Psychonomic Bulletin & Review
2003, 10 (4), 987-993

Inertial cues do not enhance knowledge of environmental layout

DAVID WALLER
Miami University, Oxford, Ohio

and

JACK M. LOOMIS and SIBYLLE D. STECK
University of California, Santa Barbara, California

Several sensory modalities besides vision are available to people as they move through an environment, learning where things are. For example, sensory information about linear and angular acceleration (i.e., inertial information) has been shown to be useful for maintaining orientation in a room-sized space. Because noise in inertial systems can compound over time and over extended travel, it is an open question whether inertial information is important for acquiring knowledge about large-scale environments. We addressed this issue in an experiment in which people learned the spatial layout of a large environment under conditions that varied in the degree to which valid inertial cues were available. The presence of valid inertial cues did not facilitate the acquisition of an accurate memory of the environment. Moreover, the presence of invalid inertial cues did not interfere with such acquisition. We conclude that the effect of inertial information on the acquisition of environmental knowledge is minimal.

When moving to a new locale, people quickly begin acquiring knowledge about their new surroundings. This knowledge consists of both memory for traveled routes and a more abstract mental representation of spatial layout, often referred to as *survey* or *configurational* knowledge (Siegel & White, 1975) or a *cognitive map* (Downs & Stea, 1977; Golledge, 1978). A major goal of research on environmental learning is to discover what types of sensory information are involved as people acquire knowledge of the layout of an environment.

Recent research on the navigation of room-sized spaces has shown that body-based vestibular, somatosensory, and somatogravity (Mittelstaedt & Mittelstaedt, 1996) cues that provide information about linear and angular acceleration (i.e., *inertial* cues) play an important role in maintaining orientation with respect to known locations and directions (Bloomberg, Melvill Jones, & Segal, 1991; Blouin et al., 1995; Chance, 2000; Israël, Grasso, Georges-François, Tsuzuku, & Berthoz, 1997; Klatzky, Loomis, Beall, Chance, & Golledge, 1998). For example, Klatzky et al. examined performance in a task that required people to keep track of a location as they moved away from it along two path segments connected by one turn. Some of the participants in this experiment performed this task by sitting still and wearing a head-mounted display (HMD) that simulated the optic flow involved in traversing the path. Another group of participants received the same visual information but were passively rotated by the proper amount when the visual field depicted a turn. Pointing errors for people who received only visual information were large and directly proportional to the magnitude of the turn. On the other hand, the pointing errors of people who were physically rotated during the depicted turn were quite low and unrelated to the magnitude of the turn. Klatzky et al. concluded from these results that optic flow alone is not sufficient for the updating of one's orientation in space and that additional sensory information (e.g., vestibular) is required (but see Riecke, van Veen, & Bülthoff, 2002).

All of the aforementioned research has been conducted in relatively small-scale (e.g., room-sized) spaces and has involved relatively simple tasks, such as keeping track of a starting location during movement. Yet it is important, both theoretically and practically, to determine whether their findings apply to larger scale, navigable spaces (e.g., campus-, building-, or city-sized spaces) and to relatively complex tasks that involve knowing the layout of an environment. Theoretically, such an investigation enables us to understand better how various sensory modalities are used and integrated by the central nervous system as people form cognitive maps. In particular, we note that inertial cues must be internally processed (i.e., doubly integrated) in order to specify position and orientation. If either the sensed inertial signals or the processing mechanisms that operate on them are noisy, then, without periodic correction, this noise can accumulate during travel that is extended in space. In-

This work was supported by grants from the National Science Foundation and the Office of Naval Research to J.M.L. We thank Michael Kinsella, Travis Moore, and Anthony Richardson for help with collecting the data. Correspondence concerning this article should be addressed to D. Waller, Department of Psychology, Miami University, Oxford, OH 45056 (e-mail: wallerda@muohio.edu).

