

Combining Locations from Working Memory and Long-Term Memory into a Common Spatial Image

Nicholas A. Giudice,¹ Roberta L. Klatzky,² Christopher R. Bennett,¹ and Jack M. Loomis³

¹ Spatial Informatics Program, School of Computing & Information Science, University of Maine, Orono, Maine, USA

² Department of Psychology, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA

³ Department of Psychology, University of California, Santa Barbara, California, USA

Abstract: This research uses a novel integration paradigm to investigate whether target locations read in from long-term memory (LTM) differ from perceptually encoded inputs in spatial working-memory (SWM) with respect to systematic spatial error and/or noise, and whether SWM can simultaneously encompass both of these sources. Our results provide evidence for a composite representation of space in SWM derived from both perception and LTM, albeit with a loss in spatial precision of locations retrieved from LTM. More generally, the data support the concept of a spatial image in working memory and extend its potential sources to representations retrieved from LTM.

Keywords: working memory, long-term memory, spatial image, spatial updating, individual differences, spatial cognition

Correspondence should be sent to Nicholas Giudice, Room 348 Boardman Hall, Spatial Informatics Program, School of Computing & Information Science, University of Maine, Orono, ME 04469.
E-mail: Giudice@spatial.maine.edu

1. Introduction

1.1. Short-term and long-term spatial memory. The distinction between short-term or working memory and long-term memory, so central to memory research, has also been important for studies of spatial representation and processing (e.g., Amorim, Glasauer, Corpinot & Berthoz, 1997; Easton & Sholl, 1995; Huttenlocher, Hedges & Duncan, 1991; Mou, McNamara, Valiquette & Rump, 2004; Waller & Hodgson, 2006; Wang & Spelke, 2000). Long-term spatial representations allow us to plan travel within familiar environments and to recognize known locations. Representations in spatial working memory, in contrast, allow us to actively imagine layouts and perform mental transformations of these layouts, which potentially include ourselves. Spatial working memory also enables us to navigate and perform location-directed action when immediate perceptual support is not available. The contrast between memory stores has been further linked to a dichotomy between frames of reference for spatial layout, such that the long-term representation of space tends to use allocentric (extrinsically defined) coordinates, similarly to what is called a *cognitive map* (Tolman, 1948; O'Keefe & Nadel, 1978), whereas the short-term representation is egocentric (self-referred, e.g., Milner, Paulignan, Dijkerman, Michel, & Jeannerod, 1999; Mou, McNamara, Valiquette, & Rump, 2004).

A detailed model of the relationship between short- and long-term spatial memory representations and their neural underpinnings was proposed by Byrne, Becker and Burgess (2007). They suggest that long-term spatial memory corresponds to the formation of allocentric representations in hippocampus and related medial temporal areas, using inputs from both dorsal and ventral cortical streams. Spatial working memory, localized in the precuneus, operates with egocentric coordinates. The translation between short- and long-term representations is accomplished in the posterior parietal and retrosplenial cortices. By this means, a layout retrieved from long-term memory can be used to fill in an incomplete, perceptually based representation in working memory. For such translation to occur, the frames of reference must be co-registered; that is, long-term knowledge of landmarks or other salient environmental features must be aligned with the current egocentric frame of short-term memory. According to the model by Byrne and colleagues, current head direction is taken into account in order to generate and update the egocentric description in relation to long-term memory.

An important claim of the Byrne et al. model, also generally acknowledged in other theoretical frameworks, is that representations can be transferred between working memory and long-term storage. The importance of information transfer between memory systems is that information acquired perceptually can enhance stored knowledge about a locale. Simply put, this is spatial learning. Conversely, the translation from long-term to working memory can be used to imagine oneself in familiar surroundings. Thus, for example, a person can call up a memory of a kitchen, imagine standing in the doorway and turn to face the sink (Easton & Sholl, 1995; Rieser, Garing, & Young, 1994; Wang, 2004).

1.2. Interaction between memory stores. A direct implication of transfer between memory stores is that representations in spatial working memory can arise from two sources: encoding from perceptual inputs or reading in from long-term memory. While substantial research on human spatial cognition has been directed at representation within environments that are either perceptually based (e.g., Loomis, DaSilva, Fujita, &

Fukushima, 1992) or knowledge-based (e.g., Rieser et al., 1994), the concept of spatial working memory does not require that the two sources of inputs be mutually exclusive. Accordingly, the present research investigates whether spatial working memory can simultaneously encompass both of these sources. Specifically, we ask whether an active representation held in working memory can integrate spatial information from perceptual sources with information retrieved from long-term memory. To assess integration, we compare spatial judgments across locations from different origins in perception vs. memory to judgments involving locations from a single source.

Prior research clearly indicates that integration of multiple locations into a common representation in spatial memory is not an automatic result of mere proximity or environmental grouping. One constraint is temporal: locations that are learned during discrete intervals appear not to be spontaneously combined within a single representation in spatial working memory. Yamamoto and Shelton (2008) presented participants with a room-sized layout, subdivided into two sequentially learned layouts of five objects each. Subsequently the subjects were tested with judgments of relative direction (JRDs: imagine being at object x, facing y, and point to z). The JRDs involving objects from the same set were performed faster than those from different sets, indicating that the sets were not integrated into a common working-memory representation. In contrast, when the sets were presented sequentially, but intermingled during learning trials until a criterion was reached, there was no cost of switching between them. A similar result was found by Giudice, Klatzky, and Loomis (2009), but where the separation into sets was based on input modality (haptic vs. visual). When the modalities were mixed at learning, integration was evidenced by the lack of inter-modal switching cost in a JRD task. However, segregation of modalities during learning, which also involved temporal separation, did produce such costs.

Another constraint on spatial integration appears to be common environmental scale. Even when spatial locations are encountered during the same time period, it appears that working memory fails to integrate environments at different levels within a spatial hierarchy. Brockmole and Wang (2002) asked whether two successive judgments of relative location about the same environment (both buildings vs. both within an office) were faster than when the environments changed between trials. Slower responses after a shift in scale indicated that the large- and small-scale environments were not integrated, although their relative spatial disposition was well known. Similarly, the cost of perspective change was greater within an environment than across environments (Brockmole & Wang, 2003). When people updated within a local environment they failed to update with respect to a global environment in which the former was included (Wang & Brockmole, 2003b). Moreover, the integration of contiguous locations at the same scale appears to be disrupted when clear environmental boundaries segregate sub-units (Wang & Brockmole, 2003a).

1.3. The spatial image. If spatial integration is hindered by non-contemporaneous learning or spatial segregation, can it succeed across the boundaries of memory storage systems? Relatively little work has addressed this question. Although it is possible to form an egocentric spatial representation in working memory entirely by retrieval from long-term memory (Easton & Sholl, 1995; Rieser et al., 1994; Wang, 2004), it is not clear how effectively an active representation in working memory can be populated with additional information from long-term memory. This possibility is suggested by a form

of spatial representation that we have called the *spatial image* (Giudice et al., 2009; Loomis & Klatzky, 2007; Loomis, Klatzky, & Giudice, in press; Loomis, Klatzky, Avraamides, Lippa & Golledge, 2007; Loomis, Lippa, Klatzky, & Golledge, 2002; Loomis & Philbeck, 2008). The spatial image refers to the contents of spatial working memory and is a three-dimensional representation of external space. It is flexible in origin, in that it can be derived from multiple sensory input modalities (vision, audition, touch), retrieved from long-term memory, or even constructed from language (see Bryant, 1997, for a similar conception). The spatial image is unlike a 2D image associated with vision, which has been characterized as "depictive" because it constitutes a direct mapping from a 2D display (Kosslyn, 1980, 1994). Visual images are in the mind's eye of the beholder, rather than being externalized in representational space. As a result, they do not vary with observer motion and, to the extent that they are subjectively externalized at all, their location in the world moves as the observer moves. In contrast, spatial images of stationary objects are externalized in 3D representational space. As the observer moves, the spatial image remains fixed relative to surrounding physical space (for exceptions, see Loomis & Philbeck, 2008, pp. 16), reflecting the fact that its egocentric coordinates have been updated. Our formulation of the spatial image is more explicit than, but fundamentally related to, representations found in other theoretical approaches (e.g., spatial working-memory in the theory of Byrne et al., 2007, or egocentric representations in the work of Mou et al., 2004).

