

Updating representations of learned scenes

Cory A. Finlay · Michael A. Motes ·
Maria Kozhevnikov

Received: 15 November 2005 / Accepted: 11 March 2006 / Published online: 19 October 2006
© Springer-Verlag 2006

Abstract Two experiments were designed to compare scene recognition reaction time (RT) and accuracy patterns following observer versus scene movement. In Experiment 1, participants memorized a scene from a single perspective. Then, either the scene was rotated or the participants moved (0°–360° in 36° increments) around the scene, and participants judged whether the objects' positions had changed. Regardless of whether the scene was rotated or the observer moved, RT increased with greater angular distance between judged and encoded views. In Experiment 2, we varied the delay (0, 6, or 12 s) between scene encoding and locomotion. Regardless of the delay, however, accuracy decreased and RT increased with angular distance. Thus, our data show that observer movement does not necessarily update representations of spatial layouts and raise questions about the effects of duration limitations and encoding points of view on the automatic spatial updating of representations of scenes.

Introduction

Often people find themselves coming upon a familiar scene or even a recently viewed scene from an orientation

that differs from their orientation when they initially encoded the scene (e.g., when one is forced to detour from a previously taken route). Several scene recognition studies suggest that an observer's mental representation of a scene is often viewpoint-dependent (e.g., Christou & Buelthoff, 1999; Diwadkar & McNamara, 1997; Shelton & McNamara, 2004; for an exception, see Mou & McNamara, 2002), that is, oriented to the perspective that the observer had when studying the scene. For example, Diwadkar and McNamara (1997) had observers first memorize a scene and then view rotated pictures (from 0° to 345°) of the same scene or pictures of different scenes (i.e., the spatial configuration of objects in the learned scene were rearranged). Response times to recognize whether the pictures were of the same or a different scene increased with angular distance between the orientation of the memorized and judged scenes.

Although there is evidence that representations of scenes are viewpoint-dependent, several studies suggest that processes associated with locomotion, even if the person is blindfolded, automatically update the person's representation of the scene to be spatially consistent with the person's perspective of the scene after moving (Farrell & Robertson, 1998; Farrell & Thomson, 1998; Reiser, 1989). Reiser (1989), for example, had observers memorize the locations of nine objects that surrounded them, then rotate within the array of objects to a target facing direction while blindfolded, and finally point to a target object from that facing direction. The observers were equally fast at pointing to targets after rotating to the new orientation, regardless of the magnitude of the rotation, as they were when pointing to targets from the encoding orientation. Thus, the data suggest that the observers'

C. A. Finlay · M. A. Motes
Department of Psychology, Rutgers,
The State University of New Jersey, Newark, NJ, USA

M. Kozhevnikov (✉)
Department of Psychology, George Mason University,
4400 University Drive, MSN 3F5, Fairfax, VA 22030, USA
e-mail: mkozhevn@gmu.edu

representations were transformed during the rotation to match the new orientation. Farrell and Robertson (1998) essentially replicated Rieser's (1989) findings, but also showed that when instructed to ignore the rotation, RT and error increased with angular distance, thus showing that the effects of updating are automatic in that they cannot be ignored (but see Waller, Montello, Richardson, & Hegarty, 2002).

Rather than directional judgments, Simons and Wang (1998), Wang and Simons (1999) studied the effects of observer movement and scene movement on scene recognition. On each trial, observers were shown a 5-object scene, followed by a brief delay. Observers were then shown the same scene except that one of the objects was moved to a different location, and the observers verbally reported the name of the moved object. Across a series of studies, observers were less accurate at identifying scene changes when they remained at the encoding station but the scene was rotated (47°) from the encoding orientation than when the observers remained at the encoding station and the scene remained at the encoding orientation. However, observers were equally accurate at identifying scene changes when they moved (47°) from the encoding station but the scene remained at the encoding orientation and when they moved (47°) from the encoding station and the scene was also rotated (47°) from the encoding orientation (or slightly more accurate in the former condition, Wang & Simons, 1999, Experiment 1). Thus, according to the spatial updating hypothesis, the observers' mental representations of the encoded scenes were spatially updated as the observers moved around the scene, but such spatial updating did not occur when the scenes were rotated but the observers remained at the encoding station.

Simons and Wang (1998, Experiment 2; Wang & Simons, 1999, Experiment 1), however, also reported data that were inconsistent with a strict spatial updating model. They reported lower change detection accuracy when observers moved from the encoding station and the judged scenes remained at the encoding orientation than when observers remained at the encoding station and judged scenes that remained at the encoding orientation. Although Simons and Wang argued that this effect of observer movement on scene recognition accuracy was "weaker than the effect of scene rotation," this particular comparison is an essential comparison for testing whether full spatial updating has occurred. If fully updated, then observers should be equally accurate and equally fast when judging scenes following movement as when they remain stationary (e.g., Rieser, 1989).

Other studies have also failed to find support for a strict, automatic, spatial updating model (Mou, McNamara, Valiquette, & Rump, 2004; Motes, Finlay, & Kozhevnikov, *in press*). Mou et al. (2004), for example, had observers learn a scene, move into the scene, make an orientation change (either 0° , 90° , or 225°), and finally, close their eyes and point to a target. The blindfolded observers were slower at pointing to targets after rotating 225° from the orientation in which they memorized the scene than when they did not make orientation changes or when they made 90° changes. Thus, the observers' representations were not updated, at least not fully updated, following the 225° orientation changes.

Motes et al. (*in press*) had observers learn a spatial layout of 11 objects from a single perspective. After learning the layout, observers judged whether subsequently presented scenes were the same as or different from the encoded scene (half of the trials had the same configuration as the encoded scene; half had different configurations in that the positions of two objects were switched). For the scene recognition judgments, half of the observers moved to one of ten viewing stations, from 0° to 360° in 40° increments, from the encoding station while the scene remained stationary, and the other observers remained at the encoding station but the scene was rotated to one of nine orientations, from 0° to 320° in 40° increments, from the encoding orientation. Regardless of whether the scene was moved or the observer moved, scene recognition RT increased with angular distance between the encoded and judged views and performance between the groups did not significantly differ for any of the angular distances. Thus, the data suggest that the observer movement group's representations of the scene were not automatically updated as they moved around the scene.

