Perception, 2006, volume 35, pages 1507 – 1520

DOI:10.1068/p5459

Scene recognition following locomotion around a scene

Michael A Motes, Cory A Finlay, Maria Kozhevnikov¶
Department of Psychology, Rutgers University, 333 Smith Hall, 101 Warren Street, Newark, NJ 07102, USA; e-mail: maria@psychology.rutgers.edu
Received 26 April 2005, in revised form 16 February 2006; published online 8 November 2006

Abstract. Effects of locomotion on scene-recognition reaction time (RT) and accuracy were studied. In experiment 1, observers memorized an 11-object scene and made scene-recognition judgments on subsequently presented scenes from the encoded view or different views (ie scenes were rotated or observers moved around the scene, both from 40° to 360°). In experiment 2, observers viewed different 5-object scenes on each trial and made scene-recognition judgments from the encoded view or after moving around the scene, from 36° to 180°. Across experiments, scene-recognition RT increased (in experiment 2 accuracy decreased) with angular distance between encoded and judged views, regardless of how the viewpoint changes occurred. The findings raise questions about conditions in which locomotion produces spatially updated representations of scenes.

1 Introduction

There has been debate in the literature whether object recognition can be considered viewpoint-independent or viewpoint-dependent. Viewpoint-independent models predict that object recognition is relatively unaffected when the currently perceived and previously learned views of an object differ (eg Biederman and Gerhardstein 1993; and see Biederman 2000), whereas viewpoint-dependent models predict that object recognition is more difficult when the currently perceived and learned previously views differ (eg Edelman and Bülthoff 1992; Tarr et al 1998). Scene recognition also involves matching memories of a previously learned scene to a perceptually available view of a scene, and in various contexts, the perceptually available view of a scene might be at a different orientation from the originally studied view: for example, when an observer has remained stationary but the orientation of the scene has changed, or when the scene has remained stationary but the observer has moved. Interestingly, although these two viewpoint changes might result in similar views of the scene, research suggests that scene-orientation changes might disrupt scene-recognition performance significantly (Christou and Bülthoff 1999; Diwadkar and McNamara 1997; Simons and Wang 1998; Wang and Simons 1999), whereas equivalent viewpoint changes due to observer movement might have little effect on scene-recognition performance (eg Farrell and Robertson 1998; Rieser 1989; Simons and Wang 1998; Wang 2004; Wang and Simons 1999).

Diwadkar and McNamara (1997) studied the effects of scene-orientation changes on scene recognition by having participants first learn the spatial layout of objects in a scene. Then, after demonstrating that they had learned the layout, participants were shown pictures of the same scene taken from views rotated 0° to 345° from the originally encoded view and pictures of a different scene (ie the objects in the scene arranged in different spatial configurations). For both types of pictures, participants judged whether the object-to-object spatial relations in these pictures differed from the spatial relations in the learned scene. Diwadkar and McNamara found that scene-recognition reaction time (RT) increased linearly with angular distance between the encoded and novel views (with the exception of a decrease in RT as angular distance approached 180°, suggesting a saving within this angular range). This RT pattern suggests that observers

¶ Author to whom all correspondence should be addressed.

formed viewpoint-dependent memories of the scene, and that they might have performed mental-rotation transformations to compare the perceived pictures with their memories of the learned scene (except with angular distances at or near 180°).

In contrast with Diwadkar and McNamara's (1997) finding that scene-recognition RT varied systematically with scene rotations over angles ranging from 0° to 345°, other data suggest that scene-recognition RT and accuracy will not vary when observers move around a scene because of spatial updating (Farrell and Robertson 1998, 2000; Farrell and Thomson 1998; Presson and Montello 1994; Rieser 1989). That is, cognitive processes associated with visual, kinesthetic, and auditory information available during locomotion have been hypothesized to adjust observers' egocentric representations of objects in an environment to account for physical self-to-object distance and orientation changes that occur as observers move. Furthermore, researchers have argued that spatial updating during locomotion occurs automatically, in that the updating processes operate spontaneously when observers move and cannot be ignored without significant costs (Farrell and Robertson 1998; Farrell and Thomson 1998; Wang 2004).

Simons and Wang (1998; Wang and Simons 1999) examined the effects of scene and observer movement on accuracy at recognizing scene changes (ie their observers always judged test scenes that were different from the encoded scenes and identified what was different). They found that observers were less accurate at identifying scene changes when the observers remained at the encoding position and the scene rotated 47° from the encoded view than when they remained at the encoding position and the scene also remained at the encoding orientation. However, observers were equally accurate at identifying scene changes when they moved 47° from the encoding position and the scene remained at the encoding orientation, or when they moved 47° from the encoding position and the scene was also rotated 47° (or, slightly more accurate in the former condition—Wang and Simons 1999, their experiment 1). On the basis of these findings, Simons and Wang argued that observers' representations of the scenes were spatially updated when they moved, and that the updated representations facilitated observers in identifying the changes to the scenes.

Although Simons and Wang (1998; Wang and Simons 1999) found evidence consistent with the hypothesis that observer movement will produce spatially updated representations of scenes, no research has been conducted to examine scene-recognition RT and accuracy patterns following observer movement over a broad range of angles while the scene remains stationary. In other spatial-updating studies researchers have examined the effects of observer-orientation changes over a broad range of angles on blind directional judgments (Farrell and Robertson 1998; Mou et al 2004; Rieser 1989; Wang 2004; Wang and Spelke 2000; Wraga 2003; Wraga et al 2004). In contrast to scene-recognition judgments, observers in the blind directional-judgment studies indicated, without visual support and from either the studied orientation or after making an orientation change, the direction to a target object in a previously studied array. These studies, however, have produced a mixed set of results. Rieser (1989), for example, had observers memorize from a particular orientation the location of 9 objects that encircled them. Then, whilst blindfolded and remaining at the encoding orientation (or after rotating in place from 40° to 360° in 40° increments from the encoding orientation), observers pointed to a target object in the array. Rieser found that the observers' RTs and accuracies at pointing to targets did not vary with the angular distance between the new orientation and the encoding orientation. Thus, the data suggest that the observers' representations were transformed during the rotation to be consistent with their new orientations. Mou et al, on the other hand, found that observers were slower at making blind directional judgments after making 225° body-orientation changes than when they did not make orientation changes or when they made 90° orientation changes (they were equally fast in the last two conditions). Thus, the observers' representations of the scenes do not appear to have been updated, or at least not fully updated, following the 225° orientation change.

