

Simultaneous spatial updating in nested environments

RANXIAO FRANCES WANG and JAMES R. BROCKMOLE
University of Illinois, Champaign, Illinois

When one moves, the spatial relationship between oneself and the entire world changes. Spatial updating refers to the cognitive process that computes these relationships as one moves. In two experiments, we tested whether spatial updating occurs automatically for multiple environments simultaneously. Participants turned relative to either a room or the surrounding campus buildings and then pointed to targets in both the environment in which they turned (updated environment) and the other environment (nonupdated environment). The participants automatically updated the room targets when they moved relative to the campus, but they did not update the campus targets when they moved relative to the room. Thus, automatic spatial updating depends on the nature of the environment. Implications for theories of spatial learning and the structure of human spatial representations are discussed.

Animals, including humans, constantly move from one place to another and turn from one direction to another. As an animal's position and orientation change, the spatial relationship between the animal and the surrounding environment changes. For example, a cup in front of someone will move to his or her left as he or she turns to the right. Thus, in order to reach for the cup, he or she has to keep track of the cup's position as he or she moves. This cognitive process is often referred to as *spatial updating* (Amorim & Stucchi, 1997; Farrell & Robertson, 1998; Loomis, Da Silva, Philbeck, & Fukusima, 1996; Wang & Spelke, 2000).

The nature of the spatial-updating process is receiving increasing attention in the visual cognition literature. In various studies, the relative ease with which different types of spatial updating can be performed has been investigated. For example, Rieser (1989; see also Huttenlocher & Presson, 1979; May, 1996; Presson & Montello, 1994; Rieser, Garing, & Young, 1994; Simons & Wang, 1998; Wang & Simons, 1999) showed that egocentric direction judgments were much faster and more accurate following physical body rotation than following imagined body rotation. Farrell and Robertson (1998) further showed that spatial updating occurs automatically: Once people move, they *update* the locations of target objects around them and have great difficulty pointing to the objects as if they did not move. Thus, spatial updating is one of the fundamental cognitive processes that operates constantly to ensure

that we can get the things that we need and go where we want.

The finding that spatial updating is automatic raises an important theoretical question—namely, does the spatial-updating system operate on everything in the world simultaneously? With every movement we make, we change the spatial relationship between everything in the world relative to ourselves. Thus, if the spatial-updating system can keep track only of a limited number of objects and places as one moves, it will lose track of where everything else is after a few turns. Moreover, if some of the spatial relationships are updated and some of them are not, one's knowledge of the spatial relationships among objects themselves may be destroyed. For example, suppose there is a cup in front of someone and a telephone to his or her right. If he or she turns right, the cup will be to his or her left, and the telephone will be in front of him or her. If he or she automatically updates these relationships, he or she will have proper knowledge of the direction of both objects. However, if he or she updates the direction of the telephone, but not the cup, not only will he or she lose track of where the cup is after he or she turns, but also his or her knowledge of the relative position of the cup and the telephone will suffer. Therefore, maintaining a coherent spatial knowledge system may require simultaneous updating of all known places. Thus, it seems plausible not only that spatial updating may be automatic, but also that it may operate on all objects and locations simultaneously.

In addition to the intuitive appeal, studies of cognitive maps provide a plausible mechanism by which such simultaneous updating may occur (e.g., Gallistel, 1990; O'Keefe & Nadel, 1978). Cognitive map theories suggest that animals learn the target locations relative to an external reference frame. Thus, the animal's spatial knowledge of these locations is referred to as an allocentric cognitive map. While navigating, the animal continuously updates a vector that indicates its position in the cognitive map (path in-

This research was supported by a Research Board Grant to R.F.W. from the University of Illinois and an NSF Graduate Research Fellowship to J.R.B. Thanks to Elizabeth Spelke, Steven Pinker, Mary Potter, and Randy Gallistel for discussion and to Rashad Abdul-Salaam and Benjamin Gillespie for data collection. Correspondence concerning this article should be addressed to R. F. Wang, Department of Psychology and Beckman Institute, 603 E. Daniel St., Room 533, University of Illinois, Champaign, IL 61820 (e-mail: francesw@s.psych.uiuc.edu).

tegration). Thus, updating involves calculation of a single vector—namely, one's own position in an external reference frame. According to this view, spatial updating either occurs relative to the cognitive map as a whole or does not occur at all. That is, an animal should update its position relative to everything in the cognitive map simultaneously.

