Body-based senses enhance knowledge of directions in large-scale environments

DAVID WALLER
Miami University, Oxford, Ohio

JACK M. LOOMIS
University of California, Santa Barbara, California

and

DANIEL B. M. HAUN
Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands
F. C. Donders Centre for Cognitive Neuroimaging, Nijmegen, The Netherlands

Previous research has shown that inertial cues resulting from passive transport through a large environment do not necessarily facilitate acquiring knowledge about its layout. Here we examine whether the additional body-based cues that result from active movement facilitate the acquisition of spatial knowledge. Three groups of participants learned locations along an 840-m route. One group walked the route during learning, allowing access to body-based cues (i.e., vestibular, proprioceptive, and efferent information). Another group learned by sitting in the laboratory, watching videos made from the first group. A third group watched a specially made video that minimized potentially confusing head-on-trunk rotations of the viewpoint. All groups were tested on their knowledge of directions in the environment as well as on its configural properties. Having access to body-based information reduced pointing error by a small but significant amount. Regardless of the sensory information available during learning, participants exhibited strikingly common biases.

People often need to know where important places are in environments with which they have little experience. For example, imagine visiting a university for the first time by taking a walking tour of the campus. Midway through the trip, the tour guide points to a building and says "Remember where this building is, because you are scheduled to meet someone there later." In such a situation, you must quickly learn the layout of a large, previously unknown environment in order to arrive at your meeting on time, after the tour is over. Fortunately, when people acquire knowledge about environments by traveling through them, several sources of sensory information are available. In this article, we distinguish two broad types of spatial information that are available as people acquire knowledge about an environment: *vision-based* and *body-based* information. In most environments, the visual system provides people with information about the layout of their immediate environment as well as information (optic flow) about current changes in heading and position. Body-based information derives from multiple sources, each of which involves the internal sensing of one's movements in space. These sources include information from the *vestibular* system, which senses linear and angular accelerations of the head, *proprioception*, which provides information about movement of one's body parts, and, in the case of active movement, *efference copy*, which represents the commands to the musculature issued by the central nervous system. In this article, we are concerned with the contribution of these body-based sensory modalities in spatial learning. Specifically, we are interested in the degree to which body-based information facilitates the visual system as we acquire knowledge of large-scale natural environments.

Although humans rely predominantly on vision to perform a wide variety of important behaviors, many results in the psychological literature have highlighted the importance of body-based senses in spatial perception, in spatial knowledge acquisition, and in the performance of rudimentary navigation tasks. Several studies have shown, for example, that information acquired from the vestibular system alone is sufficient for allowing people to maintain their orientation and to determine their position when they are passively transported along simple trajectories while blindfolded (Beritoff, 1965; Israel, Chapuis, Glasauer, Charade, & Berthoz, 1993; Potegal, 1982). Other studies have shown that active bodily movement in an environment, which provides proprioceptive and efferent information in addition to vestibular information, enhances the accuracy and precision of spatial perceptions relative to

those derived from vestibular information alone (Jürgens, Boß, & Becker, 1999; Ohmi, 1996; Sholl, 1989; Telford, Howard, & Ohmi, 1995; Yardley & Higgins, 1998). Thus, it is now commonly believed that as people move through space, even when optic flow is available, additional information from the body senses is necessary for people to maintain an accurate awareness of their orientation (Chance, Gaunet, Beall, & Loomis, 1998; Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Peruch, May, & Wartenberg, 1997; but see Kearns, Warren, Duchon, & Tarr, 2002; Riecke, Van Veen, & Bülthoff, 2002).