Copyright 2003 Psychonomic Society, Inc.

creased noise in the inertial system can, thus, result in increasingly degraded estimates of position and orientation as path complexity increases. If this is true, it is possible that in large environments, inertial information is not particularly helpful for maintaining orientation and learning a spatial layout. Accordingly, we hypothesize that, contrary to their reliance on inertial information in small-scale spaces, people rely relatively little on inertial information when acquiring knowledge of an environmental layout.

Questions about the importance of inertial information in the acquisition of environmental knowledge are also important from an applied perspective. Recently, there has been great interest in using computer simulations of large-scale environments (i.e., virtual environments) in order to conduct basic research on human wayfinding and navigation, as well as to train people on tasks requiring knowledge of geographic spaces (Durlach et al., 2000; Loomis, Blascovich, & Beall, 1999). In general, interfaces with these systems do not convey all of the inertial information about self-motion that would normally arise from users' movements through these spaces. For these applications, it is important to understand the degree to which inertial information must be available from the computer interface in order to facilitate the acquisition of spatial knowledge.

In one of the few studies in which the role of inertial information in acquiring knowledge of large-scale environments has been addressed, Goldin and Thorndyke (1982) asked participants to learn the spatial characteristics of an 8-km route through a Los Angeles neighborhood. Half of the participants in their experiment (the *tour* group) were driven in a bus along the route. The other half learned the same environment by watching a motion picture of the trip (the *film* group). Both groups were subsequently given a variety of tests that assessed different aspects of their spatial knowledge of the environment. The only measure for which the tour group performed significantly better than the film group was a test of relative directions, with tour participants approximately 10º more accurate at estimating directions than the film participants. Goldin and Thorndyke's results appear to be somewhat consistent with the laboratory findings, cited above, that have revealed an important role of vestibular involvement in maintaining one's orientation in space. However, in addition to the presence of vestibular inertial information, other factors in Goldin and Thorndyke's experiment may have accounted for the difference between the tour and the film groups. For example, the degree of active control of attention was different for the two groups, as was the field of view, resolution, and stereoscopy afforded by the learning experience. In the present experiment, we controlled these alternate possibilities in order to examine specifically the effect of inertial cues on the acquisition of environmental knowledge.

Participants were tested on their knowledge of a previously unknown large-scale environment that they had learned from a car trip through the area. One group (full cue) learned the environment as normal passengers from the front seat of the car. Another group (inertial) observed the environment from the back seat of the car while wearing an HMD. The HMD received video input from a camera that was mounted in the car, recording the trip as it occurred. A third group of participants (nonmatching) watched these videos through the same HMD while they were seated in a car traveling through an environment different from the one depicted in the video. The fourth group of participants (video) watched the videos through the HMD, but while sitting still in a laboratory. If inertial cues are important for learning large-scale spaces, we would expect the inertial group to form a more accurate representation of the environment than would either the video or the nonmatching group. Comparisons between the inertial and the nonmatching groups would offer a particularly strong test of the role of inertial information in environmental knowledge acquisition, because in both conditions, inertial cues were available for the perceptual system to use (or misuse). Finally, comparing the performance of the full-cue group with that of the inertial group could provide information about the role of additional sources of information (e.g., increased visual fidelity, field of view, and the ability to look around freely) used when people acquire knowledge of an environment. This comparison would serve as a replication of Goldin and Thorndyke's (1982) experiment.

METHOD

Participants

The participants were 124 college students (62 men, 62 women) with a mean age of 18.9 years ($SD = 2.06$). All the participants reported having normal or corrected-to-normal vision. After the experiment, none reported having had prior familiarity with the environment they had learned.

Environment and Materials

The environment to be learned consisted of a 1,600-m route through a business district in Goleta, California (see Figure 1). Five places along the route were chosen as locations for the participants to learn. Each of the locations was close to a prominent landmark: a large liquid nitrogen tank, stop signs, or curves in the road. None of the locations or landmarks was visible from the others.