However, some recent research has suggested that the opposite might be true. For example, Wang and Spelke (2000; see also Wang, 1999) showed that disruption of the spatial-updating process (disorientation) impairs people's knowledge of the relationship among a set of targets. They argued that people update the egocentric positions of the surrounding objects individually, instead of updating their own geocentric position within the environment as a whole. According to this view, accurate knowledge of the relationship among a set of targets is maintained by the coherent spatial updating of multiple targets, and disruption of the spatial-updating process also disrupts the accurate relational knowledge. According to these findings, Wang (1999) further suggested that the human spatial-updating system has a capacity limitation. That is, because the egocentric vector of each target needs to be updated individually as one moves, the more targets one updates, the greater the number of computational processes that become involved. Thus, there may be a limit on the number of targets whose egocentric positions can be computed at the same time. As a result, people may not be able to update everything all at once; only the most relevant objects are updated, such as things in the immediate surroundings, whereas more remote targets and places are not.

Thus, the mechanism of spatial updating across environments has important implications for theories of spatial learning and the structure of spatial representations. In the present study, we examined whether people simultaneously update multiple targets both in their immediate surroundings and from a more remote environment. Participants were asked to move relative to one of the two nested environments and then point to object locations in either the *updated* environment or the *nonupdated* environment. Note that when they moved, they physically moved relative to both environments simultaneously. Thus, if spatial updating occurs automatically for both environments, people should be able to respond to targets in both environments quickly and accurately, regardless of whether the environment was explicitly updated or not. On the other hand, if only one environment is updated at a time, they should be quick in responding to the *foregrounded* environment, which they explicitly considered when moving to face targets in that environment, and slow in responding to the targets from the environment in the *background*.

EXPERIMENT 1

Method

Participants. Sixteen undergraduate students enrolled in an introductory psychology class at the University of Illinois participated for course credit.

Apparatus. The study was conducted in a rectangular room with five objects (see Figure 1). An overhead video camera was mounted directly above a swivel chair to record the participants' physical movements in the room and the pointing responses they made dur-

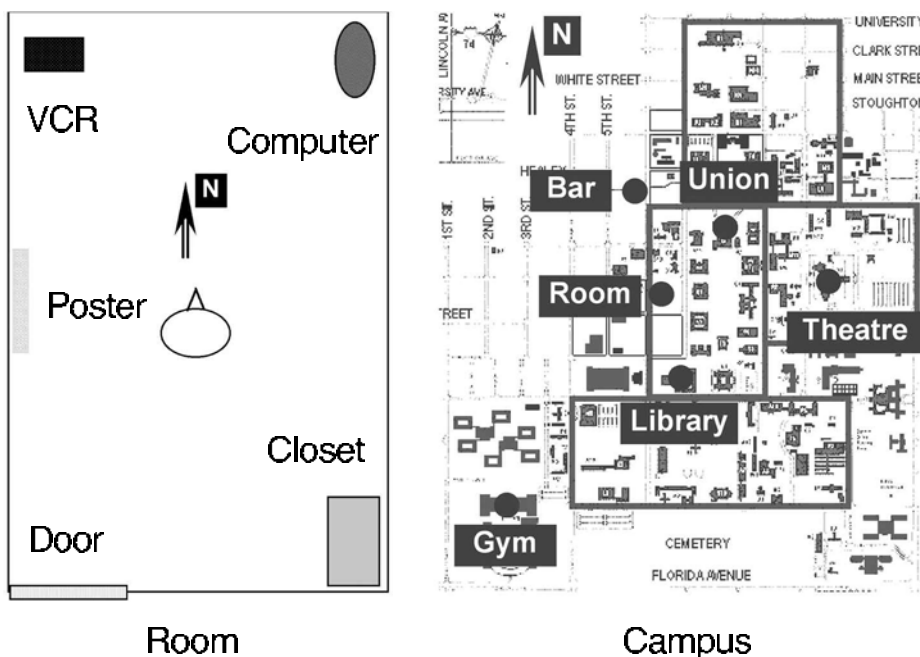


Figure 1. The apparatus and two environments: a room (left) and the surrounding campus buildings (right).

ing the experiment (see below). The chair was placed in the middle of the room and was fixed to the floor so that the participants could sit in the chair and turn freely but could not change their geographic position.