In all of the studies cited above, researchers have examined the contributions of body-based sensory information to spatial perception and learning in carefully controlled laboratory experiments. Although these experiments have permitted firm conclusions about the comparative effects of various combinations of available sensory information, the environments used have typically been extremely sparse, and the participants' task has generally been quite simple. As a result, the degree to which these findings might apply more generally to larger, more natural spaces and to more complex tasks is not clear. However, another area of spatial cognition research has examined people's ability to learn information about complex, large-scale natural spaces from purely visual sources such as photographs (Allen, Siegel, & Rosinski, 1978; Hock & Schmelzkopf, 1980), maps (Landau, 1986), scale models (Allen, Kirasic, Dobson, Long, & Beck, 1996), motion pictures (Goldin & Thorndyke, 1982), and desktop computer simulations (Rossano, West, Robertson, Wayne, & Chase, 1999; Waller, Hunt, & Knapp, 1998). Much of this literature has shown that people can readily acquire reasonably accurate spatial knowledge from such sources, despite the fact that the media through which the environment is learned do not provide all of the body-based sensory information that would be available from direct experience. Although these studies offer evidence about people's ability to acquire spatial knowledge in the absence of body-based sensory information, they are relatively uninformative about whether body-based information would facilitate such acquisition if it were available during learning.

In an effort to examine the facilitative effect of body-based information on the acquisition of knowledge of large-scale, natural environments, Waller, Loomis, and Steck (2003) asked participants to learn the layout of several landmarks from a 1,600-m car trip through a previously unknown environment. All of the participants received similar visual information about the trip, and the experimenters varied the amount of valid body-based (inertial, primarily vestibular) information available by testing three separate groups of participants. Waller et al. (2003) found that whether people had access to valid, invalid, or no body-based inertial cues did not affect their ability to learn the environment. The researchers concluded that many of the prior laboratory results illustrating the importance of body-based information were limited to smaller scale spaces.

It is important to note that Waller et al.'s (2003) results concerned learning that resulted from *passive* transport through an environment. Perhaps the proprioceptive and efferent components of spatial learning that are available during *active* navigation are more important for maintaining orientation and learning the layout of the space than the (primarily) vestibular information studied by Waller et al. (2003). The present experiment was conducted to examine this possibility. We followed the general methods of Waller et al.'s (2003) study by comparing groups of participants who experienced an environment visually, but were allowed different amounts of body-based information during learning. The participants were instructed to learn the relative locations of five landmarks along a route through a previously unknown environment, and all groups viewed this environment in a head-mounted display (HMD). One group of participants (*walk*) wore the HMD (with a small video camera attached) while walking through the environment. The camera provided on-line real-time visual information about the participant's trip. Video recordings of these trips were subsequently shown to a second group of participants (*sit*), who watched them while sitting stationary in the laboratory. After learning, all participants were measured on their knowledge of the directions between the landmarks. If body-based sensory information is important for acquiring knowledge about large spaces, then we would expect the group who physically moved through the environment to perform with greater accuracy than the group that had only vision available during learning.

We also tested a third group of participants to examine more specifically the components of visual information that influence the acquisition of environmental knowledge. In this condition, we attempted to control the optic flow that derives from head rotations (relative to the body). It is known that accurate perception of one's orientation in the dark requires vestibular information to be integrated with neck proprioception (Blouin et al., 1995). It is thus possible that visual information that changes as a result of head rotation interferes with spatial learning if it is not accompanied by concomitant proprioceptive information that signals head movement. So that we could examine this possibility, a third group of participants watched a specially prepared video of the route while sitting still in the laboratory. The video shown to this group (*smooth*) minimized the amount of simulated head rotation by generally facing in the direction of travel. If visual flow based on head-on-trunk rotations is difficult to interpret in the absence of neck proprioception, then we would expect participants in the *smooth* group (who were not distracted by head movements) to form a more accurate mental representation of the environment than would participants in the *sit* group (who had to watch head movements that they were not controlling).