Spatial images are conceived of as one form in which one's surroundings can be represented egocentrically, but are by no means the only one. Specifically, they may co-exist with modality-specific perceptual traces. Tasks that are intended to illuminate the properties of spatial images should therefore extend several seconds or more, beyond the expected duration of a perceptual trace (e.g., Sperling, 1960). One such task is spatial updating, where a participant is first exposed to one or more locations, then walks to them directly or indirectly without further perceptual feedback from the source. Successful updating during locomotion is indicated when the participant arrives at the same location, whether proceeding by the direct or indirect routes. Evidence that spatial updating is supported by amodal spatial images has been found in studies demonstrating equivalent spatial updating across a variety of source modalities, including vision, audition, touch, and spatial language (Giudice, Betty, & Loomis, 2011; Klatzky, Lippa, Loomis & Golledge, 2003; Loomis et al., 2002). People also show equivalent performance across source modalities when making judgments of relative direction from imagined perspectives (Avraamides, Loomis, Klatzky & Golledge, 2004; Giudice et al., 2009; Kelly, Avraamides, & Giudice, 2011).

Whereas spatial images represent locations in 3D space external to the observer, they are not necessarily accurate reflections of that world. To the contrary, systematic bias or noise in a percept which is the source of a spatial image will be inherited by the image itself. Studies in which images are formed by hearing sounds emitted at various points in space make this clear: The distance of sounds is systematically compressed in auditory perception (Loomis, Klatzky, Philbeck & Golledge, 1998), and the locations of the sources are accordingly distorted in the resulting spatial image held in working memory. When an observer walks to a sound source without vision, he or she will walk to the location designated by its spatial image, which tends not to be veridical (Klatzky et al., 2003; Loomis et al., 2002).

1.4 Effects of long-term storage on spatial images

In the case where a spatial image is formed by retrieval from long-term memory, an important consideration is whether the process of retrieval will itself add systematic error or noise. It has been suggested that the process of memory retrieval transforms information much like the process of perception, and there has been some effort to characterize the effects of retrieval in terms of a memory-specific "psychophysical" relation (see Algom, 1992, for review). Any such effects are presumably added on to any perceptual distortions that occurred when memories were initially stored. Alternative effects of long-term storage might also produce systematic errors, for example, a bias to pull locations toward hierarchical boundaries (Huttenlocher, Newcombe, & Sandberg, 1994) or to modify scale.

In addition to systematic bias, spatial memories retrieved from long-term memory are likely to be less reliable than those encoded perceptually. There is evidence for a loss in precision of perceptual traces during the storage process (e.g., Amorim et al., 1997; Huttenlocher et al., 1991; Waller & Hodgson, 2006), which would affect measures of variability across trials or individual subjects.

1.5 Individual differences and the spatial image.

Little research has focused on how spatial image formation and usage might vary across individuals, especially by using measures of actions directed toward target locations in 3D space. Individual differences in spatial processing are most often assessed with tasks of mental manipulation of shapes, including mental rotation (e.g., Just & Carpenter, 1985; Peters, Laeng, Latham, et al., 1995). Such mental transformations involving objects, or spatial *visualization*, have been suggested to involve fundamentally different processes from mental transformations involving egocentric perspective, or spatial *orientation* (Amorim & Stucchi, 1997; Simons & Wang, 1998). However, psychometric testing has not always supported a separation into two components of visual ability (e.g. Carroll, 1993). Kozhevnikov, Motes, & Hegarty (2007) found that spatial-abilities tests including mental rotation predicted the ability to coordinate frames of reference in a physics problem involving imagined perspective taking. Using psychometric tests more closely modeled after behavioral experiments, Kozhevnikov and Hegarty (2001) reported a dissociation between a perspective-taking test involving imagined rotation and tests of object manipulation, including rotation. Hegarty and Waller (2004) developed this approach further and found support for a two-factor model separating perspective-taking from mental rotation, although the dimensions were strongly correlated. They also found a strong relationship between paper-and-pencil tests of perspective-taking and a similar measure from a memorized environment, supporting the idea of a general ability to mentally assume novel headings within a configuration.

Relatively few studies have examined the relation between mental object transformations and the ability to adopt novel perspectives within real, 3D environments under conditions that motivate the use of a spatial image. Kozhevnikov, Motes, Rasch, and Blazhenkova (2006) assessed performance on a computerized perspective-taking task, mental rotation, and navigation within a large-scale environment. Perspective-taking, but not mental rotation, was correlated with navigation tasks of finding a shortcut and pointing to unseen targets. Although individual differences in spatial ability were not a major focus of the present study, the inclusion of a mental rotation test here allowed us

to correlate a test of spatial visualization with tasks that are assumed to draw on a 3D spatial image, either directly or after imagined perspective taking.

1.6. Present experiment: Composite spatial images from LTM and WM.

The present experiment assessed people's ability (a) to retrieve locations from LTM and instantiate them in WM as an egocentric spatial image, and (b) to form a composite representation in WM resulting from the retrieved information and locations directly stored after perception. Research reviewed above indicates that enduring spatial information can be called into working memory in the form of a spatial image. However, few studies have examined how spatial images are affected by a period of long-term storage of the original information. Moreover, empirical studies have tended to concentrate on either spatial representations encoded perceptually or retrieved from LTM, and not on composite images from the two sources. The present study was intended to fill these gaps by addressing two basic questions: (i) Does a spatial image formed by retrieving locations from long-term memory differ from a spatial image derived from the same content, but continuously maintained in WM? If so, does the period of long-term storage produce systematic bias and/or noise? (ii) Can a composite spatial image in WM incorporate locations from perceptual and LTM sources, to the extent that people can report cross-source spatial relations? More specifically, is there a "switching" cost when judgments involve relative directions between locations encoded from different sources?

To address these questions we used a multi-phase design that explicitly measured systematic error and noise arising from encoding, memory storage, and switching cost, and took each of these factors into account when assessing subsequent information flow. In the first phase, while standing at a viewing position, participants saw three locations in sequence, each presented as an illuminated object in a dark room. The objects learned and retrieved in this way will be called LTM targets. After five such exposures, they walked to each target without vision (Walk 1), by which means they demonstrated its location as represented in spatial working memory. A ten-minute delay was then instituted, during which the participants performed a competing spatial task (a mental rotation test), precluding retention of the initial spatial array in working memory. Following the delay period, the participant returned to the original viewing position and again walked to each LTM target (Walk 2), indicating its location in spatial working memory after retrieval from long-term memory. This second walk allowed us to assess the consequences of a period of LTM storage on the representation of the target locations. Without further delay, the participant moved to a new location and learned three new objects, called the WM targets, by the same procedure as before. This active movement between the viewing positions required the participants to engage in spatial updating of the spatial images of the LTM target locations. To motivate integration of the two object sets into a common representation, and to further assess consequences of LTM storage, the participant was then asked to point to both WM and LTM targets from the second viewing position (egocentric pointing).

