the distance of the spatially updated spatial image following a walk. The result is that the reported locations shift forward and parallel with the walk, with the shift magnitude equal to the length of the walk multiplied by 1-b (here 0.25). Figure 1.14A shows the predicted shifts for the same factor (b=0.75) for all of the target locations used in Experiment 2A of Loomis et al. (1998). In this experiment, observers made verbal estimates of distance and direction while viewing or listening to targets. They then moved forward 5 m while updating and then made verbal reports of the target locations. The centroids of the reported locations of the perceived and updated visual targets are shown in Figure 1.14B.[1] The general pattern of forward shifts in reported location, predicted in the left panel, can be seen in the data in the right panel, albeit with superimposed noise. At the time these data were published, the authors found this pattern of shifts enigmatic (Loomis et al., 1998, pp. 974–975).

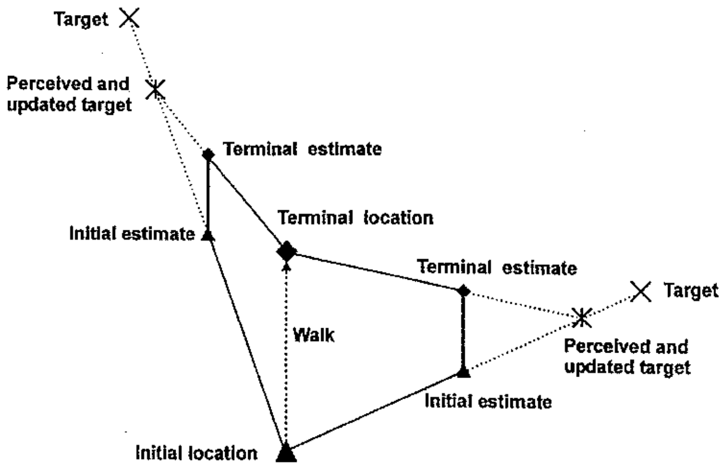

Figure 1.13 The consequences of verbally under reporting the perceived distance of a target. The physical targets are represented by the Xs. The perceived and updated targets are closer to the initial location than the physical targets. The initial estimates represent the verbally under reported locations of the perceived targets as judged from the initial location prior to the walk. After the observer walks, the terminal estimates are the verbally under reported locations of the updated targets as judged from the terminal location after the walk. When the verbally reported distance is less than the perceived and updated target locations, as depicted here, the estimated location of the target moves with the observer along a parallel path.

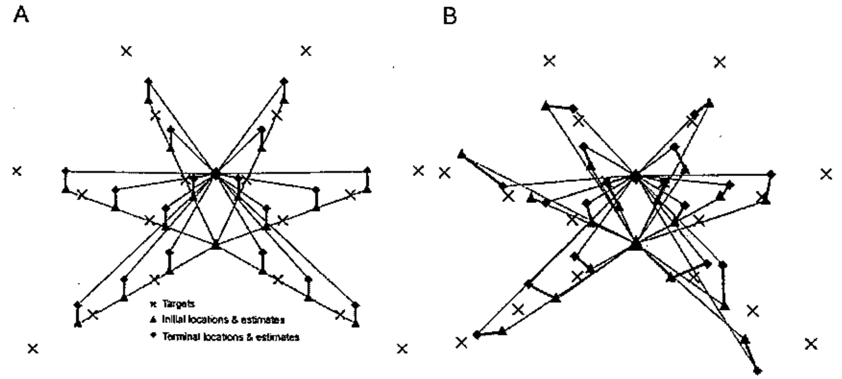

Figure 1.14 (A) The predicted pattern of shifts in the verbally reported locations for the target locations in Experiment 2 of Loomis et al. (1998), assuming that verbally reported location is only 0.75 as great as the perceived location. The targets are represented by the Xs. The terminal location from which judgments were made was forward of the initial location following a 5 m walk. The terminally reported locations are all displaced equally from the initially reported locations in the direction of the walk, as depicted by the thick lines. (B) The observed pattern of verbally reported locations to visual targets before and after walking 5 m forward in Experiment 2 of Loomis et al. (1998). The displacements are depicted by the thick lines.