Furthermore, in two of Simons and Wang's studies (1998, their experiment 2; Wang and Simons 1999, their experiment 1), change-detection accuracy following observer movement was lower than change-detection accuracy when both the observer remained at the encoding position and the scene remained at the encoding orientation. This lower accuracy following observer movement suggests that the observers' representations following locomotion were not fully updated and raises the question whether RT and error might increase with larger angular distances between the encoded and tested views. Thus, we conducted two experiments to examine whether scene-recognition RT and error would indeed increase with angular distance as observers moved around an encoded scene.

2 Experiment 1

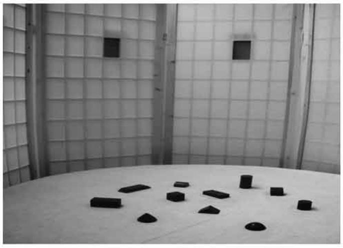
Our goals in experiment 1 were to examine scene-recognition RT and accuracy following observer movement over a range of angles (from 0° to 360°) around a scene and to compare scene-recognition RT and accuracy patterns following observer movement to scene recognition RT and accuracy patterns following scene rotations that produced equivalent viewing angle changes. Similar to Diwadkar and McNamara (1997), observers in our study learned a scene layout, and then, after demonstrating that they had learned the layout, they judged whether object-to-object spatial relations in subsequently shown scenes were identical to those in the encoded scene. For the scene-recognition test, half of the observers moved from 0° to 360° around the scene while the scene was occluded and remained stationary, and half of the observers stood at the encoding station and viewed scenes rotated from 0° to 320° from the encoded orientation. Thus, when our observers moved from the encoding position, the conditions were analogous to an observer approaching a learned (encoded) scene but being forced to detour around the scene to revisit it from a new orientation (eg approaching a previously viewed display case from a new angle or preparing to enter a room through a door that one usually uses only to find it locked or blocked and then having to use an alternative entrance).

2.1 *Method*

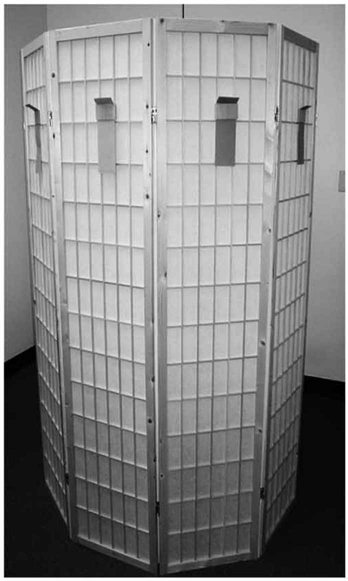
2.1.1 *Participants.* Forty-one students from Rutgers University participated in the experiment in exchange for course credit or monetary compensation. Twenty-two (ten males) were randomly assigned to the scene-movement condition, and nineteen (ten males) were randomly assigned to the observer-movement condition. Four additional observers failed to encode the scene accurately after ten attempts and thus did not complete the remainder of the experiment.

2.1.2 *Stimuli and apparatus.* The scene consisted of 11 geometric objects that were arranged on a circular table (see figure 1a). The circular table (radius = 54 cm) could be rotated 360° and was completely surrounded by an occluding screen (see figure 1b). The screen consisted of nine panels (44.6 cm × 180 cm). Each panel had a horizontally centered viewing window (5.8 cm × 12 cm) with a removable cover. The windows were 140 cm from the floor, 66 cm from the top of the table, and 40° apart. The testing room remained darkened throughout the duration of the experiment, with the only light source being directly above the array of objects. Finally, observers entered their responses via the buttons on a Gyromouse Pro (a handheld, wireless, radio-frequency computer mouse).

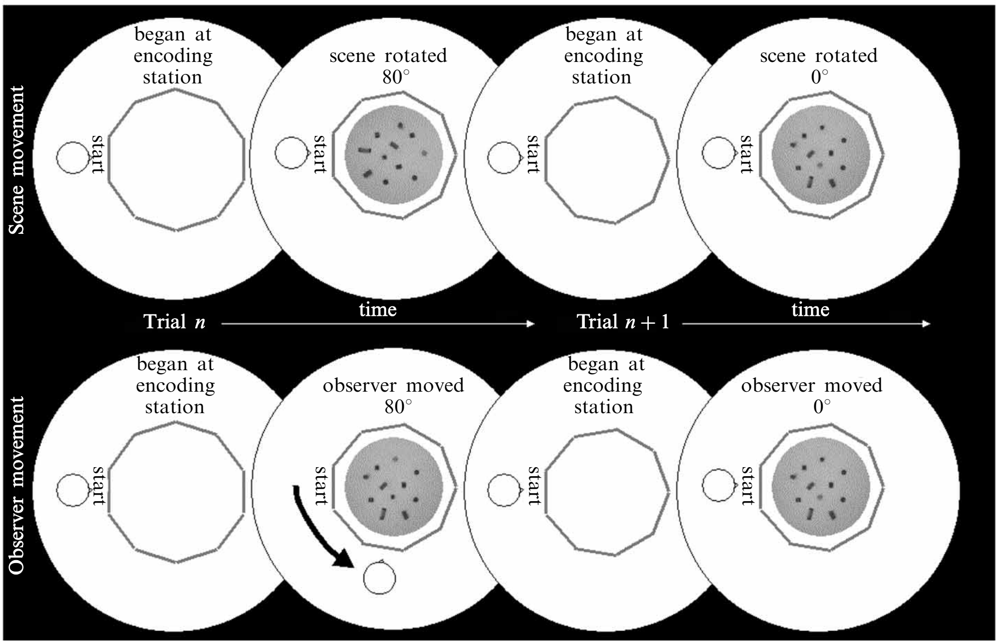
2.1.3 *Scene encoding procedure.* Observers were initially greeted in a waiting room outside of the testing room. They were told that the purpose of the study was to investigate memory for scenes. They were then taken to the testing room and began the scene-encoding phase. During this phase, observers learned the layout of the 11 objects.