The final phase of the experiment was intended to assess people's ability to integrate across targets encoded perceptually into working memory and those retrieved from long-term memory. It comprised judgments of relative direction (JRDs), in which the participant first imagined facing one object and then pointed to another, destination object. Such judgments are often used to assess the accessibility of a representation in

working memory that conveys the spatial relations among objects in an environment (Giudice et al., 2009; Mou et al., 2004; Shelton & McNamara, 2001; Waller & Hodgson, 2006). The important question for assessing integration is whether the JRDs differed, according to whether the objects were from the WM set, the LTM set, or crossed between them. Effects of LTM storage should emerge in JRDs involving targets from that set. The hypothesis that the contents of working memory in this task comprise spatial representations that are abstracted from their perceptual or memory source, and hence can be integrated into a common representation, makes a clear prediction: It should be possible to make JRDs between old and recently learned targets.

At an extreme, one could predict that there would be no cost attributable to switching between spatial representations of WM and LTM targets during these judgments. This criterion was, in fact, used in studies cited above comparing the effects of spatial hierarchies and sensory modalities. However, caution should be used when applying this extreme criterion when the origins of spatial images reside in long-term memory. Just as comparisons across sensory modalities must consider encoding biases (e.g., distance compression in audition), comparisons across different sources of entry into spatial working memory must consider the effects of the processes involved. In addition to encompassing encoding-induced bias, the concept of the spatial image does not preclude either systematic bias or loss of precision as a consequence of long-term memory storage and retrieval into working memory. If either bias or noise results, the stringent requirement of no switching cost may not be appropriate. Indeed, the dual model of memory representations described by others (Amorim et al., 1997; Huttenlocher et al., 1991; Waller & Hodgson, 2006) specifies that the LTM representations of space are less precise, if more enduring, than perceptually encoded spatial images. As will be described further, our design allows us to assess the consequences of LTM processing and to take these possible sources of error into account when evaluating switching costs in the JRD task. We can then test a further prediction of the hypothesis that spatial representations in working memory are abstracted from their perceptual source: The cost of switching between memory stores in JRDs should be no greater than would be expected from the loss attributable to LTM storage *per se*. If the cost is in fact greater, than the hypothesis would be disconfirmed.

Finally, as noted above, the present study allowed us to test the same individuals across a number of spatial subtasks, including encoding spatial images, storing and retrieving spatial information from long-term memory, and judging relative directions between spatial images in working memory. The delay period used a mental rotation task, providing a measure of object-transformation imagery that could be related to performance with spatial images.

2. Method

2.1. Participants. Twenty-four participants (13 male and 11 female), age 19-31 ($M = 23$, $SD = 3.5$), took part in the study. The research was approved by the University of Maine's local ethics committee and written informed consent was received for all participants, who received monetary compensation for their time.

2.2. Apparatus and Stimuli. The target objects consisted of six pictures: car, clock, comb, fish, kite, and tie, chosen because all had monosyllabic, high-imagery names with recall latencies between 700 and 800 ms (Snodgrass & Yuditsky, 1996). To

ensure that nothing but the image was illuminated during the experimental trials, each target was created by laser-cutting (Epilog laser engraver, Zing 24) the outline of the image (12.7 x 12.7 cm) from a piece of black paper (20.3 x 20.3 cm), which was mounted on the front of a 20.3 x 20.3 x 31.8 cm wood box. Each box was lined with a string of red-colored, battery-powered LEDs and equipped with a custom-fabricated dimming circuit. The light shone through the etching on the paper to display the image but illuminated nothing else in the room. The stimulus box was mounted on a microphone stand, with the center point of the image at a height of 1.5 m. The 6 x 6 m lab was completely dark except for the dim light coming from each target.

Figure 1 shows the layout of the six objects and two viewing positions in the room. Initial learning of the three LTM targets occurred from a position designated A, and subsequent learning of the three WM targets occurred from a position designated B, 1.5 m to the right of A. The LTM targets were placed in the room so that their azimuth / distance pairings relative to the viewing position and a vertical reference axis were 2.12 m at 27 degrees, 2.14 m at 217 degrees, and 3.05 m at 122 degrees. The three WM targets were placed relative to the viewing position and vertical axis at 1.69 m at -73 degrees, 2.60 m at 35 degrees, and 3.00 m at 150 degrees. These positions were chosen so that in the absence of errors due to encoding and memory storage (for LTM targets), the WM and LTM targets would have fairly comparable distributions of angles around the B position, from which judgments of relative direction (JRDs) were made.

Insert Figure 1 about here

Virtual visual stimuli used only during a practice session were delivered through a head-mounted display (eMagin, Z800 HMD), and auditory instructions during testing were delivered through a pair of wireless headphones (Creative labs, HS-1200). An infrared LED mounted on the headphones was used to track user movement during the experimental trials by means of a four-camera PPT optical tracking system (Worldviz Inc., Santa Barbara, CA). Recording of tracking data and sequencing of experimental trials was done using the Vizard 3D rendering suite (version 3.17, Worldviz). A Nintendo Wiimote was used for making responses in the blind walking trials, and pointing judgments were made using a custom-built haptic pointing device (HPD) based on a high precision commercial joystick (CH Products, Vista CA) topped with a response button, with a constant deflection length in all directions. To enhance haptic cues about the stick's angular position, it was attached to a 1-m aluminum extension pole and mounted on a 38.1 x 60.1 cm board.

2.3. Procedure. The experiment consisted of eight phases and adopted a within-subjects design, with each participant being exposed to each phase in sequence.

2.3.1. Phase 1: Practice. Familiarization with experimental procedures and equipment took place in an office adjacent to the main testing room. In order to demonstrate the blind walking task and three target distances, participants were given three practice trials where they looked at a taped marker on the floor, walked to the point with eyes closed, and then opened their eyes to get corrective feedback about differences between walked and actual distance. To familiarize participants with the HPD, they were brought into the completely darkened lab room, donned the HMD, and grasped the HPD, which was placed in front of them on the floor. They then took part in trials where they

were shown a clock face, followed by a direction (e.g. 3 o'clock), and they were to point the HPD in the given direction. Corrective feedback showing the difference between their response and the correct position was shown on the display after each trial. All 12 primary clock face directions had to be reported within 15 deg of absolute error for the entire sequence before the participant could continue (none took more than two run-throughs to meet criterion). Meeting the pointing criterion ensured that participants had correctly calibrated their pointing judgments to center on the joystick rather than some part of the body, from which the joystick was slightly offset. The HMD was used only during this phase of the experiment as a convenient means of training on the HPD.

2.3.2. Phase 2: LTM Learning. This phase was intended to establish representations of the initial targets for storage in LTM. The blindfolded participant was brought to a viewing position in the experimental room (Position A), positioning his or her feet against a T-shaped toe-rest that defined the 0 deg direction. Music was played through wireless headphones while the three LTM targets were placed in the room. Once arranged, the room was darkened and participants removed the blindfold for target exposure during the learning phase. The three targets were presented for 3 sec each in a randomized order, with a 2-sec interval between them. Each triplet was repeated five times for a total of 15 target exposures.

2.3.3. Phase 3: Blind Walk 1. The importance of this blind walking test was to measure the target locations as represented in working memory, since perceived rather than physical locations were used to calibrate subsequent responses. The participant replaced the blindfold, and the targets were silently removed. Facing 0 deg at the A position, the participant heard one of the three target names through the wireless headphones. The task was to walk directly to the remembered target location and to push a button on the Wiimote (dominant hand) once he/she reached this position. The button press logged the distance and direction values (from the origin) as provided by the optical tracking system. After responding, the participant was led back to the origin and performed the same task for the other two targets (target order was randomized).

