Lack of Set Size Effects in Spatial Updating: Evidence for Offline Updating

Eric Hodgson and David Waller
Miami University

Four experiments required participants to keep track of the locations of (i.e., update) 1, 2, 3, 4, 6, 8, 10, or 15 target objects after rotating. Across all conditions, updating was unaffected by set size. Although some traditional set size effects (i.e., a linear increase of latency with memory load) were observed under some conditions, these effects were independent of the updating process. Patterns of data and participant strategies were inconsistent with the common view of spatial updating as an online process. Instead, the authors concluded that participants formed enduring, long-term memory representations of the layouts at learning that were used to reconstruct spatial information about the layouts as needed (i.e., offline updating). These results support M. Amorim, S. Glasauer, K. Corpinot, and A. Berthoz's (1997) 2-system model of spatial updating that includes both online and offline updating.

Keywords: spatial cognition, spatial updating, spatial memory, set size, capacity limits

As people move through the world, the spatial relationships between themselves and objects in the environment are constantly changing. For example, as you walk into your office, the doorway shifts from being in front of you to being immediately behind you. After sitting at your desk, the distance and direction to the doorway have changed again. The phenomenon of tracking the changing relations between oneself and locations in the environment is referred to as *spatial updating* (or *egocentric updating*). Spatial updating is generally considered to be an online process in which the spatial relations to one or more objects in one's immediate environment are continually adjusted or recomputed on a moment-by-moment basis (Farrell & Robertson, 1998; Farrell & Thomson, 1998, 1999; Klatzky, Loomis, Beall, Chance, & Golledge, 1998; Lindberg & Gärling, 1981a; May & Klatzky, 2000; Rieser & Rider, 1991). The online aspect of updating has been demonstrated by Farrell and Thomson (1999), who asked participants to walk blindfolded to a previously learned target and examined the degree to which participants modified their gait en route. Participants in their study adjusted both the number of steps taken and their stride length as necessary to arrive accurately at the target destination. These adjustments were observed when participants were walking to the target with vision and without vision, but not when participants were asked to walk the same distance to the target but in the opposite direction. These observations suggested that participants were not merely preplanning their movements but were instead updating the perceived location of the target in real time and adjusting their movement online.

In contrast to the idea that spatial updating is always an online process, it is also possible for updating to be performed offline, using spatial representations in long-term memory (LTM). According to this account, people monitor minimal information about their motion during movement and subsequently apply the consequences of this motion to relations stored in an enduring and relatively rich LTM representation. For example, given an enduring and relatively comprehensive representation of the local environment, such as a cognitive map (Gallistel, 1990; Tolman, 1948), it is possible to update by simply monitoring one's acceleration through space. At the terminus of the movement, a new position and orientation may be calculated via path integration (Loomis, Klatzky, Golledge, & Philbeck, 1999), which can then be applied to a stored representation to estimate the self-to-object relations that are necessary for a given task. Although the storage requirements for offline updating are relatively large compared with those of online updating, it can enable people to derive the spatial information necessary to complete a task without continually having to keep track of objects that are irrelevant to the current situation, are in a distal environment, or are simply beyond their capacity to monitor online.

Evidence for offline updating (as well as online updating) was shown by Amorim, Glasauer, Corpinot, and Berthoz (1997), who used an instructional manipulation to induce participants to update either online as they moved or, alternatively, offline after walking. In their experiment, participants were asked to study a three-dimensional capital letter "F" and then to update both its position and orientation relative to themselves while walking along a path. While walking, participants were required either to report continually the number of steps they had walked (inducing them to focus on the path) or to report continually which side of the F was closest to them (inducing them to focus on the target). After walking, participants were asked to turn and face the target as accurately as possible, and also to rotate a small model of the F to indicate its orientation. The different patterns of errors between the two groups

led Amorim et al. to conclude that the participants who focused on the target updated online, whereas those who focused on the path updated offline, after reaching their destination. More specifically, offline updating was associated with greater error in facing the target and with occasional gross errors in estimating its orientation (e.g., misjudging the orientation of the F by 90°).

Although Amorim et al. (1997) showed that two distinct types of updating may be induced as a result of deliberate instructions (and methods that enforce these instructions), it is unclear which means of updating is relied on in the absence of such instructions. The primary aim of the present research is to examine this issue in the context of a simple updating task that is common in the spatial cognition literature: pointing to a set of surrounding targets before and after a rotation (e.g., Brou & Doane, 2003; Farrell & Robertson, 1998, 2000; Féry, Magnac, & Israël, 2004; Holmes & Sholl, 2005; Rieser, 1989; Wang, 1999; Wang & Brockmole, 2003a, 2003b; Wang & Spelke, 2000; Woodin & Allport, 1998; Wraga, 2003). In our experiments, participants learned a layout of target objects and were asked to point to each target without vision. Then, while still blindfolded, participants were asked to rotate a small amount (i.e., less than half a revolution) and then point to each target again. Although such a task is often conceptualized as involving online updating (e.g., Wang & Spelke, 2000), we were curious whether performance could be better explained as involving offline processes.

Our primary means of assessing the type of updating that participants used was to examine the effect on performance of the number of targets (set size) in one's environment. Because online and offline updating rely on different sources of information, they would be expected to have different characteristics and limitations. More specifically, we would expect any form of online updating to be sensitive to set size, while offline updating should remain unaffected. Online updating is generally thought to require the simultaneous (and independent) monitoring of each egocentric spatial relation as one moves (see, e.g., Wang & Spelke, 2000), and there is good evidence that online updating requires—at least to some degree—the resources of working memory (Böök & Gärling, 1981; Lindberg & Gärling, 1981a, 1981b, 1983; Sholl & Bartels, 2002; for a review of working memory involvement in updating, see Sholl & Fraone, 2004). Even if online updating is conceptualized as being wholly automatic or obligatory (e.g., Farrell & Robertson, 1998, 2000; Farrell & Thomson, 1998), online updating would still generally be expected to exhibit set size effects (Wang et al., in press), because simultaneously updating large numbers of landmarks is both functionally unnecessary and computationally expensive—even for an automatic process. When people update using online processes, we expect that error in updating will be relatively low for small set sizes and will either become progressively higher for larger set sizes (i.e., a linear function) or make a discrete rise from subcapacity sets to supracapacity sets (i.e., a stair-step function). If updating in our tasks is performed online, then determining its capacity limit (e.g., the point at which updating error sharply increases) represents a secondary aim of the present research. Additionally, because processing for online updating occurs in real time during movement, latencies for pointing responses should not be different before and after rotating. Thus, online updating predicts an effect of set size for pointing error, but no differences in latencies between pre- and postrotation estimates.

On the other hand, if updating is performed offline, it would not be expected to show set size effects for pointing error, because it is largely supported by LTM. Given sufficient encoding, relatively large sets of targets could be updated as easily as smaller sets. Processing for offline updating would be expected to occur after the rotation, perhaps on a trial-by-trial basis, as information about each target's location is recovered from LTM. Because of this, latencies for pointing responses would be expected to be slower after rotating if participants are updating offline. Thus, offline updating predicts no effect of set size for pointing error, but a difference between latencies before and after rotating. These effects for error and latency will be the primary means used to determine whether participants in our experiments relied on online or offline updating processes.

To date, research on the effect of number of the targets on updating performance has been limited. Yet, across the literature, there is a wide diversity in the number of targets that participants are asked to update (see Table 1), ranging from a single target (e.g., Amorim et al., 1997; Rieser & Rider, 1991) to more than 10 locations (e.g., Mou, McNamara, Valiquette, & Rump, 2004; Wang & Brockmole, 2003a, 2003b). In one of the few updating studies to manipulate set size, Rieser and Rider (1991) asked participants to update a layout of one, three, or five target objects while walking blindfolded to an unknown endpoint within the layout. Pointing estimates were significantly less accurate after walking than before but did not differ (either before or after walking) as a function of the number of targets. On the other hand, Wang et al. (in press) had participants update the locations of one, two, or three objects in a computer-simulated environment, and they found that the accuracy with which participants could place the simulated objects in their original locations after walking 120° around the perimeter of the layout was degraded as the number of objects increased. Close examination of Wang et al.'s figures suggests that much of this effect was due to the difference between updating a single object versus updating multiple objects.

In our first experiment, we examined the effect of number of targets for sets of size 4, 6, 8, and 10. Experiments 1B and 1C extended the range tested up to 15 targets and down to a single target, respectively. Because participants frequently reported relying on elaborate strategies in Experiments 1A–C, Experiment 2 sought to examine how people update when they do not encode their environment deliberately. In this experiment, participants encoded and pointed to targets incidentally in the course of com-

Table 1

A Sample of the Diversity of Memory Set Size Across the Updating Literature

Study	No. of to-be-remembered targets
Amorim et al. (1997)	1
Rieser and Rider (1991)	1, 3, or 5
Klatzky et al. (1998)	3
Waller et al. (2002)	4
Wang and Spelke (2000)	4, 6
Féry et al. (2004)	5
Brou and Doane (2003)	6
Farrell and Robertson (2000)	7
Mou et al. (2004)	9, 10
Wang and Brockmole (2003b)	11 (local and global combined)

pleting another task. Using the criteria outlined above, we examined patterns of error and latency as indicators of the type of updating that was used by default in both deliberate-learning and incidental-learning situations. To preview, no indications of a capacity-limited updating process was found in any of the experiments, and evidence is shown suggesting that participants relied on offline updating via effortful reconstruction of spatial information from an enduring representation. These findings have both theoretical and methodological implications that are discussed more fully below.

