

Spatial Updating of Locations Specified by 3-D Sound and Spatial Language

Jack M. Loomis and Yvonne Lippa
University of California, Santa Barbara

Roberta L. Klatzky
Carnegie Mellon University

Reginald G. Golledge
University of California, Santa Barbara

Blind and blindfolded sighted observers were presented with auditory stimuli specifying target locations. The stimulus was either sound from a loudspeaker or spatial language (e.g., “2 o’clock, 16 ft”). On each trial, an observer attempted to walk to the target location along a direct or indirect path. The ability to mentally keep track of the target location without concurrent perceptual information about it (spatial updating) was assessed in terms of the separation between the stopping points for the 2 paths. Updating performance was very nearly the same for the 2 modalities, indicating that once an internal representation of a location has been determined, subsequent updating performance is nearly independent of the modality used to specify the representation.

Research has shown that spatial language (SL) can be used to create spatial images of layouts of objects (e.g., Bosco, Filomena, Sardone, Scalisi, & Longoni, 1996; Ferguson & Hegarty, 1994; Perrig & Kintsch, 1985; Taylor & Tversky, 1992; Wilson, Tlauka, & Wildbur, 1999). The question we address here is whether spatial images based on language can be used to guide a person’s action through space with the same precision with which perceptually based images can be used to guide action. In particular, we ask whether a verbal description, such as “1 o’clock, 3 m,” results in a spatial image functionally similar to that produced by a sound source perceived to be at the same direction and distance.

The task of interest here is called *spatial updating*. The term refers to the ability of a moving person to mentally update the location of a target initially seen, heard, or touched from a stationary observation point. A multitude of studies have shown that when people view a visual stimulus in space and are then prevented from receiving further information about the stimulus, they can update an internal representation of the stimulus while moving about in space (e.g., Amorim, Glasauer, Corpinot, & Berthoz, 1997; Bök & Gärling, 1981; Easton & Sholl, 1995; Farrell &

Robertson, 1997; Farrell & Thomson, 1999; Fukusima, Loomis, & Da Silva, 1997; Huttenlocher & Presson, 1979; Loarer & Savoyant, 1991; Loomis, Da Silva, Fujita, & Fukusima, 1992; Loomis, Klatzky, Philbeck, & Golledge, 1998; Philbeck, Klatzky, Behrmann, Loomis, & Goodridge, 2001; Philbeck, Loomis, & Beall, 1997; Presson & Montello, 1994; Rieser, 1989; Rieser, Ashmead, Talor, & Youngquist, 1990; Rieser, Frymire, & Berry, 1997; Rieser & Rider, 1991; Simons & Wang, 1998; Thomson, 1983). Other studies have demonstrated spatial updating of auditory targets (Ashmead, DeFord, & Northington, 1995; Loomis et al., 1998; Speigle & Loomis, 1993) and haptic targets (Barber & Lederman, 1988; Hollins & Kelley, 1988). The question of primary interest in the present study is whether a spatial image of an external location specified by SL can be updated and, if so, whether it can be updated with the same precision as one specified by 3-D sound. Prior to conducting this study, we had no idea if an internal representation created by SL could be updated at all.

Because we were using two different input modalities, assessing whether the two differ in terms of spatial updating required that we rule out any differences in performance due to differences in perceiving (encoding) the stimuli. The possibility of experimentally distinguishing between errors in spatial updating and errors in stimulus encoding is made possible by an updating task first developed by Philbeck et al. (1997), later used in a study comparing vision and audition (Loomis et al., 1998), and now being used in the present study. The task involves having observers travel along one of two different paths while updating a specified location. Figure 1 gives the stimulus layout and results of the Loomis et al. (1998) study. The visual and auditory targets, indicated by the ×s, were 3 or 10 m from the origin and had azimuths of 30° or 80° to the left and right of the observer’s facing direction. After the observer saw or heard the target, information about the target was removed, and the observer then attempted to walk to its location along a direct path or an indirect path; the latter involved

Jack M. Loomis and Yvonne Lippa, Department of Psychology, University of California, Santa Barbara; Roberta L. Klatzky, Department of Psychology, Carnegie Mellon University; Reginald G. Golledge, Department of Geography, University of California, Santa Barbara.

This research was supported by National Eye Institute Grant EY09740. A preliminary version of this work was reported at the 40th Annual Meeting of the Psychonomic Society in Los Angeles, April 2000. We are grateful to Mike Kinsella for assistance in conducting the two experiments, to Pavel Zahorik for helpful comments on the work, and to Nancy Collins for statistical consultation.

Correspondence concerning this article should be addressed to Jack M. Loomis, Department of Psychology, University of California, Santa Barbara, California 93106-9660. E-mail: loomis@psych.ucsb.edu

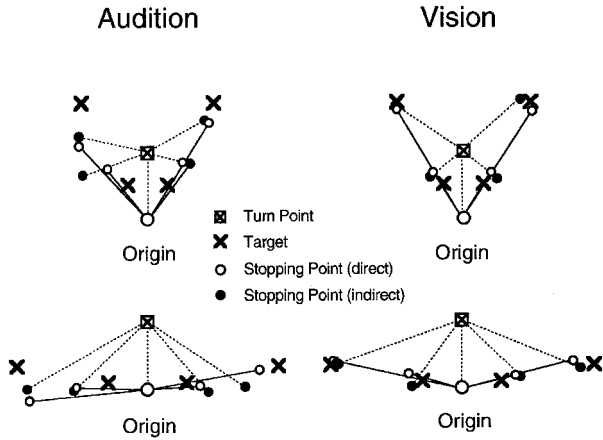


Figure 1. Stimulus layout and results of an experiment on spatial updating of visual and auditory targets. The observer stood at the origin and saw or heard a target located either 3 or 10 m distant at an azimuth of -80° , -30° , 30° , or 80° . Without further perceptual information about the target, 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 seven observers, and the small closed circles are the centroids of the indirect path stopping points. From “Human Navigation by Path Integration,” by J. M. Loomis, R. L. Klatzky, R. G. Golledge, and J. W. Philbeck. In *Wayfinding Behavior: Cognitive Mapping and Other Spatial Processes* (p. 147, Figure 5.6), by R. G. Golledge (Ed.), 1999, Baltimore: Johns Hopkins University Press. Copyright 1999 by Johns Hopkins University Press. Reprinted with permission.

first walking straight ahead 5 m and then turning to walk the rest of the way to the specified location. Figure 1 shows the centroids (two-dimensional means) of the stopping points, averaged over observers, for the direct and indirect paths to each target in each modality. The fact that the centroids were much closer to the visual targets than to the auditory targets means that visual distance perception was more accurate than auditory distance perception, a result confirmed by other research performed under similar conditions (for reviews, see Loomis, Klatzky, & Golledge, 1999 [audition]; Loomis & Knapp, in press [vision]). What is more important for our present concern is that regardless of the initial perceptual error, the near congruence of the centroids for direct and indirect paths to the various targets for both vision and audition indicates that spatial updating was performed well for the two modalities.

To be more precise about how to assess spatial updating performance, we present a two-process model of the task involving direct and indirect paths to a target. The model is given in Figure 2. In what follows, we focus only on the modalities used in the current research: 3-D sound and SL.