2.3.4. Phase 4: Delay Period. This phase was intended to eliminate any trace of the LTM targets in working memory and also provided a test of individual differences. The participant was brought to an adjoining lab room, removed the blindfold, and performed a paper-and-pencil mental rotation test (MRT-A from Peters et al., 1995, based on the stimuli of Shepard and Metzler, 1971). For each array, the participant marked the two stimuli that could be mentally rotated to match a target. Twenty-four trials were done over a 6-min session, split into two 3-min blocks with twelve trials per block. The total time needed for this delay period was 10 minutes.

2.3.5. Phase 5: Blind Walk 2. To measure memory-induced shift in the LTM representation, a second blind walking test was given. The blindfold was replaced, and the participant was brought back to the main lab room and aligned with the orientation block at the A starting position. He or she then followed the same procedure as for the Walk-1 trials in Phase 3. The LTM targets were not re-exposed before testing, and test order was randomized.

2.3.6. Phase 6: WM Learning. The purpose of this phase was to establish a representation of the second set of targets in working memory. Immediately after completing the second set of blind walking trials, the participant side-stepped 1.5 m to the right from the A viewing position to a second toe-block marking the B viewing position

(see Figure 1). The importance of this movement procedure is that the participant had to update his/her self-position relative to the LTM targets, with the result that the report of target locations could not rely on direct sensory input. Klatzky et al. (2003) found that target positions are successfully updated after sidestepping, so any interference from this task should be minimal. The participants then removed the blindfold, and the WM targets were visually exposed following the same procedure as was used during learning in Phase 2 with the LTM targets. After the 15 exposure trials with the WM targets, the participant replaced the blindfold and stood in place at the B viewing position. Note that no further assessment of the representation of WM target locations, such as by blind walking, was made at this point, because we did not want to interfere with the representation that had just been formed visually and was presumably held in spatial WM.

2.3.7. Phase 7: Egocentric pointing. To motivate integration of the target sets, as well as to measure the directional bias and noise in the WM and LTM target representations, the participant was tested on the target locations. Using the HPD at the B viewing position, the participant was instructed to point to all six target locations, blocked by array (WM or LTM; block order was balanced between participants). On each trial, the participant heard through the headphones “Point to the [target name],” oriented the HPD, and pushed the response button.

2.3.8. Phase 8: JRDs. In the final experimental phase, the participant used the HPD at the B viewing position to make judgments of relative direction (JRDs) between all pairs of unique target locations, constituting 30 trials in total. Each trial began with an instruction given through the wireless headphones in the form, “Imagine you are facing the [target name]”. The participant pushed the response button upon completing this task, the time was logged, and the next instruction was played in the form, “Point to the [target name]”. The participant pushed the button a second time to indicate completion of the pointing response and the computer logged the response time and angle. Of note, the facing and destination targets could be within or between arrays.

3. Results

To address the specific issues raised in the introduction, a set of analyses was implemented, as follows.

3.1. Accuracy of target learning. The accuracy with which the targets were learned after five exposures was assessed from the endpoints of subjects' walks to LTM targets without vision, immediately after learning (Walk 1). Figure 2 shows the centroids of the stopping locations for Walk-1, as well as the actual locations and the corresponding data for the walk following the delay period (Walk-2). It is apparent that target distances were under-estimated in Walk-1. The walked values in relation to the physical values were well fit by a linear function with slope 0.42 and intercept of 0.85 m. The significance of this trend was confirmed by finding the confidence interval (CI) around the mean signed distance errors in Walk-1, averaging over targets, to be $0.55 \text{ m} \pm 0.14$, which excludes the value of zero. Because isolated targets closer than 3 m, near eye level, and viewed in an otherwise dark room tend to be perceived further away instead of closer (e.g., Ooi, Wu, & He, 2006; Philbeck & Loomis, 1997), the underestimation here at such a distance is unexpected, perhaps reflecting our use of back-illuminated pictures. The egocentric direction errors in Walk-1, which are apparent in Figure 2, were not

systematically related to the target locations, and as can also be seen, were not further observed in Walk 2 (without further feedback), indicating that they were simply the result of random variation. Moreover, because directions over this range have been found to be perceived quite accurately in previous research, e.g., Loomis et al., 1998, direction errors will not be considered further in any detail.

Insert Figure 2 about here

3.2. Effect of LTM storage on represented target locations. To assess the effect of memory retention on systematic distance errors for LTM targets after the 10 minute delay, we computed the signed difference between the distance of the stopping point for Walk-2 and the distance of the group centroid of the Walk-1 stopping point. Recall that the latter is taken to be the represented location of the target immediately after learning. The CI for the signed distance error, averaging over targets, was found to be $-0.11 \text{ m} \pm 0.14$. Importantly, this CI included zero, indicating that the retention interval added no further systematic error in the representation of target distance. On this basis we used the Walk-1 centroids as the represented target locations for further analyses.

3.3. Localization in spatial working memory for old and new targets. To compare WM and LTM target localization from the B viewing position with respect to bias and precision, we compared signed and absolute errors in the egocentric pointing task that preceded the JRDs. For LTM targets, the direction of pointing was compared to the direction of the centroid of Walk-1 stopping points from the B viewing position, computed as if response directions in Walk-1 were without error. Signed error was defined as the lesser of the two values of the difference between the correct and response angle, with clockwise being positive. A t-test comparing the signed errors for LTM and WM targets (1.0 deg vs. -0.6 deg) was not significant, $t(23) = 0.28$, $p = 0.78$, two-tailed. This indicates that once the adjustment for bias from encoding was made, and given the absence of additional bias from LTM storage, no further systematic error was introduced by the movement from viewing position A to B. There was, however, a difference in absolute error, $t(23) = 2.75$, $p < 0.01$, one-tailed, (12.6 deg for LTM vs. 24.6 deg for WM targets). In principle, absolute error incorporates both noise and systematic error tendencies. As the signed error was small and equivalent across the two target sets, the absolute error difference can be taken to indicate that WM targets were reported with greater precision than LTM targets.

3.4. Judgments of relative direction for old and new targets. The JRD task is intended to assess participants' ability to integrate targets from LTM and WM into a common spatial image. Ideally, the azimuths of the facing object, relative to the physical facing direction of the participant, and the JRD angles would be matched across judgments that use LTM stimuli only, WM only, or mixed conditions. However, stimulus locations in this study were selected under multiple constraints: using no more than six targets to accommodate WM capacity limitations, placing targets at discriminable azimuths, and attempting to achieve comparable distributions of WM and LTM azimuths around the final report location (the B viewing position). Furthermore, although we could manipulate the objective JRD angles, we could not fully control the correct JRD values, because they depended on subjects' representation of target locations after encoding and memory storage. Table 1 reports the locations of the centroids of the

targets as represented from B, based on the true angles and the response distances from Walk-1. Although the distances of the WM targets from B do not affect the JRDs, for purposes of comparison to LTM targets they have been adjusted in the table by the compression observed in Walk-1 distances. Ultimately, the angles for JRDs from WM to LTM targets were completely matched with those in the reverse direction, since the same target pairs were used but switched with respect to facing and destination objects. However, the angles used in JRDs were not balanced across other combinations of LTM and WM objects (as can be seen in the x axis of Figure 3), leading to subsidiary analyses by item, as described below.

Insert Table 1 about here

We considered several measures when evaluating JRDs. The first measure was the time to retrieve the facing object (imagination time), which was 1.74 sec vs. 1.70 sec for LTM and WM targets. These did not reliably differ, $t(23) = 0.31$, $p = 0.38$, one-tailed. This result indicates that by the time of the JRD test, the LTM targets were sufficiently reinstated so as to be accessed as quickly as those encoded directly into WM.