Experiment 1A

To date, previous investigations of updating capacity have tested rather limited set sizes (i.e., 1–3 targets tested by Lindberg & Gärling, 1981b; 1–5 targets tested by Rieser & Rider, 1991; 1–3 targets tested by Wang et al., in press). Because working memory processes have been implicated in online updating, and because some traditional accounts of short-term memory storage have advocated a capacity of around seven items (e.g., Banks & Fariello, 1974; Burrows & Okada, 1975; Miller, 1956), we thought that the lack of capacity effects reported by Rieser and Rider (1991) and Lindberg and Gärling (1981b) may have been the result of not testing large enough sets. Additionally, we wanted to use a relatively broad range of set sizes that spanned those commonly used in the updating literature (see Table 1).

Specifically, in Experiment 1A, we asked people to learn and update sets of 4, 6, 8, and 10 targets. This range was expected to be especially informative, given the conclusions drawn in two influential investigations of spatial updating (Mou et al., 2004; Wang & Spelke, 2000). In several experiments, Wang and Spelke (2000) tested participants' ability to update a layout of either 4 or 6 objects (separate experiments) after a small rotation and a subsequent disorientation procedure. Patterns of error led the authors to conclude that performance was governed by an online, constantly updated, egocentric representation of the environment. These results were also interpreted as providing evidence against enduring spatial representations (Wang & Spelke, 2000, 2002; but see Waller & Hodgson, 2006). Alternatively, Mou et al. (2004) had participants update a layout of either 9 or 10 objects (separate experiments) and make judgments of relative direction (e.g., imagine being at the banana, facing the wood, point to the pan). Patterns of error in these experiments were interpreted as indicating that participants used an enduring representation of the spatial layout with a preferred orientation (i.e., the learned view). Although these two studies both confirmed humans' ability to update self-to-object relationships during self-motion, they came to nearly opposite conclusions about the nature of spatial representations. One possible reason for this difference is that participants in Wang and Spelke's (2000) study never updated more than 6 targets to complete their task, whereas those in Mou et al.'s (2004) study never updated fewer than 9. These represent very different memory loads, and participants in these studies may have been relying on different types of updating. For example, if online updating is subject to a capacity limit falling between 6 and 9 locations, then the contrasting results of Mou et al. (2004) and Wang and Spelke (2000) might be expected. Mou et al. (2004) addressed this as a possibility, but no firm conclusions could be drawn because updating capacity was not manipulated or tested in either study. By focusing on the range between sets of 4 and 10 objects, we will be able to extend the ranges previously tested, and we can determine if differences in set sizes contributed to the discrepant conclusions drawn by Mou et al. (2004) and Wang and Spelke (2000).

Method

Participants. Twenty-six students (13 female, 13 male) from Miami University's psychology subject pool participated in exchange for course credit. All of the participants were tested individually in 45-min sessions. Two participants were omitted for failing to follow directions (e.g., lifting up the blindfold to look at the layout of objects), leaving a sample of 24 participants (12 female, 12 male).

Materials. During the experiment, each participant learned four different sets of real-world target objects: one each of size 4, 6, 8, and 10. The sets were composed of different thematically related sets of objects (kitchen objects, office objects, stuffed animals, and sports equipment) to reduce interference between sets. Different combinations of themes and set sizes were used for each participant so that each combination of a set size and theme occurred equally often.

For each participant and each set size, the locations of the targets were pseudorandomized[1] from a set of 15 predetermined locations that surrounded the participant in a 5.94 × 3.43 m space (see Figure 1). The distance and orientation from the participant to each of the 15 locations was staggered to create an irregular array.

Throughout the experiment, participants were seated on a rotating stool in the center of the layout. During testing, participants wore a V8 head-mounted display (HMD) from Virtual Research Systems, Inc. (Aptos, CA), that obstructed any vision of the layout and displayed either a black screen or simple text instructions (e.g., "Point to the Cup") at 72 Hz. Responses were made with a gun-shaped pointing device (ACT Labs PC USB Light Gun) equipped with an Intersense InertiaCube2 that provided online (180 Hz) measurement of the pointing direction (pitch, roll, and yaw) with a resolution of 0.01°, accurate to within 1° *MSE*. Headphones mounted on the HMD played white noise that masked ambient noise during testing and prevented participants from hearing the experimenter arranging the next layout. A Pentium IV computer running the Vizard virtual reality toolkit (WorldViz, 2003, 2004) controlled stimulus presentation and data collection.

Procedure. Following an informed-consent procedure and a brief introduction to the experiment, participants were asked to sit on the stool in the center of the room (see Figure 1). The experimenter then pointed to and named each object in the first layout, and participants were given as much time as they desired to learn the locations of the objects. During learning, participants were free to rotate on the stool, with the foreknowledge that they would have to point to each object while blindfolded. Participants were also instructed that at one point during each of the four layouts, they would be required to rotate on the stool while blindfolded and that they should try to keep track of all of the objects while rotating.

After learning the first layout, the participant was instructed on the procedure for using the pointing device. To start each trial, participants were required to lay the pointer flat in their lap (the resting position, enforced by the computer program). When the participants placed the pointer in its resting position, the next stimulus was presented (e.g., "Point to the Stapler"), and a latency timer was started. The latency timer was stopped when the pointer was lifted from the participant's lap (within 20° of vertical, as measured by the roll of the inertial tracker). All participants were given a set of practice trials with the pointing procedure and were

---

[1] Purely random arrangements were used in pilot testing but occasionally led to unusual configurations (e.g., with a set size of four, all four objects might be placed right next to each other, 20° apart, directly behind the participant). This would not be a comparable task to learning a layout that was more dispersed. Thus, the layouts were randomly selected from subsets of configurations that were of roughly equal irregularity and dispersion to balance difficulty across participants.
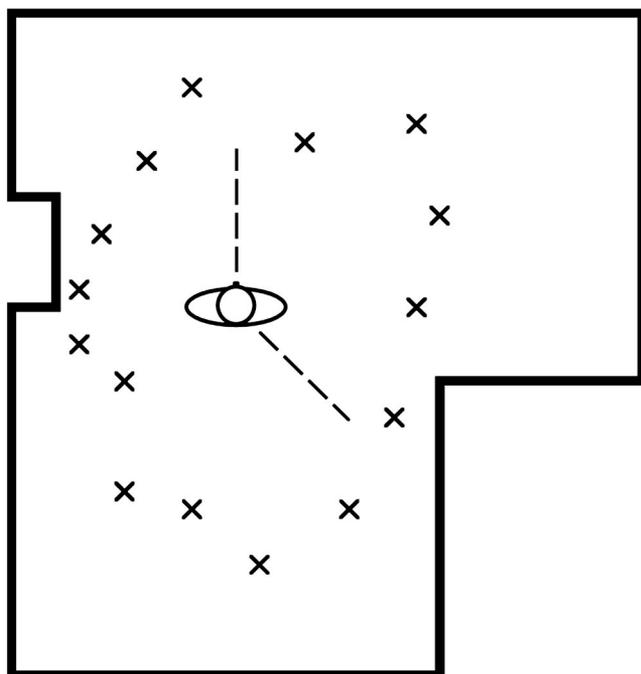
Figure 1. Schematic layout of the testing room, with the participant in the center of an irregular layout. Target locations were selected from the 15 potential object locations marked with *X*s. Sample participant facing directions before and after rotation are indicated by dotted lines.

corrected if they had any problems holding the pointer in the correct positions (no correction or feedback was given regarding participants' accuracy). Equal emphasis was given to speed and accuracy during the pointing instructions.

For each trial, a message such as "Point to the Stapler" appeared in the center of the HMD display (or on a computer monitor during practice) and remained on the screen until the participant pointed to the target and clicked the button of the pointing device. After the participant returned the pointing device to its resting position, the next trial was begun following a 2-s delay.

For each set size, participants pointed to the objects in each of three phases. After studying the layout, participants completed a *practice phase* with full vision, pointing to each target once as prompted. This was followed by a blindfolded, *prerotation phase*, in which participants made additional pointing estimates to each target. For the practice and prerotation phases, a learning criterion (described below) was imposed to ensure that each layout was adequately learned prior to updating. After giving their prerotation estimates, participants were reminded to keep track of all of the objects in the set as best they could and were prompted to rotate slowly either to the right or to the left (counterbalanced across sets for each participant) until the experimenter asked them to stop. All rotations were approximately 135° and were followed by a final, *postrotation phase* of pointing. During testing (i.e., the pre- and postrotation phases), each object was pointed to twice per phase, with the order of objects being randomized in two blocks (i.e., point to each of the objects once in a random order, and then a second time in a newly randomized order). These procedures were repeated until the participant completed all four sets.