The first stage of the model is stimulus encoding. Stimulus encoding, in turn, consists of two substages: processing of the stimulus and creation of an internal spatial representation (spatial image). In the case of 3-D sound, the first substage involves developing a percept of the spatial location of the source. This percept is generally not at the location of the sound source because of the aforementioned error in perceived distance. The second substage of encoding involves creation of a spatial image of the

source location, an image that continues to exist after the sound is no longer present. We assume that the spatial image corresponds to the location of the percept. In the case of SL, the utterance is linguistically converted into meaning, which may or may not involve a spatial image. If it does not, a second subprocess is assumed to result in the creation of a spatial image. Thus, for both 3-D sound and SL, the result of stimulus encoding is a spatial image. This image might well be amodal; that is, once created, its functional characteristics, such as its spatial precision and decay time, may be independent of the input modality (i.e., visual space perception, auditory space perception, haptic space perception, and SL).

The second stage of the model is spatial updating. As the observer rotates and translates, the spatial image is continually updated with respect to the person’s changing orientation and location within the environment. Ideally, the updating process would cause the updated image to remain fixed in relation to the surrounding space as if it were a stationary object. The updating process obviously interacts with the person’s spatial behavior—updating guides spatial behavior and spatial behavior causes updating.

Associated with each of the two stages of processing are two distinct forms of error: bias (systematic error) and noise (variability about the mean). We assume that the output of each process results in internal representations that can be referred to the physical space. Because of noise in the processing, this represented location varies from trial to trial (within observers) or from one observer to another; in what follows, we generally do not distinguish between these sources of noise. If we think of an idealized population of represented locations for a fixed input to the process, the dispersion of represented locations about the mean reflects the noise. Panel A of Figure 3 depicts the variation in a small number of encoded locations (of spatial images) in response to encoding of a fixed stimulus, and Panel B depicts the variation in a small number of stopping locations following spatial updating in response to a fixed encoded location.

The central tendency of the idealized population of represented values in two dimensions is given by its centroid. When the perceived location of a stimulus differs systematically from its physical location, this reflects a perceptual bias, which is more likely an error in perceiving distance than in perceiving direction (e.g., Loomis & Knapp, in press). Thus, when there is systematic perceptual error, the centroid of encoded locations of the spatial images differs from the target location; Panel C of Figure 3 conveys this idea for the case of misperceived distance. When there is bias associated with spatial updating, the centroid of the stopping points for that path will differ from the centroid of the encoded locations (which cannot be directly observed); accordingly, the centroids of the stopping points for the direct and indirect paths will differ by the vector sum of the two updating biases (see Panel D).

To make the model more explicit, we present equations below for the stopping-point error. This is the deviation of the stopping

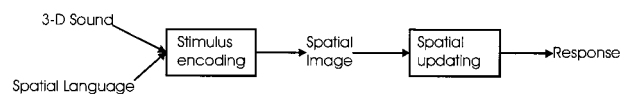


Figure 2. Two-stage model of stimulus encoding and spatial updating.

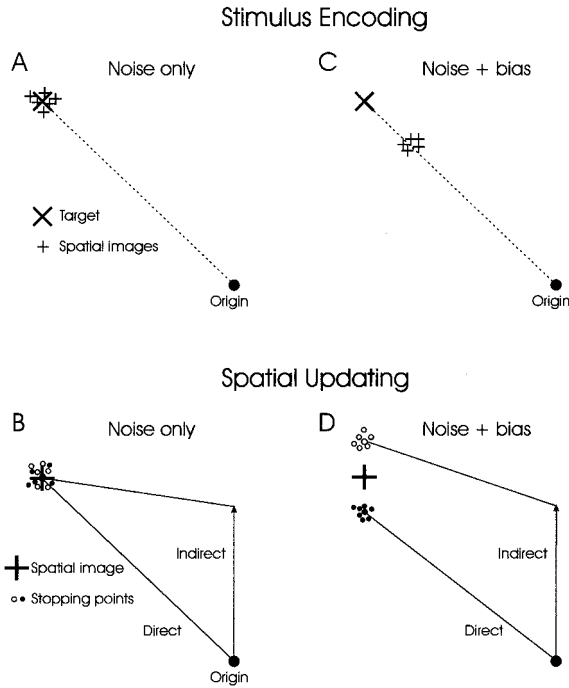


Figure 3. Representation of noise and bias for the two stages of the model: stimulus encoding and spatial updating. Stimulus encoding of a target results in a spatial image. Spatial updating in response to a spatial image results in a stopping point for direct and indirect paths. Updating bias is assumed to be symmetric for the direct and indirect paths.

point from the target on each of the two spatial dimensions (which might be either Cartesian coordinates or polar coordinates relative to the origin of locomotion). Although stopping-point error has these two vector components, we write the equations for the direct and indirect paths without denoting the component axes:

$$\text{error}_{\text{direct}} = \text{bias}_{\text{encoding}} + \text{noise}_{\text{encoding}} + \text{bias}_{\text{direct}} + \text{noise}_{\text{direct}}$$

and

$$\text{error}_{\text{indirect}} = \text{bias}_{\text{encoding}} + \text{noise}_{\text{encoding}} + \text{bias}_{\text{indirect}} + \text{noise}_{\text{indirect}}$$

where the two noise terms are assumed to be Gaussians with means equal to zero (for both spatial dimensions). The terms *direct* and *indirect* denote the bias associated with spatial updating along the direct and indirect pathways to the encoded image. The expected values of error are given in the next two equations:

$$E(\text{error}_{\text{direct}}) = \text{bias}_{\text{encoding}} + \text{bias}_{\text{direct}}$$

and

$$E(\text{error}_{\text{indirect}}) = \text{bias}_{\text{encoding}} + \text{bias}_{\text{indirect}}$$

Figure 4 shows how the centroids of the stopping points (for infinite samples) vary with different combinations of encoding bias and updating bias. Panel A depicts the case where this is neither type of bias; here, the centroids of the direct and indirect paths coincide with the target location. Panel C depicts the case of encoding bias alone. Because the encoding bias is independent of

the path to be walked, the direct and indirect centroids coincide with the spatial image. Panel B depicts the case where there is only updating bias; here the updating biases of the two paths are assumed to be symmetric about the target (see below). Finally, Panel D depicts the case where there is bias in both stimulus encoding and spatial updating; here the centroids differ by the vector sum of the two biases. Thus, the centroid for either the direct path or indirect path is the sum of the encoding bias relative to the target and the updating bias relative to the encoded location.

In evaluating whether 3-D sound and SL differ in terms of spatial updating, we consider only differences in bias, because we are not able to decompose the noise into its encoding and updating components. We also devote some of the analysis to differences between the two modalities in terms of encoding (cf. updating) bias, but this is only of secondary interest for the current article.

Because we have no means of measuring the encoded locations of the spatial images directly, we have no way of separately estimating the updating biases of the direct and indirect path. Consequently, we assume that the two updating biases are symmetric about the encoded location of the spatial image. That is, the errors that are unique to the direct and indirect paths will be equidistant from the encoded location but opposite in direction (see Panels B and D in Figure 4). Under this assumption, the estimated location of the spatial image for a given target is the centroid of the direct path and indirect path stopping points. The import of this assumption is that any updating bias common to the two paths is interpreted as encoding bias. For example, if both direct and indirect updating lead to a common drift outward from the origin, the analysis will assume that the encoded image is biased outward rather than attributing the error to a common

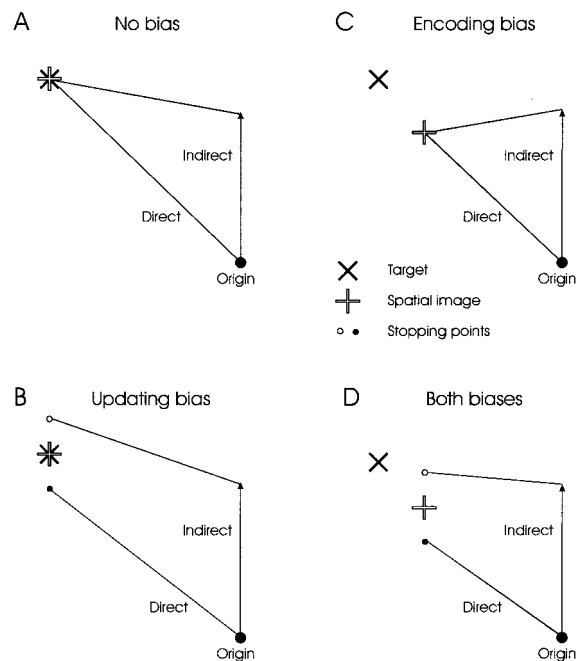


Figure 4. Representation of the idealized direct and indirect paths (no noise) for various combinations of encoding bias and updating bias. When there is no encoding bias, the spatial image is coincident with the target. When there is no updating bias, the direct and indirect paths intersect at the spatial image (or target).

updating bias. Some of the results we report justify this assumption that biases common to the two paths are indeed encoding biases.