These failures to find evidence of automatic spatial updating might have been due to observers relying on allocentric rather than egocentric representations when judging the scenes. According to Mou et al. (2004), the allocentric, environment-based system holds enduring representations of relatively familiar environments. The system codes the geometry of the surroundings and the spatial relations of objects with respect to other objects. The allocentric system also codes self-to-object spatial relations, but the self is merely treated as another object in the environment. The egocentric system, on the other hand, holds transient representations of one's relatively immediate environment. The system codes the spatial relations between oneself and objects in the environment with respect to one's prominent body axes: left-right, front-back, above-below. Furthermore, Mou et al. suggested that egocentric representations begin to

fade with delays as brief as 10 s between encoding and locomotion, and thus, their maintenance requires on-line sensory feedback. Mou et al. further argued that automatic spatial updating requires these egocentric representations and that observers would be forced to rely on enduring, allocentric representations (if available) if the egocentric representations fade.

Observers in our previous study (Motes et al., *in press*) memorized a single scene at the start of the experiment, and they were required to rely on their memory of that scene throughout the duration of the experiment (i.e., they were not explicitly given the opportunity to refresh their memories of the scene). Therefore, the observers' egocentric representations might have faded over the duration of the scene recognition trials. If so, observers were then forced to use enduring, allocentric representations of the scene that were not updated when they moved, and thus showed viewpoint-dependent response patterns. In fact, Mou et al. (2004) also argued that observers in their study might have been forced to use allocentric rather than egocentric representations because the observers were delayed approximately 10 s from pointing to a target after moving within the scene. Furthermore, studies that have found evidence of observer movement leading to spatially updated representations have not had such delays (e.g., Avraamides, 2003; Farrell & Robertson, 1998; Reiser, 1989). Thus, the goal of our present work was to systematically examine the effects of delaying observer movement, after encoding a scene, on the observers' scene recognition RT and accuracy profiles over a range of angles, and we compared these effects to the effects of delaying observers from viewing equivalent rotated versions of the same scenes.

Experiment 1

Prior to systematically examining the effects of delaying locomotion, we re-examined the effect of observer movement, over a range of angles around a scene, on scene recognition RT and accuracy. We also compared the effects of observer movement to the effects of equivalent viewpoint changes produced by scene movement. Similar to our previous work (Motes et al., *in press*), observers learned a spatial layout of objects from a single perspective, and after learning the layout, observers made scene recognition judgments on subsequently viewed scenes. The views of the judged scene were varied by having observers either walk to viewing stations (0° – 360°) placed around the scene or by having observers remain at the encoding station and rotating the scenes (0° – 360°) from the encoding orientation.

In our previous experiment (Motes et al., *in press*), observers exited and re-entered the room between trials so that they would not hear adjustments being made to the scene. However, Wang and Spelke (2000) have shown that self-rotation (compared to not rotating) for a short period after learning a scene negatively affects self-to-object pointing judgments, that is, impairs egocentric representations. Although we had observers begin each trial at the encoding station in order to facilitate re-orientation, the use of the encoding station as a re-orientation landmark might not have been effective if exiting and re-entering was akin to self-rotation in that it impaired the observers' egocentric representations. To rule out this plausible alternative explanation, in the current study we changed our previous methodology so that observers remained in the testing room throughout the duration of the experiment.

Method

Participants

Fifty-one students from Rutgers University participated in the experiment in exchange for course credit or monetary compensation. Twenty-six were randomly assigned to the scene movement condition, and 25 were randomly assigned to the observer movement condition.

Stimuli and apparatus

The experiment was conducted in a darkened room. A Dell 17" LCD flat-panel monitor was laid horizontally on a circular table (radius = 54 cm). A black cloth covered the monitor and table, except for a circle (radius = 11 cm) that was cut out of the black cloth that was centered over the viewing screen of the monitor, within which all of the scenes were shown. The black cloth prevented the outer edges of the monitor and the inner edges of the viewing screen from being used as frames of reference when learning and judging the scenes.

Viewing stations were created with ten occluding screens (44.6×180 cm) that surrounded the monitor and table. Each screen had a horizontally centered viewing window (5.8×12 cm) cut into the screen 140 cm from the floor and 59 cm from the top of the viewing screen of the monitor. The windows were 36° apart. From the center of a viewing window, the visual angle of the display was approximately 12° . To eliminate the potential use of the ceiling and the walls of the room as frames of reference, the ceiling of the room was covered with black cloth, and black cloth,

hanging from the ceiling to the floor, encircled the occluding screens and an approximately 58 cm wide path through which the experimenter and observers could walk.

On one of the screens, a *Start* sign was placed above the viewing window. For both the scene rotation and observer movement groups, this was the station from which the scene that was to-be-remembered throughout the duration of the experiment was encoded. In the scene rotation condition, observers remained at this station throughout the study, whereas in the observer movement condition, observers began each trial at this station, and on trials in which they judged the scenes after moving, they moved counterclockwise from this station. The encoded scene consisted of a computerized display of ten objects (see Fig. 1), and to test whether observers had indeed learned the scene, the observers recreated the scene with cutouts of the objects on a magnetic clipboard.

Sixty-six scenes were judged by the observers. Thirty-three variations of the 10-object encoded scene were created. For these scenes the locations of two randomly selected objects were switched with each other (e.g., see Fig. 1). Each of these 33 scenes and each of the 33 copies of the encoded scene were then randomly assigned to 1 of 11 viewing angles (0° – 360°). That is, for the scene rotation condition, 27 *different* and 27

same scenes were rotated, clockwise between 36° and 324° from the encoded scene orientation. The other 12 scenes (6 *same* and 6 *different*) were kept at the encoded scene orientation and served as 0° (3 *same* and 3 *different*) observer and scene movement scenes and 360° (3 *same* and 3 *different*) observer and scene movement scenes. For the observer movement condition, all 66 scenes were kept at the same orientation as the encoded scene. The same scenes were used for the equivalent scene movement and observer movement viewing angles.