It was assumed that participants could not correctly update an object's location after self-rotation if they did not know its position in the layout. Therefore, a learning criterion was imposed on the prerotation phase to ensure that participants had adequately learned the layout. To proceed to the postrotated phase, participants were required to point to each object in the layout within 45° of absolute angular error and maintain a mean

absolute error below 25° for the entire phase. The cutoff values for this criterion were determined in pilot testing to be reasonably passable, while still catching major errors (i.e., switching two objects or general sloppiness in pointing) prior to updating. If a participant failed to meet the criterion, he or she reviewed the layout and repeated the practice and prerotation phases until he or she was able to pass successfully.

At the conclusion of the experiment, informal exit interviews were conducted and participants were debriefed about the purpose of the experiment. Participants were given the opportunity to specify how they approached the task and what, if any, strategies they used. These interviews provided insights that complemented our formal dependent measures, outlined below.

Design and analyses. Experiment 1A had a 4 (set size: 4, 6, 8, 10) × 2 (phase: before or after rotation) factorial design with both factors manipulated within participants. Responses were measured in terms of both latency and error. For each trial, pointing error was measured as the signed difference between the orientation (yaw) of the pointing device when the button was pressed and the actual bearing from the center of the stool to the location of the target. Latency was measured as the time between the onset of instructions (e.g., point to the stapler) in the HMD and the time at which the participant began to move the pointing device (described above). Two dependent measures were analyzed in the following experiments: mean absolute updating error (which we referred to simply as *updating error*) and mean latency.

Of primary interest for this research was updating error. This variable was computed as the absolute difference in signed errors of pointing to a target before and after rotation, and it indicates how well the participant kept track of the remembered target location (regardless of the physical location of that target). For example, if a participant missed the stapler by −10° before rotating and +15° after rotating, the participant's updating error for that object would be 25°. In a perfectly accurate updating process, updating error should be 0°, reflecting the fact that pre- and postrotation estimates did not differ. Because updating error is an indication of the difference in estimates between the two testing phases, it is analyzed in the context of a one-way, repeated measures analysis of variance (ANOVA). Specifically, the criteria that we outlined above for detecting a capacity limit indicative of online updating calls for us to test one of two specific patterns: a linear function or a step function. The former was examined with a linear contrast (−3 −1 1 3), whereas the latter was tested with a series of three contrasts (−3 1 1 1; −1 −1 1 1; −1 −1 −1 3), each of which specifies a different point of increase for the step function. For example, the first contrast (−3 1 1 1) tests for low error with a set size of 4 and equivalently high error for sets of 6, 8, and 10 and would indicate an online-updating capacity limit of between 4 and 6 landmarks. Note that these 4 contrasts are nonorthogonal but provide the most direct test of our hypotheses. Conversely, offline updating predicts no effect of set size for updating error. Power analyses for null effects will be conducted using G*Power (Faul & Erdfelder, 1992) according to the guidelines provided in its documentation (Buchner, Erdfelder, & Faul, 1997).

Latency was measured as the time between stimulus onset and the beginning of the participant's response (described above). This represents a measure of "thinking time" and excludes the amount of time it took the participant to move his or her arm to the desired bearing and depress the response button. Latency was analyzed with a 4 (set size) × 2 (phase) repeated measures ANOVA with both factors manipulated within participants. The Greenhouse–Geisser correction for sphericity was used when analyzing within-participant main effects (Greenhouse & Geisser, 1959). Of particular interest for our hypotheses is the main effect of phase. As outlined earlier, offline updating predicts longer latencies after rotation, whereas online updating predicts a null effect.

Gender of the participant was included as a factor in all initial analyses. However, because it did not yield any systematic effects or interact with any variables of interest, gender was dropped from all analyses reported in this article. Additionally, in all of our experiments, we examined (both on a trial-by-trial basis and a participant-by-participant basis) whether partic-

ipants may have traded speed for accuracy. In no case was there an indication of a speed–accuracy trade-off.

Results

Updating error. Updating error in Experiment 1A averaged 20.87° (95% confidence interval [CI] ± 2.72°) across layouts (see Figure 2) and was constant across the number of targets (4, 6, 8, 10). Linear and step-function contrasts were conducted in the context of a repeated measures ANOVA to test for a set size effect in updating error. Set size was found to have no appreciable effect on updating error (all $Fs < 1$).

To maximize the power to detect an effect of set size, we combined the data from Experiment 1A with data from two other experiments not reported in this article. The design, methods, and results of these two experiments were the same as those of Experiment 1A in all major respects, and additional manipulations in the unreported experiments did not interact with set size. These three experiments yielded a combined data set with 70 participants who each updated sets of 4, 6, 8, and 10 target objects. A repeated measures ANOVA tested the effect of set size on updating error. This analysis still failed to yield any main effect, $F(3, 197) = 1.07$, $p > .35$, and none of our contrasts approached significance ($Fs < 1.51$, $ps > .22$). We used the variance and covariance estimates from the combined data to estimate the power of the omnibus ANOVA to detect a main effect of set size (the least powerful test). The power of this test to detect a medium-sized main effect of set size ($\mathbf{f} = 0.25$) was found to be .99. This is good evidence that the manipulation of set size had no effect on updating error.

Latency. The latency results for Experiment 1A are presented in Figure 3. Response times were significantly slower after rotation (mean increase = 151 ms), $F(1, 23) = 20.77$, $p < .001$, $\mathbf{f} = 0.95$. Additionally, latencies increased with the number of target objects, yielding a significant linear trend, $F(1, 23) = 27.41$, $p < .001$, $\mathbf{f} =$
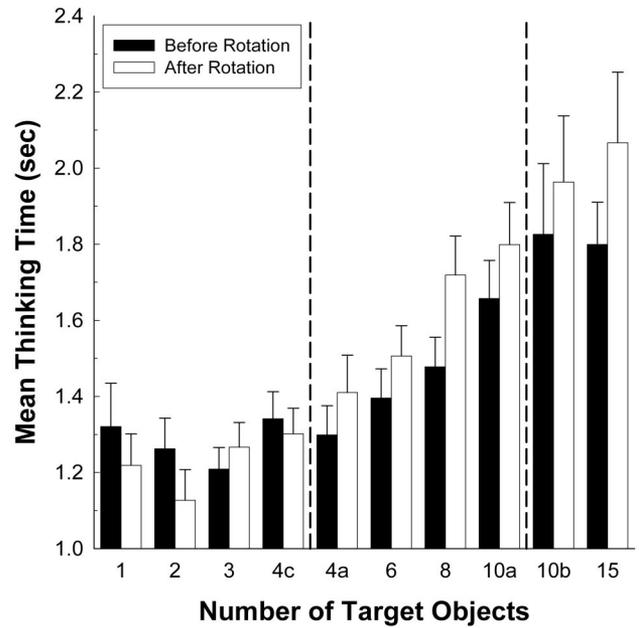
Figure 3. Thinking time (in seconds) from Experiments 1A, 1B, and 1C. Dashed lines separate results from each experiment. Error bars represent standard errors that include between-participants variation.

1.09. This linear increase was the same before and after the rotation and did not interact with the test phase (before or after rotation).

Discussion

Of primary interest in Experiment 1A were the findings that no differences in updating error were observed as the number of target objects was increased (see Figure 2). If updating had occurred online and had a capacity limit within the tested range, one of two patterns would have been expected. First, updating error might be expected to increase linearly as the number of targets was increased, reflecting the increased processing load. Alternatively, a sizable increase in updating error may have been expected at the point that participants' online updating capacity was exceeded, indicating a switch from a capacity-limited online process to an offline process better able to handle the larger number of targets. The absence of either of these trends casts doubt on the presence of online updating in the current experiment.

One counterargument to our claims that participants in this experiment engaged in offline updating is that online updating is sufficiently automatic and efficient so as to be unaffected by increases in memory load (at least up to 10 items). However, we do not think that this is a valid interpretation of our results for three reasons. First, as mentioned above, there is compelling evidence of working memory involvement in updating (see Sholl & Fraone, 2004). For example, participants walk slower if forced to update online (Amorim et al., 1997), are able to ignore their movement with sufficient processing time (Farrell & Robertson, 1998; Waller, Montello, Richardson, & Hegarty, 2002), and are inhibited in updating by a dual task (Lindberg & Gärling, 1981a, 1981b). Because working memory is involved in online updating, it is logical to suppose that the manipulation of set size should have
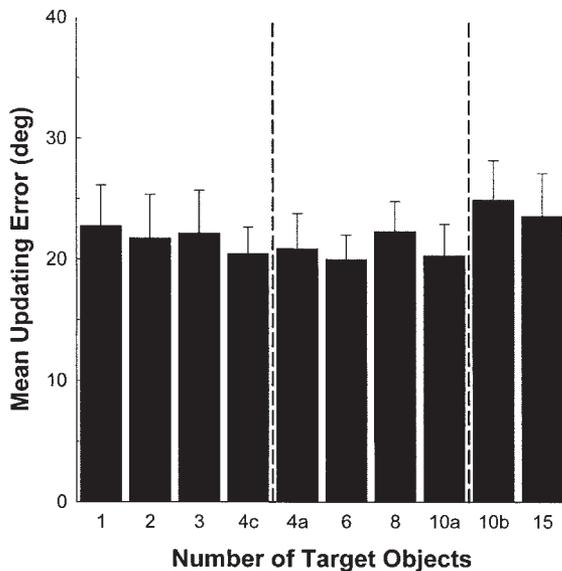

Figure 2. Updating error (in degrees) for Experiments 1A, 1B, and 1C. Dashed lines separate results from each experiment. Error bars represent standard errors that include between-participants variation. No indication of a capacity effect was observed between 1 and 15 targets.

affected online updating performance. Additionally, it has been argued that even an online process that is assumed to be automatic would not be expected to update an unlimited number of targets and thus would be expected to show set size effects (Wang & Brockmole, 2003b; Wang et al., in press).