In evaluating potential differences between 3-D sound and SL in terms of updating bias, we start with the vector difference between the group centroids for the direct and indirect path stopping points for each target and modality. We then determine whether these vector differences differ for 3-D sound and for SL using multivariate analysis of variance (MANOVA).

Ideally, we would also evaluate whether the two modalities differ in terms of updating noise (cf. bias). It might seem that the correct approach is to assess the variability of the individuals' stopping points relative to the group centroids for the same target and path and to determine whether the variability is greater for one modality than the other. However, because the walking response along either the direct or the indirect path reflects both stages of processing (encoding and updating), variability in the individuals' stopping points reflects both encoding noise and updating noise. Unfortunately, without independent estimates of encoding noise, there is no way to estimate the updating noise for each path.

Our principal interest in this research is in the results obtained with 10 sighted, blindfolded observers. However, a subsidiary issue here is whether spatial updating of targets specified by 3-D and SL is exhibited by blind observers. Accordingly, after running the 10 blindfolded sighted observers, we ran 6 blind observers using a nearly identical procedure but obtained only half as much data from each observer.

Method

Sighted Observers

Thirteen undergraduate students from the University of California, Santa Barbara (UCSB), volunteered to serve as observers to fulfill a course requirement. Two observers were excluded because they did not pass a directional hearing test (see below) and 1 observer was excluded because she did not pass the two practice trials in the SL condition (see below). The remaining 10 observers were 3 women and 7 men, with a mean age of 19.1 years (range from 18 to 25).

Blind Observers

Six blind observers (two women and four men) referred to us by the Braille Institute of Santa Barbara, served as observers and were paid \$40 each. Each had been blind from the first year of life. Because of technical problems, one observer had to be rerun and received \$80 for the second session. The observers' mean age was 47.7 years (range from 34 to 54). Two of these observers suffered from congenital glaucoma, two from retinitis pigmentosa, one from optic nerve damage, and one from retinal blastoma. Generally, observers had good travel skills. When asked to rate on a 5-point scale (1 = *very unsure*, 2 = *unsure*, 3 = *average*, 4 = *confident*, and 5 = *very confident*) how confidently they travel (a) at home, (b) on local streets, (c) on busy roads, (d) when crossing at traffic lights, (e) when crossing at unsignaled crossings, (f) in new environments, (g) when making detours around unknown hazards, (h) when shortcutting, and (i) when exploring away from a known route, five observers had mean ratings of 3.4 or higher and one observer had a mean rating of 2.1.

Experimental Setup, Stimuli, and Apparatus

The experiment took place in a flat grassy field without obstructions on the UCSB campus. As Figure 5 shows, there were 12 target locations that were arranged in a semicircle around an origin located at 0°, 0 cm. Because

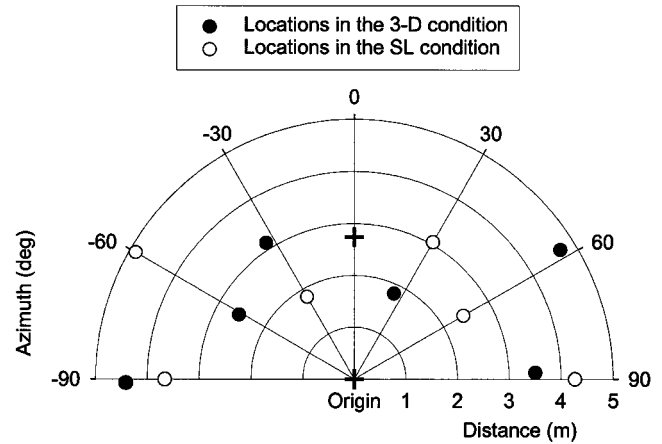


Figure 5. The spatial layout of the targets as a function of distance and azimuth from the origin. Filled and open circles indicate the locations of the targets in the 3-D and spatial language (SL) conditions, respectively. When observers walked indirectly to a target, they first walked forward to the turn point (indicated by +), turned, and then walked the rest of the way to the updated target location. Deg = degrees.

our verbal descriptions were expressed in feet, the unit of measurement most familiar to our observers, we used nominal target distances of 6, 8, 10, 12, 14, and 16 ft, but below we give the corresponding metric distances. Six locations were used in the 3-D sound condition and marked on the grassy field. Nominally, the targets had azimuths of -90° , -60° , -30° , 30° , 60° , and 90° and distances of 4.27, 2.44, 3.05, 1.83, 4.88, and 3.66 m, respectively, with positive azimuths representing angles to the right of the origin and negative azimuths representing angles to the left of the origin. Because of inaccuracy in constructing the target locations, the actual locations marked in the grassy field had azimuths of -91° , -61° , -33° , 25° , 58° , and 88° and distances of 4.41, 2.55, 3.13, 1.82, 4.70, and 3.51 m. For the SL condition, mirror-reversed azimuth-distance assignments were used, that is, target locations of azimuths -90° , -60° , -30° , 30° , 60° , and 90° and distances 3.66, 4.88, 1.83, 3.05, 2.44, and 4.27 m, respectively. In addition, the location at azimuth 0° , 2.74 m distant was used for the turn point when observers walked along an indirect path to the location.

The stimulus in the 3-D sound condition was recorded speech of a synthesized male voice saying "Speaker 3"; it was used in the study by Loomis et al. (1998). The utterance was repeatedly recorded on audiotape and played back in the experiment on a Realistic (Fort Worth, TX) CTR-66 cassette recorder with a loudspeaker 5.5 cm in diameter. The cassette recorder was positioned on a tripod at approximately ear level and oriented upward, so that the sound radiated uniformly in all directions within the horizontal plane of the speaker; the speaker was oriented upward rather than toward the observer to prevent unwanted variations in sound intensity resulting from errors in aiming the directional speaker toward the observer. The recorded speech string was played two times during a trial, for a total duration of 3 s. The speech sound level was 69 dB (A weighting¹) as measured from 1 m away using a Realistic sound-level meter. The stimulus in the SL condition was an utterance by the experimenter specifying the target location. The experimenter first spoke the azimuth in clock position (e.g., "1 o'clock") and then the distance in feet. Between trials and while walking, the observer wore hearing protectors (AOSafety [Indianapolis, IN], Model 1000 Earmuff). Their Noise Reduction Rating (a standard measure of attenuation used by the hearing protection industry) was reported by the manufacturer as 20.

¹ A sound meter using the A weighting gives different sound frequencies weights comparable to that of human hearing at middle values of intensity.