If, as hypothesized, the general pattern of shifts is the result of systematically under reporting perceived distance, the reported perceived locations prior to walking and the reported updated locations after walking should be brought into better congruence by scaling the reported distances, before and after walking, by the inverse of the reporting bias b. Two measures of congruence are (1) mean distance between the initial and terminal locations, averaged over all targets, and (2) the degree of directional alignment between the initial and terminal locations from the origin, averaged over all targets. Figure 1.15A gives the results of the scaling analysis for both measures of congruence, as applied to the visual data. A scaling factor of 1.3 produces the maximum congruence. (The auditory data were several times noisier; the resulting scaling factor was 1.7.) Figure 1.15B shows the revised initial and terminal reported locations after the rescaling of reported distances, for both the perceived and updated targets. Because of other unknown sources of error, the congruence is still far from perfect, but the two measures of congruence provide a similar estimate of 1.3 for the scaling factor. By hypothesis, this means that verbal estimates are biased toward underreporting of perceived distance by 1/1.3=0.77. The adjusted reported locations are now on average quite close to the target locations (Figure 1.15B) in terms of distance from the origin. Interestingly, this bias is close to the average factor by which the verbal reports summarized in Figure 1.1 differ from the action based measures of Figures 1.2 and 1.7. Although more research needs to be done, the analysis here suggests that spatial updating can be used to correct for the bias in verbal reports,[2] provided that path integration is properly calibrated. An implication of this analysis is that corrected verbal reports and action measures of perceived egocentric distance agree in showing linear and accurate perception of distance over the ground surface under full cue conditions.

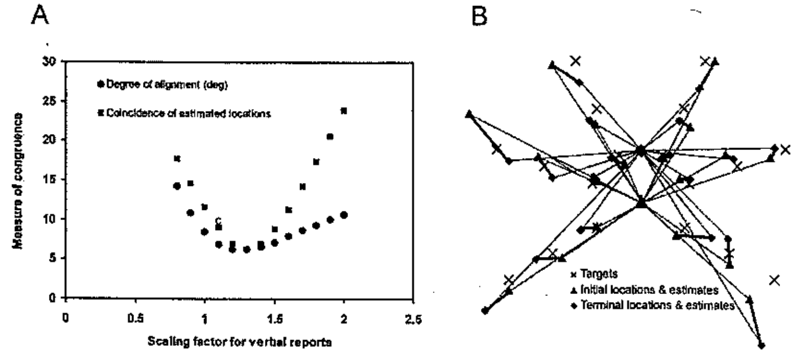

Figure 1.15 (A) Two measures of congruence of the initial and terminal estimated locations (from verbal report) in response to visual targets in Experiment 2 of Loomis et al. (1998), as a function of the scaling factor used to correct verbal reports of distance. The two measures of congruence are mean distance between the initial and terminal locations, averaged over all targets, and the degree of directional misalignment between the initial and terminal locations from the origin, averaged over all targets. Low values indicate high congruence for both measures. The results for both measures indicate that multiplying the verbal reports by a factor of 1.3 produces maximum congruence. (B) The data of Figure 14B after rescaling the verbal reports of distance using the scale factor of 1.3. In addition to the initial and terminal estimates being maximally congruent, the estimated locations are now, on average, quite close to the physical targets in terms of distance from the origin. The displacements are depicted by the thick lines.

Distortions of Perceived Exocentric Distance and Perceived Shape

The earlier conclusion that visual distance perception is proportional to target distance out to 25 m on the ground plane under full-cue conditions seems to contradict other evidence that exocentric distances and visual shapes are systematically misperceived, even within 2 m (Beusmans, 1998; Foley et al., 2004; Kudoh, 2005; Levin & Haber,1993; Loomis et al., 1992; Loomis & Philbeck, 1999; Loomis, Philbeck, & Zahorik, 2002; Ooi et al., 2006; Toye, 1986; Wagner, 1985; Wu et al., 2004).[3] Generally speaking, exocentric depth extents are perceived as smaller than physically equal exocentric extents in the frontoparallel plane and distant exocentric depth extents are perceived as smaller than physically equal exocentric depth extents that are close. A full discussion of this topic is beyond the scope of this chapter. Here, we mention five recent papers that go a long way in reconciling the linearity of visual distance perception with the large distortion of exocentric distance and shape within the same range. Foley et al. (2004) have put forth a mathematical model of visual space in which large distortions in perceived exocentric extents are consistent with a linear function (or slightly nonlinear function, as they maintain) relating perceived egocentric distance to physical egocentric distance. Although the processes underlying the model have yet to be elucidated, the model has great promise for explaining a variety of distortions in visual space. Wu et al. (2004) deal with the