Second, our participants required an average of 151 ms (95% CI ± 69 ms) to make their pointing estimates after rotating, indicating that additional processing time was required to make each updated response. This additional latency is indicative of time-consuming, effortful processing that is uncharacteristic of an automatic process (Hasher & Zacks, 1979). By our criteria, this is an indication of effortful offline updating.

Finally, participants were frequently observed forming elaborate mnemonic strategies to memorize the layout of objects, and then deliberately using these strategies during testing. Some participants, for example, reported making up a sentence with the first letter of each object, or verbally repeating the order of targets around the room until they felt confident in their knowledge of the layout. During testing, pointing responses could thus be made by counting the correct number of places down the list, rather than automatically accessing the correct bearing to the target. Those who used this type of strategy were often observed making counting motions during their pointing responses (i.e., pointing to each object in order and iterating through the list until the designated target was reached). Exit interviews confirmed the prevalent use of these types of strategies. In the entire experiment, only 1 participant failed to offer a precise, verbalizable strategy for completing the task. Of the 23 participants who did offer a strategy, only 5 (20.8%) reported something that could be construed as a visual or spatial strategy (i.e., visualizing the layout revolving as they turn, or remembering objects close to cardinal directions). Thus, it seems that many participants treated the layouts as mere lists, memorizing the target objects in a particular sequence at learning, using mnemonic techniques to commit the sequence to LTM, and then iterating through the list of targets at test. The analysis of latencies in Experiment 1A supports this interpretation. The linear increase in thinking time as the number of targets was increased was not moderated by the updating process and did not differ before or after rotation. This type of effect is consistent with traditional list-length effects (Atkinson, Holmgren, & Juola, 1969; Banks & Fariello, 1974; Flexser, 1978; Holmgren, Juola, & Atkinson, 1974), in which response time increases with the number of to-be-remembered items. Because participants used such deliberate mnemonics, required additional processing time after rotation, and showed no effect of set size, we believe that positing an automatic online process to explain participants' performance in Experiment 1A is untenable.

In summary, the results of Experiment 1A indicate that participants engaged in offline updating in this particular updating task. For errors, no capacity effects were observed between 4 and 10 target objects despite a power to detect these effects of at least .99. Moreover, participants required additional processing time after rotating. One immediate implication of these results is to rule out a capacity-limited updating process as an explanation of the disparate conclusions reached by Wang and Spelke (2000) and Mou et al. (2004). Whereas participants in their respective studies were asked to update differing numbers of targets, Experiment 1A indicated that—when updating is done offline—there is no appreciable difference between participants updating 4 or 6 targets (as in Wang & Spelke, 2000) versus updating 9 or 10 targets (as in Mou

et al., 2004). Experiments 1B and 1C extend the range of targets tested in this paradigm to provide a more comprehensive test of updating capacity.

Experiment 1B

As mentioned earlier, one potential limitation of previous investigations of updating capacity is the relatively narrow range of set sizes that were tested (Lindberg & Gärling, 1981b, 1–3 targets; Rieser & Rider, 1991, 1–5 targets; Wang et al., in press, 1–3 targets). It could be argued that, in these past studies, the relatively small number of items tested may not have exceeded humans' updating capacity. By the same token, it may be that humans have a high capacity for spatial updating—one that was not exceeded by the 10 locations tested in Experiment 1A. Experiment 1B extended this paradigm to test people's ability to update 10 and 15 targets.

Method

Participants. Twelve students (6 female, 6 male) from Miami University's psychology subject pool participated in exchange for course credit. All of the participants were tested individually in 45-min sessions. None of the participants had participated in Experiment 1A.

Materials, procedure, and analysis. Materials and procedures were the same as those used in Experiment 1A, with the exception that each participant was asked to learn and update two layouts—one comprised of 10 targets and another of 15. Because there were only two layouts, the four thematic sets of objects used in Experiment 1 were consolidated into two larger groups. Office and kitchen supplies were combined to form a "household objects" set. Sports equipment and stuffed animals composed a set of "toys." Set size order and theme were counterbalanced across participants. The design of Experiment 1B was similar to that of Experiment 1A with the exception that two fewer levels of set size were used. Thus, in place of the specific contrasts used to test for capacity limits in the other experiments in this article, updating error was analyzed with a paired-samples *t* test.

Results

Updating error. The results for updating error in Experiment 1B are depicted in Figure 2. As in Experiment 1A, the change in set size had no significant effect on updating error, $t(11) = 0.24$, $p > .80$. Collapsed across set sizes, the overall magnitude of updating error was 24.26° ± 4.06 (for a 95% CI), which is comparable with that in Experiment 1A.

Latency. Latency results were analyzed using a 2 (set size) × 2 (phase: before or after rotation) repeated measures ANOVA. As in the previous experiment, participants were 202 ± 132 ms slower after rotation, yielding a significant main effect of phase, $F(1, 11) = 11.32$, $p < .01$, $\mathbf{f} = 1.01$. Set size did not interact with phase ($p > .25$), and the linear increase in thinking time across set sizes that was observed in Experiment 1A was not significant in Experiment 1B, $F(1, 11) < 1$, $p > .65$. The lack of this effect seems largely due to slower responses with 10 targets in this experiment compared with those in Experiment 1A (a difference of 167 ms).

Discussion

The results of Experiment 1B closely paralleled those of Experiment 1A. There was no indication of a set size effect between 10 and 15 targets, and both error and latency levels were roughly equivalent to those of Experiment 1A. Updating error was again

relatively large (24.26° ± 4.06) and equivalent to that found with 4 to 10 targets in Experiment 1A (20.87° ± 2.72°). Additionally, participants in Experiment 1B also required additional processing time after rotating (increased = 202 ms ± 132). Each of these findings suggests that participants in Experiment 1B were updating offline.

The elaborate strategies observed in Experiment 1A were also prevalent in Experiment 1B. In exit interviews, 7 of the 12 participants (58.3%) reported that they approached the task by trying to memorize the order of target objects around the room. Only 2 participants (16.7%) reported trying to form a "mental map" or trying to visualize the layout. As in Experiment 1A, all but 1 participant reported some deliberate, verbalizable strategy. The list-length effect observed in the latency measure of Experiment 1A was not significant in Experiment 1B, but the lack of effect is likely attributable to the somewhat slower response times with 10 targets in Experiment 1B compared with those in Experiment 1A. Across experiments, the linear increase in response times—both before and after rotation—is consistent between 4 and 15 targets (see Figure 3), indicating that participants with these larger sets were probably also treating the layouts as sequential lists of targets.

In summary, the results of Experiment 1B replicate and extend those of Experiment 1A. No effects of capacity limitations in updating were apparent, and participants required additional time to respond after rotating. Taken together, the results of Experiments 1A and 1B indicate that our participants used offline updating when updating between 4 and 15 target locations.

Experiment 1C

The results of Experiments 1A and 1B gave no indications of capacity-limited, online updating of between 4 and 15 objects. The rationale for Experiment 1C was similar to that of Experiment 1B: to further extended the range of set sizes by testing for a capacity limit with relatively small sets of targets. Whereas Lindberg and Gärling (1981b) and Rieser and Rider (1991) found no effect of memory load for updating 1, 3, or 5 targets, the evidence presented by Wang et al. (in press) suggests that there may be a difference between updating a single object and updating multiple objects. Thus, in Experiment 1C, participants were asked to learn and update sets of 1, 2, 3, and 4 target objects.

Method

Participants. Twenty-four students (12 female, 12 male) participated in Experiment 1C. Undergraduate participants were drawn from the departmental subject pool in return for course credit. None of the participants had taken part in any of the other experiments.

Materials, procedures, and analyses. All materials, procedures, and analyses were the same as those in Experiment 1A, with the exception that participants learned and updated sets of one, two, three, or four target objects. Capacity limits in updating error were analyzed using the contrasts specified above to test for linear or step functions in the context of a repeated measures ANOVA. Of particular interest in this experiment was the contrast testing for a difference in updating a single target versus multiple targets, which represents the effect found in Wang et al. (in press).

Results

Updating error. Results for updating error in Experiment 1C are displayed in Figure 2. Like the previous experiments, updating error was unaffected by the number of targets. None of the planned contrasts approached significance (all Fs < 1). Collapsed across set sizes, the mean updating error for Experiment 1C was 21.26° ± 2.64, which is equivalent to the mean updating error observed in Experiment 1A (20.87°). A power analysis was conducted using G*Power (Faul & Erdfelder, 1992) to estimate the power of our tests to detect a medium effect ($\mathbf{f} = 0.25$)[2] of set size in this experiment. Specifically, it is important that we had sufficient power to detect a difference between updating a single versus multiple targets, as reported by Wang et al. (in press). The power of each contrast was found to be .84.