For the scene recognition task, observers in both conditions used a handheld, wireless, radio-frequency, computer mouse to indicate whether the scenes were the same as the encoded scene (right mouse button = *same* and left = *different*). An E-Prime (Psychological Software Tools, Version 1.1) program presented the scenes and recorded RT and accuracy. When recording RT, the program started a software clock when the judged scene was shown and stopped the software clock when one of the mouse buttons was pressed. Pressing the *same* button triggered the program to play a digital audio file saying “same”. Pressing the *different* mouse button triggered the program to play a digital audio file saying “different; which objects were switched?” The experimenter recorded the observers’ verbal responses on a trial record sheet, and

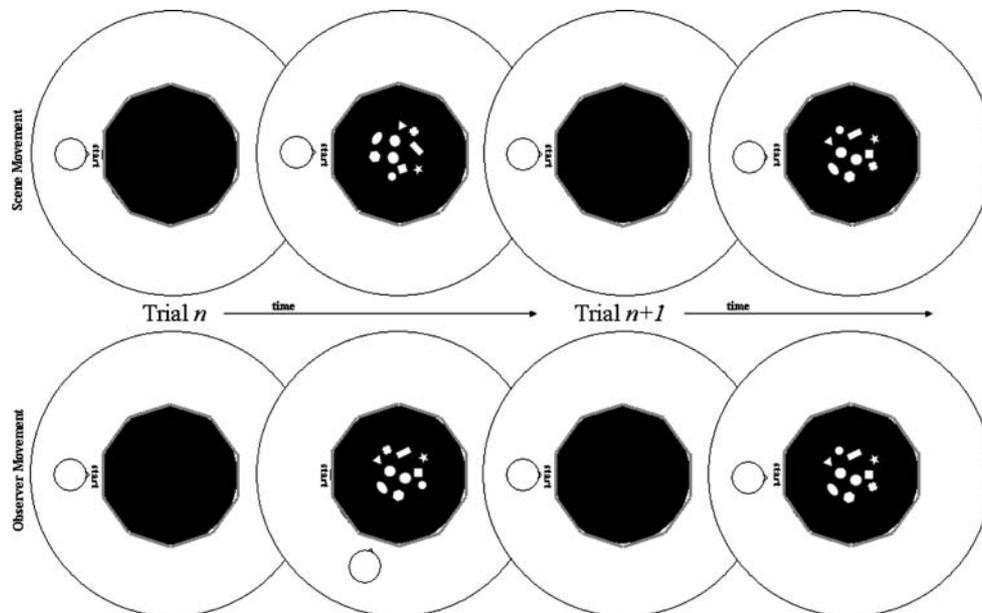


Fig. 1 Diagram of scene recognition trials for Experiment 1. The diagram depicts two viewpoint-equivalent trials (Trial n and Trial $n + 1$) for the observer and scene movement groups. On each trial, observers initially waited at the encoding station (*start*) for 3 s. Observers in the scene movement group remained at the encoding station throughout the experiment and judged scenes that were rotated (0° – 324° in 36° increments) from the encoded scene

orientation. Observers in the observer movement group walked (0° – 360° in 36° increments) to a new viewing station on each trial and made scene recognition judgments. Trial n depicts observer and scene movement viewpoint-equivalent, 72° , *different* trials, and Trial $n + 1$ depicts subsequent observer and scene movement viewpoint-equivalent, 0° , *same* trials

the experimenter used a small flashlight to see when recording responses, for the observers to see and arrange the cutouts, and to maneuver around the path.

Procedure

Both the scene and observer movement groups were told that they would first have to learn the spatial arrangement of a scene composed of geometric objects. After entering the testing room but before learning the encoded scene, the observers were disoriented to reduce the possibility of their using the door or other stimuli external to the testing room as spatial reference points. The observers closed their eyes and were led counterclockwise by the experimenter partially around the occluding screens, then rotated in place three times, led clockwise partially back, rotated three times, led counterclockwise partially around again, rotated three times, and then opened their eyes and attempted to point to the door. However, none accurately pointed to the door.

Observers were then led by the experimenter to the encoding station where they learned the layout of the scene that they were to remember for the duration of the study. Observers were shown the scene for 1 min. They counted “1-2-3” aloud to interfere with their use of verbal coding strategies. After 1 min, the scene disappeared, and the observers were asked to recreate the scene. They repeated this procedure until they accurately reproduced the scene.

After learning the scene, the observers were told that they would then be shown scenes that either would be the same as the encoded scene or different from the encoded scene in that the locations of two of the objects in the scene would be switched (this was demonstrated with the cutouts). Observers in the scene movement group were then told that the judged scenes might be rotated from the orientation of the encoded scene (this was also demonstrated with the cutouts) but that they were to judge the scene as different only if the locations of two objects were switched, regardless of whether the entire scene was rotated or not. Observers in the observer movement condition were told that the orientation of the scene would remain the same throughout the duration of the experiment.

Observers completed five practice trials randomly selected from the 66-scene/angle combinations and then completed the 66 scene recognition trials. Each observer worked through the trials in one of two fixed random orders, and the observers were instructed to respond both quickly and accurately. Testing lasted 1–1.5 h.

At the beginning of each trial, observers in the observer movement group waited at the encoding

station for 3 s (see Fig. 1). A tone then played cueing the experimenter to escort the observer to the scene judgment station listed on the experimenter’s trial record sheet. After reaching the judgment station, the observer pressed the left mouse button to view the judged scene and then entered a scene recognition judgment. The observer returned to the encoding station, and pressed the mouse button to begin the next trial.

Observers in the scene movement group remained at the encoding station throughout the duration of the experiment (see Fig. 1). They also waited through a 3 s delay between each trial. After 3 s, a tone played, and the experimenter directed the observer to press the left mouse button to view the scene. The observer then entered the scene recognition judgment.

Results

Scene recognition accuracy

For accuracy, responses with RTs below 750 ms (0.09%, 3 of 3,366, of the total trials) were deleted; these were trials in which the observers inadvertently pressed the response button before actually viewing the judged scenes. For the remaining data, the proportion of correct responses for each viewing angle was then calculated for each observer, and a mixed-model ANOVA (2 Movement Group X 11 Angular Distance) was calculated using these accuracy data. Mean accuracy as a function of movement group and angular distance is shown in Fig. 2a.