The participants in the inertial, nonmatching, and video conditions viewed the environment with a Virtual Research V8 HMD. The HMD provided identical images to the two eyes in a 50º horizontal × 38º vertical field of view. The inertial participants watched the video as it was being made from a JVC GR-AX420 VHS camcorder mounted, on a tripod inside the car, 24 cm from the windshield between the driver and the passenger seats.[1] The camera employed an additional lens that provided an approximately 50º horizontal field of view. The video signal from this camera was converted from NTSC to 640 horizontal × 480 vertical VGA (by a Digimedia VD-300 video box) before being displayed in the HMD. The participants in the nonmatching condition viewed these tapes while being driven on a route through the University of California in Santa Barbara. The route used for the nonmatching condition generally went in a clockwise direction (as opposed to the generally counterclockwise orientation of the route for the other groups) and was driven at approximately the same speed as that in the full-cue and inertial conditions.

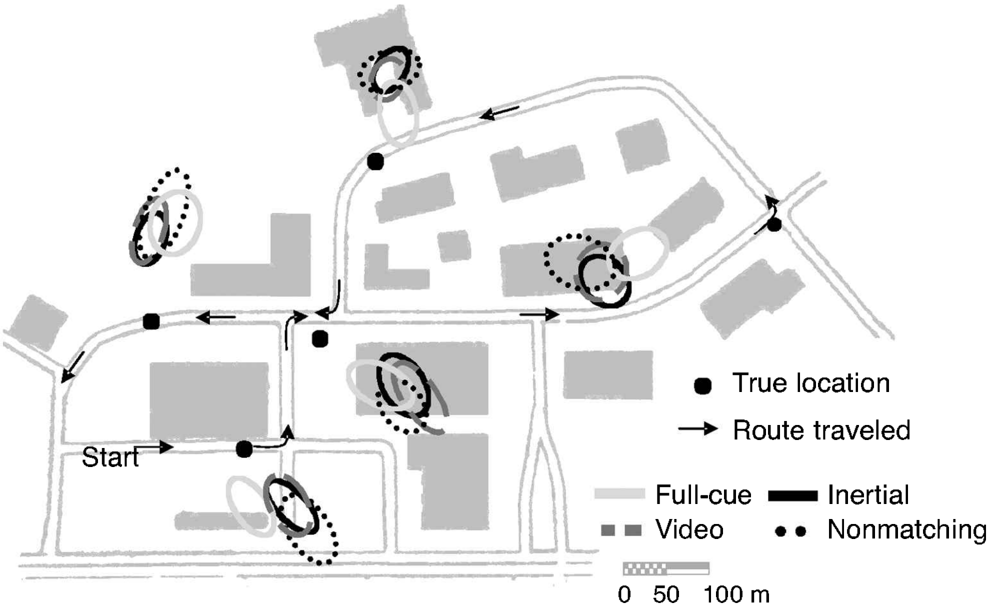

Figure 1. Map of the environment learned and the route traveled. The environment consisted of five locations along an approximately 1,600-m route through a secluded area in Goleta, California. In addition to other tasks, the participants configured five small pieces on a sheet of paper to represent a map of the learned environment. The *x*- and *y*-coordinates of the map placements were recorded and averaged (along with the actual locations) with Procrustes superimposition (see Dryden & Mardia, 1988). Shown are 95% confidence ellipses (see Batschelet, 1981) for the estimated locations of each group in the map placement task.