Latency. The latency results of Experiment 1C (shown in Figure 3) were in stark contrast to those of Experiments 1A and 1B. Participants performed equivalently before and after rotation, and no set size effects were observed. A 4 (set size) × 2 (phase) repeated measures ANOVA indicated no effect of set size ($F < 1$); phase, $F(1, 23) = 2.83, p > .10$; or interaction between set size and phase, $F(3, 55) = 1.55, p > .20$. The absolute level of thinking times collapsed across set size and phase in Experiment 1C (1.26 s ± 0.11) was equivalent to that observed with four targets prior to rotation in Experiment 1A (1.30 s ± 0.16).

Discussion

The results of Experiment 1C further extended the conclusions drawn from the previous experiments. Across these experiments, the number of to-be-updated targets—ranging from 1 to 15—had no effect on updating accuracy (as can clearly be seen in Figure 2), and in two of the three experiments participants required additional processing time after rotating (see Figure 3). We interpret these findings as evidence that participants engaged primarily in offline updating in this paradigm. It is interesting to note that the absolute level of updating error with very small sets (21.26° ± 2.64) was comparable with updating error with the larger sets used in Experiments 1A (20.87° ± 2.72) and 1B (24.26° ± 4.06); participants' performance was roughly equivalent with a single object or with 15 objects.

The null effects of phase and set size for latencies in Experiment 1C present an interesting case when compared with their significant effects in Experiments 1A and 1B. Taken as a whole, the global trend in latencies demonstrates a point of inflection at a set size of around four. Specifically, with four or fewer objects, there is no additional processing time needed after rotating (i.e., no main effect of phase), and latency is not dependent on the number of targets (i.e., no main effect of set size). With larger sets, however, latencies rise monotonically with the number of objects, and some additional processing time is required after rotating. It is possible to interpret this pattern as indicative of an online process with a capacity of around four objects. However, this account does not provide an adequate explanation of the increasing latencies with more than four objects that occurs *prior* to rotation. If the point of inflection in Figure 3 is due to a capacity limit of updating, then there is no reason why the same pattern of latencies should be apparent prior to any updating taking place. Rather than exhibiting

---

[2] Although Wang et al. (in press) did not report an effect size or the statistics necessary to compute one, their data figure is most consistent with a large effect. Given these findings, hypothesizing a medium effect size may thus be somewhat conservative.

clear evidence for online updating, we feel that the pattern of latency results across experiments can be parsimoniously explained by assuming not that set sizes differed in their manner of *updating* but rather in the ease with which they were *encoded*.

Additional evidence that latency effects resulted from encoding differences comes from the prevalence of self-reported strategies in each of the experiments. Indeed, the presence of main effects for both phase and set size was wholly coincident with the use of deliberate encoding strategies by our participants. Specifically, with the small sets of Experiment 1C, only 3 participants (12.5%) reported using a deliberate, verbalizable strategy (compared with 95.8% and 91.7% of participants in Experiments 1A and 1B). Most participants instead reported that they did not feel a strategy to be necessary, because the layouts were easy to remember. Furthermore, a close examination of Figure 3 shows us that, of all the participants who were tested with four targets, those tested in the context of larger sets (Experiment 1A) required extra processing time after rotating, whereas those tested in the context of smaller sets (Experiment 1C) did not. Thus, it seems likely that the pattern of latencies across these experiments is indeed an indication of a capacity limit, but it is a capacity limit of *memory* and *encoding*, not a capacity limit of *updating*. The capacity limit at learning dictated what type of encoding and memory strategies were used, which then had a direct influence on response latencies.

In Experiment 2, we sought to provide further evidence for this interpretation by testing participants with sets of 4 to 10 objects in an incidental learning paradigm in which encoding strategies were unavailable. If the pattern of latencies across Experiments 1A–C are an indication of a capacity-limited updating process (and not encoding strategies), then participants should exhibit a similar pattern of latencies to Experiment 1A (i.e., additional processing time after rotating and a linear increase of latency with set size). Conversely, if these effects were simply an artifact of encoding strategy, then preventing deliberate encoding strategies should remove the main effects of phase and set size, yielding a pattern of latencies more similar to that in Experiment 1C.

Experiment 2

In Experiments 1A–1C, we provided evidence that participants used offline processes (i.e., post hoc reconstruction of enduring spatial information) for spatial updating. It could be argued, however, that this tendency was the result of the explicit and deliberate nature of the learning task, which caused participants to use high-level conscious strategies and thus reduced their tendency to use a more perceptually driven online updating process. In Experiment 2, we examine whether offline updating is used by participants, even when learning is not deliberate and does not afford elaborate strategies. To this end, Experiment 2 introduced an incidental-learning paradigm that was designed to prevent elaborate LTM encoding techniques (i.e., memorizing the order of the targets around the room). Such a paradigm should not prevent any LTM encoding that may occur from natural interaction with the environment. However, it should prevent deliberate LTM strategies that are unrepresentative of spatial updating in day-to-day life, such as making up a sentence with the first letter of each landmark's name. If online updating is found in this type of task, then the offline updating observed in the deliberate-learning paradigm may be attributed to task-specific demands—perhaps because participants attempted to maximize accuracy at testing by forgoing normal processes.

Because no difference was found in deliberately updating between 1 and 15 targets in our previous experiments, we returned to testing the range of set sizes used in Experiment 1A. Returning to this range of set sizes should provide us with additional information that may shed light on the disparate conclusions reached by Wang and Spelke (2000) and Mou et al. (2004) and encompasses a range that has been used in many studies of spatial updating (see Table 1). Thus, separate groups of participants in Experiment 2 were asked to interact with (but not to learn or update) either 4, 6, 8, or 10 target objects by pointing to and "selecting" each object to rate it on various nonspatial characteristics. Pointing and ratings were conducted with and without vision of the layout under the pretence of comparing desktop and "immersed" computer interfaces, and a rotation was introduced by a feigned equipment failure. Thus, participants completed the same pointing behaviors and rotation as those in Experiment 1A, but they had no knowledge that they were being tested on the accuracy with which they could point to each target. The effectiveness of this paradigm was examined by the relative prevalence of strategy use as compared with the previous experiments.[3]

Method

Participants. Seventy participants from Miami University's psychology subject pool participated in exchange for course credit. Four participants were removed for failing to follow directions. Additionally, data from a relatively high number of participants (10) had to be removed in Experiment 2 for reasons related to the incidental paradigm. Specifically, there was some concern that participants who had participated in other experiments in our lab might guess the true nature of the experiment and make deliberate attempts to encode the target locations if they suspected that they were to be tested on their spatial abilities. Participants were questioned about this during the postexperiment debriefing, and 1 participant was removed for reporting prior familiarity with the type of research done in our lab. Additionally, 9 participants were removed for failing to learn the object locations adequately (i.e., exhibiting levels of pointing error greater than 2/3 of chance before rotating). Because we were interested in measuring updating performance, it was important that participants knew where the target objects were located before the rotation. In the end, 56 participants (28 female, 28 male) were included in the final analysis. None had participated in any of the previous experiments.

Materials. Materials were the same as those used in Experiment 1 with the following exception. Because the number of targets was manipulated between participants in this experiment, only one thematic set of target objects was required. Thus, all targets were drawn from the set of office supplies. Participants interacted with and updated either 4, 6, 8, or 10 targets.

Procedure. All of the participants were tested individually in 20-min sessions and were fully debriefed as to the true nature of the experiment at the conclusion of testing. As in the other experiments, the layout was

---

[3] The effectiveness of the incidental-learning paradigm was also tested by decomposing updating error into components that represented perceptual error (i.e., misperceiving the rotation) and the coarseness of our participants' spatial representations. Although overall updating error was equivalent across experiments, the portion of error attributable to representational coarseness was substantially higher in Experiment 2 than in Experiments 1A–C ($10.88° \pm 3.24$ compared with $3.48° \pm 0.92$, $4.92° \pm 2.71$, and $1.78° \pm 0.79$, respectively). If one grants that incidental learning leads to a coarser representation, it is interesting to note that the coarseness of the spatial representation contributed substantially to updating error.

arranged on the floor of the lab room in a different configuration for each participant, and the participant was seated on a rotating stool in the center of the layout. All references to spatial cognition were omitted from the introduction and informed-consent procedures. Instead, the experiment was introduced as user-testing for a product-rating interface. The interface (see Figure 4) allowed participants to examine a randomized list of "products" and rate them on dimensions such as usability, esthetics, and appropriateness for all age groups. Each to-be-rated object was pointed to and named by the experimenter, but participants were not given time to study the layout deliberately.

Prior to each phase, participants received instructions about the attribute on which they were to rate the objects, as well as the scale anchors (e.g., 1 = *not visually appealing,* 5 = *very visually appealing*). Participants pointed to each object with the pointing device and then "selected" it by depressing the button on the pointing device. Ratings were given by tilting the pointing device up or down, which moved an indicator along a scale (see left side of interface in Figure 4) in the corresponding direction on the interface. An integer between 1 and 5 was visible on the screen, showing the current scale value. After moving the pointer to the desired rating, participants depressed the button of the pointing device to enter their rating and were prompted to select the next object on the list.