(a)



(b)



(c)

Figure 1. The apparatus and stimuli used in experiment 1. An 11-object scene (a) was placed on the round rotating table (radius = 54 cm) that was surrounded by nine, 44.6 cm × 180 cm, occluding screens (b). Each screen had a 5.8 cm × 12 cm covered viewing window, horizontally centered 140 cm from the floor and 66 cm from the top of the rotating table. After learning the scene, observers in the observer-movement and scene-movement groups began each trial at the encoding station (c). Observers in the observer-movement group then walked to the judging station with an open window, whereas observers in the scene-movement group remained at the encoding station. Observers in both groups then indicated whether the scene was the same as or different from (the locations of two objects were switched, as shown in Trial *n*) the encoded scene. Trial *n* shows equivalent scene and observer movement 80° viewpoint changes, and Trial $n + 1$ shows equivalent 0° (no-change) viewpoint changes.

Observers viewed the array of objects through one of the viewing windows (the *encoding* station) for 1 min while repeating orally "1, 2, 3" to interfere with their verbal encoding of the objects' locations. Then, to assess accurate memory for the array, observers were led to a separate room and were instructed to reproduce the encoded arrangement with a set of matching objects. Observers repeated this encoding procedure until they correctly reproduced the arrangement. They were told that they would be comparing subsequently presented scenes to their memories of the encoded scene, and, after the encoding phase, they were not explicitly given the opportunity to refresh their memories.

2.1.4 *Scene recognition procedure.* Observers were randomly assigned either to the scene movement or to the observer-movement condition. In both conditions, observers judged whether the scene interobject distances were different from the interobject distances learned during the scene-encoding phase (see figure 1c). On trials in which the arrangement was different, one of 28 pairs of objects was randomly chosen (three of the objects were never moved), and the positions of the objects were switched. In the scene-movement condition, the observers always viewed the judged scenes from the encoding station. On each trial, the table on which the scene was placed was rotated to one of nine possible orientations (0° to 320° in 40° intervals) from the encoded orientation. To maintain a balanced design between the scene-movement and observer-movement conditions, the 0° orientation occurred twice as often as the other viewing orientations, and for the data analyses half of these trials were coded as 360° scene rotations. In the observer-movement condition, observers began at the scene encoding station on each trial. On the 0° trials, viewers remained at the encoding station, but on all of the other trials they walked counterclockwise from the encoding station to the station with a viewing window left open by the experimenter. On the 360° trials, viewers walked around the entire occluding screen back to the encoding station.

For both the scene and observer-movement conditions, RT, measured as the time an observer began viewing the scene to the time the observer entered a judgment, and accuracy were recorded via an E-Prime (Psychological Software Tools version 1.0) computer program. When the observer arrived at the viewing station and verbally indicated that he/she was ready to look at the scene, the experimenter gave a verbal "OK" prompt and pressed a key on the keyboard to start a software clock. The observer then looked at the scene and indicated whether it was the same as or different from the learned scene by pressing a correspondingly labeled button on the wireless mouse. The button press stopped the software clock.

Observers were asked to respond as quickly and as accurately as possible, but they did not receive feedback on accuracy or RT. After responding, observers from both the scene-movement and the observer-movement conditions exited and remained outside of the room while the experimenter rearranged the objects and table for the next trial. There were 60 trials—3 same and 3 different trials at each of the ten viewing orientations for the scene-movement condition and at each of the ten viewing stations for the observer-movement condition. The order of the 60 trials was randomly selected for each observer.

2.2 *Results*

2.2.1 *Preliminary analyses.* The RT data were screened for outliers. Trials in which the observers inadvertently pressed the response button before actually viewing the judged scenes (RTs below 750 ms = 2.56%; 63 out of the 2460 trials), and trials in which they took longer than 30 s to respond (0.16%, 4 out of the 2460 trials) were deleted. Only the RTs for correct responses were analyzed. RTs ±2.5 SDs from an individual observer's mean RT were deleted (2.25%; 0 at −2.5 SD and 42 at +2.5 of 1868 correct responses). After deleting the outlier RTs, the mean RT and the proportion of correct responses at each viewing angle were calculated for each observer. Three observers in the scene-movement condition and two in the observer-movement condition had mean accuracies at or near chance (< 0.56). Separate mixed-model ANOVAs (2 movement condition $\times$ 10 angular distance) were calculated for the RT and the accuracy data both with and without these observers. The general patterns of results did not differ. The results of the analyses without the data from these observers are reported below.[(1)]

2.2.2 *Scene-recognition RT.* For scene-recognition RT (see figure 2a), an ANOVA revealed a significant effect of angular distance ($F_{9,306} = 3.97$, $p < 0.001$), but neither the main effect of movement condition ($F_{1,34} < 1$) nor the interaction ($F_{9,306} = 1.18$, $p = 0.31$) were significant. Trend analyses of the effect of angular distance revealed that only the quadratic component was significant ($F_{1,34} = 15.44$, $p < 0.001$), and follow-up analyses revealed that the linear trend components both for the effect of angular distance from 0° to 160° ($F_{1,35} = 8.43$, $p < 0.01$) and from 200° to 360° ($F_{1,35} = 9.58$, $p < 0.04$) were significant. Thus, regardless of whether the scene was moved or the observer moved, the analyses showed that RT increased with angular distance from the encoded view. Therefore, the data suggest that both groups had viewpoint-dependent representations of the scenes and suggest that both groups compared the novel and remembered views by performing mental transformations that required more time as the angular distance between the novel and remembered views increased. Additionally, the data did not reveal any advantage of observer movement over scene movement across any of the angular distances.