All the participants' knowledge of directions and distances was tested in the laboratory with immersive computer-generated simulations of each of the learned locations. These simulations were created from 360º panoramic digital photographs of each of the five locations that were applied to the inside surface of a computer-modeled cylinder. The user's viewpoint was placed in the center of the cylinder. The participants viewed these panoramas wearing the V8 HMD, on which was secured an inertially based orientation tracker (Intersense Model IS-300). The tracker was used to update the orientation of the visual image (relative to the orientation of the participant's head), as well as to record facing directions when the participant was pointing to the various targets. The computer rendered these scenes by using a Pentium III chipset and an NVIDIA GeForce2 MX graphics card that updated the graphics at more than 100 frames per second. Thus, the displayed graphics update was limited by the refresh rate of the HMD, which was 72 Hz. The participants interacted with these environments much as they would in the real world—simply by turning their heads to view different directions. Randomization and presentation of the stimuli, as well as the collection of pointing data and distance 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 (Beall, Blascovich, & Loomis, 2001).

Procedure

The participants were instructed that they would be required to learn the layout of five locations in an unfamiliar environment. They were shown photographs of the to-be-learned locations and were told that they would later be asked to point at and estimate distances to and from these locations. Equal numbers of men and women were assigned to the full-cue and the inertial groups. These participants were blindfolded as they were driven to and from the to-be-learned environment (the HMD served as the blindfold for the inertial group). At the start of the route, the full-cue participants removed their blindfold, and the experimenter turned on the inertial participants' HMD. The participants were then driven twice around the route while the experimenter announced each location as it was passed. The participants who learned with the HMD were instructed to keep their head orientation fixed (because moving their heads did not change their view of the environment). In general, the entire exposure to the environment lasted about 6 min. Tapes of each trip were subsequently shown to gender-matched participants in the nonmatching and the video groups.

For testing, the participants pointed and estimated distances between all possible (20) pairs of locations. These trials were blocked by testing location, with directions and distances being judged sequentially for each target. The participants were instructed to estimate straight-line distances, using any unit they wished, and were told to try to maintain relative accuracy across all of their estimates (as opposed to absolute accuracy). The order of testing locations, as well as the order of targets for each location, was randomized separately for each participant. After pointing and estimating distances, the participants were asked to construct a map of the environment by placing small cardboard pieces representing landmarks on a blank sheet of grid paper. The *x*- and *y*-coordinates of their map placements were recorded for further analysis.

RESULTS AND DISCUSSION

Three people in each of the inertial and nonmatching conditions and two participants in the video group elected to stop during the learning portion of the experiment, complaining of mild symptoms of simulator or motion sickness. These participants (and their yoked counterparts in the full-cue condition) were subsequently replaced. Pointing error was first examined by calculating the mean absolute difference between the indicated and the actual directions, averaged over all trials. Accuracy with judging relative distances was measured as the correlation between indicated and actual distances. Map accuracy was measured as the bidimensional correlation between the actual configuration of locations and people's estimated locations. The bidimensional correlation is an index of the similarity between two two-dimensional configurations and can be interpreted similarly to Pearson's r (see Tobler, 1994). In general, the participants were able to estimate directions and distances and to construct maps with reasonably high accuracy. Mean pointing error across all participants averaged 38.7° (SD = 13.2°). Mean distance estimation accuracy across all participants was .39 and was significantly greater than chance performance [$t(123) = 29.65, p < .01$]. Similarly, mean map accuracy was quite large, averaging .71 (SD = .16) across all participants.

The participants in the inertial, video, and nonmatching conditions did not differ significantly from each other on any measure of spatial knowledge (see Figure 2). These effects were tested with pairwise contrasts on pointing accuracy, distance estimation accuracy, and map construction accuracy. For each of these tests, t values were generally less than 1; all were less than 1.09. Estimates of effect size among these three groups (Cohen's f; see Cohen, 1988) were 0.16 for pointing error, 0.03 for distance estimation accuracy, and 0.12 for map construction accuracy. The effect of inertial information was also tested among the inertial, nonmatching, and video conditions in a multivariate analysis of variance (MANOVA) that used pointing error, distance estimation accuracy, and map placement accuracy as dependent variables. The MANOVA revealed no significant effect of inertial information [$F(6,176) = 0.6, p = .73$]. Thus, if inertial information facilitates the acquisition of environmental knowledge, its effect appears to be minimal.