ANOVA revealed a significant effect of angular distance, $F(10, 490) = 3.71$, $p < 0.01$, but neither the main effect of movement group nor the interaction were significant, $F(1, 49) < 1$ and $F(10, 490) < 1$, respectively. However, upon examining the means per group and angle, we found that the means were significantly lower in the 0° condition than in the equivalent 360° for both the scene and observer movement groups, $t(25) = 3.08$, $p < 0.01$, and $t(24) = 3.07$, $p < 0.01$, respectively. Upon further investigation, we discovered below chance performance for only one of the 0° condition *different* scenes (6 of 13 observers in the observer movement group and 6 of 13 observers in the scene movement group answered correctly), a scene in which the octagon and hexagon were switched. This scene was also used on the other stimulus list, but the scene was viewed 216° from the encoded orientation. At the 216° viewing angle, scene recognition performance was also below chance (4 of 12 observers in the observer movement group and 4 of 13 observers in the scene movement group answered correctly). No other scenes

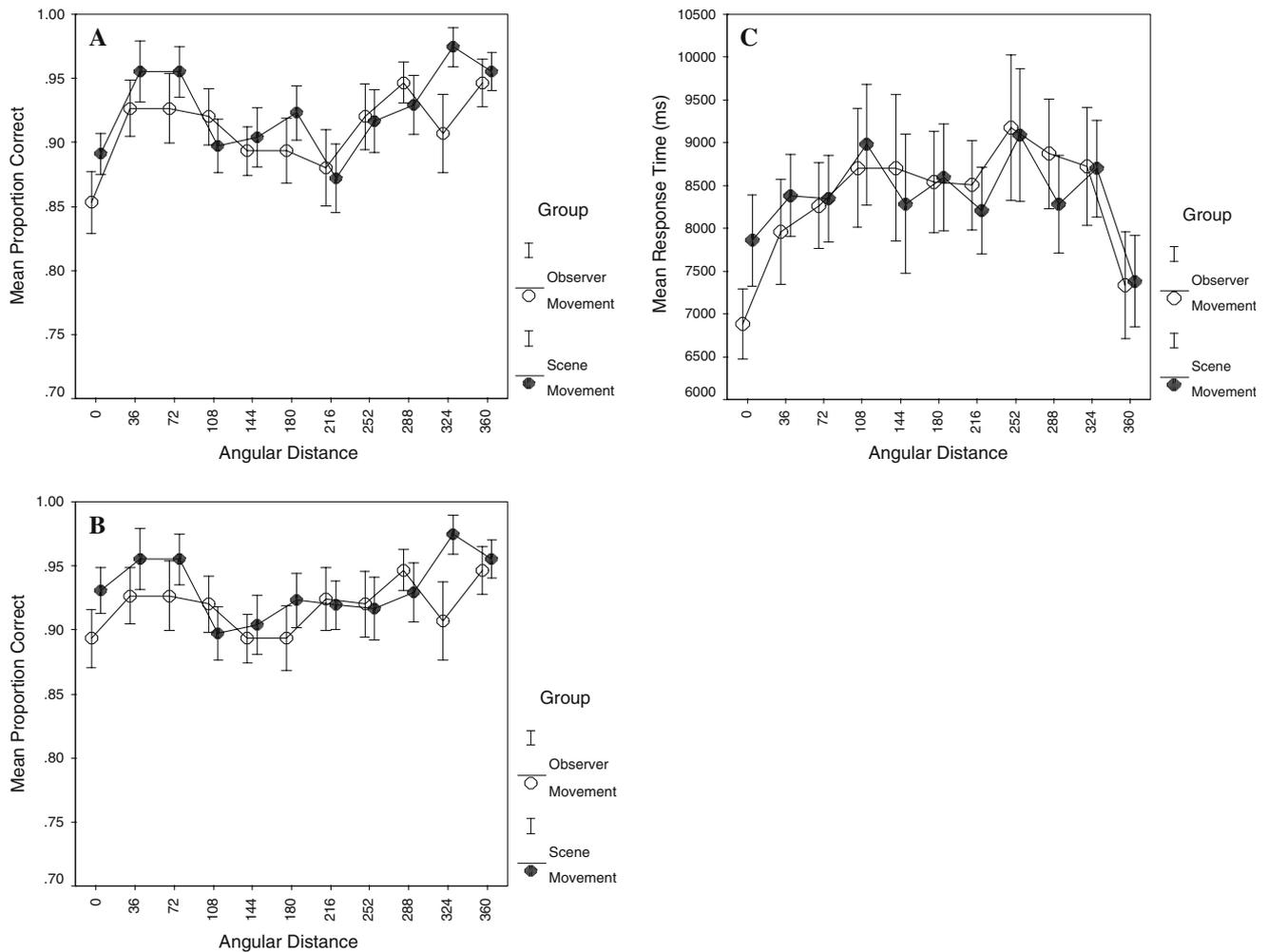


Fig. 2 Scene recognition accuracy and response time as a function of movement group and angular distance. The *open circles* show the data for observers in the observer movement group, and the *filled circles* show the data for observers in the scene move-

ment group. **a** Shows scene recognition accuracy for all trials, **b** shows scene recognition accuracy after deleting trials in which the octagon and hexagon were switched, and **c** shows scene recognition response time

showed similar below chance performance patterns. Together, these data suggested that the mean for the 0° condition and for the 216° condition were artificially low due to the inclusion of these difficult to discriminate trial types. Thus, we removed these trials and reanalyzed the accuracy data. Mean accuracy as a function of movement group and angular distance is shown in Fig. 2b.

After removing the trials in which the octagon and hexagon were switched, neither the main effect of angular distance, the main effect of movement group, nor the interaction were significant, $F(10, 490) = 1.59$, $p > 0.1$, $F(1, 49) < 1$, and $F(10, 490) < 1$, respectively. Thus, consistent with the results from our previous experiment (Motes et al., *in press*), accuracy did not significantly differ with angular distance for either the scene or observer movement groups.

Scene recognition RT

Scene recognition RTs for each observer were screened for outliers. First, RTs below 750 ms (0.09%, 3 of 3,366, of the total trials) and above 30 s (1.16%, 39 of 3,366, of the total trials) were deleted. Then, RTs ± 2.5 SDs from an individual observer's mean RT were deleted (observer movement = 2.14%, 35 of 1,632, and scene movement = 2.42%, 41 of 1,692). After deleting outlier RTs, mean RTs for correct responses for each viewing angle were calculated for each observer, and a mixed-model ANOVA (2 Movement Group X 11 Angular Distance) was calculated using these data. Mean RT as a function of movement group and angular distance is shown in Fig. 2c.