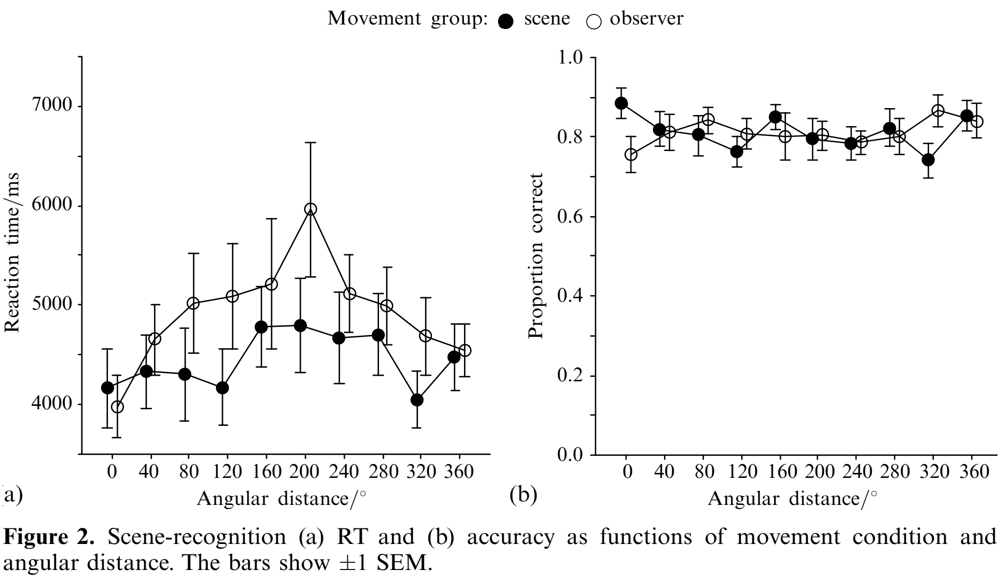


Figure 2. Scene-recognition (a) RT and (b) accuracy as functions of movement condition and angular distance. The bars show ±1 SEM.

2.2.3 *Scene-recognition accuracy.* For scene-recognition accuracy (see figure 2b), an ANOVA did not reveal any significant main effects (both Fs < 1) or a significant interaction ($F_{9,306} = 1.59$, $p = 0.12$). Thus, accuracy did not vary significantly with angular distance for either the scene-movement or observer-movement conditions. The lack of a difference in accuracy, however, did not appear to be due to ceiling or floor effects (for the scene-movement condition: $M = 0.81$; for the observer-movement condition: $M = 0.81$), but there was considerable variability in both conditions (for the scene-movement condition: SD $= 0.11$; for the observer-movement condition: SD $= 0.10$).

[(1)] For the scene-movement and observer-movement conditions, there were positive correlations between the proportion of correct responses given and RT, after averaging across the angular distances ($r_{\text{scene movement}\ 19} = 0.52$, $p = 0.02$, and $r_{\text{observer movement}\ 17} = 0.35$, $p = 0.17$). Thus, there was evidence that some participants sacrificed accuracy for speed. However, this trade-off cannot explain the overall pattern of results.

2.3 *Discussion*

The data showed that regardless whether the scene was moved or the observer moved from the view in which the scene was encoded, scene-recognition RT increased as a function of angular distance, and the data did not reveal any advantage of observer movement over scene movement across any of the angular distances. The increases in RT with angular distance both for the observer-movement and the scene-movement conditions suggest that both groups of observers relied on viewpoint-dependent representations when judging the scenes and, for the observer-movement group, the data suggest that as observers moved to new viewing stations, their mental representations of the encoded scene were not perfectly updated to match how the array would look from the new viewing angles. Finally, in addition to RT, we also examined scene-recognition accuracy, but we did not find that accuracy varied with angular distance from the encoded view for either the observer-movement or for the scene-movement group.

One possible explanation for our finding that RT increased with angular distance for both groups is that observers in both groups relied on allocentric (object-to-object) rather than egocentric (self-to-object) representations when making scene-recognition judgments. Research suggests that there are two spatial-coding systems that can be used for representing scenes (Mou et al 2004). The egocentric representational system codes for the spatial relations of objects with respect to one's own body axes (eg left – right, front – back, up – down), and the allocentric representational system codes for the spatial relations of objects with respect to other objects. The object-to-object relations represented in the allocentric system are represented in a particular orientation, either from the viewer's encoding perspective or along a salient scene axis (Mou and McNamara 2002; Shelton and McNamara 2001), and locomotion does not automatically spatially update these allocentric representations. In fact, Mou et al (2004) argued that automatic spatial updating is based on egocentric, transient spatial representations that require on-line sensory feedback, but that such egocentric representations begin to decay with delays between encoding and action as brief as 10 s, depending on the task and the complexity of the array. Furthermore, they argued that, after the egocentric representation fades, an observer is forced to use enduring, orientation-specific, allocentric representations that are not automatically spatially updated by the observer's movement, if such representations are available.

Thus, as Diwadkar and McNamara (1997), we had observers learn a single scene at the beginning of the study and remember that scene throughout the duration of the study, and this delay between encoding the scene and making scene-recognition judgments might have disrupted the use of egocentric representations and led to the use of allocentric representations—representations that were not spatially updated by the observer's movements. Therefore, in experiment 2, we reexamined the effects of observer movement on scene-recognition RT and accuracy patterns over a range of angles (from 0° to 180° in 36° increments), but we used a scene-recognition paradigm that did not include such delays in the encoding, locomotion, and scene-recognition sequence.