The participants in the full-cue condition pointed to targets with about 29.8° of absolute error, whereas, on average, the participants in the other three groups erred about 41.7°. This 11.9° difference between the full-cue group and the others was significant [$t(120) = 4.66, p < .01$]. The full-cue participants were also more accurate than the other participants at judging relative distances, and this difference was also significant [$t(120) = 2.94, p < .01$]. These results suggest that noninertial factors, such as the field of view, the visual fidelity, and the degree to which people are able to look around actively (thus receiving proprioceptive information), exert a large and significant influence over people's directional knowledge in large-scale environments.

Failure to encode or remember curves or turns is likely to result in a left or right bias in pointing from a location to different targets. Mean signed error at each location was used as a measure of pointing bias and is illustrated in

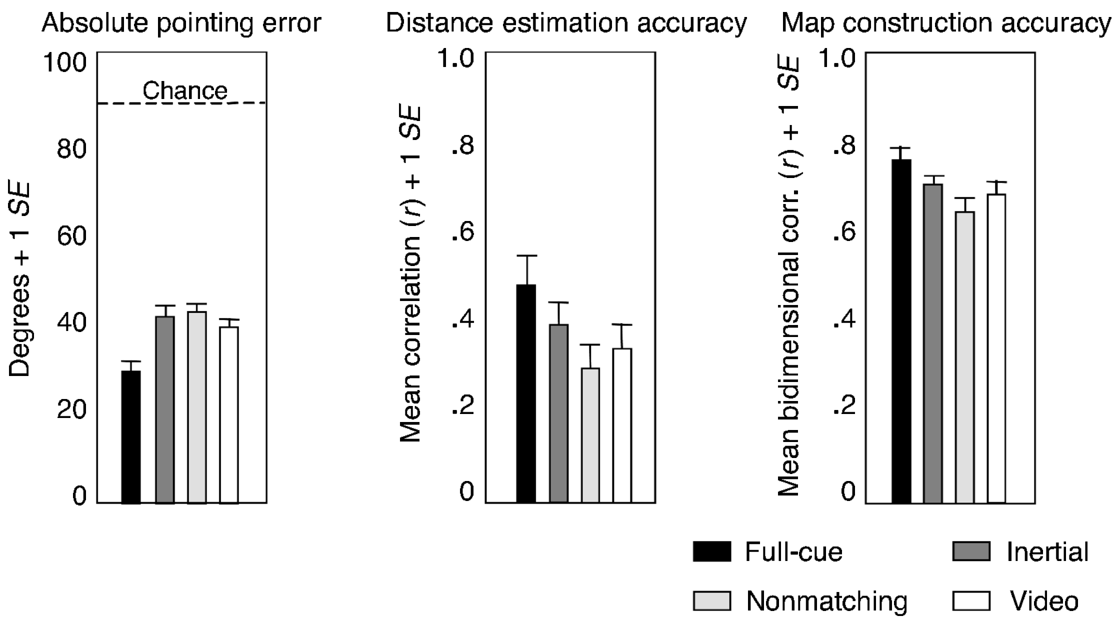

Figure 2. Means and standard errors for three measures of spatial knowledge for the four experimental groups. (Left) The full-cue participants pointed with greater accuracy than did the participants in the other three groups. The inertial, nonmatching, and video groups did not differ statistically. The same pattern of results obtained for distance estimation accuracy (center) and map construction accuracy (right), both measured as a correlation (or bidimensional correlation) between actual and estimated distances or locations.