METHOD

Participants

Seventy-two undergraduate students (36 men and 36 women) at the University of California in Santa Barbara participated in the experiment in order to satisfy a requirement for their introductory psychology course. After testing, all participants reported having had no prior familiarity with the learned environment.
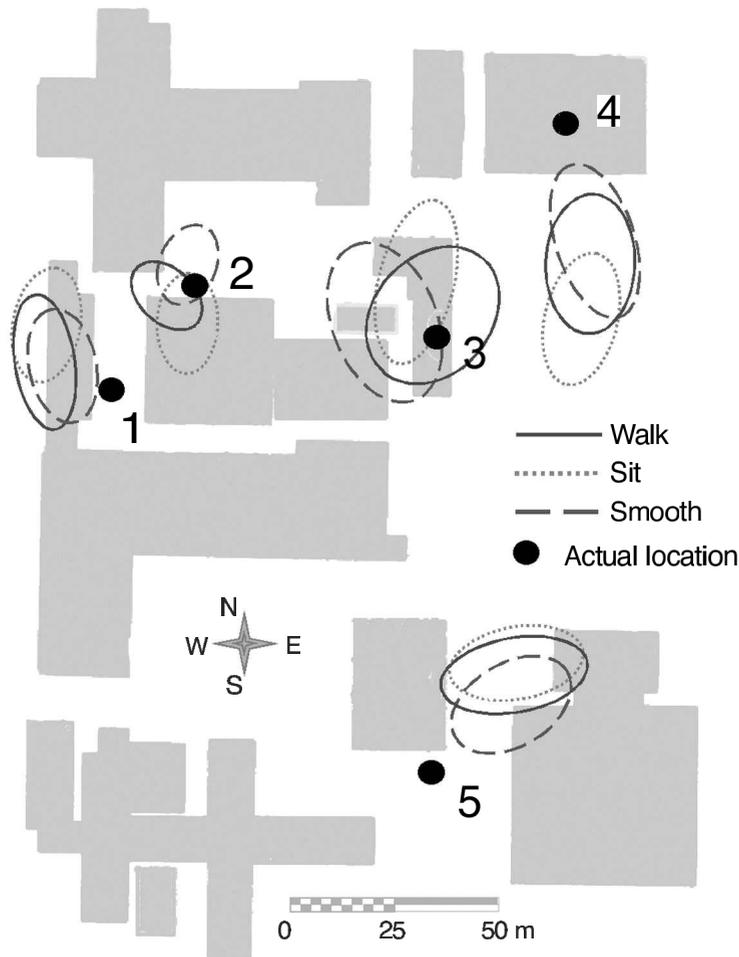
Figure 1. Map of the environment, showing the locations on which the participants were tested. Three groups of participants learned the environment with varying amounts of visual and body-based information available and were subsequently asked to construct a map of the five locations. Shown also are 95% confidence ellipses for the participants' map construction in each group. These ellipses were created by averaging the *x*- and *y*-coordinates of each participant's map placements in each condition with Procrustes superimposition (Dryden & Mardia, 1998) and summarizing the resultant point clouds with ellipses that cover the population centroid with 95% confidence (see Batschelet, 1981).

Environment and Materials

The environment to be learned consisted of an 840-m route through a secluded section of the UCSB campus (see Figure 1). Five locations along this route (a courtyard, a fence, a water fountain, a restroom, and a set of stairs) were selected as landmarks for the participants to learn. None of the five locations was visible from any of the others.

All participants learned about the environment by viewing it in a Virtual Research V8 HMD. The HMD provided identical images to both eyes at a resolution of 640 × 480 and a 48° horizontal field of view. For participants in the *walk* group, an Opcom CVC-514C color board camera (measuring 3.81 × 4.45 × 2.54 cm and weighing 51.03 g) was securely mounted to the front of the HMD. The camera provided color video images, with a 50° horizontal field of view. The images from this camera were converted to VGA and then split, running both to the participant's HMD and to a video recorder that

made a tape of each trip. The video equipment, power supply, and signal converters were carried by the experimenter in a framed backpack that was tethered to the participant's HMD with a 1.6-m cable. (see Figure 2).

We made an additional video of the environment to show to the participants in the *smooth* condition. To make this video, we used a JVC GR-AX420 camcorder with a 61° horizontal field of view. We minimized jitter, as well as any unnecessary rotations, by attaching two 1.1-kg counterweights to steady the camera and dampen the slight momentary movements that typically arise from the cameraperson's gait. During taping, the camera generally faced in the direction of travel and panned only while stopped at each of the five learning locations.