3 Experiment 2

In experiment 2, the observers were allowed to move to the judging station immediately after encoding a scene so that their egocentric representations of the scene should have been updated during locomotion. The procedure and scene set-sizes were consistent with previously used procedures and to set-sizes in studies that reported finding locomotion-induced spatial-updating effects (Simons and Wang 1998; Wang and Simons 1999). Observers viewed a different 5-object scene on each trial. Observers then viewed a judged scene from either the same perspective or from a different perspective (ie observers walked around the scene) and made scene-recognition judgments (ie identified which object in the scene had been moved).

In experiment 2, we did not include a scene-rotation condition. Although faster RT and greater accuracy in a locomotion condition relative to an equivalent scene-rotation condition might be indicative of an updating benefit, comparisons between scene-recognition judgments at the encoded view versus novel views that result from locomotion are the most stringent tests of the claim of spatial updating. If processes associated with locomotion *fully* update observers' representations, then RT, in particular, and accuracy should not systematically vary between such conditions (Rieser 1989). Furthermore, RT or accuracy differences between a locomotion condition and an equivalent scene-rotation condition might result from other scene-recognition processes. Examinations of the processes involved in the recognition of rotated figures, for example, have shown that observers initially engage in a search for corresponding parts, then they rotate seemingly related parts into congruence, and finally they confirm recognition by rotating other seemingly related parts into congruence (Just and Carpenter 1976). A similar set of processes might be involved in scene recognition. Thus, differences in RT and accuracy between equivalent observer-movement and scene-rotation conditions could reflect differences in the use of any of these (or other) processes.

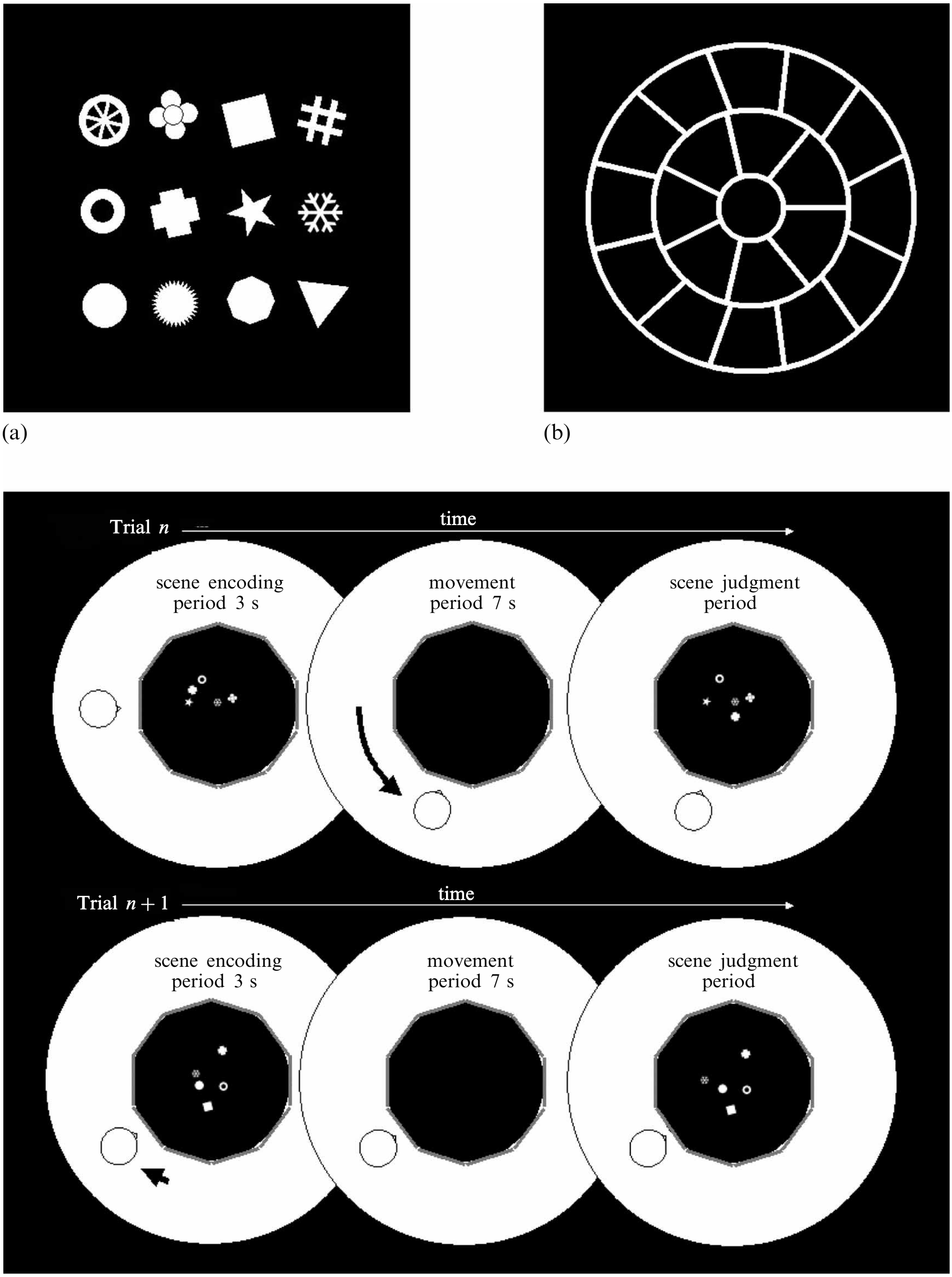
3.1 *Method*

3.1.1 *Participants.* Ten students (four males) from Rutgers University participated in the experiment in exchange for course credit.