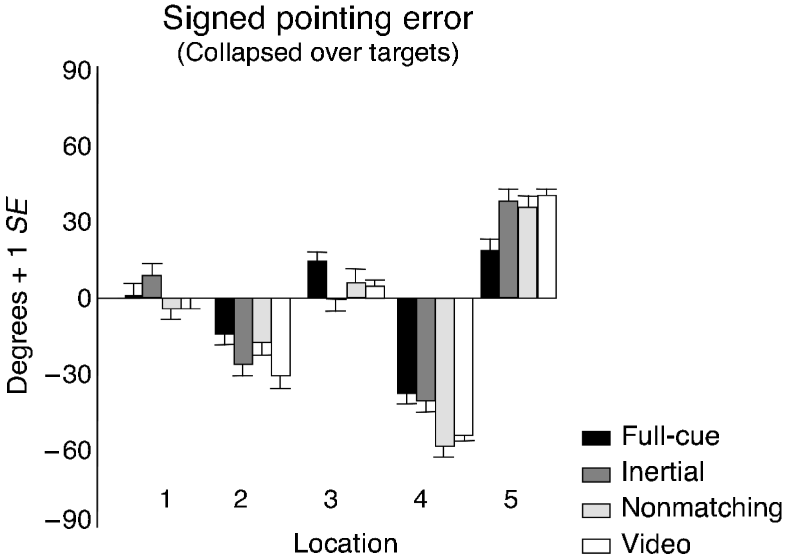

Figure 3. Mean signed pointing error (and standard errors) from each location for each experimental group. Despite significant overall pointing bias at Locations 2, 4, and 5, there is no evidence of a difference among the inertial, the nonmatching, and the video groups. Patterns of biases were more influenced by geographic location than by the mode of learning.

Figure 3. At three locations, mean signed error was significantly different from zero. However, at none of these locations did the participants in the inertial condition exhibit signed errors that differed significantly from those made by the participants in either the video or the nonmatching condition. Notably, Figure 3 also shows a relatively consistent pattern of biases across groups at each location. This suggests that aspects of the environment exerted a larger influence on the participants' mental representations than did the learning conditions that we manipulated. Such a finding offers promise for future experimental work that would aim to link environmental variables to the processes involved in creating and maintaining mental representations of large-scale spaces.

The map placement data provided converging evidence for these conclusions. As with all other measures, the inertial, the video, and the nonmatching groups did not differ significantly on the accuracy of their constructed maps (see Figure 2). The participants in the full-cue group constructed significantly more accurate maps than did the people in the other three groups [$t(120) = 3.70, p < .01$]. Yet all the groups exhibited similar biases in their constructed maps. Figure 1 illustrates 95% confidence ellipses (see Batschelet, 1981) for each estimated location that the participants in each group made when constructing their maps. The figure shows that, regardless of experimental condition, the participants tended to shorten longer distances and underestimate the degree of certain curves and turns, particularly those experienced toward the end of the route.

Very similar biases were also present in maps that we generated from the participants' pointing estimates. To generate such maps, we used a variant of multidimensional scaling (MDS) that uses interpoint bearing estimates (instead of distance estimates) in order to determine a two-dimensional map of the environment that maximally fits each participant's data (see Waller & Haun, 2003). Figure 4 illustrates 95% confidence ellipses for the configurations derived from all the participants in both the pointing and the map placement tasks. In general, the maps derived from the participants' pointing estimates exhibited a very high similarity to the participants' explicitly constructed maps (bidimensional correlation of .90). This was greater similarity than the .80 (explicitly constructed) and .84 (generated) correlations between these maps and the actual configuration of locations in the environment. For these biases to persist across assessment techniques strongly suggests that we measured a unitary, stable mental representation of the environment.