To measure the actual values of the marked locations in the 3-D sound condition and the stopping positions occupied by the observers, two SONIN (Scarsdale, NY) Combo PRO ultrasonic distance-measuring devices were used. Each device consists of an ultrasonic transmitter and an ultrasonic receiver. The ultrasonic receiver sends out an infrared pulse to the transmitter, which then sends back an ultrasonic pulse. When calibrated, the device is accurate to within 1 or 2 cm over a 10-m range. We positioned two receivers on either side of the origin of locomotion and one transmitter at the location to be measured. Using trigonometry, we converted the two distances into polar coordinates relative to the origin.² We confirmed that the distance measures provided by the ultrasonic devices were accurate and reliable. Across 10 observations under different conditions of temperature and humidity, for an actual distance of 1.22 m, as constructed with a metal tape measure, the measured mean distance was 1.21 m with a standard deviation of 0.01 m. The measured mean distance for 6.1 m was 6.09 m with a standard deviation of 0.03 m.

Procedure and Design

The observer's task was to listen to an auditory stimulus indicating a location in the grassy field and then to walk to the location. For the sighted, blindfolded observers, there were two sessions, which were separated by 4 to 11 days. In one session the auditory stimulus was the speech from the cassette recorder (3-D sound condition), and in the other session the stimulus was a verbal description of the location (SL condition). For the blind observers, the four experimental conditions were completed within one session lasting 90 min. For both groups of observers, order of stimulus condition was counterbalanced across observers.

Before the experiment started, observers were familiarized with the stimulus. In the 3-D condition, observers from both groups stood next to the cassette recorder with their eyes closed and reached out to touch the recorder while the speech signal was played several times. The purpose of this was to allow the observer to learn the source intensity for a known distance so as to be better able to later use sound intensity at the ears for judging source distance. In the SL condition, the experimenter named several clock positions, illustrating each by moving to the appropriate positions around the observer. In addition, observers were asked to walk a few distances expressed in feet and were corrected if they mistook number of steps for number of feet.

Thereafter, observers were screened for their ability to orient themselves toward the azimuth specified by the stimulus. Screening was performed with a blindfold to eliminate the use of vision. In the 3-D condition, observers completed a directional hearing test (during which they wore hearing protectors between trials). At the beginning of a trial, the observer's head and body were oriented to face 0° azimuth. Then the cassette recorder was placed at some other azimuth at a distance of 1.83 m from the observer. The speech signal "Speaker 3" was played twice and the observer was allowed to oscillate the head slightly ($\pm 10^\circ$) to facilitate localization. After the sound ended, the observer turned with the body and attempted to face the azimuth of the source. The facing azimuth was measured with a mechanical compass (Brunton [Riverton, WY] 8097 Eclipse, 1° graduation) worn on the body at waist height. Seven azimuths were tested in this order: 90°, 120°, 300°, 120°, 30°, 180°, and 330°. If the observer's mean absolute deviations exceeded 20°, the observer was dismissed from the experiment; two sighted observers failed the test and were dismissed. For the 10 sighted observers who passed, the mean absolute deviation from the true azimuths was 10.9° (range from 6.6° to 15.4°); for the 6 blind observers, the mean absolute deviation was 11.8° (range from 8.7° to 16.1°). In the SL condition, observers completed the same tasks, except that the experimenter was positioned at 0° azimuth and announced the respective azimuths as clock positions (30° was given as 1 o'clock). Seven azimuths were tested in this order: 120°, 330°, 210°, 30°, 300°, 90°, and 180°. If the observer's mean absolute deviations exceeded 20°, the observer was dismissed from the experiment; one sighted observer failed the

test and was dismissed. For the 10 sighted observers who passed this test, the mean absolute deviation from the true azimuths was 16.5° (range from 7.6° to 25.9°). The mean absolute deviation for the blind observers was 15.1° (range from 8.0° to 23.4°).

The experiment was conducted by an experimenter, who remained just behind the observer, and an assistant, who remained in the field to position the target tripod with the cassette recorder and the target unit of the measuring device. Throughout the experiment, the observer was blindfolded; in addition the observer wore hearing protectors except during stimulus presentations and during short breaks.

A trial started with positioning the observer at the origin, facing 0° azimuth. By tapping the observer on the shoulder, the experimenter signaled the observer to remove the hearing protectors and to listen to the stimulus. In the 3-D condition, the sound signal was played twice and the observer oscillated the head slightly to facilitate localization. In the SL condition, the experimenter, standing behind the observer, spoke the verbal description, expressed as a clock position and number of feet. The observer was allowed to take as much time for the localization process as desired. In the 3-D condition, he or she was asked to "get a clear image in mind of where the real loudspeaker is." In the SL condition, he or she was instructed to "imagine an object at that location, such as a loudspeaker" and to "get a clear image in mind of where the imagined object (loudspeaker) is."³ When ready, the observer put the hearing protectors back on. The experimenter then gave one of two instructions, which could be heard through the hearing protectors. Following the instruction "direct," the observer attempted to walk directly to the identified location (direct-walking condition). Following the instruction "forward," the observer began walking forward until the experimenter said "turn" (after 2.7 m). Then the observer turned and attempted to walk the rest of the way to the identified location (indirect-walking condition). As the observer walked, the experimenter followed close behind as a spotter to reduce the slight risk of falling, while the assistant silently removed the target tripod. After the observer had arrived at the stopping point, it was measured and the observer was guided back to the origin. Observers never received feedback about the accuracy of their performance.

Each of the sighted observers completed 24 trials in each of the two sessions (for the two stimulus conditions), with each of the six target locations being repeated two times for both the direct- and the indirect-walking condition. Each of the blind observers participated in only one session, which contained both stimulus conditions. In this case, each observer completed just 12 trials, with each of the six target locations being presented only once for both the direct and indirect walking conditions. For both groups, each observer received a different, randomized order of trials, with no repetitions of the same target-path combination on consecutive trials.

² We were able to accurately measure a given target location using trigonometric calculations based on the two distances provided by the ultrasonic devices. However, attempting to construct a target location using the same technique proved most difficult; consequently, we opted to construct the target locations in the field using a large protractor and tape measure, a technique that is less accurate. As a result, the constructed target locations, as subsequently measured by the ultrasonic devices, differed slightly from the nominal locations we desired.

³ A pilot study with several observers indicated that they had difficulty performing the walking tasks in the SL condition when they were asked to imagine the mere location. We attributed this finding to the difficulty in creating and externalizing a mental image of a point in space that has no other properties than its position. We therefore supplemented the instruction in the experiments by asking observers to imagine an object at that location. We believe that this enriched the representation and facilitated the understanding of the walking tasks in the SL condition. A formal study to investigate this issue is now underway.

To ensure that observers understood the task in the direct- and indirect-walking conditions, practice trials were completed before each session started. For the sighted observers, two practice trials involved a visual stimulus (the loudspeaker on the tripod) and were used to illustrate the difference between a direct path and an indirect path toward the target location. The loudspeaker was placed at 60° , 3.1 m. The observers looked at it and after memorizing its position, closed their eyes and walked either directly to the location of the loudspeaker or indirectly after having walked approximately 4.6 m forward. Two subsequent trials for the indirect condition used the sound signal and the verbal description, respectively, as the stimulus. The test locations were at -60° , 1.2 m and at 90° , 3.1 m, and observers had to approach them after having walked approximately 4.6 m forward. If in these two trials observers tracked backward after turning, we considered the task understood and the experiment was started. One person (out of 11) who did not track backward was excluded from this experiment.

A similar procedure was used with the blind observers. To illustrate and practice the direct- and indirect-walking tasks, a haptic target (one leg of the tripod) was used. The instruction in the SL condition to imagine an object at the target location was phrased in terms of a sound that emanates from the imagined location. There was no criterion of exclusion, and the task was practiced until the observers showed a clear understanding of it.