ANOVA revealed the main effect of angular distance was significant, $F(10, 490) = 5.77$, $p < 0.01$, but

neither the main effect of movement group nor the interaction were significant, $F(1, 49) < 1$ and $F(10, 490) < 1$, respectively. Trend analyses of the effect of angular distance revealed that the quadratic component was significant, $F(1, 49) = 29.90$, $p < 0.01$. To further examine this quadratic trend, the linear components from 0° to 180° and from 180° to 360° were analyzed. The linear trend components were significant for both the effect of angular distance from 0° to 180° , $F(1, 49) = 10.31$, $p < 0.01$, and from 180° to 360° , $F(1, 49) = 10.37$, $p < 0.01$, indicating that the increase in RT for both groups was greater with larger angular distances.

Thus, consistent with our previous findings (Motes et al., *in press*), regardless of whether the scene was moved or the observer moved, the analyses showed that RT increased with angular distance from the encoded view. Furthermore, although the movement group by angular distance interaction was not significant, we explicitly examined RT differences between the scene and observer movement groups at each angular distance, but none of the differences were significant, all $ps > 0.05$. Given the increase in RT with angular distance for both groups and the failure to find differences between the groups for any of the angular distances, the data suggest that both groups formed viewpoint-dependent representations of the encoded scene and possibly performed mental rotation transformations to compare their viewpoint-dependent representations of the encoded scene to the judged scenes when the angular distances varied.

Experiment 2

In Experiment 2, we systematically investigated the effect of delays following scene encoding and observer movement. As in Experiment 1, observers were assigned to either an observer or scene movement group. Different from Experiment 1, however, observers viewed a different 4-object scene on each trial. Then, following a 0, 6, or 12 s delay, observers viewed the scene from either the same or a different perspective (i.e., observers in the observer movement group walked around the scene or observers in the scene movement group viewed rotated scenes) and judged which object in the scene had been moved. To examine the effects of delays above and below 10 s, we used a short-term retention procedure rather than the long-term retention procedure used in Experiment 1. The scene set-size was also reduced to four objects, from 10 used in Experiment 1, to be consistent with the constraints of visual working memory (Luck & Vogel,

1997). The procedure and set-size were consistent with those used in studies that previously reported finding locomotion-induced spatial updating effects on scene recognition (Simons & Wang, 1998; Wang & Simons, 1999).

Method

Participants

Thirty students from Rutgers University participated for course credit. Fifteen participants were randomly assigned to the scene movement condition and fifteen participants were randomly assigned to the observer movement condition.

Stimuli and apparatus

The arrangement of the testing room and apparatus were identical to Experiment 1, with the exception of the stimuli. Thirty-six different 4-object, computerized, encoded scenes were designed. For each scene, four objects were randomly selected from a set of 12, easily identifiable objects (snowflake, pound sign, square, flower, wheel, cross, star, triangle, octagon, sun, circle and donut, see Fig. 3a, each measuring from 2.5 to 3.5 cm in height and 2.2 m to 3.5 cm in width). Each object was then placed in one randomly selected location of 21 predefined locations (see Fig. 3b). For each encoded scene, a judged scene was created by randomly selecting one object from the encoded scene and moving that object to one of the randomly selected 17 unoccupied spaces. Each of the 36 pairs of encoded and judged scenes were then randomly assigned to one of six viewing angles (0° – 180° in 36° increments). For the scene movement condition, 30 scene pairs were rotated clockwise to one of the five rotated viewing angles. The remaining six pairs served as 0° scene and observer movement scenes.

Three randomly arranged lists of the 36 trial pairs were then created. Observers in both groups worked through three blocks of 36 trials (i.e., each list), one block for each delay condition. Trial lists and delay conditions were counterbalanced using a 3×3 Latin Square design. Testing lasted approximately 1 h.

Procedure

Observers in the scene movement group were told that the judged scenes might be rotated from the orientation of the encoded scenes but that they were to identify the single object moved within the scene, and on each trial, they were told the magnitude and direction

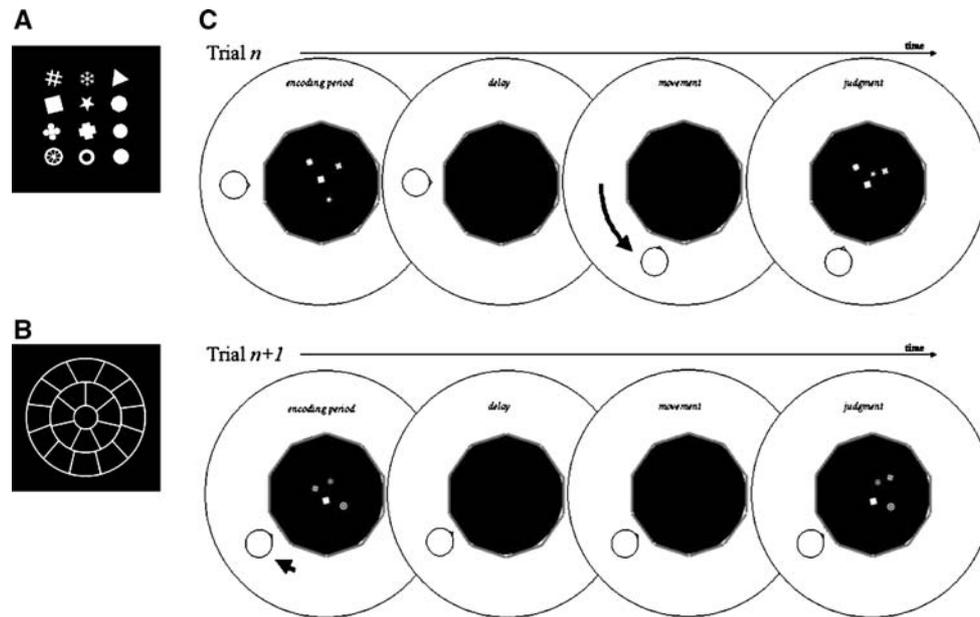


Fig. 3 Diagram of stimuli and observer movement trials for Experiment 2. **a** Shows the 12 objects used to create the 4-object scenes, and **b** shows the template used to define the object locations. **c** Depicts two trials for the observer movement group. On each trial, an observer had 3 s to encode a scene that was followed by a delay period of 0, 6, or 12 s. Then, the observer walked (0° – 180° in 36° increments) to a new viewing station and made a scene

of the scene rotation. Observers in the observer movement group were told that the orientation of the scene would remain the same regardless of whether they moved from the encoding station or not. Observers were then taken into the testing room, put through the disorientation procedure used in Experiment 1, and then led to the first encoding station.