People who watched a video of a car trip, whether while sitting still or while moving in ways that the video did not depict, were not statistically distinguishable from those who were in the car, watching the video as it was being made, on any measure of their knowledge of the environment's layout. In contrast to recent findings conducted in sparse small-scale environments (Bloomberg et al., 1991; Blouin et al., 1995; Chance, 2000; Israël et al., 1997; Klatzky et al., 1998), our results suggest that the effect of inertial information on the acquisition of environmental layout is quite small. We suggest two possible reasons for this small effect. First, the environment we asked the participants to learn was much larger than the environments used in previous studies. As we mentioned in the introduction, the use of inertial information

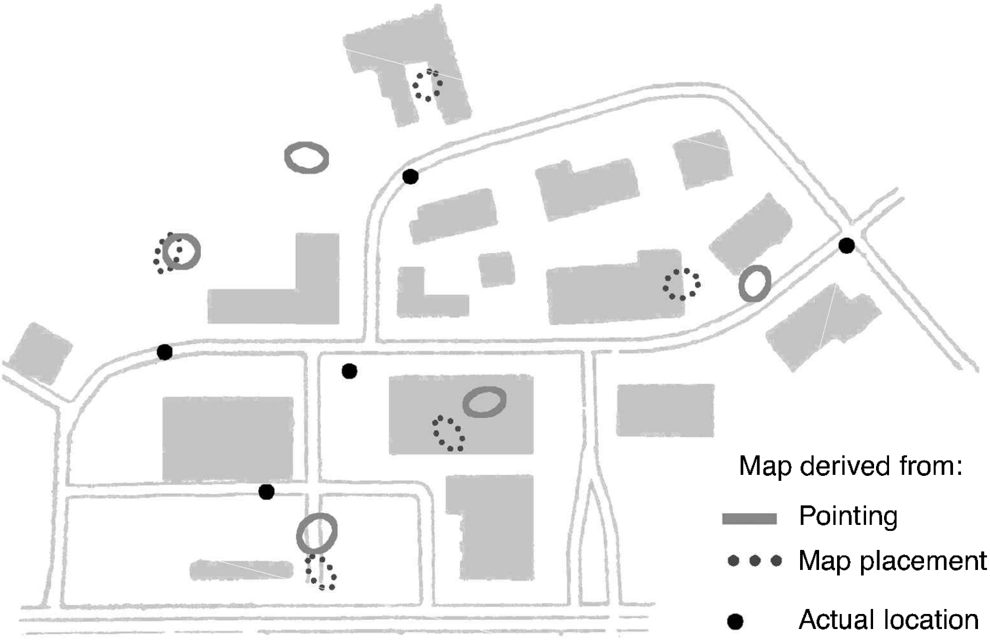

Figure 4. Ninety-five percent confidence ellipses (see Batschelet, 1981) for location estimates from both the map placement (dashed) and the pointing (solid) tasks. A technique similar to multidimensional scaling was used to derive maps from pointing responses (see Waller & Haun, 2003). Explicitly constructed and derived maps were averaged separately across all participants with Procrustes superimposition (see Dryden & Mardia, 1988).

can result in increasingly noisy estimates of position as travel is extended in space. Because of the relatively great amount of noise in inertial systems, it may be adaptive for organisms navigating in large spaces to rely relatively little upon it. A second explanation for the minimal role of inertial cues in our experiment involves the nature of the task. Most previous studies have not investigated people's ability to acquire knowledge of layout but, rather, their ability to update their relationship to a single location in space. It is possible that acquiring knowledge of layout, regardless of its scale, relies much more heavily on visual information than do the updating tasks that are often investigated.

Perhaps most interesting, patterns of biases in spatial judgments among all the groups were remarkably similar, indicating that the information that the participants used during learning was internally processed in similar ways across different conditions. In addition to furthering our understanding of the sensory bases underlying environmental cognition, these findings have important ramifications for such applications as computer-generated (virtual) environments that simulate large spaces for people to learn. Our results suggest that interfaces for many of these applications need not require inertial (e.g., vestibular) involvement and, thus, support a growing number of studies showing that desktop virtual environments can be useful for enabling relatively accurate spatial knowledge of large spaces (Richardson, Montello, & Hegarty, 1999; Rossano, West, Robertson, Wayne, & Chase, 1999; Ruddle, Payne, & Jones, 1997). The finding that absolute pointing error is relatively low after only visual exposure to a complex real-world environment confirms that people are able to use visual information alone to acquire knowledge of layout and that vestibular and other inertial cues do not necessarily facilitate this acquisition.