The participants' knowledge of the environment was tested in the laboratory in computer simulations of each of the five learned locations (see Waller, Beall, & Loomis, 2004). These simulations were created from 360° panoramic photographs of the locations that were

Figure 2. Photograph of the experimenter and participant in the *walk* condition during learning. The experimenter's backpack contained video recording equipment and a power supply. It was connected to a head-mounted display worn by the participant. A small camera, mounted to the front of the display, provided on-line visual information to the participant.

applied to the inside of a computer-modeled cylinder. The user's simulated viewpoint was placed at the center of the cylinder. The participants viewed the panoramas in a Virtual Research V8 HMD, on which was mounted an Intersense IS-300 inertial tracking sensor. The tracker updated the orientation of the visual image (relative to the user's head) and was also used to record facing directions when participants pointed to the various targets. The computer rendered the simulations using a Pentium III chipset and an NVIDIA GeForce 2MX graphics accelerator, updating the graphics and display at 72 Hz. Randomization and presentation of the stimuli as well as the collection of the pointing estimates were controlled through a scripting facility in the Python programming language, supplemented with a utility module written by Andrew Beall specifically for virtual environment applications.

Procedure

The participants were run individually through the experiment. After a brief practice session with the pointing task (described below), the participants learned the environment in one of three ways. People in the *walk* condition were asked to sit in a wheelchair and don the HMD. The HMD was turned off (thus serving as a blindfold), and the experimenter then wheeled the participant out of the lab, along a circuitous path to the start of the route. At the start, the participant's HMD and camera were turned on. The participant then stood up and was led by the experimenter along the route. The experimenter guided the participant by the forearm, saying only "turn here" when approaching a turn in the route. At each of the five learning locations, the experimenter announced the location and gave the participant a few moments to look around. At the end of the trip, the participant's HMD was turned off, and he or she was wheeled along a circuitous route back to the laboratory for testing.

Participants in the *sit* and *smooth* conditions learned the environment by sitting stationary in the lab. Videos from gender-matched *walk* participants were shown in the HMD to participants in the *sit* condition; participants in the *smooth* condition all watched our specially made video in the HMD. During learning (as in the *walk* condition), the experimenter said "turn here" when the video depicted a turn in the route, and announced each of the locations as it occurred.

All participants were tested by pointing to and from all possible pairs of the five locations. These pairs were given in five blocks of four—one block for each location—resulting in 20 direction estimates. For each pointing judgment, the participants were shown the appropriate panorama in the HMD and told to pretend that that they were at the depicted location. They were then asked to face each of the other locations (targets), and their facing directions were written to an external computer file for later analysis. The presentation order of the five simulated locations as well as the ordering of the four targets at each location were randomized for each participant.

Finally, the participants constructed a map of the environment by placing small cardboard pieces representing the five locations on a blank sheet of grid paper. The *x*- and *y*-coordinates of the participants' map placements were recorded for later analysis.

RESULTS

We examined pointing accuracy by computing, for each participant, the unsigned (absolute) error in the pointing task, averaged over the 20 trials in the experiment. By this measure, the participants who learned the environment by physically walking through it were able to point to targets with significantly greater accuracy (mean absolute error =