On each trial, observers viewed a 4-object scene for 3 s. Then, the screen went blank for a delay period of 0, 6, or 12 s, depending on the trial block. After which, a tone played to indicate that the delay period had passed. Observers in the scene movement group then pressed the right mouse button to view the judged scene. Observers in the observer movement group were led by the experimenter to the judgment station and instructed to press the right mouse button to view the judged scene. Observers in both groups pressed the right mouse button to indicate when they recognized which object had been moved, and they verbally reported the object to the experimenter. Observers were told whether there would be a delay before beginning each block, but they were not told the duration of the delay. Observers completed four practice trials at the beginning of each block.

Observers in the observer movement group did not return to the original encoding station after judging the scene but instead moved one viewing station to the left

recognition judgment. At the conclusion of each trial, the observer walked one viewing station to the left to begin the next trial. Observers in the scene movement group also changed viewing stations between trials to encode the scenes from the same stations as observers in the observer movement group, thus working through viewpoint-equivalent sets of trials

of the judgment station to encode the next scene. This change was made to reduce the total testing time. Observers in the scene movement group also did not remain at the original encoding station throughout the experiment. They, instead, changed viewing stations between trials to encode the scenes from the same stations as observers in the observer movement group, thus working through viewpoint-equivalent sets of trials (see Fig. 3c).

Results

Scene recognition accuracy

For accuracy, responses with RTs below 750 ms (0.22%, 7 of 3,240, of the total trials) were deleted; these were trials in which the observers inadvertently pressed the response button before actually viewing the judged scenes. For the remaining data, the proportion of correct responses for each viewing angle was then calculated for each observer, and a mixed-model ANOVA (2 Movement Group X 3 Delay X 6 Angular Distance) was calculated using these accuracy data. Mean accuracy as a function of movement group by delay by angular distance is shown in Fig. 4a, b.

ANOVA revealed a significant effect of angular distance, $F(5, 140) = 20.77$, $p < 0.01$, and a significant

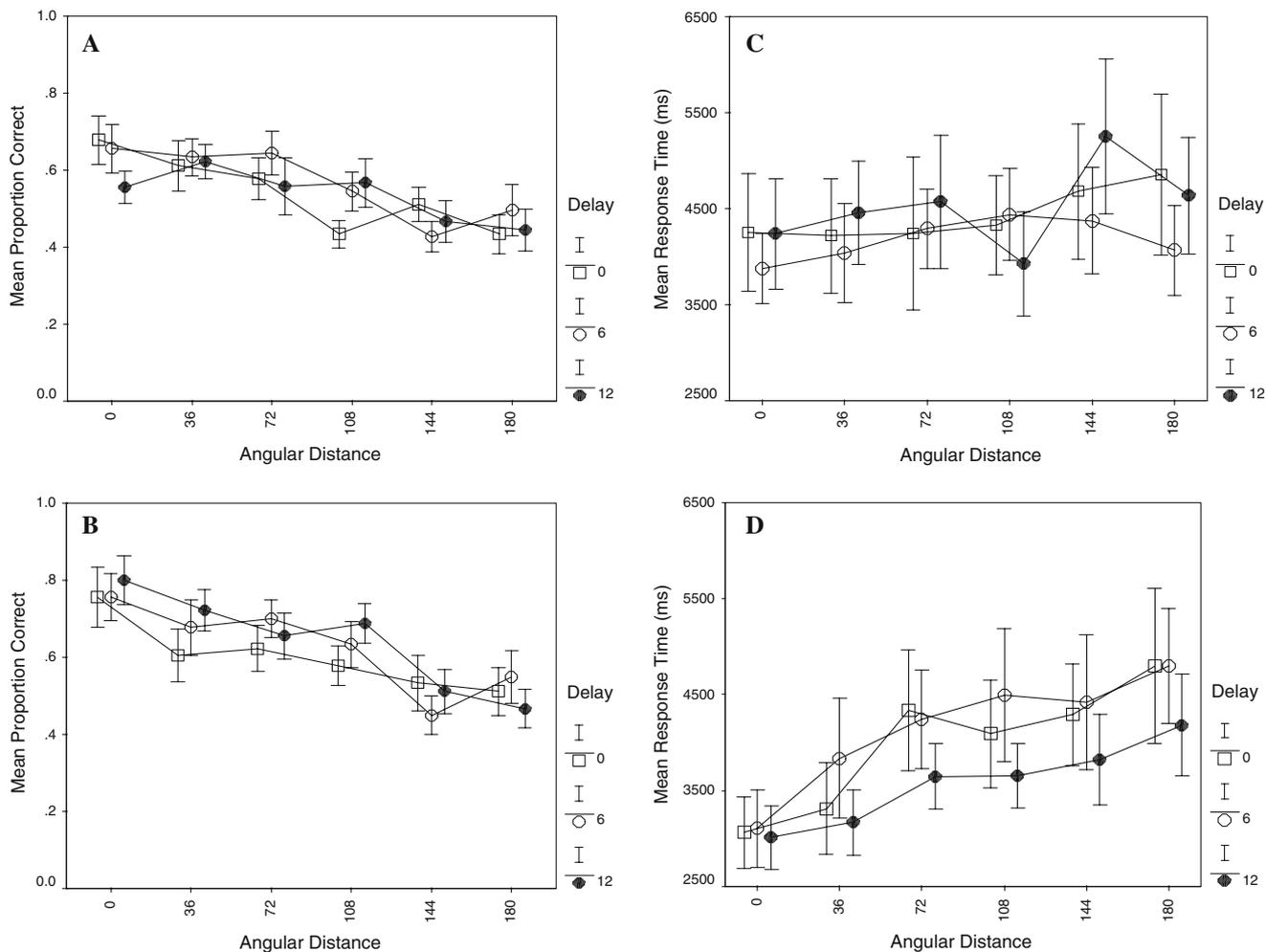


Fig. 4 Scene recognition accuracy and response time as a function of movement group, delay, and angular distance. **a, b** Show the accuracy for the scene and observer movement groups, respectively, and **c, d** show the RT for the scene and observer