Design and Procedure. The participants were randomly divided into two groups: half were in the *update-campus* condition, and half were in the *update-room* condition. In both conditions, the participants were tested in three stages: a learning stage (before updating), a movement stage, during which the participants turned to face various targets, and a testing stage (after updating). In the learning stage, each participant was familiarized with the identity and location of the five room targets while the experimenter pointed to them one by one. This learning stage was untimed. The participants were then blindfolded and sound masked and turned to face east (see Figure 1). They were then asked to point to the five objects in the room in a random sequence as the experimenter named them. If they made an error ($>20^\circ$ from the correct direction), they were corrected by the experimenter by moving their hand to point to the correct direction. If errors were made, the participants were retested. This procedure ensured that they had correctly learned the room environment. Next, the participants were asked to point to the five selected buildings on campus; if errors were made, incorrect responses were corrected by the experimenter, and the participants were then retested on the campus target locations.¹ After this learning stage, the participants were instructed to move relative to one of the environments. For the *update-room* group, the participants were asked to turn, while sitting in the swivel chair, to face each of the five room targets one by one in a random sequence as the experimenter named them. For the *update-campus* group, they were asked to turn to face each of the five campus buildings one by one. The last turning target was the VCR for the *update-room* group and the bar for the *update-campus* group. Thus, for both groups, the participants faced approximately the same orientation at the end of the turning stage. Finally, after this turning stage, the participants pointed to the five objects in the updated environment from this final position, and then they pointed to the targets in the nonupdated environment in a random sequence.

Data analysis. Both response times (RTs) and angular errors were analyzed. The direction of the pointing responses and the turning responses were measured from the video recording by superim-

posing a transparent radial grid on the TV monitor. RT was measured as the elapsed time between the announcement of the probed target and the end of the pointing or turning response, as indicated by the stabilization of the hand or body.²

Results and Discussion

Overall, in the learning stage prior to turning, the participants' pointing responses did not differ between the two groups for either environment [$F_s(1,14) < 0.32, p_s > .58$]. Mean RTs for room and campus targets were 1.4 and 1.5 sec, respectively; mean pointing errors were 10° and 17° , respectively. In addition, during turning, the participants were able to turn relative to both environments equally well and showed no significant difference in either RT or angular error in their turning [mean RTs, 3.3 and 2.9 sec; mean errors, 24° and 25° , for turning relative to room and campus, respectively; $t_s(14) < 1.1, p_s > .31$]. Responses after the turning stage, however, did differ between the two groups (see Figure 2). Pointing responses to the room targets were not affected by whether people updated the room or the campus [$F_s(1,14) < 0.3, p_s > .59$]. However, pointing responses to the campus targets differed significantly between the two groups: Whereas the *update-campus* group could point to campus targets with little difficulty, the *update-room* group was significantly impaired. Both RTs and errors were significantly higher for the *update-room* group than for the *update-campus* group [$F_s(1,14) > 4.7, p_s < .05$], and a significant interaction between target environment and updating condition was observed [$F_s(1,14) > 4.9, p_s < .05$].

These data showed that when the participants were asked to turn to face targets on campus, they were later able to identify the location of both the campus and the room targets easily, suggesting that spatial updating acted upon both the campus and the room targets. However,

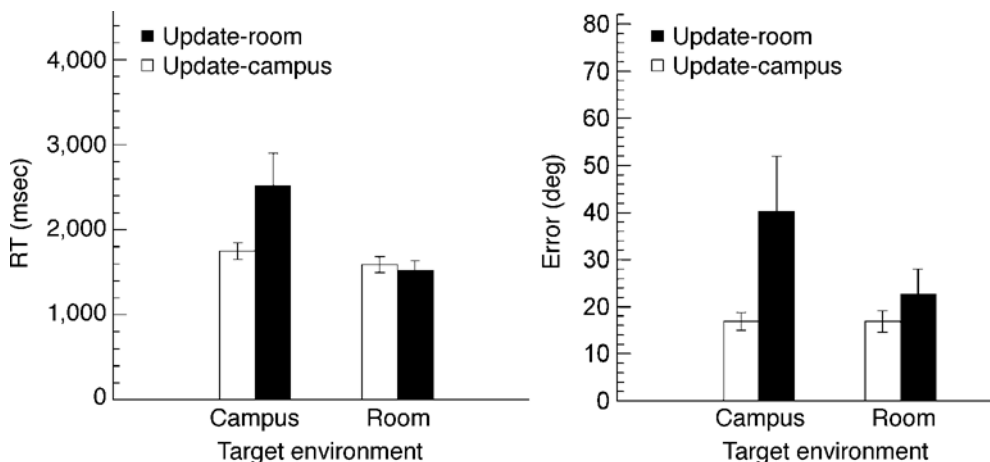


Figure 2. The response times (RTs, left) and errors (right) in Experiment 1. The error bars represent standard errors.