3.1.2 *Stimuli and apparatus.* The experiment was conducted in a darkened room. A Dell 17-inch LCD flat-panel monitor was laid horizontally on the circular table used in experiment 1. A black cloth covered the monitor and table, except for a circle (radius = 11 cm) cut out of the black cloth that was centered over the viewing screen of the monitor. All of the scenes were shown within this circle. The black cloth prevented observers using the outer edges of the monitor and the inner edges of the viewing screen as frames of reference when learning and judging the scenes. Viewing stations were created with ten occluding screens that surrounded the monitor and table. That is, one more screen was added to the nine used in experiment 1. Therefore, the viewing windows were 36° apart. The testing room remained darkened throughout the duration of the experiment, with the only light sources being the computer monitor when the scenes were shown and a small flashlight used by the experimenter to see when recording observers' verbal responses on a trial record sheet and when walking with the observers around the path. Observers entered their responses via the buttons on a Gyromouse Pro.

The scenes consisted of computerized displays of 5 objects. Thirty-six different 5-object computerized scenes were designed. For each scene, 5 objects were randomly selected from a set of 12 objects (snowflake, pound sign, square, flower, wheel, cross, star, triangle, octagon, sun, circle, and donut; see figure 3a) and measured from 2.5 cm to 3.5 cm high and 2.25 cm to 3.5 cm wide. Each object was then placed in one randomly selected location of twenty-one predefined locations (see figure 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 sixteen randomly selected unoccupied spaces. Each of the thirty-six pairs of encoded and judged scenes was then randomly assigned to one of six viewing angles (0° to 180° in 36° increments). Three blocks were then created, each consisting of a randomly arranged set of the thirty-six trial pairs. Each observer worked through the set of the three blocks (ie 108 trials, 18 trials per angular distance) in a fixed random order.

3.1.3 *Procedure.* Observers were initially greeted in a waiting room outside of the testing room. They were told that the purpose of the study was to investigate memory for scenes. Observers were then taken into the testing room and disoriented to reduce the possibility of their using the door or other stimuli external to the testing room as



(c)

Figure 3. The stimuli and design used in experiment 2. For each scene, 5 of 12 possible objects (a) were randomly chosen and each assigned to one of twenty-one locations (b). On each trial (c), an observer was given 3 s to encode a scene and was escorted to a judging station within 7 s. Trial n shows a 72° viewpoint change, and Trial $n+1$ shows a 0° viewpoint change (or no change). After the 7 s, the judged scene appeared, and the observer identified which object had been moved. After entering a response, the observers walked one viewing station to the left to begin the new trial.

spatial reference points. For the disorientation procedure, the observers closed their eyes and were led counterclockwise by the experimenter partially around the occluding screens, then they were rotated in place three times, led clockwise partially back, rotated in place three times, led counterclockwise partially around again, rotated three times, and finally asked to open their eyes and point to the door. None was able to accurately point to the door after this disorientation procedure.

The observers were then led to the first encoding station, and they were told that the orientation of the scene would remain the same throughout the experiment. On each trial (see figure 3c), observers viewed a 5-object scene for 3 s, after which a tone was played to signal to the experimenter to have the observer move to the judging station; observers always moved counterclockwise around the scene. The judged scene appeared 7 s after the tone played. After the scene appeared, the observers pressed the right mouse button on the wireless mouse to indicate when they recognized which object had been moved, and they verbally reported the object to the experimenter. The experimenter recorded the response on the trial record sheet.

Observers completed 4 practice trials at the beginning of the experiment. For both the practice and test trials, the observers were asked to respond as quickly and as accurately as possible, but they did not receive accuracy or RT feedback. An E-Prime (Version 1.1) program presented the scenes and recorded RT and accuracy. The program started a RT software clock when the judged scene was shown and stopped the software clock when the participant pressed one of the mouse buttons on the wireless mouse to indicate that they recognized which object had been moved. Observers did not return to the original encoding station after judging the scene (see figure 3c). Instead, they moved one viewing station to the left of the judging station to encode the next scene; this change was made to the procedure used in experiment 1 to reduce the total testing time.

3.2 *Results*

3.2.1 *Preliminary analyses.* The RT data were screened for outliers. Trials in which the observers inadvertently pressed the mouse button before actually viewing the judged scenes (RTs below 750 ms = 0.15%; 2 out of the 1296 trials), and trials in which the observers took longer than 30 s to respond (0.77%; 10 out of the 1296 trials), were deleted. Only the RTs for correct responses were analyzed. RTs $\pm$ 2.5 SDs from an individual observer's mean RT were deleted (3.29%; 17 out of 516 total correct responses). After deleting the outlier RTs, mean RT and the proportion of correct responses at each viewing angle were calculated for each observer. Separate repeated-measures one-way ANOVAs, with angular distance as the independent variable, were performed on the RT and the accuracy data.

3.2.2 *Scene-recognition RT.* For scene-recognition RT (see figure 4a), an ANOVA revealed a significant effect of angular distance ($F_{5,45} = 4.64$, $p < 0.01$). Trend analyses of the effect of angular distance revealed that the linear component was significant ($F_{1,9} = 20.00$, $p < 0.01$). Thus, the analyses showed that RT increased with angular distance from the encoded view and therefore suggest that the observers relied on viewpoint-dependent representations of the scenes when making scene-recognition judgments.

3.2.3 *Scene-recognition accuracy.* For scene-recognition accuracy (see figure 4b), an ANOVA revealed a significant effect of angular distance ($F_{5,45} = 9.75$, $p < 0.001$). Trend analyses of the effect of angular distance revealed that the linear and quadratic components were significant ($F_{1,9} = 22.48$, $p < 0.01$ and $F_{1,9} = 12.42$, $p < 0.01$, respectively). Separate contrasts revealed that accuracy decreased with angular distance up to 72° from linear encoded view ($F_{1,9} = 39.26$, $p < 0.001$) and then did not significantly change across 72° to 180° ($F_{1,9} = 2.37$, $p = 0.16$). Thus, the accuracy data also suggest that