Signed pointing error
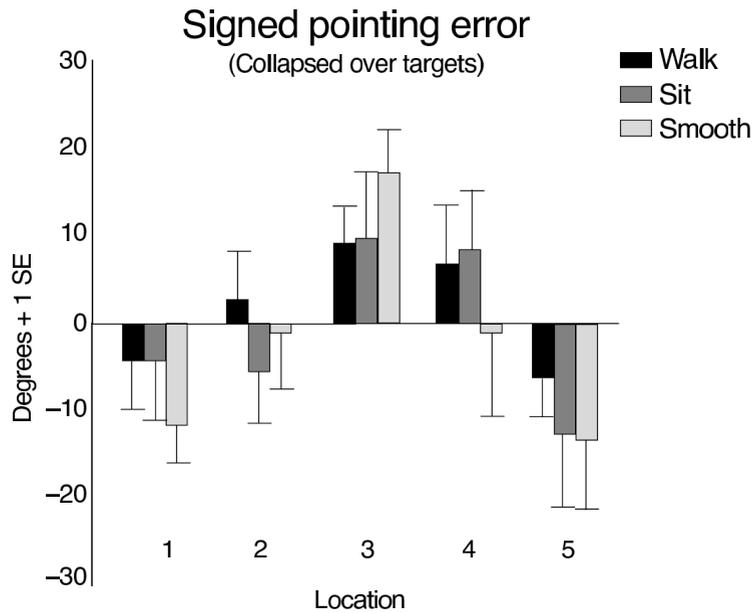### (Collapsed over targets)

Figure 3. Bias (mean signed pointing error) for each group at each of the five locations. Participants showed significant pointing biases at Locations 1, 3, and 5, yet there were no significant differences in bias between the three experimental groups at any location. Error bars represent one standard error of the mean.

$36.3°$, $SD = 17.6$) than were the participants in either of the other groups. In general, the participants in the *smooth* group (mean absolute error = $46.3°$, $SD = 16.4$) outperformed those in the *sit* group (mean absolute error = $50.8°$, $SD = 12.3$), although this difference was small and nonsignificant [$t(66) = 1.01$, $p = .317$]. These effects were tested in a 3 (condition) $\times$ 2 (gender) between-subjects analysis of variance (ANOVA) with absolute pointing error as a dependent variable.[1] The only significant effect from this model was that of condition [$F(2,66) = 5.44$, $p < .01$]. Planned contrasts showed that most of the effect of condition was accounted for by significant differences between the *walk* and *sit* groups [$t(66) = 3.22$, $p < .01$] and between the *walk* and *smooth* groups [$t(66) = 2.22$, $p = .03$].

The preceding analysis of unsigned error is useful for drawing conclusions about overall performance accuracy; however, it does not account for possible systematic sources of error that might underlie pointing performance. In particular, it is possible that participants could have acquired an accurate knowledge of the environment's layout, yet still have had high absolute pointing error if, for example, they were disoriented at some of the testing locations. In such a case, pointing responses would show a systematic left or right bias. To examine this possibility, we computed, for each participant, the bias at each location as the mean direction (see Batschelet, 1981) of the participant's signed pointing error, averaged over the four targets at the location.[2] Figure 3 illustrates mean signed error across testing locations for the three experimental groups. Locations 1, 3, and 5 tended to be associated with strong left (mean signed error

less than zero) or right (mean signed error greater than zero) pointing biases; however, these biases did not differ among the experimental groups. These observations were confirmed in a 3 (condition) $\times$ 2 (gender) $\times$ 5 (location) ANOVA with repeated measures on the last factor. The only significant effect in this model was that of location [$F(4,63) = 9.96$, $p < .01$]. There was no significant effect of condition [$F(2,66) = 0.87$, $p = .422$], nor was the simple main effect of condition significant at any of the locations.

The maps that the participants constructed also showed common biases, and little difference among the three experimental groups. Accuracy of map placements was measured as the bidimensional correlation between the actual locations and the participants' placements. The bidimensional correlation serves as an index of similarity of two 2-D configurations, independent of differences in the position, scale, and orientation of the two configurations (see Tobler, 1994). The participants in the *smooth* condition produced the most accurate maps of the environment (mean bidimensional correlation = $.66$, $SD = .29$), followed by the *walk* group (mean bidimensional correlation = $.65$, $SD = .26$) and the *sit* group (mean bidimensional correlation = $.58$, $SD = .25$). These group differences in map accuracy were not significant [$F(2,69) = 1.43$, $p = .247$]. The participants' map placements showed extremely similar biases. Figure 1 summarizes the participants' map placements in each of the three experimental groups. Most notable in this figure are the consistent biases among all participants, regardless of the means by which the environment was learned. For example, each group significantly underestimated the

distance between Locations 4 and 5, and each group underestimated the turning angle between Locations 3 and 4.