when the participants were asked to turn to face targets in the room, the later identification of the location of campus targets was impaired, suggesting that updating did not act on the campus targets in this condition. These data suggest that multiple nested environments are not always updated simultaneously and that automatic updating depends on the nature of the environment.

However, a possible alternative explanation for the results of Experiment 1 is that the differences between the update-campus and the update-room groups were due simply to the relative difficulty of switching between two environments. For example, Brockmole and Wang (2002) showed that after a direction judgment about things in one environment is made (e.g., one's office), people are slower in making a judgment about things in a different environment (e.g., the building) than in making a judgment about things in the same environment, suggesting there is a cost in changing environments. It is possible that switching from campus to room (which is required in the update-campus condition) is easier than switching from room to campus (which is required in the update-room condition), because a mental image of the campus includes the room, but not vice versa. As a result, the switching cost between the two environments may have been disproportionately greater in the update-room condition. Experiment 2 tested this possibility by requiring participants to switch environments, as in Experiment 1, in the absence of spatial updating.

EXPERIMENT 2

Method

Twenty participants from the same population as that in Experiment 1 were tested in Experiment 2. The participants followed the same procedure as that in Experiment 1, except that they did not turn. Specifically, the participants were tested in three stages. They

first familiarized themselves with the locations of the room and campus targets and pointed to the targets while blindfolded, as in Experiment 1 (learning stage). After the learning stage, instead of turning to face the targets in one of the environments, the participants either pointed to the five room targets in a random order (baseline-room condition) or pointed to the five campus targets in a random order (baseline-campus condition). Half of the participants were randomly assigned to the baseline-room condition, and half were in the baseline-campus condition. After this baseline stage, the participants pointed to the room targets, followed by the campus targets (baseline-room condition), or the campus targets, followed by the room targets (baseline-campus condition). Thus, the procedure was identical to that in Experiment 1, except that turning to face each target was changed to pointing to each target in the second stage. As a result, the participants still needed to switch from making judgments of one environment to making judgments of another environment, but they did not need to update their heading relative to the two environments. If the difficulty in pointing to campus targets after the participants had turned relative to the room was due to the difficulty in switching from a small environment to a larger one, Experiment 2 should show the same results. On the other hand, if the results in Experiment 1 were due to the nature of the spatial updating system, the difficulty should be eliminated in Experiment 2, because no spatial updating was needed.

Results and Discussion

As in Experiment 1, during the learning stage, responses did not differ between groups [$F_s(1,18) < 0.60$, $p_s > .45$]. Unlike in Experiment 1, however, performance during the testing stage did not differ between the baseline-room group and the baseline-campus group for either the room or the campus targets [$F_s(1,18) < 0.60$, $p_s > .45$] (see Figure 3), even though the participants had to switch from room to campus in the baseline-room group and from campus to room in the baseline-campus group. These data suggest that the difficulty in pointing to campus buildings after turning relative to the room in Experiment 1 was not simply a result of mentally switching from