The additional measures, signed error (defined as above), absolute error and response time, relate to the judgment of relative direction. Means and standard errors are shown in Table 2 subdivided into four cases, according to the set membership of the facing object and the destination object (WM vs. LTM). One-way repeated measures ANOVAs on this *object-set* variable were used to evaluate how well the relative locations of objects arising from different memory stores can be related, and in particular, whether there are costs of switching between WM and LTM target sets. Note that each of the measures could be affected by LTM storage. Our design, however, allows us to assess the consequences of LTM storage and to take these into account when evaluating switching costs in the JRD task, as we explain in connection with each analysis.

Insert Table 2 about here

Signed error in the JRDs indicates systematic bias in assessing angular differences. This measure could potentially be affected by any shift in the represented locations of the LTM target set over the period of memory retention; however, as noted above, comparison of Walk-2 and Walk-1 centroids indicated no further shift after initial encoding. For signed error, the object-set effect was not significant, $F(3,69) = 1.10$, $p = 0.35$. Mean signed errors in JRDs in each condition were less than 10 deg (although some items exhibited sizeable errors). This is in agreement with the equivalence in systematic error between WM and LTM targets found in the egocentric pointing task. The JRD judgment appears to have introduced no differential systematic error based on object set.

Absolute error in JRDs indicates the precision of the angular comparisons. If LTM targets are represented with less precision, as indicated by the egocentric pointing results, absolute errors in JRDs involving those targets may be correspondingly greater. In this case the cost of switching can still be assessed by comparing the average of pure cases (facing object and destination object for both LTM or WM targets) to switch cases (facing and destination object drawn from different target sets), as these averages match

the number of objects from each set. (This analysis assumes that the imprecision associated with the LTM target is essentially independent of the JRD process *per se*.)

The ANOVA on absolute error showed a significant effect of object set, $F(3,69) = 6.98$, $p < 0.001$. As our hypothesis regarding integration across stores favors a null comparison between mixed and pure conditions, we used t-tests to compare means, setting alpha to 0.05 without correction. By 2-tailed test, the LTM/LTM condition was significantly worse than all others, and the LTM/WM condition had significantly higher error than the WM/LTM condition. The LTM/WM condition was significantly greater than the WM/WM by 1-tailed test ($p = .04$). The WM/LTM and WM/WM were clearly equivalent ($p = 0.97$). These results indicate a loss of precision associated with long-term memory, particularly when both items tested in a JRD have been retrieved from LTM. Importantly, however, the absolute errors did not show an overall greater cost when the two objects come from different sets (i.e., switching cost). The average absolute JRD errors for pure pairs (LTM/LTM, WM/WM) were actually higher than for switch pairs (LTM/WM, WM/LTM), 45 deg. vs. 53 deg., respectively.

An item analysis relating absolute error to the judged JRD angle further indicated that the differential ranges of angles within the various combinations of LTM and WM were not affecting the results. The correlation of JRD angle with absolute error was $r(28) = .10$, n.s., and the means of residuals after removing the linear trend for angle showed the same pattern as the unadjusted means.¹ Note that in support of the assumption that the absolute error measures imprecision, there was a strong correlation across JRD angle between absolute error and the between-subject s.d. in signed error, which has been shown to measure noise (Avraamides et al., 2004), $r(28) = .85$. Between-subject s.d. was also notably higher for the LTM/LTM condition (average across angles = 70 deg vs. 60 deg for each other condition).

Response time for JRDs, defined as the period from naming the destination object in the JRD to the pointing response, is a critical measure of the effect of switching between sets. This measure is potentially impacted if storage in LTM degrades the WM strength of targets, in which case they may be more slowly accessed. However, any differences between times to access LTM and WM targets should be evident in the imagination times, and as that measure did not show any disadvantage for accessing LTM targets, any effect of object set on response time should indicate differences in JRD processing *per se*. The effect of object set on response time was not significant, $F(3, 69) = 1.21$, $p = 0.31$, indicating no switching cost on the JRD. Again, because the geometry of the layout resulted in different ranges of the correct angle across these comparisons, we further examined the response times for the individual items within each combination of facing object and target object, to determine whether object-set effects might be masked by different ranges of angles in the different sets. As can be seen in Figure 3, there was no trend relating response time to judged JRD angle, and the essential equivalence of pure and mixed sets is apparent across the stimulus range. The correlation of JRD angle with response time was only $r(28) = .19$, n.s., and the means of the residuals after removing any linear trend for angle showed the same pattern as the unadjusted means.

Insert Figure 3 about here

3.5. Correlations with Mental Rotation. We observed substantial variability across subjects' mental rotation scores, which ranged from zero to 83 and averaged 55.0 (s.d. = 23.6). Correlations between the MRT-A scores and other dependent variables are shown in Table 3. The only significant correlations involved the JRD task, particularly absolute error. Note that if the one obvious JRD outlier is removed, that correlation reaches .80. The correlation between rotation ability and absolute JRD error was also significant within each combination of LTM/WM targets. Figure 4 shows the strong relation between MRT-A and absolute error in the JRD task by subject, in comparison to the negligible correlation between MRT-A and absolute error in the egocentric pointing task.

Insert Table 3 and Figure 4 about here

Given the strong relation of the MRT-A to absolute error in the JRD task, we repeated the analysis of that measure while excluding nine subjects who had MRT scores below 50. This had essentially no effect on the outcome, as shown in Figure 5.

Insert Figure 5 about here

4. Discussion

The present study is motivated by the theoretical construct of a spatial image. As reviewed in the introduction, the spatial image is a working-memory representation of locations in the 3D environment so that they are accessible to action. Spatial images can arise from multiple sensory modalities and cognitive sources. In previous work, we and colleagues have shown that spatial language can be used to form representations that allow spatial updating (Loomis et al., 2002; Klatzky et al., 2003) and judgments of relative direction (Avraamides et al., 2004). Here, we extend the cognitive sources of spatial images to representations retrieved from long-term memory.

A principal goal of the present experiment was to compare spatial images encoded into working memory directly from perceptual processing, to those formed by retrieval of previously perceived spatial locations from long-term memory. This comparison particularly assessed the extent to which a period of long-term storage introduces degradation. A further goal was to test an important implication of two-store theories of spatial memory, namely, that information derived from perceptual encoding and long-term memory retrieval can be integrated within an active spatial image.

With regard to the first issue, we hypothesized initially that a spatial image formed by retrieval of location information from LTM could show systematic distortion (bias) and/or additional noise. The present results provide no indication that long-term storage adds bias. After accounting for systematic error in spatial representations resulting from perceptual encoding, we found that storage in and retrieval from long-term memory produced no further systematic error in spatial parameters. This was evident from several comparisons: Walking performance with LTM targets was equivalent before and after the delay period, signed error in egocentric pointing was equal for WM and LTM targets, and the signed errors in the JRD task were independent of the memory store from which items were drawn.

However, multiple measures indicate that storage in long-term memory reduced

the precision of spatial representations (in contrast to systematic error). The post-retrieval egocentric pointing task produced higher absolute error for LTM targets than WM targets, and JRDs involving LTM objects also showed higher absolute error levels as well as between-subject variability. The advantage for WM targets occurred despite the fact that only the LTM target presentation was followed by a blind walking test. Should the act of blind walking by itself enhance memory representations, its elimination for the WM targets could only under-estimate the magnitude of the WM advantage. The loss of precision due to storage of spatial memories is not unexpected, given the general assumption that memory strength declines over retention, and is consistent with other spatial memory studies, as was noted in Section 1.4.

The second major goal of the study was to demonstrate integration of locations drawn from LTM and perceptual encoding into a single spatial image. Multiple measures from the JRD task converge to indicate that a composite image was formed, by showing that once locations were retrieved from long-term memory and combined with perceptually based spatial images, judgments of spatial relations were insensitive to their processing origins. Importantly, the signed error in JRD was low whatever the source of the judged items; people made unbiased judgments of the angular relations between objects drawn from different stores, just as they made them between objects from a single memory source. In addition, the time to retrieve the location of the facing target in the JRD task was unaffected by whether the named object was a WM or LTM target, and similarly, response times for angular judgments were not moderated by target sources. The sole measure in the JRD task to show an effect of memory source was absolute error, but the trend was not toward higher error for mixed pairs (in fact, the means were in the reverse direction).