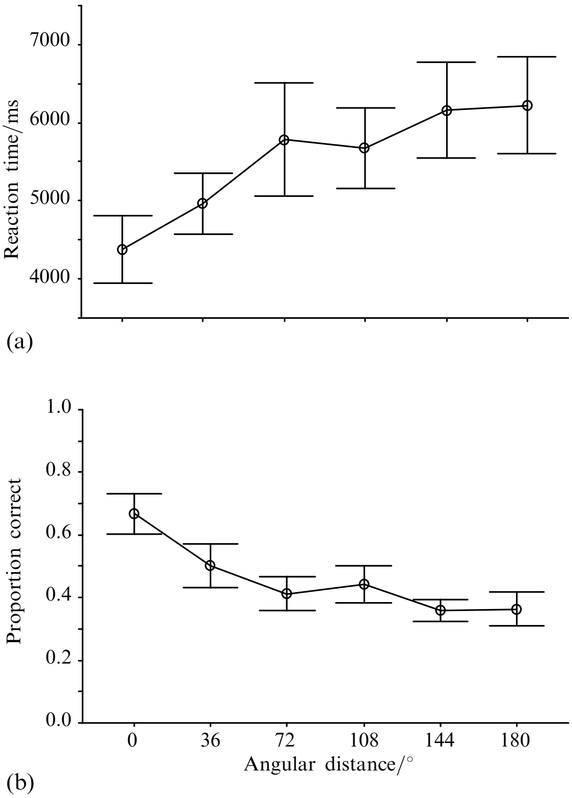


Figure 4. Scene-recognition (a) RT and (b) accuracy as functions of angular distance. The bars show ±1 SEM.

observers relied on viewpoint-dependent representations of the scenes when making the scene-recognition judgments.

3.3 *Discussion*
The data show that as observers moved from the views in which they encoded the scenes, RT increased and accuracy decreased as functions of angular distance. The increases in RT and decreases in accuracy with angular distance suggest that when observers moved to new viewing stations they relied on viewpoint-dependent representations to judge the scenes; that is, their mental representations of the encoded scenes were not perfectly updated to match how the scenes would look from the new perspectives. In fact, as in experiment 1, the RT and accuracy patterns suggest that observers might have compared the novel and encoded views by performing mental-rotation transformations that required more time as the angular distance between the novel and remembered views increased.

Thus, we hypothesized that our failure to find evidence of locomotion leading to fully updated representations of the learned scene in experiment 1 might have been due to the delay between the scene encoding and the movement and judgment phases of the experiment, but we failed to find support for this hypothesis in experiment 2. In experiment 2, we eliminated such delays but still found that RT increased with angular distance when observers judged scenes after moving from the encoding station. Thus, at least for scenarios like those tested in experiments 1 and 2, delays do not appear to moderate whether locomotion leads to full spatial updating.

Additionally, the data from experiment 2 show that our failure to find evidence of locomotion leading to fully updated representations in experiment 1 was not due to our use of a 10-object scene. We used 5-object scenes in experiment 2, but we still did not find evidence of full spatial updating. Furthermore, in other research (Motes et al, 2006), we systematically examined the effect of varying scene set-size with 4-, 5-, 6-, 8-, and 10-object scenes and, although the effects of angular distance vary with set-size, none of the set-sizes produced effects consistent with locomotion producing fully updated representations of the scenes.

Finally, one possible reason that we found evidence of systematic scene-recognition costs associated with observer movement in experiment 2 is that we used 2-D computerized displays rather than 3-D real objects. However, the pattern of subjects' responses we found in experiment 2 was very similar to the patterns of subjects' responses we found in experiment 1, which used 3-D real objects. Furthermore, Simons and Wang (1998, their experiment 2; Wang and Simons 1999, their experiment 1) also found costs associated with observer movement in two of their studies (ie change-detection accuracy was lower when observers moved but the scenes remained stationary than when both the observers and the scenes remained stationary), yet they used 3-D real objects in these studies. Therefore, it does not appear that computerized versus real-object display differences account for our finding of observer-movement costs.

4 General discussion

We designed experiments 1 and 2 to examine the effects of locomotion around a scene on scene-recognition RT and accuracy. In experiment 1 we revealed that scene-recognition RT increased as a function of angular distance between encoded and judged views of a scene, regardless of whether the scene was moved or an observer moved from the view in which the scene was encoded. Furthermore, the data did not reveal an advantage of observer movement over scene movement across any of the angular distances. In experiment 2 we revealed that scene-recognition RT increased and accuracy decreased with angular distance as observers moved from the encoded view. Thus, across the two studies, the data suggest that as observers moved to new viewing positions their mental representations of the encoded scene were not perfectly updated to match how the scene would look from the new perspectives.

Other researchers have also found costs associated with observers moving from the positions in which they encoded scenes (Mou et al 2004; Simons and Wang 1998; Wang and Simons 1999; Wraga 2003). Mou et al had observers memorize the spatial arrangement of a 10-object scene placed on a mat in front of them, walk into the center of the scene, and in the three conditions relevant to the current studies, make blind directional judgments either from the encoding orientation or after rotating 90° or 225°. Observers moved into the scene and made the orientation changes with their eyes open, and thus they had retinal and extraretinal information available to update their representations of the scene. Mou et al found that observers were as fast indicating the directions to targets when they made 90° orientation changes as when they did not make an orientation change. However, Mou et al also found that observers were slower at indicating the directions to targets when they made 225° orientation changes than when they did not make an orientation change or when they made 90° orientation changes; thus suggesting that the observers' directional judgments following the 225° orientation changes were not based on fully updated representations. Wraga had observers memorize a surrounding, equally spaced, circular array of six colored pieces of felt, and after learning the array, observers made blind directional judgments either from the encoding orientation or after being passively rotated (ie the experimenter rotated the chair in which the observers sat) from 60° to 320°. Wraga found that observers were slower and less accurate at making blind directional judgments after being passively moved from 60° to 320° than when they did not move from the encoding orientation, thus also suggesting the observers' judgments were not based on fully updated representations of the array. Furthermore, as described in section 1, Simons and Wang (1998, their experiment 2; Wang and Simons 1999, their experiment 1) found that observers who moved while the scene remained stationary were less accurate at recognizing scene changes than when both the scene and the observer remained stationary, suggesting that observers' scene-recognition judgments were not based on fully updated representations of the scenes. Likewise, our findings that scene-recognition RT increased and, in experiment 2, accuracy decreased with angular distance as observers walked around the scenes show that these observers were not relying on fully updated representations when judging the scenes from new views. Our findings also add to the above collection of studies by showing that the disruptive effects of locomotion on scene recognition increased systematically with angular distance as observers moved around the scenes.