Three rounds of ratings were conducted in the following order: perceived usability, aesthetics, and appropriateness for all age ranges. The first round of ratings was conducted by having participants interact with the interface on the computer screen. Thus, the first phase represented an incidental learning phase in which participants could interact with the layout and the object locations with full vision, but with no intention of learning or memorizing the targets' locations.

The final two rounds of ratings were conducted by having the participants interact with the interface in the HMD, under the pretence of comparing the usability of the interface on a desktop computer versus using virtual reality equipment. The HMD effectively blindfolded participants so that they no longer could see the room or the layout. Participants were instructed not to remove the HMD for the duration of the experiment and were told that, if they forgot exactly where one of the objects was located, they should give their best estimate. It is worth noting that, at this point, even if participants realized the true nature of the experiment, they no longer had perceptual access to the layout that would afford deliberate encoding.

Between the second and third rounds of ratings, parts of the interface suddenly disappeared, and a message was presented telling the participants that the orientation tracker had failed. The experimenter feigned surprise and asked the participant to rotate either to the right or to the left (counterbalanced across participants) so that he could straighten out the wires and reset the tracker. During pilot testing, it was discovered that participants would not rotate 135° without excessive prompting, so—unlike the previous experiments—participants were stopped after a 90° rotation. The experimenter pretended to reset the tracker so that the experiment could continue and asked the participants to finish the last round of ratings from their current facing direction.

Pointing estimates were recorded when the participant pointed to each object and depressed the button of the pointing device to "select" it. No instructions were given to emphasize high accuracy or speeded responses, but participants were told that if real product evaluations were being done, it would be important that people rated the correct object. As such, they were told that if they accidentally selected the wrong object, they would be prompted to try again. Thus, in the first round of ratings, while participants pointed with vision, a "Try Again" message was displayed any time the participant's pointing direction differed from the actual bearing to the target by more than 30°. No limitations or criterion for accuracy was imposed in the other rounds of ratings.

As with the previous experiments, exit interviews were conducted at the conclusion of the experiment. Participants were given an opportunity to specify what, if any, strategies they had used to complete the task. Partic-
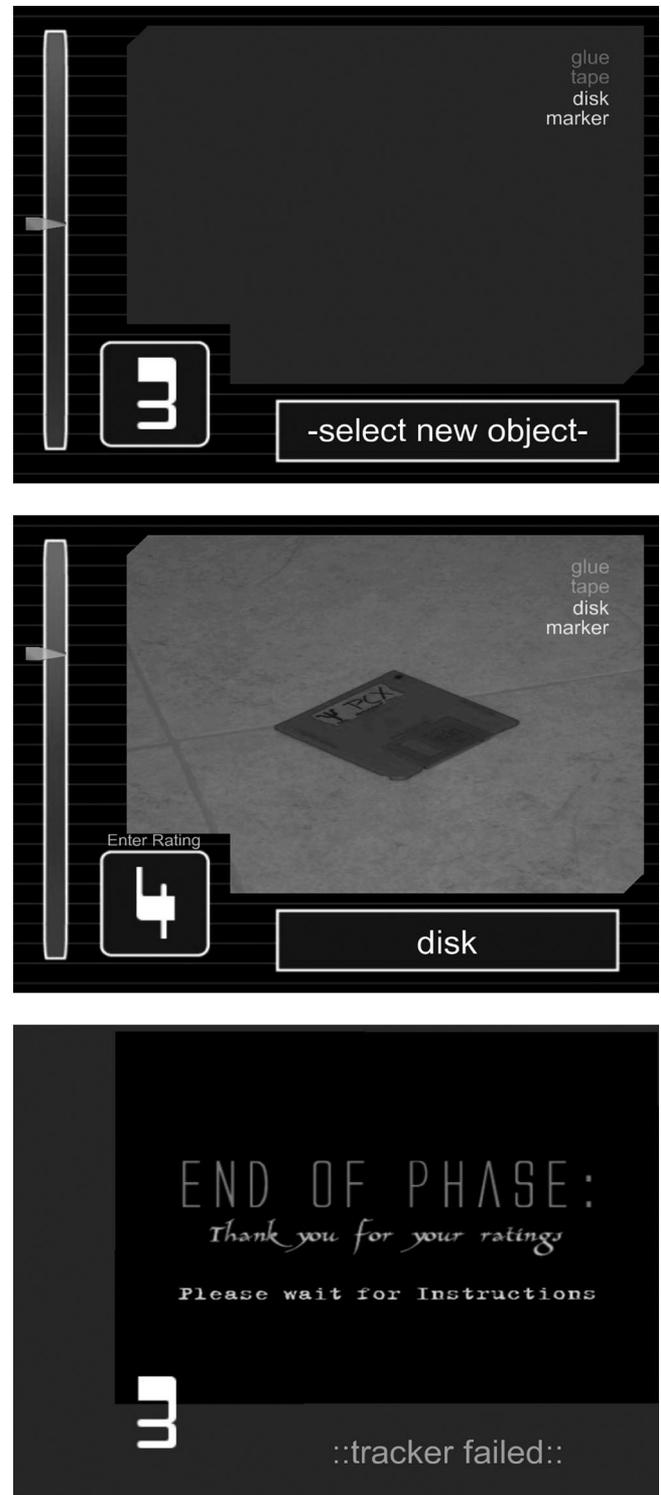

Figure 4. Screenshots of the interface that participants used in Experiment 2 under the pretense of gathering product ratings. Unlike this reproduction, the actual interface was in color. Top: Participants pointed to and "selected" each object in the list (upper-right corner of display). Middle: Once an object was selected, participants entered their rating (slider and number on the left side of the display) by adjusting the pitch of the pointing device, which caused a pointer to move along the slider. Bottom: Portions of the interface disappeared and a "tracker failed" message was displayed when it was time for the experimenter to ask participants to rotate.

ipants were also fully debriefed as to the true nature of the experiment and were asked not to discuss the experiment with any of their classmates.

Design and analyses. Experiment 2 represented a 4 (set size: 4, 6, 8, 10) × 2 (phase: before or after rotation) mixed-model design, with set size manipulated between participants and phase manipulated within participants. The initial phase—pointing with vision before rotation—was treated as practice phase and was not included in any analyses. Capacity limits for updating error were analyzed using the specific contrasts outlined earlier in the context of a one-way ANOVA. Latencies were analyzed in a mixed-model ANOVA, using the Greenhouse–Geisser adjustment where necessary (Greenhouse & Geisser, 1959).

Results

Updating error. The results for updating error in Experiment 2 are presented in Figure 5. Updating error increased slightly as number of targets was increased in Experiment 2. However, the linear trend was nonsignificant, $F(1, 55) = 2.59$, $p > .11$. Additionally, none of the contrasts specifying a step function were significant, although the contrast testing for a difference between updating 4, 6, or 8 targets and updating 10 targets did approach significance, $F(1, 52) = 2.97$, $p = .091$, $\mathbf{f} = 0.24$. Collapsing across set size, mean updating error was $25.02° \pm 3.43$, which was comparable with the updating error in Experiments 1A–1C. As before, a power analysis was conducted to estimate the power of each contrast to detect a medium-sized effect (Cohen, 1988) of set size. The power of each contrast was found to be .76.

Latency. Unlike the previous experiments, participants were not instructed to respond quickly. Moreover, the previous method of measuring "thinking times" was not available, because the incidental task did not require participants to return the pointing device to a resting position in between trials. Thus, the response latencies reported in Experiment 2 (shown in Figure 6) represent the total response time including participants' movement. Differ-
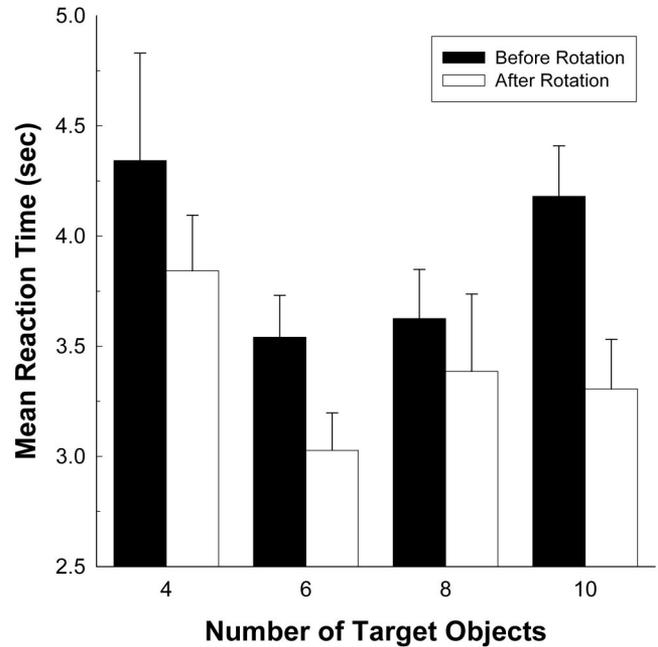
Figure 6. Reaction time (in seconds) for Experiment 2. Error bars represent standard errors that include between-participants variation. Unlike Experiment 1, no list-length effect was observed, and participants responded faster after the rotation.

ing from the previous findings, participants in Experiment 2 pointed 532 ± 342 ms *faster* after rotating than before, which was a significant difference, $F(1, 52) = 9.46$, $p < .01$, $\mathbf{f} = 0.43$. The linear increase in reaction times seen in Experiments 1A and 1B was not present in Experiment 2 ($F < 1$). Set size did not exhibit a significant main effect, $F(3, 52) = 2.37$, $p > .08$, and set size and phase did not interact ($F < 1$).