Rather than showing a switching cost, absolute JRD errors tended to be elevated for directional judgments involving LTM facing objects and were particularly high when both the facing and target object were drawn from LTM. The egocentric pointing task that preceded the JRDs also showed essentially a doubling of absolute error for LTM relative to WM items, indicating that the period of storage in LTM added substantial noise to the spatial representation. It is not surprising, then, that the JRD task would also indicate greater noise for LTM targets. However, the JRD task appears to have additional factors that affect absolute error, as an overall loss of precision for LTM items does not explain why performance suffered particularly when they constituted the facing object, as opposed to the destination object.

The present data are in accord with the proposal of Byrne et al. (2007) that long-term storage of spatial information can be used to enrich egocentric representations in spatial working memory. Their model emphasizes the need for a mechanism to align the stored traces in long-term memory with the ongoing egocentric frame of reference in working memory. Here, alignment was supported by returning the subjects to the original viewing position and heading direction after the retention interval, and then allowing them to move actively to the second viewing position. The results indicate that this active movement, and the spatial updating it afforded, was sufficient to align the two sets of targets.

Although the multiple measures in this study are derived from a small number of subjects relative to psychometric tests, they may be usefully related to spatial ability, particularly as they draw from actions toward objects in 3D space along with the standard

MRT. As discussed in the introduction, a theoretical and empirical distinction has been made between two components of spatial processing, visualization and orientation. Here, mental rotation scores (i.e., visualization) were strongly related to the JRD task, which involves imagined perspective taking, but not to egocentric pointing, which interrogates a person's relation to the environment defined by the spatial image. An essential difference between tasks that rely on spatial images and those that require imagined perspective taking is the magnitude of disparity from a coordinate system defined by the body of the observer. Klatzky and Wu (2008) suggested that as spatial tasks rely on frames of reference that are further abstracted from the body, they become more demanding. When a person stands in a room and makes egocentric pointing judgments, the operative frame of reference is that defined by the body. When the person makes JRDs that require imagined perspectives among objects in the room, a new frame of reference defined by room geometry or object layout is instantiated. The discrepancy between this frame and that of the body adds to task complexity. Still more extreme would be JRDs defined by object-centered coordinates, using objects lying within the room. The present findings suggest that a process similar to mental rotation is invoked when judgments use a frame of reference not aligned with the body, even if the frame is not object-centered.

This research offers novel evidence for the flexibility of origin of representations in spatial working memory. We have shown that a representation of layout around an observer can incorporate perceptually based and long-term-memory based locations. Importantly, the mixing of the two sources has no associated cost. There does appear to be a loss of precision for spatial information after retrieval from LTM into spatial WM; however, this does not further impede judging relations between targets from different memory sources. Finally, our results support a distinction between processes that access spatial images, localized in external space with respect to the body, from more abstract spatial thinking.

Acknowledgement

This research was supported by NIH grant 1R01EY016817. Roberta Klatzky acknowledges the support of the Alexander Von Humboldt Foundation during preparation of this manuscript. We thank Professor Michael Peters, University of Guelph, for the use of the Revised Vandenberg & Kuse Mental Rotations Test. We also thank Tim McGrath for programming assistance.

¹There are also more rearward LTM angles than WM angles. In an effort to assess whether this might affect the JRDs, we analyzed the data by the region (front, back) of the targets along with the WM/LTM status. There are only small numbers of target pairs in each cell of this 2 X 2 breakdown, and two missing cells (no LTM/LTM pairs for front/front or WM/WM pairs for back/back), but we could discern no clear effect of region on latency or absolute error. Moreover, any front/back effect would not explain the advantage of WM/LTM pairs over LTM/WM pairs, which are matched for region, nor the equivalence of mixed pairs to WM/WM pairs. Note also that if subjects effectively adopt a mental representation facing the first-named object in the JRD, the meaning of front and back change. "Back" targets are now those that have absolute JRDs greater than 90 deg. The by-angle analysis indicates no particular disadvantage for those angles; in fact, the WM/WM combination entirely relies on JRD angles greater than 90 deg but is the lowest condition in absolute error.

Table1. Perceived locations of targets relative to B viewing position (L = long-term memory, W = working memory).

Target	Distance (m)	Angle (deg. from vertical)
L1	1.70	335
L2	2.92	242
L3	1.18	164
W1	1.57	287
W2	1.95	35
W3	2.12	150

Table 2. Means (and standard errors) of measures from the JRD task by combination of facing object and destination object source.

Measure	Facing Object / Destination Object			
	LTM / WM	WM / LTM	LTM / LTM	WM / WM
Response time (sec)	5.6 (0.6)	5.4 (0.5)	5.7 (0.5)	5.2 (0.4)
Signed Error (deg)	5.6 (3.3)	-3.1 (4.7)	4.4 (5.4)	8.1 (5.0)
Absolute Error (deg)	51.3 (5.2)	45.4 (5.8)	60.7 (6.5)	45.2 (6.4)

Table 3 Correlations of MRT with Principal Dependent Variables
 *p < .05 **p<.001

Measure	Correlation with MRT
Walk-1 Absolute Distance Error	0.27
Walk-1 Absolute Angle Error	-0.18
Egocentric Pointing Absolute Error	-0.24
JRD Response time	0.07
JRD Imagination Time	-0.45*
JRD Absolute Error	-0.71**

References

- Algom, D. (1992). Memory psychophysics: An examination of its perceptual and cognitive prospects. In Algom, D. (Ed.) *Psychophysical approaches to cognition*, 441-513. Elsevier: (Advances in Psychology Series Vol. 92, pp. iii-x, 1-627.)
- 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.
- Amorim, M.A. and Stucchi, N., 1997. Viewer- and object-centered mental explorations of an imagined environment are not equivalent. *Cognitive Brain Research*, 5, 229–239.
- Avraamides, M., Loomis, J. Klatzky, R. L., & Golledge, R. G. (2004). Functional equivalence of spatial representations derived from vision and language: Evidence from allocentric judgments. *Journal of Experimental Psychology: Human Learning, Memory & Cognition*, 30, 801-814.
- Brockmole, J., & Wang, R. F. (2002). Switching between environmental representations in memory. *Cognition*, 83, 295–316
- Brockmole, J., & Wang, R. F. (2003). Changing perspective within and across environments. *Cognition*, 87, B59–B67
- Bryant, D. J. (1997). Representing space in language and perception. *Mind & Language*, 12, 239–264.
- Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the past and imagining the future: a neural model of spatial memory and imagery. *Psychological Review*, 114, 340 –375. DOI: 10.1037/0033-295X.114.2.340
- Carroll, J. (1993). Human cognitive abilities: A survey of factor-analytical studies. New York: Cambridge University Press.
- 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.
- Giudice, N. A., Klatzky, R. L., & Loomis, J. M. (2009). Evidence for amodal representations after bimodal learning: Integration of haptic-visual layouts into a common spatial image. *Spatial Cognition and Computation*, 9, 287-304.
- Giudice, N. A., Betty, M. R., & Loomis, J. M. (2011). Functional equivalence of spatial images from touch and vision: Evidence from spatial updating in blind and sighted individuals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37, 621-634. doi: 10.1037/a0022331
- Hegarty, M. N., & Waller, D. (2004). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, 32, 175-191. doi:10.1016/j.intell.2003.12.001
- Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological Review*, 98, 352–376.
- Huttenlocher, J., Newcombe, N., & Sandberg, E. H. (1994). The coding of spatial location in young children. *Cognitive Psychology*, 27, 115-147.

- Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: accounts of mental rotation and individual differences in spatial ability. *Psychological Review*, 92, 137-172. doi:10.1037/0033-295X.92.2.137
- Kelly, J.W., Avraamides, M.N., & Giudice, N.A. (2011). Haptic experiences influence visually-acquired memories: Reference frames during multimodal spatial learning. *Psychonomic Bulletin & Review*, 18(6), 1119-1125.
- Klatzky, R. L., Lippa, Y., Loomis, J. M., & Golledge, R. G. (2003). Encoding, learning and spatial updating of multiple object locations specified by 3-D sound, spatial language, and vision. *Experimental Brain Research*, 149, 48-61.
- Klatzky, R. L., & Wu, B. (2008). The embodied actor in multiple frames of reference. In R. Klatzky, M. Behrmann, & B. MacWhinney (Eds.), *Embodiment, Ego-space and Action* (pp. 145-177). New York: Psychology Press.
- Kosslyn, S. M. (1994). *Image and brain*. Boston, MA: MIT Press.
- Kosslyn, S. M. (1980). *Image and mind*. Cambridge, MA: Harvard University Press.
- Kozhevnikov, M., & Hegarty, M. (2001). A dissociation between object-manipulation spatial ability and spatial orientation ability. *Memory and Cognition*, 29, 745–756.
- Kozhevnikov, M., Motes, M. A. & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science: A Multidisciplinary Journal*, 31, 549 – 579.
- Kozhevnikov, M., Motes, M. Rasch, B. & Blajenkova, O. (2006). Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance. *Applied Cognitive Psychology*, 20, 397–417.
- Loomis, J. M., Da Silva, J.A., Fujita, N., & Fukusima, S. S. (1992) Visual space perception and visually directed action. *Journal of Experimental Psychology: Human Perception and Performance*, 18, 906-921.
- Loomis, J. M., & Klatzky, R. L. (2007). Functional equivalence of spatial representations from vision, touch and hearing: Relevance for sensory substitution. In J. Rieser, D., Ashmead, F. Ebner, & A. Corn (Eds.), *Blindness, brain plasticity and spatial function* (pp. 155-184). Mahwah, NJ: Erlbaum.
- Loomis, J. M., Klatzky, R., Avraamides, M., Lippa Y., & Golledge, R. (2007). Functional equivalence of spatial images produced by perception and spatial language. In F. Mast & L. Jäncke (Eds), *Spatial processing in navigation, imagery, and perception* (pp. 29-48). New York: Springer.
- Loomis, J.M., Klatzky, R.L., & Giudice, N.A. (in press). Representing 3D space in working memory: Spatial images from vision, touch, hearing, and language. S. Lacey & R. Lawson (Eds). *Multisensory Imagery: Theory & Applications*. Springer.
- Loomis, J. M., Klatzky, R. L., Philbeck, J. W., & Golledge, R. G. (1998). Assessing auditory distance perception using perceptually directed action. *Perception & Psychophysics*, 60, 966-980.
- Loomis, J. M., Lippa, Y., Klatzky, R. L., & Golledge, R. G. (2002). Spatial updating of locations specified by 3-D sound and spatial language. *Journal of Experimental Psychology: Human Learning, Memory, and Cognition*, 28, 335-345.
- Loomis, J. M. & Philbeck, J. W. (2008). Measuring spatial perception with spatial updating and action. In R. L. Klatzky, M. Behrmann, & B. MacWhinney (Eds.), *Embodiment, ego-space, and action* (pp 1-43). New York: Taylor & Francis.

- Milner, A., Paulignan, Y., Dijkerman, H., Michel, F., & Jeannerod, M. (1999). A paradoxical improvement of misreaching in optic ataxia: New evidence for two separate neural systems for visual localization. *Proceedings of the Royal Society of London B: Biological Sciences*, 26, 2225–2229.
- 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.
- O'Keefe, J., & Nadel, L. (1978). *The hippocampus as a cognitive map*. London: Oxford University Press.
- Ooi, T. L., Wu, B., & He, Z. J. (2006). Perceptual space in the dark affected by the intrinsic bias of the visual system. *Perception*, 35, 605-624.
- Peters, M., Laeng, B., Latham, K., Jackson, M., Zaiyouna, R. and Richardson, C. (1995). A redrawn Vandenberg & Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition*, 28, 39-58.
- Philbeck, J. W. & Loomis, J. M. (1997). Comparison of two indicators of visually perceived egocentric distance under full-cue and reduced-cue conditions. *Journal of Experimental Psychology: Human Perception and Performance*, 23, 72-85.
- Rieser, J.J., Garing, A.E, & Young, M.F. (1994). Imagery, action, and your children's spatial orientation: It's not being there that counts, it's what one has in mind. *Child Development*, 65, 1262-1278.
- Shelton, A. L., & McNamara, T. P. (2001). Systems of spatial reference in human memory. *Cognitive Psychology*, 43, 274 –310.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, 171, 701-703. DOI: 10.1126/science.171.3972.701
- Simons, D.J. and Wang, R.F., 1998. Perceiving real-world viewpoint changes. *Psychological Science*, 9, 315–320.
- Snodgrass, J. G., & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, 28(4), 516-536.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74, 1-29.
- Tolman, E. (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. DOI: 10.1037/0278-7393.32.4.867
- Wang, R. F., & Spelke, E. S. (2000). Updating egocentric representations in human navigation. *Cognition*, 77, 215 – 250.
- Wang, R. F. (2004). Between reality and imagination: When is spatial updating automatic? *Perception & Psychophysics*, 66, 68-76.
- Wang, R., & Brockmole, J. (2003). Human navigation in nested environments. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 398 – 404. (a)
- Wang, R., & Brockmole, J. (2003). Simultaneous spatial updating in nested environments. *Psychonomic Bulletin & Review*, 10, 981–986. (b)
- Yamamoto & Shelton (2008). Integrating object locations in the memory representation of a spatial layout. *Visual Cognition*, 16, 140-143.

Figure Captions

Figure 1. Arrangement of targets in the experimental room. A and B are the first and second viewing positions, respectively. W and L index the WM and LTM targets.

Figure 2. Physical locations of the LTM targets relative to the first viewing position and the centroids of subjects' walks to the targets immediately after learning (Walk-1) and after the delay period (Walk-2).

Figure 3. Pointing response time in the JRD task for individual items within each combination of facing object and target object.

Figure 4. Relation between individual subjects' mental rotation score and absolute angular error in two tasks: egocentric pointing and judgments of relative direction. Data are averaged over LTM and WM targets. Lines show least-squares fit.

Figure 5. Comparison of mean absolute error in JRD task for three subject groups: high mental rotation score ≥ 50 (high MRT, N = 15), < 50 (low MRT, N = 9), and the average of the two groups.