In summary, the experimental design consisted of the between-observer variable of visual status and within-observer variables of location (-90° through 90°), stimulus condition (3-D vs. SL), and path (direct vs. indirect).

Results

The observer's stopping points, as measured by ultrasonic devices, were converted into polar coordinates (azimuth and distance) relative to the origin. For some of the analyses involving the sighted observers, the stopping points of the two replications for a given target and path (direct vs. indirect) were averaged over the two polar coordinates to give each observer's centroid for that target and path; for the blind observers, we considered the single stopping point for each target and path to be a centroid, to simplify the exposition. We then computed the group centroid for each target and path by averaging across all of the individual centroids in each group. Figure 6 gives the individual and group centroids, computed in polar coordinates, for the sighted observers. This figure conveys the variability of the individual responses for each

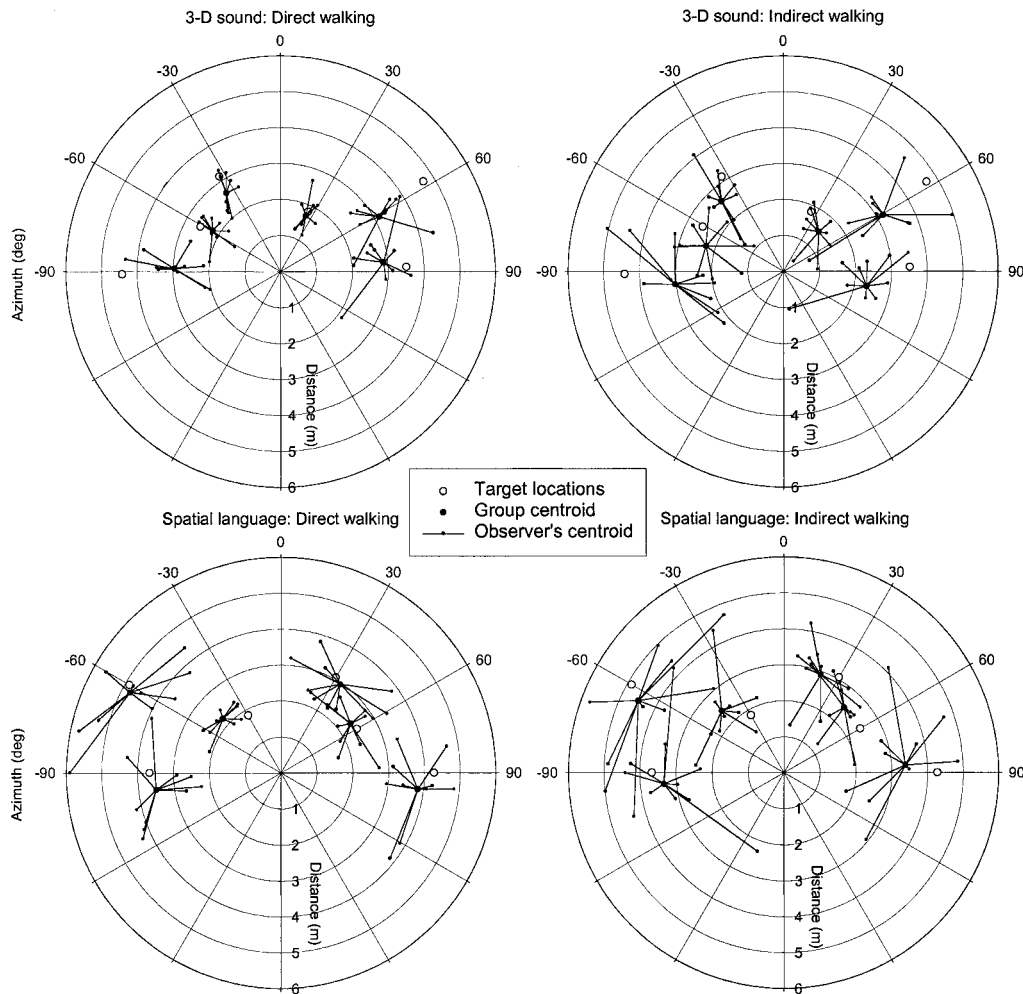


Figure 6. The results for the sighted observers as a function of distance and azimuth from the origin for the four combinations of modality and path. Filled small circles indicate the observer's centroids, averaged over the two stopping points for each target, whereas the filled large circles indicate the centroid of the group. Open circles indicate the target locations. Deg = degrees.

path and modality. Figure 7 gives the group centroids for each path, target, modality, and observer population. What is most apparent is the tendency for direct and indirect path centroids to be quite close to one another, especially for the sighted observers. (Because the order of stimulus modality did not significantly modulate the effects of interest, we take no further account of order in the analyses that follow.)

Biases in Stimulus Encoding

In the analysis of encoding bias, we began by computing the centroid of the direct and indirect stopping points and then computing the distance and azimuth of this centroid. Figures 8 and 9 give the mean distances and azimuths, respectively, averaged over observers, relative to the target values. These mean distances and azimuths are assumed to reflect the mean encoded positions; in the presentation of the model, we made the assumption that when the mean stopping point, averaged over direct and indirect paths, differs from the target, this deviation reflects encoding bias only, there being no updating bias common to the two paths. This assumption is supported by the results of the 3-D condition, for there are significant systematic errors in the stopping points that fit a particular pattern—the systematic errors in this condition are largely ones of distance (see Figure 8), with azimuth being responded to quite accurately (see Figure 9). Although it is possible that the direct and indirect paths could share a common bias, it would be highly coincidental that this common bias would result in an average stopping point that differs from the target in distance only. It is much more plausible to attribute this systematic error in distance solely to encoding error, specifically error in perceived auditory distance (Philbeck et al., 1997; Loomis et al., 1998).

Distance—sighted observers. An analysis of variance (ANOVA) carried out on walked distance, using the within-observer variables of stimulus condition (3-D vs. SL) and target distance, showed a main effect of distance, $F(5, 45) = 72.51$, $MSE = 1.44$, $p < .01$, a main effect of stimulus condition, $F(1, 9) = 7.22$, $MSE = 21.43$, $p < .05$, and an interaction between these variables, $F(5, 45) = 5.31$, $MSE = 4.21$, $p < .01$. In the SL condition, observers

were more accurate in traversing the required distances across the entire range (Figure 8). A linear function fit the SL data with an intercept of 0.76 m and a slope of 0.73 and accounted for 94% of the variance. In the 3-D condition, in turn, observers considerably undershot farther distances (see Figure 8), indicating the expected perceptual error in auditory distance perception, on the basis of prior studies (e.g., Loomis et al., 1998). A linear function fit the 3-D data with an intercept of 0.83 m and a slope of 0.51 and accounted for 95% of the variance.

Distance—blind observers. An ANOVA carried out on the walked distance, using the within-observer variables of stimulus condition (3-D vs. SL) and target distance, showed a main effect of distance, $F(5, 25) = 22.59$, $MSE = 3.22$, $p < .01$, and a Stimulus Condition \times Distance interaction, $F(5, 25) = 3.01$, $MSE = 3.18$, $p < .05$. In the SL condition, observers slightly overshoot the distances across the entire range in a consistent manner (see Figure 8). A linear function fit the SL data with an intercept of 0.79 m and a slope of 0.88 and accounted for 99.4% of the variance. In the 3-D condition, observers overshoot the two shortest distances and undershot the remaining farther distances (see Figure 4), again indicating perceptual error in auditory distance perception. A linear function fit the 3-D data with an intercept of 1.57 m and a slope of 0.45 and accounted for 89% of the variance.