Discussion

Experiment 2 was designed to examine spatial updating under incidental-learning conditions, when participants would be unable to use mnemonic techniques and other deliberate LTM encoding strategies at learning. The magnitude of updating error in Experiment 2 ($25.02° \pm 3.43$) was roughly equivalent to that of Experiments 1A–1C (21°, 24°, and 21°, respectively), and again was not appreciably affected by the number of target objects. It is possible that our experiment was not sensitive enough to detect an effect of set size and that additional participants may help to bolster the effects that did approach significance in this experiment. However, because of the consistency of Experiments 1A–C, and the converging evidence from other measures in Experiment 2, we believe that participants in this experiment were primarily updating offline and not engaging in a capacity-limited online updating process.

Latencies in Experiment 2 showed a distinct change from those in the deliberate-learning paradigm. Specifically, in Experiments 1A and 1B, latencies exhibited a linear increase as the number of target items was increased, and participants were significantly slower after rotation—a difference that we presumed to reflect the additional processing time of an offline updating process. How-
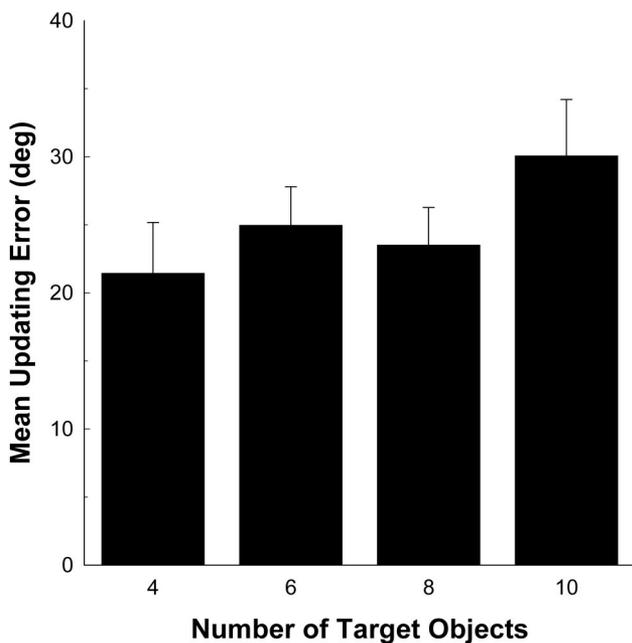

Figure 5. Updating error (in degrees) for Experiment 2. Error bars represent standard errors that include between-participants variation. As in the previous experiments, no indication of a capacity limit was observed.

ever, the reaction times in Experiment 2 did not exhibit either of these trends. While there was a significant main effect of phase for the latencies in Experiment 2, the effect was in the opposite direction as previous experiments; participants in Experiment 2 pointed *faster* after rotation. Earlier, we interpreted an increase in latency after rotation as evidence that additional processing time was required for an effortful, offline updating process whereby spatial information was recovered from an enduring representation. By the same token, it could be expected in Experiment 2 that some spatial information would have needed to be recovered once vision of the layout was occluded. Given that participants had no prior knowledge that they were required to remember the locations of targets that were no longer visible, it is reasonable to assume that some of this processing was performed after being blindfolded but before rotating. Exit interviews suggested that many of the participants, on donning the HMD, realized that they would need to remember where the objects were located in order to select them. Thus, we conclude that there was a certain amount of offline reconstruction of object position information from LTM that occurred prior to rotation, and this facilitated later reconstructions.

Additionally, the lack of a linear increase (also observed in Experiment 1C) suggested that participants in Experiment 2 were not treating targets in the layout like items in a to-be-remembered list. This was corroborated by the lack of strategies reported in exit interviews in Experiment 2. Of the 56 participants included in the final analyses, only 13 (23.2%) reported using any type of strategy. Furthermore, the types of strategies reportedly used by participants in the incidental-learning paradigm were, for the most part, qualitatively different from those reported in our previous experiments. While remembering the order was a primary strategy in the deliberate-learning paradigm, no participant reported attempting to memorize the order of objects around the room in Experiment 2, and only 3 of the 56 participants (5.4%) reported using what could be considered a deliberate strategy (e.g., grouping objects together). This represented a major shift from previous experiments in which nearly all participants reported using some type of deliberate strategy, and it further indicated that the incidental-learning paradigm was effective in eliminating the use of deliberate encoding strategies.

General Discussion

The present research was undertaken with two goals. First, we wanted to determine how participants updated during simple rotations—whether they used online or offline processes. Second, if online processes were implicated and set size effects were observed, we wanted to estimate the capacity limit of online updating. Online updating was assumed to be capacity-limited and perceptually supported, whereas offline updating was assumed not to be capacity-limited but to require time for effortful processing and to be supported by enduring memory representations. No capacity effects were found in Experiments 1A–C for participants deliberately learning and updating layouts of between 1 and 15 targets. Even under the incidental-learning paradigm of Experiment 2, participants' updating performance showed no indication of capacity-limited updating. Additionally, exit interviews with participants indicated that those in the deliberate-learning conditions (particularly Experiments 1A–B) reported using elaborate mnemonic techniques as necessary to aid LTM encoding of the layout. However, these mnemonics were apparently unnecessary

for adequate encoding with 4 or fewer targets (Experiment 1C) and were unavailable to participants in the incidental-learning paradigm (Experiment 2). This tendency probably led participants in Experiments 1A and 1B to treat the layout as a list of to-be-remembered targets rather than a spatial array, which was reflected in the list-length effects for response latencies in the two experiments in which those strategies were prevalent. Thus, the pattern of latencies across Experiments 1A–C cannot be attributed to a capacity-limited updating process with a capacity around 4 objects, as the observed effects were not reliant on set size but rather on encoding strategy. Finally, with the exception of Experiment 1C, additional processing time was required either after rotation (for deliberate-learning experiments) or upon participants' realizing that it was necessary to point to now-unseen targets (for the incidental-learning experiment). In sum, offline updating was implicated by (a) the lack of set size effects for accuracy in each experiment, (b) the additional processing time required in most conditions, and (c) the explicit attempts made by participants in the deliberate-learning paradigm to form elaborate, enduring memory representations that could facilitate updating processes at testing.

The results of these experiments have several implications for better understanding the phenomenon of human spatial updating and for understanding how to investigate it. Although spatial updating is commonly considered to be an online process (Farrell & Robertson, 1998; Farrell & Thomson, 1998; May & Klatzky, 2000), the present results are highly incompatible with any form of online updating of multiple targets in this paradigm (although they are compatible with online monitoring of the magnitude that one has rotated). We believe that online (possibly automatic) updating does in fact occur (see, e.g., Amorim et al., 1997; Farrell & Thomson, 1998) but that the experiments reported here simply did not produce it in our participants. Indeed, it seems highly plausible that both online and offline updating occur, as proposed by Amorim et al. (1997). This conclusion raises an important methodological issue with respect to much of the research conducted on spatial updating. The paradigm we used in these experiments involved commonly used tasks in the spatial updating literature (e.g., Brou & Doane, 2003; Farrell & Robertson, 1998, 2000; Féry et al., 2004; Holmes & Sholl, 2005; Rieser, 1989; Waller & Hodgson, 2006; Wang, 1999; Wang & Brockmole, 2003a, 2003b; Wang & Spelke, 2000; Woodin & Allport, 1998; Wraga, 2003). Investigators who use this paradigm may often consider themselves to be examining an online updating process. However, this research shows that participants engaging in this type of task likely use offline updating to recover self-to-object relations rather than updating those relations in real time. If one intends to study online updating, care must be taken to ensure that the process being used by participants is, in fact, an online updating process and not an offline reconstructive process.

With this in mind, an important avenue of research will be to understand the conditions under which updating is governed by online processes and the conditions under which updating is governed by offline processes. Experiment 2 provided the beginning of some separation between these two types of updating inasmuch as an incidental-learning paradigm was necessary to prevent elaborate strategies that would likely foster LTM representations and better support offline updating. However, even under these conditions, offline updating provided the most parsimonious explanation of our results.

Amorim et al.'s (1997) work provides further insight into the conditions that may elicit online updating. These investigators showed that participants did update a target object online when they were required to report continuously on the changing self-to-object relation as they moved. Additionally, Farrell and Thomson (1999) showed that participants adjusted their stride length as they approached a target location while walking without vision, suggesting that their participants were updating the self-to-target distance online. Thus, we speculate that online updating is only used in tasks that *require* moment-by-moment calculations, such as walking to a desired location, continually reporting the self-to-object spatial relationship to a landmark, or avoiding an obstacle while moving.