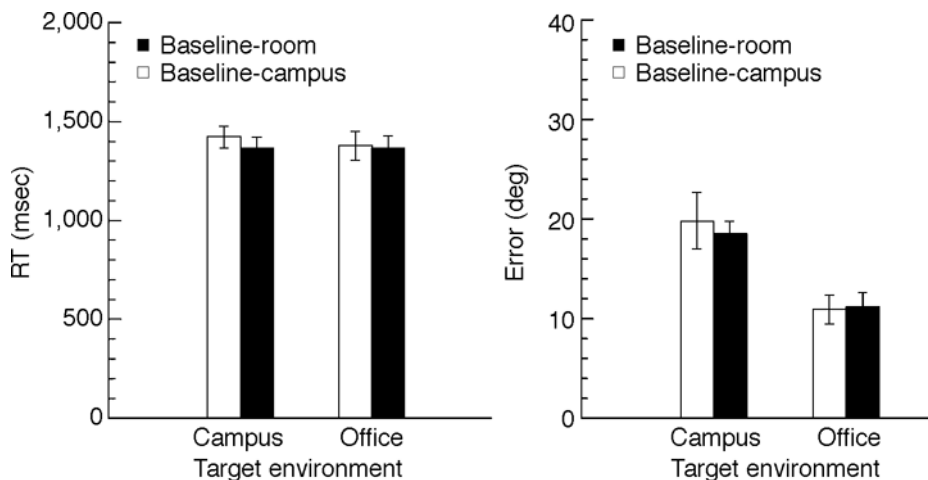


Figure 3. The response times (RTs, left) and errors (right) in Experiment 2. The error bars represent standard errors.

room to campus but, rather, of the need to spatially update target locations during turning.

GENERAL DISCUSSION

In two experiments, we tested whether people can simultaneously update their spatial relationship relative to two nested environments. People automatically updated object locations in their immediate surroundings. However, a more remote environment was not automatically updated; responses to targets in the remote environment were significantly impaired (with respect to both accuracy and RT) when people were instructed to move relative to their immediate surroundings. Thus, it appears that spatial updating does not always operate automatically with respect to all known places simultaneously; what is automatically updated is determined by the nature of the environment.

The room environment and the campus environment differed in several ways. First, the room targets were closer to the participants than were the campus targets. Second, the room targets were small and relevant for one's direct actions, whereas the campus targets were large and relevant for navigation. Third, the campus environment "includes" the room environment, but not vice versa. Finally, the participants learned the room targets visually by looking at several targets at a time, whereas the campus buildings were learned either from a map or from multiple navigation trips between them. Thus, the difference in spatial updating in these two environments could be due to any of these differences or to a combination of them. In ongoing research, it is being investigated whether spatial updating occurs automatically for near targets, for objects relevant for actions, for environments included in the one being updated, or because of the nature of the learning process.

Regardless of the exact source of the observed disparity in spatial updating for the immediate and remote environments tested, these findings have important implications for theories of spatial learning, navigation, and environmental representations. For example, learning the spatial relationship between two places through daily navigation often requires continuous updating of both places as one moves between them: If the spatial-updating system loses track of Place A when one goes to Place B, the navigator will have great difficulty in learning the spatial relationship between the two. Thus, a failure to simultaneously update multiple environments may lead to a failure to learn the spatial relationship across environments and to construct larger integrative mental maps. Studies have shown that spatial learning is indeed particularly difficult for the orientation of an interior room relative to the outside world or the relative orientation of different sections of a building, even after years of continuous navigational experience, when spatial updating is required to connect them (Brockmole & Wang, 2002; Moeser, 1988).

If the spatial relationship between oneself and locations in multiple environments is not updated at the same time, the relationship between the updated environments and the nonupdated environments may change as updating

continues. Our data suggested that failure to simultaneously update multiple environments could indeed result in errors in the relationship between the updated and the nonupdated environments. For example, after turning relative to the room targets, the participants not only were slower in pointing to the campus buildings, but also made more errors. However, their pointing responses to the room targets were not impaired. If the participants preserved the spatial relationship between the two environments, they should have pointed to the targets in both environments with equal accuracy. Thus, the increase in their pointing errors for the campus targets, but not for the room targets, suggests that their pointing responses failed to preserve the correct spatial relationship between the two environments when they failed to update both simultaneously.

These findings raise a theoretical question concerning human spatial representations of complex environments. Do people represent the spatial relationship between environments that they do not update simultaneously? We suggest two possibilities. One possibility is that people may have a long-term memory of the relationships between environments. However, these representations are not directly used to guide actions and, therefore, are not reflected in people's pointing responses. This hypothesis suggests that humans have two systems of spatial representation: one long-term, stable, and more comprehensive representation and one dynamic representation that is updated as one moves. Moreover, the dynamic representation system has a limited capacity. Thus, although people fail to update multiple environments simultaneously and the failure leads to inconsistency in their pointing responses to the two environments, some knowledge of the correct relationship between these environments is, nonetheless, preserved in long-term memory.