Azimuth—sighted observers. An ANOVA conducted on response azimuth, using the within-observer variables of stimulus condition (3-D vs. SL) and target azimuth, yielded a main effect of azimuth only, $F(5, 45) = 383.21$, $MSE = 789.09$, $p < .01$. Performance in the 3-D condition was almost perfect for all azimuths (see Figure 9). A linear function fit the 3-D data with an intercept of 0.9° and a slope of 1.06 and accounted for 99.8% of the variance. In the SL condition, observers slightly overestimated azimuths in left hemisphere (see Figure 9). A linear function fit the SL data with an intercept of -6.3° and a slope of 1.03 and accounted for 99.3% of the variance.

Azimuth—blind observers. An ANOVA conducted on response azimuth, using the within-observer variables of stimulus condition (3-D vs. SL) and azimuth, yielded a main effect of

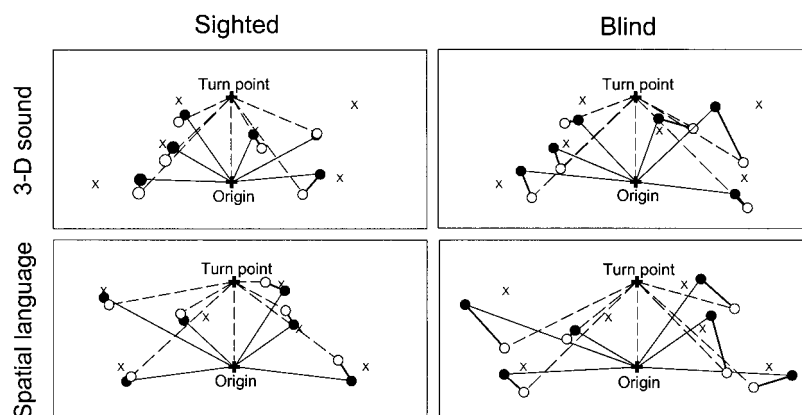


Figure 7. Group centroids as a function of path, stimulus modality, and observer group. The open circles are the centroids of the indirect path stopping points, the closed circles are the centroids of the direct path stopping points, and the \times s are the targets. Thin solid lines connect the origin with the centroids of the direct path, and dashed lines connect the turn point with the centroids of the indirect path. Direct and indirect path centroids to the same target are connected by thick solid lines.

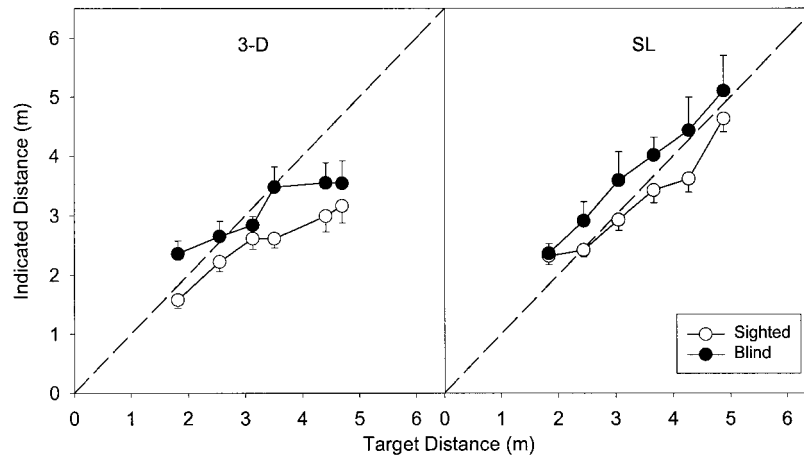


Figure 8. Mean indicated distance as a function of target distance, stimulus modality, and observer group. Error bars represent standard errors. SL = spatial language.

azimuth only, $F(5, 25) = 101.58$, $MSE = 3,044.32$, $p < .01$. Azimuthal judgments were fairly accurate in both the 3-D and SL conditions (see Figure 9). A linear function fit the 3-D data with an intercept of 0.1° and a slope of 1.12 and accounted for 99.2% of the variance. A linear function fit the SL data with an intercept of -3.2° and a slope of 1.19 and accounted for 97% of the variance.

Biases in Spatial Updating

Our primary interest in this research was whether the 3-D and SL conditions differed in terms of updating bias. As mentioned in discussion of the model, the vector difference between the centroids for the direct and indirect walking paths is an unbiased estimate of the vector difference of the updating biases for the two paths.⁴ Thus, we computed the vector difference (azimuth and distance) between the centroids for the direct and indirect paths as a function of observer, target, modality, and observer group. For each group of observers, a within-observers MANOVA was carried out on the vector difference between the indirect and direct centroids, with target and modality as independent variables.⁵ For the sighted observers, a main effect of modality was found, $F(2, 8) = 6.03$, $p < .05$, $\Lambda = 0.40$; there was no main effect of target nor a Target \times Modality interaction. The univariate tests showed that the modality effect was solely due to azimuth, $F(1, 9) = 12.29$, $MSE = 647.02$, $p < .01$. For the 3-D condition, there was a mean outward shift of 9.9° , and for the SL condition, there was a mean inward shift of 6.4° ; these opposing shifts are apparent in Figure 7. This means that there is a significant but small difference in the updating biases of 3-D sound and SL for the sighted observers. For the blind observers, the MANOVA revealed a Modality \times Target interaction, $F(10, 48) = 2.09$, $p < .05$, $\Lambda = 0.49$, but the univariate tests resulted in no significant results. The clear pattern of shifts obtained with the sighted observers (see Figure 7) was not at all evident with the blind observers.

Processing Noise

Ideally, we would be able to decompose the processing noise in the experiment into its encoding and updating components. Un-

fortunately, we cannot do so, because we have no knowledge of the encoding locations for individuals. Thus, the best we can do is to measure the combined encoding and updating noise for each modality, path (direct and indirect), observer group, and target. Even though our analysis does not focus on updating noise alone, we include it here for interest. The combined processing noise (between observers) is given by the mean Euclidean distance of the individual stopping-point centroids from the group centroid. To allow a comparison between the blind and sighted observers, this analysis of noise is based on only the first of the two responses by the sighted observers to each target, because the blind observers responded only once to each target along each path. Processing noise (between-observer variability) is plotted in Figure 10 as a function of modality, path, observer group, and target; target is represented by the distance of its group centroid from the origin. For each group of observers, a standard multiple regression analysis was performed on between-observer variability with modality, path, and distance from origin as independent variables. For the sighted observers, both path, $\beta = .44$, $t(20) = 4.30$, $p < .01$, and distance, $\beta = .82$, $t(20) = 7.15$, $p < .01$, were significant predictors; whereas modality was not ($\beta = -.02$, $t < 1$). For the blind,

⁴ If we were to instead compute the Euclidean distance between the centroids of the two paths for the different observers and then average these distances to obtain the mean distance for the group, this distance would be larger than the distance between the group centroids. To understand this, consider two observers, A and B. If the direct path stopping point of A is equal to the indirect path stopping point of B and vice versa, the centroids averaged over A and B would be identical for the two paths; in contrast, the separations between the direct and indirect stopping paths for the two observers would be equal and nonzero.

⁵ Because the targets used in the SL condition were mirror images of those in the 3-D condition, the comparison of the two modalities involved the symmetrically corresponding targets, matched in terms of distance from the origin. Also, because there was a tendency for the centroids for the indirect paths to be closer to the midline (0° azimuth line) than those for the direct path, for targets both to the left and right of the midline, we gave algebraic signs to the azimuth difference in the MANOVA so that shifts away from the midline were positive, both for left and right targets.