It is also possible that, because offline updating is largely memory-supported, online updating may be the preferred method in situations that do not allow a sufficiently rich LTM representation to be formed. This conjecture may explain the differing results found by Rieser and Rider (1991) and Wang et al. (in press). Wang and her colleagues found an effect of the number of to-be-updated objects by testing participants in a sparse virtual environment—containing only the target object(s) and a finely textured ground plane. Rieser and Rider—like the present experiments—failed to find the same effect when testing participants in a relatively rich, real-world environment. Such real-world environments offer visual structure and additional cues or landmarks that may be used to organize targets. For example, research has shown that people readily use salient environmental cues (McNamara, Rump, & Werner, 2003; Werner & Schmidt, 1999), straight walls and flooring edges (Shelton & McNamara, 1997, 2004), or intrinsic axes within a layout (Mou & McNamara, 2002; Mou et al., 2004) to organize spatial representations. Rich, real-world environments such as the ones used by Rieser and Rider (1991) and by the present experiments offer these types of cues. However, the only external cues available to participants in Wang et al.'s (in press) study were the target objects themselves. It is possible that updating under such conditions was carried out online because the environment did not easily afford LTM encoding.

Conversely, the conditions that foster offline updating also warrant investigation. For example, research on nested environments (Wang & Brockmole, 2003a, 2003b) shows that people fragment larger environments into smaller units, and although they do not actively update distal environments unless specifically instructed to do so, they can make roughly accurate estimates to distal landmarks if asked, and they are able to recover accurate spatial information about an environment—presumably by perceptually recalibrating a coarse, enduring spatial representation—upon reentering it. From a pragmatic perspective, segmenting the global environment provides an intuitive differentiation of updating functions; continually updating every known location would exceed functional necessity (e.g., one does not need to actively update the direction to the cereal aisle unless one is in a grocery store), whereas it is highly useful to store spatial information about an environment that can be retrieved and used when necessary. Thus, online updating would be predicted to be used within one's local environment, whereas offline processes would be preferred when updating landmarks or spatial relations in environments that are not perceptually available (i.e., imagined/distal environments).

The present research has two additional theoretical implications that are worth mentioning. First, demonstrating offline updating in the absence of instructions to do so supports the differentiation of online and offline updating proposed by Amorim et al. (1997) and adds converging support to two-system models of human spatial cognition that differentiate between transient (online) and enduring (offline) spatial systems (Amorim et al., 1997; Easton & Sholl, 1995; Mou et al., 2004; Waller & Hodgson, 2006). Second, our assertion of offline updating via enduring spatial memory presents a substantial challenge to recent models of spatial cognition that describe the updating process as governed predominantly by online, transient, and egocentric representations (Wang & Spelke, 2000, 2002), as well as to models that assume updating to be wholly automatic (Farrell & Robertson, 1998; Farrell & Thomson, 1998). Because participants in our experiments appeared to rely on effortful, offline processes and enduring representations for updating required additional processing time, the present results are difficult to reconcile with either of these positions.

References

Amorim, M., Glasauer, S., Corpinot, K., & Berthoz, A. (1997). Updating an object's orientation and location during nonvisual navigation: A comparison between two processing modes. *Perception & Psychophysics, 59,* 404–418.

Atkinson, R. C., Holmgren, J. E., & Juola, J. F. (1969). Processing time as influenced by the number of elements in a visual display. *Perception & Psychophysics, 6,* 321–326.

Banks, W. P., & Fariello, G. R. (1974). Memory load and latency in recognition of pictures. *Memory & Cognition, 2,* 144–148.

Böök, A., & Gärling, T. (1981). Maintenance of orientation during locomotion in unfamiliar environments. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 995–1006.

Brou, R. J., & Doane, S. M. (2003). Individual differences in object localization in virtual environments. *Spatial Cognition and Computation, 3,* 291–314.

Buchner, A., Erdfelder, E., & Faul, F. (1997). How to use g*power [Computer software manual]. Retrieved September 1, 2005, from http://www.psycho.uni-duesseldorf.de/aap/projects/gpower/how_to_use_gpower.html

Burrows, D., & Okada, R. (1975). Memory retrieval from long and short lists. *Science, 188,* 1031–1033.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 483–500.

Farrell, M. J., & Robertson, I. H. (1998). Mental rotation and the automatic updating of body-centered spatial relationships. *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 impairment due to right posterior cortical lesions. *Neuropsychologia, 38,* 585–595.

Farrell, M. J., & Thomson, J. A. (1998). Automatic spatial updating during locomotion without vision. *Quarterly Journal of Experimental Psychology, 51A,* 637–654.

Farrell, M. J., & Thomson, J. A. (1999). On-line updating of spatial information during locomotion without vision. *Journal of Motor Behavior, 31,* 39–53.

Faul, F., & Erdfelder, E. (1992). Gpower: A priori, post-hoc, and compromise power analyses for ms-dos (Version 2.0) [Computer software]. Bonn, Germany: Bonn University, Department of Psychology.

Féry, Y., Magnac, R., & Israël, I. (2004). Commanding the direction of passive whole-body rotations facilitates egocentric spatial updating. *Cognition, 91,* B1–B10.

Flexser, A. J. (1978). Long-term recognition latencies under rehearsal-controlled conditions: Do list-length effects depend on active memory?

Journal of Experimental Psychology: Learning, Memory, and Cognition, 4, 47–54.

Gallistel, C. R. (1990). *The organization of learning.* Cambridge, MA: MIT Press.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika, 24,* 95–112.

Hasher, L., & Zacks, R. T. (1979). Automatic and effortful processes in memory. *Journal of Experimental Psychology: General, 108,* 356–388.

Holmes, M. C., & Sholl, M. J. (2005). Allocentric coding of object-to-object relations in overlearned and novel environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1069–1087.

Holmgren, J. E., Juola, J. F., & Atkinson, R. C. (1974). Response latency in visual search with redundancy in the visual display. *Perception & Psychophysics, 16,* 123–128.

Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9,* 293–298.

Lindberg, E., & Gärling, T. (1981a). Acquisition of locational information about reference points during blindfolded and sighted locomotion: Effects of a concurrent task and locomotion paths. *Scandinavian Journal of Psychology, 22,* 101–108.

Lindberg, E., & Gärling, T. (1981b). Acquisition of locational information about reference points during locomotion with and without a concurrent task: Effects of number of reference points. *Scandinavian Journal of Psychology, 22,* 109–115.

Lindberg, E., & Gärling, T. (1983). Acquisition of different types of locational information in cognitive maps: Automatic or effortful processing? *Psychological Research, 45,* 19–38.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., & Philbeck, J. W. (1999). Human navigation by path integration. In R. G. Golledge (Ed.), *Wayfinding: Cognitive mapping and other spatial processes* (pp. 125–151). Baltimore: Johns Hopkins University Press.

May, M., & Klatzky, R. L. (2000). Path integration while ignoring irrelevant movement. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 169–186.

McNamara, T. P., Rump, B., & Werner, S. (2003). Egocentric and geocentric frames of reference in memory of large-scale space. *Psychonomic Bulletin & Review, 10,* 589–595.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63,* 81–97.

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.

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.

Rieser, J. J., & Rider, E. A. (1991). Young children's spatial orientation with respect to multiple targets when walking without vision. *Developmental Psychology, 27,* 97–107.

Shelton, A., & McNamara, T. P. (1997). Multiple views of spatial memory. *Psychonomic Bulletin & Review, 4,* 102–106.

Shelton, A., & McNamara, T. P. (2004). Spatial memory and perspective taking. *Memory & Cognition, 32,* 416–426.

Sholl, M. J., & Bartels, G. P. (2002). The role of self-to-object updating in orientation-free performance on spatial-memory tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28,* 422–436.

Sholl, M. J., & Fraone, S. K. (2004). Visuospatial working memory for different scales of space: Weighing the evidence. In G. Allen (Ed.), *Human spatial memory: Remembering where* (pp. 67–100). Mahwah, NJ: Erlbaum.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review, 55,* 189–208.

Waller, D., & Hodgson, E. (2006). Transient and enduring spatial representations under disorientation and self-rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32,* 867–882.

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, and Cognition, 28,* 1051–1063.

Wang, R. F. (1999). Representing a stable environment by egocentric updating and invariant representations. *Spatial Cognition and Computation, 1,* 431–445.

Wang, R. F., & Brockmole, J. R. (2003a). Human navigation in nested environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29,* 398–404.

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

Wang, R. F., Crowell, J. A., Simons, D. J., Irwin, D. E., Kramer, M. S., Ambinder, M. S., et al. (in press). Spatial updating relies on an egocentric representation of space: Effects of the number of objects. *Psychonomic Bulletin and Review.*

Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition, 77,* 215–250.

Wang, R. F., & Spelke, E. S. (2002). Human spatial representation: Insights from animals. *Trends in Cognitive Sciences, 6,* 376–382.

Werner, S., & Schmidt, K. (1999). Environmental reference systems for large-scale spaces. *Spatial Cognition and Computation, 1,* 447–473.

Woodin, M. E., & Allport, A. (1998). Independent reference frames in human spatial memory: Body-centered and environment-centered coding in near and far space. *Memory & Cognition, 26,* 1109–1116.

WorldViz. (2003). Vizard (Version 2.13) [Computer software]. Santa Barbara, CA: Author.

WorldViz. (2004). Vizard (Version 2.15) [Computer software]. Santa Barbara, CA: Author.

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.