DISCUSSION

People who had access to vestibular, proprioceptive, and efferent information while acquiring knowledge about a large-scale complex environment were able to point subsequently to and from targets in the environment with approximately 10° to 15° more accuracy than people who learned the environment through purely visual means. Although this is not a particularly large difference, it is highly significant, and it indicates that body-based sensory information is used during (and facilitates) active learning of large, natural environments. Moreover, the performance of the participants in the *smooth* condition relative to those in the *sit* condition suggests that visual information based on head-on-trunk movements that are not coupled with corresponding neck proprioception may interfere with the acquisition of spatial knowledge.

These results offer an important contrast to those of Waller et al. (2003), who found no evidence of body-based (primarily vestibular) information on spatial knowledge acquisition during passive transport through a large environment. Most of the important aspects of the procedures and measures in the present study were replicated from those of Waller et al.'s (2003) experiment. One critical difference, however, was that in the present experiment, participants learned the environment by actively walking through it, instead of by being passively driven through it. The similarity of the methods between the present study and that of Waller et al. (2003), as well as the contrasting results for pointing error, provides some evidence that the effect of body-based information on spatial knowledge acquisition derives primarily from the proprioceptive and efferent information available as a result of active movement through the environment.

What is perhaps most striking with our results is the relatively small magnitude of the effect of body-based information, its failure to exert an influence on other measures of spatial knowledge, and the remarkable consistency of the pointing and mapping biases exhibited by all participants, regardless of their experimental condition. For example, Figure 3 makes clear that the structure of the environment had a much stronger impact on people's pointing errors than did the presence or absence of body-based information. The lack of group differences in map construction accuracy, as well as the high degree of similarity in the shapes of the maps constructed, indicates that the effect of body-based information on developing complex configural knowledge of spatial layout (as opposed to knowledge of self-to-object relations) may be minimal.

Our findings have important ramifications for many current applications of computer-simulated ("virtual") environments. Computer simulations of large-scale environments offer investigators an opportunity to conduct basic research on human wayfinding and navigation as well as to train people on tasks requiring knowledge of geographic spaces. In general, however, interfaces with these systems—particularly "desktop" systems—do not convey all of the body-based information about self-motion that would normally arise from users' movement through these spaces. The relatively small facilitative effect of body-based information in our experiment suggests that virtual environments that do not incorporate physical movement into their interface may be nearly as useful for teaching people about large-scale spaces as those that do. On the other hand, the high statistical significance of this small effect implies that these applications cannot fully substitute for active learning.

REFERENCES

ALLEN, G. L., KIRASIC, K. C., DOBSON, S. H., LONG, R. G., & BECK, S. (1996). Predicting environmental learning from spatial abilities: An indirect route. *Intelligence*, **22**, 327-355.

ALLEN, G. L., SIEGEL, A. W., & ROSINSKI, R. R. (1978). The role of perceptual context in structuring spatial knowledge. *Journal of Experimental Psychology: Human Learning & Memory*, **4**, 617-630.

BATSCHELET, E. (1981). *Circular statistics in biology.* London: Academic Press.

BERITOFF, J. S. (1965). *Neural mechanisms of higher vertebrate behavior.* Boston: Little, Brown.

BLOUIN, J., VERCHER, J. L., GAUTHIER, G. M., PAILLARD, J., BARD, C., & LAMARRE, Y. (1995). Perception of passive whole-body rotations in the absence of neck and body proprioception. *Journal of Neurophysiology*, **74**, 2216-2219.

CHANCE, S. S., GAUNET, F., BEALL, A. C., & LOOMIS, J. M. (1998). Locomotion mode affects the updating of objects encountered during travel: The contribution of vestibular and proprioceptive inputs to path integration. *Presence: Teleoperators & Virtual Environments*, **7**, 168-178.

DRYDEN, I. L., & MARDIA, K. V. (1998). *Statistical shape analysis.* Chichester, U.K.: Wiley.