movement groups, respectively. *Open squares* represent the 0 s delay condition, *open circles* represent the 6 s delay condition, and *filled circles* represent the 12 s delay condition

angular distance by delay interaction, $F(10, 280) = 2.09$, $p < 0.05$, but the main effects of delay, $F(2, 27) < 1$, and movement group, $F(1, 28) = 1.94$, $p > 0.05$, and the angular distance by movement group, $F(5, 140) = 1.30$, $p > 0.05$, delay by movement group, $F(2, 27) < 1$, and angular distance by delay by movement group, $F(10, 280) < 1$, interactions were not significant. Trend analyses of the effect of angular distance revealed that the linear trend component was significant, $F(1, 28) = 98.30$, $p < 0.01$. Thus, unlike our previous findings (Experiment 1 above and Motes et al., *in press*), regardless of whether the scene was moved or the observer moved, accuracy decreased with increasing angular distance from the encoded view. Thus, the type of retention paradigm, short-term versus long-term, appeared to affect the representation available for making the scene recognition judgments. This difference across studies, however, must be treated with

caution because the judgment tasks also differed. Finally, the finding that the impairment effects of observer movement on scene recognition increased with angular distance suggests that previously reported mixed findings (Simons & Wang, 1998; Wang & Simons, 1999) might have been due to the relatively short angular distances used (i.e., impairment effects at 47° should be weaker than impairment effects at larger angular distances).

Follow-up analyses of the angular distance by delay interaction revealed that the effects of angular distance for the three delay conditions were significant, 0 s $F(5, 145) = 7.63$, $p < 0.001$, 6 s $F(5, 145) = 11.80$, $p < 0.001$, and 12 s $F(5, 145) = 9.51$, $p < 0.001$. The linear trend components for all three delay conditions were significant, 0 s $F(1, 29) = 31.27$, $p < 0.001$, 6 s $F(1, 29) = 53.99$, $p < 0.001$, and 12 s $F(1, 29) = 32.59$, $p < 0.001$. Thus, regardless of the length of the delay and whether the

observer or scene moved, accuracy decreased with angular distance. Therefore, the analyses of the accuracy data did not reveal that delays affected the observers' representations. Further analyses of the angular distance by delay interaction revealed a significant difference only at 108°, $F(2, 58) = 4.40$, $p < 0.05$. At 108°, observers in both groups were less accurate in the 0 s delay condition ($M = 0.50$) than in either the 6 s ($M = 0.59$) or 12 s ($M = 0.63$) conditions, $t(29) = 2.09$, $p < 0.05$, and $t(29) = 3.29$, $p < 0.01$, respectively. This finding, however, was not a part of a systematic pattern, making the interpretation of this result unclear.

Scene recognition RT

Scene recognition RTs for both groups were screened for outliers. RTs below 750 ms (0.22%, 7 of 3,240, of the total trials) and above 30 s (0.09%, 3 of 3,240, of the total trials) were deleted. RTs ± 2.5 SDs from an individual observer's mean RT were deleted (observer movement = 2.23%, 36 of 1,615, and scene movement = 2.6%, 42 of 1,615). After deleting outlier RTs, mean RTs for correct responses for each viewing angle and delay period were calculated for each observer and a mixed-model ANOVA (2 Movement Group X 3 Delay X 6 Angular Distance) was calculated. Four observers were excluded from further analyses, two from each group, due to incomplete data sets (i.e., incorrectly responding to all six trials for at least one of the delay by angular distance conditions). Mean RT as a function of movement group by delay by angular distance is shown in Fig. 4c, d.

ANOVA revealed a significant effect of angular distance, $F(5, 120) = 5.79$, $p < 0.001$, and a marginally significant angular distance and movement group interaction, $F(5, 120) = 1.91$, $p = 0.10$, but the main effects of delay, $F(2, 48) < 1$, and movement group, $F(1, 24) < 1$, and the angular distance by delay, $F(10, 240) < 1$, the delay by movement group, $F(2, 48) = 1.07$, $p > 0.10$, and angular distance by delay by movement group, $F(10, 240) < 1$, interactions were not significant. Trend analyses of the effect of angular distance revealed that the linear component was significant, $F(1, 24) = 19.75$, $p < 0.001$. However, follow-up tests of the marginally significant angle by movement group interaction revealed that the effect of angular distance was significant for the observer movement group, $F(5, 60) = 9.32$, $p < 0.001$, but not the scene movement group, $F(5, 60) < 1$. Trend analyses for the observer movement group revealed that the linear trend component was significant, $F(1, 12) = 31.17$, $p < 0.001$ (for the scene movement group, $F(1, 12) = 2.02$, $p = 0.18$). Thus, consistent with our previous findings (Experiment 1 and

Motes et al., [in press](#)), when the observers moved around the scenes, RT increased with angular distance from the encoded view.

Importantly, the length of delay between scene encoding and observer movement did not have a detectable effect on scene recognition accuracy or RT. Thus, the data do not support our hypothesis that our previous failure to find evidence of locomotion-induced spatial updating was due to delays between scene encoding and locomotion (Motes et al., [in press](#)). In the present study, we did not find evidence of full spatial updating even when there was not a delay. In Experiment 2, the retention interval differed with whether the observer moved or the scene was moved.

Observers in both groups waited 0, 6, or 12 s after encoding the scene, but observers in the observer movement group had to maintain their representations of the scenes for the duration of the locomotion period, thus adding additional time to the retention interval for this group. However, if maintenance duration alone were relevant, this additional time should have led to even greater differences between the 0 s condition and the 6 and 12 s conditions for the observer movement group, but in fact, no significant differences were detected for any of the angular distances. With respect to the effects of locomotion delays on automatic spatial updating in general, future research will have to examine whether delays affect automatic spatial updating in other research paradigms (e.g., tasks like those used by Farrell & Robertson, 1998; Farrell & Thompson, 1998; Rieser, 1989).

Additionally, the data from Experiment 2 show that our previous failures to find evidence of locomotion leading to spatially updated representations of scenes were not due to our use of 10 or 11-object scenes (Experiment 1 and Motes et al., [in press](#), respectively). Although we only used 4-object scenes in Experiment 2, we still did not find evidence of full spatial updating. Furthermore, in other research (Motes, Finlay, & Kozhevnikov, 2006), we have systematically examined the effect of varying scene set-size using 4-, 5-, 6-, 8-, and 10-object scenes, and although the effects of angular distance varied with set-size, none of the set-sizes yielded clear full spatial updating effects. The data from Experiment 2 also suggest that our previous failures were not due to our methods of having observers reproduce the scenes to show that they had learned them. In Experiment 2, the costs associated with locomotion were similar to our previous findings, but observers did not reproduce the scenes.