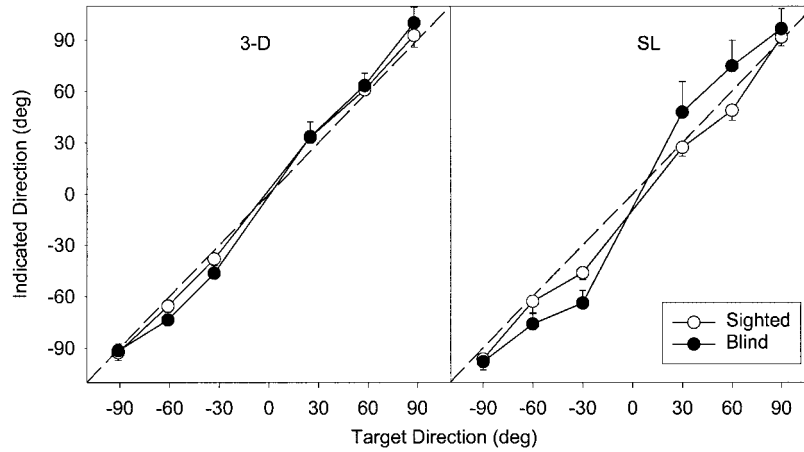


Figure 9. Mean indicated azimuth as a function of target azimuth, stimulus modality, and observer group. Error bars represent standard errors. SL = spatial language; deg = degrees.

all three variables were significant predictors: path, $\beta = .43$, $t(20) = 3.16$, $p < .01$, distance, $\beta = .54$, $t(20) = 3.67$, $p < .01$, and modality, $\beta = .31$, $t(20) = 2.11$, $p < .05$. We conclude that processing noise generally increases with distance from the origin, that the indirect path results in greater noise, and that the difference in processing noise between the two modalities is small.

Discussion

The most important finding of this study is the occurrence of spatial updating in the SL condition. This is readily apparent in the

small separations between the group centroids of the direct and indirect paths in Figure 7, especially for the sighted observers. It is not surprising that observers were able to walk directly to locations specified by SL, because this condition makes use of explicit directions of how to walk to the target. What is surprising is performance for the indirect path. The fact that the centroids for the indirect path are close to those for the direct path is strong evidence for spatial updating. Because the verbal description does not explicitly state how to arrive at the target, observers need to convert the verbal description into a spatial image and then update this spatial image while walking and turning.

Previous research has already demonstrated that observers are able to spatially update locations specified by 3-D sound with ease (Ashmead et al., 1995; Loomis et al., 1998; Speigle & Loomis, 1993). Besides providing further evidence of this, the 3-D condition of the current study, in conjunction with the SL condition, allows us to address the question of whether the spatial image created by SL is functionally equivalent to that created by 3-D sound. Using a model in which there is bias and noise for both encoding and updating, we were able to analyze whether the two modalities differed in terms of updating bias. Our analysis revealed a small but significant difference for the 10 sighted observers—for 3-D sound, the stopping points for the indirect paths were shifted slightly away from the midline relative to those for the direct paths; whereas, an opposite shift was observed for SL. However, there is no hint of such opposing shifts for the 6 blind observers. Accordingly, we conclude that over the two groups, the difference in updating bias between 3-D sound and SL is minimal. Because of this near equivalence of updating bias in the two conditions, we conclude that once an internal representation of a location has been determined, subsequent updating performance is very nearly independent of whether 3-D sound or SL was used to specify the representation.

An important subsidiary issue addressed by this study is whether individuals with essentially no visual experience are able to perform spatial updating of locations specified by 3-D sound and SL. Given the ubiquity of 3-D sound and its importance for blind individuals, it would be very surprising if they could not perform spatial updating of 3-D sound. The experiment shows not only that

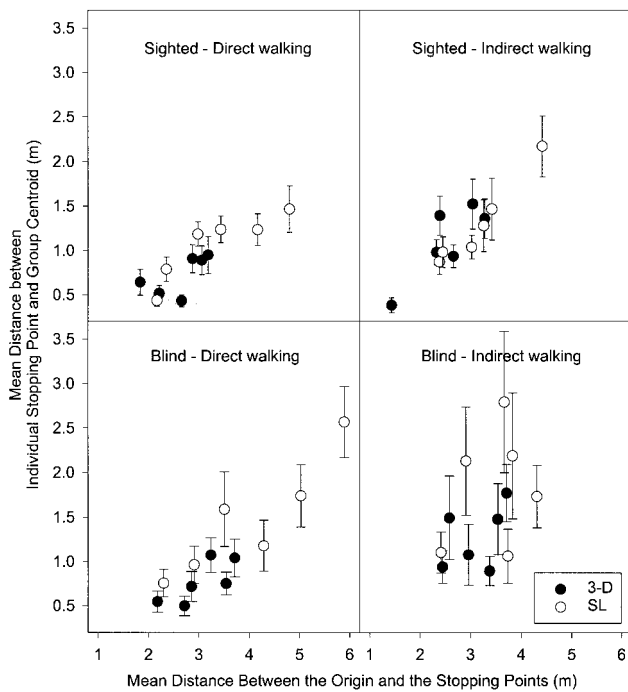


Figure 10. Processing noise (mean distance between individual stopping point and group centroid) as a stimulus modality, path, observer group, and the distance between the origin and the group centroid. Error bars represent standard errors. SL = spatial language.

the 6 blind observers were indeed capable of updating locations specified by 3-D sound but that they were also able to do so with locations specified by language. This indicates that visual experience is not a necessary condition for developing skill in spatially updating locations specified by these two modalities. This is not to say, however, that visual experience is unimportant for updating ability. Rieser, Guth, and Hill (1986) have found that the early blind, on average, were considerably poorer than the sighted and late blind in performing a spatial updating task that relied on haptic and proprioceptive input for forming a spatial representation. In addition, the larger literature on spatial abilities suggests that the early blind, on average, perform more poorly than the other groups (Millar, 1994; Thinus-Blanc & Gaunet, 1997), although our own prior research has found little difference between the groups (Klatzky, Golledge, Loomis, Cicinelli, & Pellegrino, 1995; Loomis et al., 1993).

A related issue concerns auditory distance perception by blind and sighted observers. Believing as we do that the distance functions in Figure 8 for 3-D sound reflect auditory perceived distance (Loomis et al., 1998), it is interesting that both the blind and sighted in Experiment 1 yielded best-fitting linear distance functions with slopes of 0.45 and 0.51, respectively. These values are close to those of all of the functions measured under similar conditions in other studies using open-loop walking (for a review, see Loomis et al., 1999).⁶ The fact that the blind observers, all of whom were skilled travelers, showed the same two-fold compression of range in perceived distance that the sighted observers showed suggests that the perceptual functions do not change significantly with experience. This is interesting, for it might have been expected that the blind observers would have used their extensive experience in walking to sound sources to calibrate their walking behavior and thus compensate for the systematic errors in perceiving auditory distance.

Finally, our conclusions about spatial updating are relevant to the design of navigation systems for the visually impaired. For well over a decade, we have been working on the development of such a system (Golledge, Klatzky, Loomis, Speigle, & Tietz, 1998; Loomis, 1985; Loomis, Golledge, & Klatzky, 2001). From the outset, our preferred design has been to inform the traveler of environmental locations of interest using 3-D sound, presented by way of earphones and a virtual acoustic display. SL generated by speech synthesizer is the natural alternative. In the navigation aids being developed by other researchers, SL is used to indicate the traveler's current location in terms of streets and cardinal directions. However, SL can also be used to convey information about the locations of off-route landmarks and other points of interest ("A public phone is at 3 o'clock and 30 ft"). The current study indicates that SL and 3-D sound are comparable in effectiveness in allowing a user to spatially update locations in the environment.