GOLDIN, S. E., & THORNDYKE, P. W. (1982). Simulating navigation for spatial knowledge acquisition. *Human Factors*, **24**, 457-471.

HOCK, H. S., & SCHMELZKOPF, K. F. (1980). The abstraction of schematic representations from photographs of real-world scenes. *Memory & Cognition*, **8**, 543-554.

ISRAEL, I., CHAPUIS, N., GLASAUER, S., CHARADE, O., & BERTHOZ, A. (1993). Estimation of passive horizontal linear whole-body displacement in humans. *Journal of Neurophysiology*, **70**, 1270-1273.

JÜRGENS, R., BOß, T., & BECKER, W. (1999). Estimation of self-turning in the dark: Comparison between active and passive rotation. *Experimental Brain Research*, **128**, 491-504.

KEARNS, M. J., WARREN, W. H., DUCHON, A. P., & TARR, M. J. (2002). Path integration from optic flow and body senses in a homing task. *Perception*, **31**, 349-374.

KLATZKY, R. L., LOOMIS, J. M., BEALL, A. C., CHANCE, S. S., & GOLLEDGE, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science*, **9**, 293-298.

LANDAU, B. (1986). Early map use as an unlearned ability. *Cognition*, **22**, 201-223.

OHMI, M. (1996). Egocentric perception through interaction among many sensory systems. *Cognitive Brain Research*, **5**, 87-96.

PERUCH, P., MAY, M., & WARTENBERG, F. (1997). Homing in virtual environments: Effects of field of view and path layout. *Perception*, **26**, 301-311.

POTEGAL, M. (1982). Vestibular and neostriatal contributions to spatial orientation. In M. Potegal (Ed.), *Spatial abilities: Development and physiological foundations* (pp. 361-387). New York: Academic Press.

RIECKE, B. E., VAN VEEN, H. A. H. C., & BÜLTHOFF, H. H. (2002). Visual homing is possible without landmarks—a path integration study in virtual reality. *Presence: Teleoperators & Virtual Environments*, **11**, 443-473.

ROSSANO, M. J., WEST, S. O., ROBERTSON, T. J., WAYNE, M. C., & CHASE, R. B. (1999). The acquisition of route and survey knowledge from computer models. *Journal of Environmental Psychology*, **19**, 101-115.

SHOLL, M. J. (1989). The relation between horizontality and rod-and-

frame and vestibular navigational performance. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 110-125.

TELFORD, L., HOWARD, I. P., & OHMI, M. (1995). Heading judgments during active and passive self-motion. *Experimental Brain Research*, **104**, 502-510.

TOBLER, W. (1994). Bidimensional regression. *Geographical Analysis*, **26**, 187-212.

WALLER, D., BEALL, A. C., & LOOMIS, J. M. (2004). Using virtual environments to assess directional knowledge. *Journal of Environemental Psychology*, **24**, 105-116.

WALLER, D., HUNT, E., & KNAPP, D. (1998). The transfer of spatial knowledge in virtual environment training. *Presence: Teleoperators & Virtual Environments*, **7**, 129-143.

WALLER, D., LOOMIS, J. M., & STECK, S. D. (2003). Inertial cues do not enhance knowledge of environmental layout. *Psychonomic Bulletin & Review*, **10**, 987-993.

YARDLEY, L., & HIGGINS, M. (1998). Spatial updating during rotation: The role of vestibular information and mental activity. *Journal of Vestibular Research*, **8**, 435-442.

NOTES

1. Although the yoked design permitted analyses of difference scores between *walk* and matched *sit* participants, the correlation of absolute error between these two groups was extremely low [$r(24) = -.09$, $p = .69$], rendering the power of such an approach dubious.

2. The counterpart to mean signed error (also called "constant" error) is *variable* error, which can be computed as the standard deviation of each participant's signed errors at a given location. In the present experiment, variable error correlated extremely highly with absolute error [$r(72) = .98$, $p < .01$]. Thus, conclusions based on the statistical tests that we report are the same, whether they are based on absolute, or on variable, error.