Finally, although we found evidence of systematic scene recognition costs associated with observer movement while Simons and Wang (1998, Experiment 1) did

not, there were differences between the materials used in these studies, and such differences might explain the contrasting results. For example, we used computerized displays of relatively simple, easily identifiable, geometric figures, and our display space was relatively small (radius = 11 cm; covering an approximate 12° visual angle), whereas Simons and Wang used three-dimensional, real objects, and their display space was larger (radius = 61 cm). However, in Motes et al. (*in press*), we used three-dimensional, real geometric objects placed across a circular table similar in size (radius = 54 cm) to the table used by Simons and Wang (1998), yet we found locomotion induced costs similar to those reported in the present Experiments 1 and 2. Furthermore, Simons and Wang (1998, Experiment 2; Wang & Simons, 1999, Experiment 1) also found costs associated with observer movement in two of their studies (i.e., change detection accuracy was lower when observers moved but the scene remained stationary than when both the observer and the scene remained stationary). Thus, it does not appear that differences in the properties of the displays, per se, account for the costs. Other differences, like the range and number of angular distances used or differences in the availability of the walls of the room as a frame of reference, might be relevant. For example, the walls were visually available in Simons and Wang's (1998, Experiment 1) study where they did not find that observer movement led to scene recognition costs, but the walls were neither visually available in our studies nor in Simons and Wang's studies where observer movement costs occurred.

Discussion

Across Experiments 1 and 2, we investigated whether processes associated with locomotion would automatically update an observer's representation of a scene to be spatially consistent with the observer's perspective of the scene after moving (e.g., Farrell & Robertson, 1998; Reiser, 1989). In Experiment 2, we systematically investigated whether delays between encoding a scene and moving around that scene would affect whether the observer movement would produce spatially updated representations of the scene. Across the two experiments, we failed to find that locomotion about a scene clearly produced spatially updated representations of the scene.

Thus, the data from the two studies question the generality of locomotion producing updated spatial representations. Based on a dynamic model of spatial updating offered by Mou et al. (2004), we argued, following Experiment 1, that our observers' egocentric

representations of the scene, formed when the observers learned the scene, might have faded over the duration of the scene recognition portion of the experiment, possibly forcing the observers to use more enduring, allocentric representations of the scene that were not updated when they moved. However, in Experiment 2, we did not find evidence to support this hypothesis.

One of the possible explanations for our results is that observer movement might lead to spatially updated representations in the case of externally viewed scenes only when the orientation change is limited to short angular distances around the scene (i.e., at least, not beyond 72° based on Experiment 2) or within the scene (i.e., at least not beyond 90° based on Mou et al., 2004). Viewing a scene externally might result in an observer processing a psychological separation of oneself from the other objects in the environment. We suspect that such a psychological separation occurs when a physical boundary (e.g., partially occluding panels) is imposed between the observer and the scene, but we are not suggesting that psychological separation occurs only in the presence of physical boundaries. Internally viewed scenes, on the other hand, might lead to spatially updated representations of scenes over a larger range of orientation changes (e.g., at least 360°, Rieser, 1989, barring other forms of interference like spinning quickly enough to make oneself dizzy). Internally viewed scenes are those in which an observer does not perceive a psychological boundary between oneself and other objects in the scene (e.g., when one is surrounded by objects in the scene [Farrell & Robinson, 1998; Rieser, 1989] or when a layout of objects encompasses a large expanse of the observer's visual field [Farrell & Thompson, 1998]). Internally viewed scenes might then elicit stronger egocentric coding of the spatial relations of objects in a scene or sensorimotor representations (Avraamides, 2003). Thus, in Experiments 1 and 2 and our previous work (Motes et al., *in press*), the scenes were externally viewed due to the occluding screens. Therefore, the representations of the scenes were not updated as observers moved around the scenes, regardless of whether locomotion after encoding the scenes was delayed or not.

In conclusion, the present experiments revealed that scene recognition following observer movement around scenes does not necessarily lead to automatic spatially updated representations of those scenes and add to the existing conditions in which automatic spatial updating has not been found. The present experiments showed that our previous failure to find evidence of spatial updating (Motes et al., *in press*) was not due to disorientation caused by our previous procedure and that

delays between encoding a scene and movement about that scene do not affect spatial updating. Furthermore, the present experiments raise important questions concerning the influence of encoding points of view and the effects of locomotion on spatial updating.

References

- Avraamides, M. N. (2003). Spatial updating of environments described in texts. *Cognitive Psychology*, *47*, 402–431.
- Christou, C. G., & Buelthoff, H. H. (1999). View dependence in scene recognition after active learning. *Memory & Cognition*, *27*, 996–1007.
- Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science*, *8*, 302–307.
- Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and automatic updating of body-centered spatial relationships. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *24*, 227–233.
- Farrell, M. J., & Thomson, J. A. (1998). Autonomic spatial updating during locomotion without vision. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, *51A*, 637–654.
- Luck, S., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. *Nature*, *390*, 279–281.
- Motes, M. A., Finlay, C. A., & Kozhevnikov, M. (in press). Scene movement versus observer movement in scene recognition: A test of the spatial updating hypothesis. *Perception*.
- Motes, M. A., Finlay, C. A., & Kozhevnikov, M. (2006). Effects of set-size on scene recognition following locomotion. Presented at the Vision Sciences Society Annual Meeting.
- Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 162–170.
- Mou, W., McNamara, T. P., Valiquette, C. M., & Rump, B. (2004). Allocentric and egocentric updating of spatial memories. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 142–157.
- 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.
- Shelton, A. L., & McNamara, T. P. (2004). Orientation and perspective dependence in route and survey learning. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 158–170.
- Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science*, *9*, 315–320.
- Waller, D., Montello, D. R., Richardson, A. E., & Hegarty, M. (2002). Orientation specificity and spatial updating of memories for layouts. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 1051–1063.
- 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.

Copyright of Psychological Research is the property of Springer Science & Business Media B.V. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.