⁶ The fact that in the present study the intercepts for the best fitting linear functions were different for blind and sighted observers is of no great consequence, because previous research found large variations in intercept whereas the slopes were all close to 0.5 (see Loomis et al., 1999).

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. *Perception & Psychophysics*, 59, 404–418.
- Ashmead, D. H., DeFord, L. D., & Northington, A. (1995). Contribution of listeners' approaching motion to auditory distance perception. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 239–256.
- Barber, P. O., & Lederman, S. J. (1988). Encoding direction in manipulatory space and the role of visual experience. *Journal of Visual Impairment and Blindness*, 82, 99–106.
- Böök, A., & Gärling, T. (1981). Maintenance of orientation during locomotion in unfamiliar environments. *Journal of Experimental Psychology: Human Perception and Performance*, 7, 995–1006.
- Bosco, A., Filomena, S., Sardone, L., Scalisi, T. G., & Longoni, A. M. (1996). Spatial models derived from verbal descriptions of fictitious environments: The influence of study time and the individual differences in visuo-spatial ability. *Psychologische Beiträge*, 38, 451–466.
- Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 483–500.
- Farrell, M. J., & Robertson, I. H. (1997). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 227–233.
- Farrell, M. J., & Thomson, J. A. (1999). On-line updating of spatial information during locomotion without vision. *Journal of Motor Behavior*, 3, 39–53.
- Ferguson, E. L., & Hegarty, M. (1994). Properties of cognitive maps constructed from text. *Memory & Cognition*, 22, 455–473.
- Fukushima, S. S., Loomis, J. M., & Da Silva, J. A. (1997). Visual perception of egocentric distance as assessed by triangulation. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 86–100.
- Golledge, R. G., Klatzky, R. L., Loomis, J. M., Speigle, J., & Tietz, J. (1998). A geographic information system for a GPS based personal guidance system. *International Journal of Geographical Information Science*, 12, 727–749.
- Hollins, M., & Kelley, E. K. (1988). Spatial updating in blind and sighted people. *Perception & Psychophysics*, 43, 380–388.
- Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology*, 11, 375–394.
- Klatzky, R. L., Golledge, R. G., Loomis, J. M., Cicinelli, J. G., & Pellegrino, J. W. (1995). Performance of blind and sighted in spatial tasks. *Journal of Visual Impairment & Blindness*, 89, 70–82.
- Loarer E., & Savoyant A. (1991). Visual imagery in locomotor movement without vision. In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 35–46). The Hague: Elsevier.
- Loomis, J. M. (1985). *Digital map and navigation system for the visually impaired*. Unpublished manuscript, University of California, Santa Barbara.
- Loomis, J. M., Da Silva, J. A., Fujita, N., & Fukushima, S. S. (1992). Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 906–921.
- Loomis, J. M., Golledge, R. G., & Klatzky, R. L. (2001). GPS-based navigation systems for the visually impaired. In W. Barfield & T. Caudell (Eds.), *Fundamentals of wearable computers and augmented reality* (pp. 429–446). Mahwah, NJ: Erlbaum.
- Loomis, J. M., Klatzky, R. L., & Golledge, R. G. (1999). Auditory distance perception in real, virtual, and mixed environments. In Y. Ohta & H. Tamura (Eds.), *Mixed reality: Merging real and virtual worlds* (pp. 201–214). Tokyo: Ohmsha.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., & Fry, P. A. (1993). Nonvisual navigation by blind and sighted: Assessment of path integration ability. *Journal of Experimental Psychology: General*, 122, 73–91.
- Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Way-*

- finding behavior: *Cognitive mapping and other spatial processes* (p. 147). Baltimore: Johns Hopkins University Press.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, *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, NJ: Erlbaum.
- Millar, S. (1994). *Understanding and representing space. Theory and evidence from studies with blind and sighted children*. Oxford, England: Clarendon Press.
- Perrig, W., & Kintsch, W. (1985). Propositional and situational representations of text. *Journal of Memory and Language*, *24*, 503–518.
- Philbeck, J. W., Klatzky, R. L., Behrmann, M., Loomis, J. M., & Goodridge, J. (2001). Active control of locomotion facilitates nonvisual navigation. *Journal of Experimental Psychology: Human Perception and Performance*, *27*, 141–153.
- Philbeck, J. W., Loomis, J. M., & Beall, A. C. (1997). Visually perceived location is an invariant in the control of action. *Perception & Psychophysics*, *59*, 601–612.
- Presson, C. C., & Montello, D. R. (1994). Updating after rotational and translational body movements: Coordinate structure of perspective space. *Perception*, *23*, 1447–1455.
- Rieser, J. J. (1989). Access to knowledge of spatial structure at novel points of observation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1157–1165.
- 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.
- Rieser, J. J., Frymire, M., & Berry, D. (1997, November). *Geometrical constraints on imagery and action when walking without vision*. Paper presented at the 38th Annual Meeting of the Psychonomic Society, Philadelphia.
- Rieser, J. J., Guth, D. A., & Hill, E. W. (1986). Sensitivity to perspective structure while walking without vision. *Perception*, *15*, 173–188.
- Rieser, J. J., & Rider, E. A. (1991). Young children's spatial orientation with respect to multiple targets when walking without vision. *Developmental Psychology*, *27*, 97–107.
- Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science*, *9*, 315–320.
- Speigle, J. M., & Loomis, J. M. (1993). Auditory distance perception by translating observers. *Proceedings of IEEE Symposium on Research Frontiers in Virtual Reality* (pp. 92–99). San Jose, CA.
- Taylor, H., & Tversky, B. (1992). Spatial mental models derived from survey and route descriptions. *Journal of Memory and Language*, *31*, 261–282.
- Thinus-Blanc, C., & Gaunet, F. (1997). Representation of space in blind persons: Vision as a spatial sense? *Psychological Bulletin*, *121*, 20–41.
- Thomson, J. A. (1983). Is continuous visual monitoring necessary in visually guided locomotion? *Journal of Experimental Psychology: Human Perception and Performance*, *9*, 427–443.
- Wilson, P. N., Tlauka, M., & Wildbur, D. (1999). Orientation specificity occurs in both small- and large-scale imagined routes presented as verbal descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 664–679.

Received October 11, 2000

Revision received July 9, 2001

Accepted July 12, 2001 ■

ORDER FORM

Start my 2002 subscription to *Journal of Experimental Psychology: Learning, Memory, and Cognition!*

ISSN: 0278-7393

_____ \$125.00, APA MEMBER/AFFILIATE _____
 _____ \$251.00, INDIVIDUAL NONMEMBER _____
 _____ \$554.00, INSTITUTION _____
 In DC add 5.75% / In MD add 5% sales tax _____
TOTAL AMOUNT ENCLOSED \$ _____

Subscription orders must be prepaid. (Subscriptions are on a calendar basis only.) Allow 4-6 weeks for delivery of the first issue. Call for international subscription rates.



AMERICAN
PSYCHOLOGICAL
ASSOCIATION

SEND THIS ORDER FORM TO:

American Psychological Association
Subscriptions
750 First Street, NE
Washington, DC 20002-4242

Or call (800) 374-2721, fax (202) 336-5568.
TDD/TTY (202)336-6123.
Email: subscriptions@apa.org

Send me a FREE Sample Issue

Check Enclosed (make payable to APA)

Charge my: VISA MasterCard American Express

Cardholder Name _____

Card No. _____ Exp. date _____

Signature (Required for Charge)

BILLING ADDRESS: _____

City _____ State _____ Zip _____

Daytime Phone _____

SHIP TO:

Name _____

Address _____

City _____ State _____ Zip _____

APA Member # _____

XLMA12