After experiment 1, we proposed that a possible explanation for our failure to find evidence of locomotion producing fully updated representations of the scene was that the delay between encoding the scene and moving and making scene-recognition judgments allowed for the observers' egocentric representations of the scene to fade and forced them to rely on viewpoint-dependent, allocentric representations when making scene-recognition judgments. However, such delays were not a part of experiment 2, though here we also did not find evidence that locomotion produced fully updated representations of the scenes. Thus, for scenarios analogous to ours, locomotion does not appear to produce fully spatially updated representations regardless of whether such delays were present.

Another possible explanation is that the effect of locomotion on representations of scenes depends on whether the scenes were encoded from an internal or an external point of view. Locomotion might only lead to fully updated representations of externally viewed scenes when the movement is limited to short angular distances. However, locomotion may well lead to fully updated representations of internally viewed scenes regardless of the magnitude of the distance change. Internal views of scenes might more strongly elicit the use of the egocentric representational system than external views, and processes associated with locomotion might then be more effective at updating these representations over greater orientation changes. Thus, based on this model, we did not find evidence of full spatial updating because our participants studied the scenes from external points of view, which did not elicit strong self-based coding. Rieser (1989; and see Farrell and Robertson 1998), on the other hand, found evidence of full spatial updating following observer movement because his observers had internal study views of the scene (ie they were encircled by the objects) and these strongly elicited self-based coding. Furthermore, Mou et al (2004) might not have found evidence of full spatial updating because their observers learned the scene from an external point of view before moving into and making blind directional judgments from within the scene. Thus, the observers' predominant memories of the scene might have been viewpoint-dependent, allocentric representations because of the external study view.

Although the results from several studies have supported locomotion-induced, automatic, spatial-updating predictions, research is beginning to reveal conditions in which locomotion-induced spatial updating does not occur, or at least full spatial updating does not occur. Wang (2004) has shown some conditions in which observer movement does not automatically update imagined spatial environments, and Wang and Brockmole (2003) have shown that movement within an environment nested within a larger surrounding environment (eg a room in a building on a campus) produced updated representations of the nested environment but not of the surrounding environment. Our data also show that observer movement does not necessarily fully update representations of spatial layouts. Thus, they add to the existing qualifications of the automatic spatial-updating hypothesis, and when viewed in the context of other spatial-updating studies, our data raise important questions about the effects of encoding points of view on the automatic spatial updating of representations of scenes. Finally, given that the scene-recognition costs associated with locomotion increased with angular distance in experiments 1 and 2, our data show that future researchers examining the effects of locomotion on scene recognition should avoid using a narrow range of angular distances between the encoded and judged views of scenes.

References

Biederman I, 2000 "Recognizing depth-rotated objects: A review of recent research and theory" *Spatial Vision* **13** 241 – 253

Biederman I, Gerhardstein P C, 1993 "Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance" *Journal of Experimental Psychology: Human Perception and Performance* **19** 1162 – 1182

Christou C G, Bülthoff 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

Edelman S, Bülthoff H H, 1992 "Orientation dependence in the recognition of familiar and novel views of three-dimensional objects" *Vision Research* **32** 2385 – 2400

Farrell M J, Robertson I H, 1998 "Mental rotation and automatic updating of body-centered spatial relationship" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **24** 227 – 233

Farrell M J, Robertson I H, 2000 "The automatic updating of egocentric spatial relationships and its importance due to right posterior cortical lesions" *Neuropsychologia* **38** 585 – 595

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

Just M A, Carpenter P A, 1976 "Eye fixations and cognitive processes" *Cognitive Psychology* **8** 441 – 480

Motes M A, Finlay C A, Kozhevnikov M, 2006 "Effects of set-size on scene recognition following locomotion", paper presented at the Vision Sciences Society Annual Meeting, Sarasota, FL, 5 – 10 May

Mou W, McNamara T P, 2002 "Intrinsic frames of reference in spatial memory" *Journal of Experimental Psychology: Learning, Memory, and 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, and Cognition* **30** 142 – 157

Presson C C, Montello D R, 1994 "Updating after rotational and translational body movements: coordinate structure of perspective space" *Perception* **23** 1447 – 1455

Rieser J J, 1989 "Access to knowledge of spatial structure at novel points of observation" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **15** 1157 – 1165

Shelton A L, McNamara T P, 2001 "Systems of spatial reference in human memory" *Cognitive Psychology* **43** 274 – 310

Simons D J, Wang R F, 1998 "Perceiving real-world viewpoint changes" *Psychological Science* **9** 315 – 320

Tarr M J, Williams P, Hayward W G, Gauthier I, 1998 "Three-dimensional object recognition is viewpoint dependent" *Nature Neuroscience* **1** 275 – 277

Wang R F, 2004 "Between reality and imagination: When is spatial updating automatic?" *Perception & Psychophysics* **66** 68 – 76

Wang R F, Brockmole J R, 2003 "Simultaneous spatial updating in nested environments" *Psychonomic Bulletin & Review* **10** 981 – 986

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

Wraga M, 2003 "Thinking outside the body: An advantage for spatial updating during imagined versus physical self-rotation" *Journal of Experimental Psychology: Learning, Memory, and Cognition* **29** 993 – 1005

Wraga M, Creem-Regehr S H, Proffitt D R, 2004 "Spatial updating of virtual displays during self- and display rotation" *Memory & Cognition* **32** 399 – 415

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 35 2006

www.perceptionweb.com