REFERENCES

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

Beall, A. C., Blascovich, J., & Loomis, J. M. (2001). *PC-based immersive virtual environment technology as a behavioral science research tool.* Unpublished manuscript, University of California, Santa Barbara.

Bloomberg, J., Melvill Jones, G., & Segal, B. (1991). Adaptive modification of vestibularly perceived rotation. *Experimental Brain Research*, **84**, 47-56.

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. (2000). How we use sensory information to keep oriented as we locomote. *Dissertation Abstracts International*, **61** (12-B), 6728.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Erlbaum.

Downs, R. M., & Stea, D. (1977). *Maps in minds: Reflections on cognitive mapping.* New York: Harper & Row.

Dryden, I. L., & Mardia, K. V., (1988). *Statistical shape analysis.* Chichester, U.K.: Wiley.

Durlach, N, Allen, G., Darken, R., Garnett, R. L., Loomis, J., Templeman, J., & von Wiegand, T. E. (2000). Virtual environments and the enhancement of spatial behavior: Toward a comprehensive research agenda. *Presence: Teleoperators & Virtual Environments*, **9**, 593-615.

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

Golledge, R. G. (1978). Learning about urban environments. In T. Carlstein, D. Parkes, & N. Thrift (Eds.), *Timing space and spacing time: Vol. 1. Making sense of time* (pp. 76-98). London: Arnold.

Isra l, I., Grasso, R., Georges-Fran ois, P., Tsuzuku, T., & Berthoz, A. (1997). Spatial memory and path integration studied by self-driven passive linear displacement: I. Basic properties. *Journal of Neurophysiology*, **77**, 3180-3192.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Updating an egocentric spatial representation during real, imagined, and virtual locomotion. *Psychological Science*, **9**, 293-298.

Loomis, J. M., Blascovich, J. J., & Beall, A. C. (1999). Immersive virtual environment technology as a basic research tool in psychology. *Behavior Research Methods, Instruments, & Computers*, **31**, 557-564.

Mittelstaedt, M. L., & Mittelstaedt, H. (1996). The influence of otoliths and somatic graviceptors on angular velocity estimation. *Journal of Vestibular Research*, **6**, 355-366.

Richardson, A. E., Montello, D. R., & Hegarty, M. (1999). Spatial knowledge acquisition from maps and from navigation in real and virtual environments. *Memory & Cognition*, **27**, 741-750.

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.

Ruddle, R. A., Payne, S. J., & Jones, D. M. (1997). Navigating buildings in "desk-top" virtual environments: Experimental investigations using extended navigational experience. *Journal of Experimental Psychology: Applied*, **3**, 143-159.

Siegel, A. W., & White, S. H. (1975). The development of spatial representations of large-scale environments. In H. Reese (Ed.), *Advances in child development and behavior* (Vol. 10, pp. 10-55). New York: Academic Press.

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

Waller, D., & Haun, D. B. M. (2003). Scaling techniques for modeling directional knowledge. *Behavior Research Methods, Instruments, & Computers*, **35**, 285-293.

NOTE

1. The placement of the camera introduced a very slight discrepancy between the inertial cues experienced by the inertial participants (in the back seat) and the visual cues they received from the camera (in the front seat). There are two reasons to believe that this discrepancy was inconsequential in this experiment. First, the maximum physical difference between the inertial information at the video camera and that at the participant's head during the trip was extremely small—approximately 0.0014 m/sec^2 (or roughly 0.0001 g). Researchers will likely agree that this difference is too small to be sensed by the vestibular system (note that detection thresholds are around 0.005 g). Second, the results from the nonmatching condition show that an extremely high degree of spatial learning can occur even when inertial and visual information are completely uncorrelated.

(Manuscript received April 10, 2002;
revision accepted for publication October 8, 2002.)