Alternatively, humans may represent each environment individually and not encode the spatial relationship between them. That is, the system of spatial knowledge in humans is essentially fragmented (Wang & Brockmole, 2003). It is possible that fragmentation of the spatial representations leads to one environment being updated and the other ones not. However, it is also possible that failure to simultaneously update multiple environments is the fundamental reason that the mental map system that people acquire through navigation is divided into units, instead of being integrative. If the spatial-updating system does not provide one with information about how one part of the world is related to another part, one may have individual mental maps of each part but may not be able to create the appropriate link between them.

The finding that spatial updating does not always occur for multiple environments simultaneously is partly consistent with hierarchical models of human spatial representations (Stevens & Coupe, 1978; see also Hirtle & Jonides, 1985; Huttenlocher, Hedges, & Duncan, 1991; McNamara, 1986; McNamara, Hardy, & Hirtle, 1989; Taylor & Tversky, 1992), in the sense that spatial information of different environments is encoded separately. However, these models in their current format do not ex-

plain the asymmetry between the two environments. In fact, one might expect the opposite: Because the higher level representations are coarser, knowing one's position in a campus map may not tell one exactly where one is in the room map. However, knowing where one is in the room map gives one more precise knowledge, and thus, it is easier to find out where one is on campus. Moreover, the asymmetry is specific to spatial updating, but not to simple switching between the environments. Thus, new assumptions and modifications may be needed in existing hierarchical models to account for the present findings.

REFERENCES

- AMORIM, M. A., & STUCCHI, N. (1997). Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognitive Brain Research*, *5*, 229-239.
- BROCKMOLE, J. R., & WANG, R. F. (2002). Switching between environmental representations in memory. *Cognition*, *83*, 295-316.
- FARRELL, M. J., & ROBERTSON, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 227-233.
- GALLISTEL, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
- HIRTLE, S. C., & JONIDES, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition*, *13*, 208-217.
- HUTTENLOCHER, J., HEDGES, L. V., & DUNCAN, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, *98*, 352-376.
- HUTTENLOCHER, J., & PRESSON, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology*, *11*, 375-394.
- LOOMIS, J. M., DA SILVA, J. A., PHILBECK, J. W., & FUKUSIMA, S. S. (1996). Visual perception of location and distance. *Current Directions in Psychological Science*, *5*, 72-77.
- MAY, M. (1996). Cognitive and embodied modes of spatial imagery. *Psychologische Beiträge*, *38*, 418-434.
- MCMANARA, T. P. (1986). Mental representations of spatial judgments. *Cognitive Psychology*, *18*, 87-121.
- MCMANARA, T. P., HARDY, J. K., & HIRTLE, S. C. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *15*, 211-227.
- MOESER, S. D. (1988). Cognitive mapping in a complex building. *Environment & Behavior*, *20*, 21-49.
- O'KEEFE, J., & NADEL, L. (1978). *The hippocampus as a cognitive map*. Oxford: Oxford University Press, Clarendon Press.
- 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, & Cognition*, *15*, 1157-1165.
- RIESER, J. J., GARING, A. E., & YOUNG, M. F. (1994). Imagery, action, and young children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Development*, *65*, 1262-1278.
- SIMONS, D. J., & WANG, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science*, *9*, 315-320.
- STEVENS, A., & COUPE, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology*, *10*, 422-437.
- TAYLOR, H. A., & TVERSKY, B. (1992). Descriptions and depictions of environments. *Memory & Cognition*, *20*, 483-496.
- WANG, R. F. (1999). Representing a stable environment by egocentric updating and invariant representations. *Spatial Cognition & Computation*, *1*, 431-455.
- WANG, R. F., & BROCKMOLE, J. R. (2003). Human navigation in nested environments. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 398-404.
- WANG, R. F., & SIMONS, D. J. (1999). Active and passive scene recognition across views. *Cognition*, *70*, 191-210.
- WANG, R. F., & SPELKE, E. S. (2000). Updating egocentric representations in human navigation. *Cognition*, *77*, 215-250.

NOTES

1. The campus buildings were familiar to the participants through daily navigation and were invisible from the room. The pointing-correction procedure ensured that the campus targets were learned to criteria.

2. The angular errors after the turning stage were corrected for their turning errors. For example, after making five turns, some errors would have accrued. Thus, the participants might think that they were facing the VCR, but in reality, they were 30° off. This error would be present in every trial in the last two sessions. This common turning error was removed from the angular error reported here.

(Manuscript received May 28, 2002;
revision accepted for publication October 2, 2002.)