Figure 1

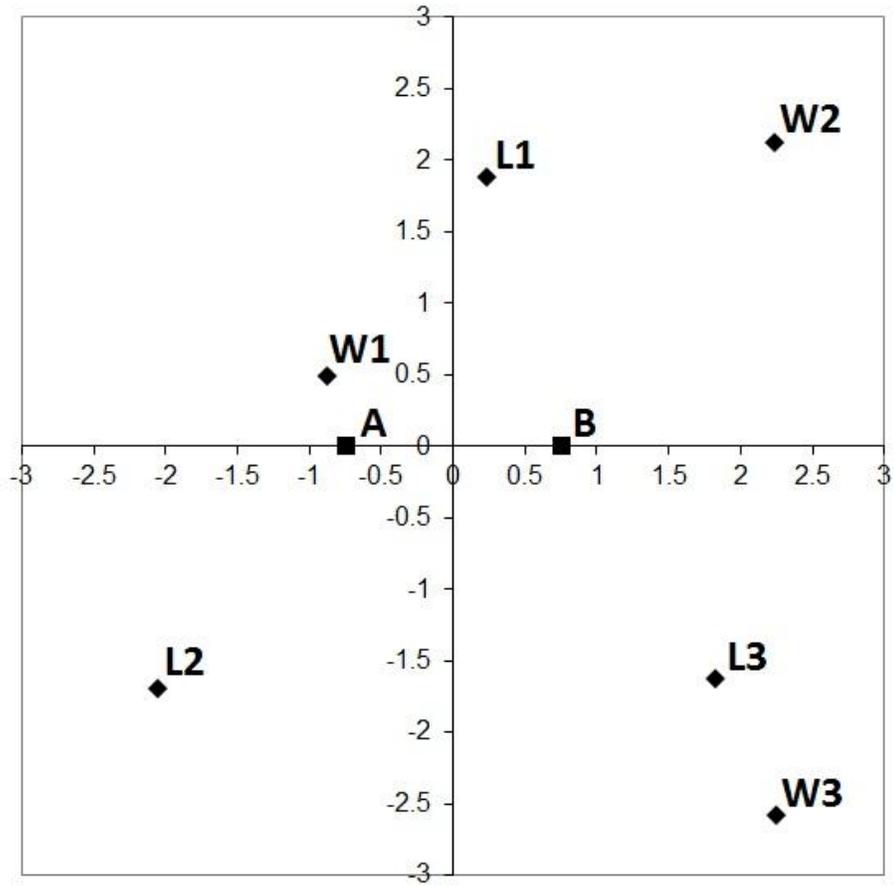


Figure 2

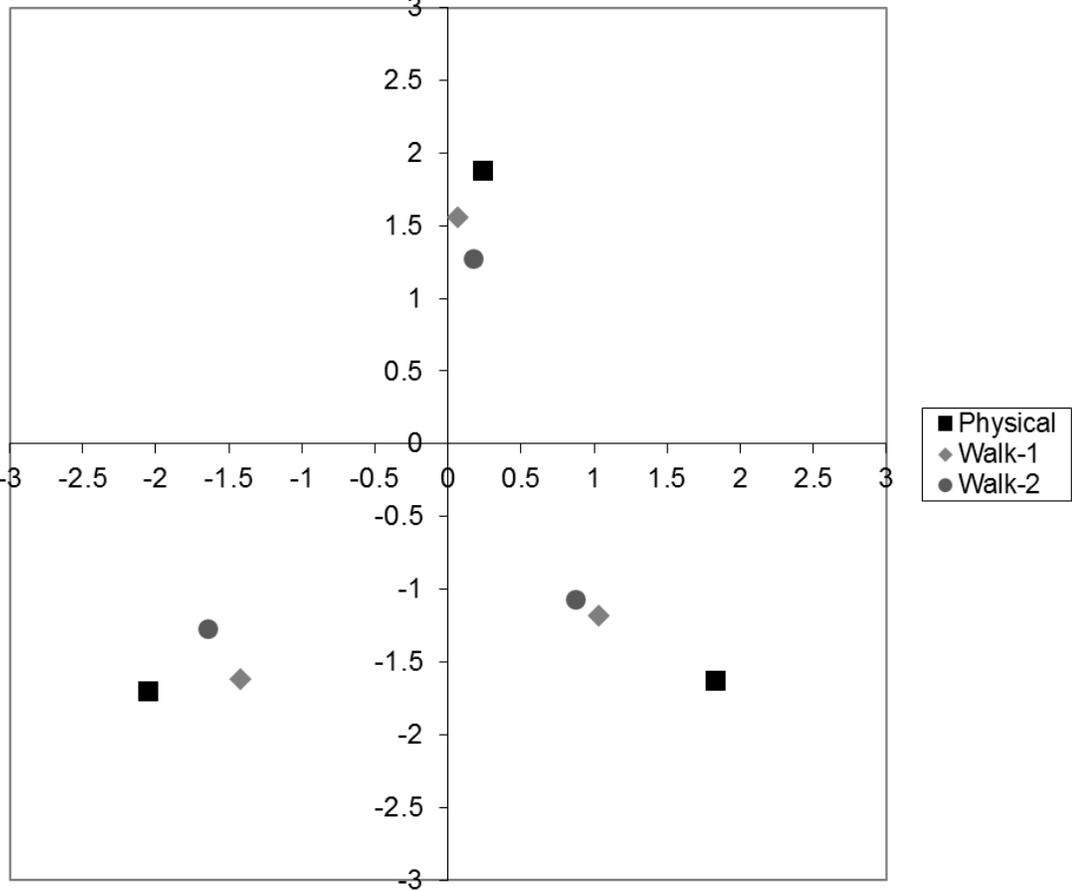


Figure 3

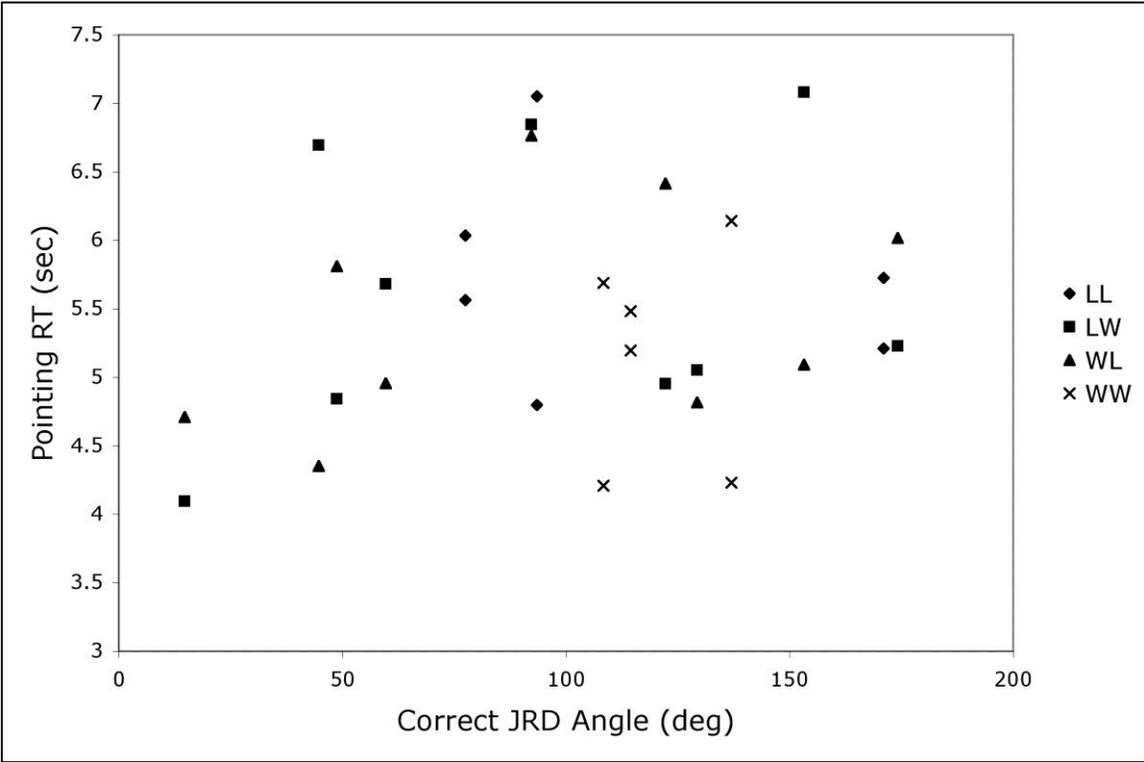


Figure 4

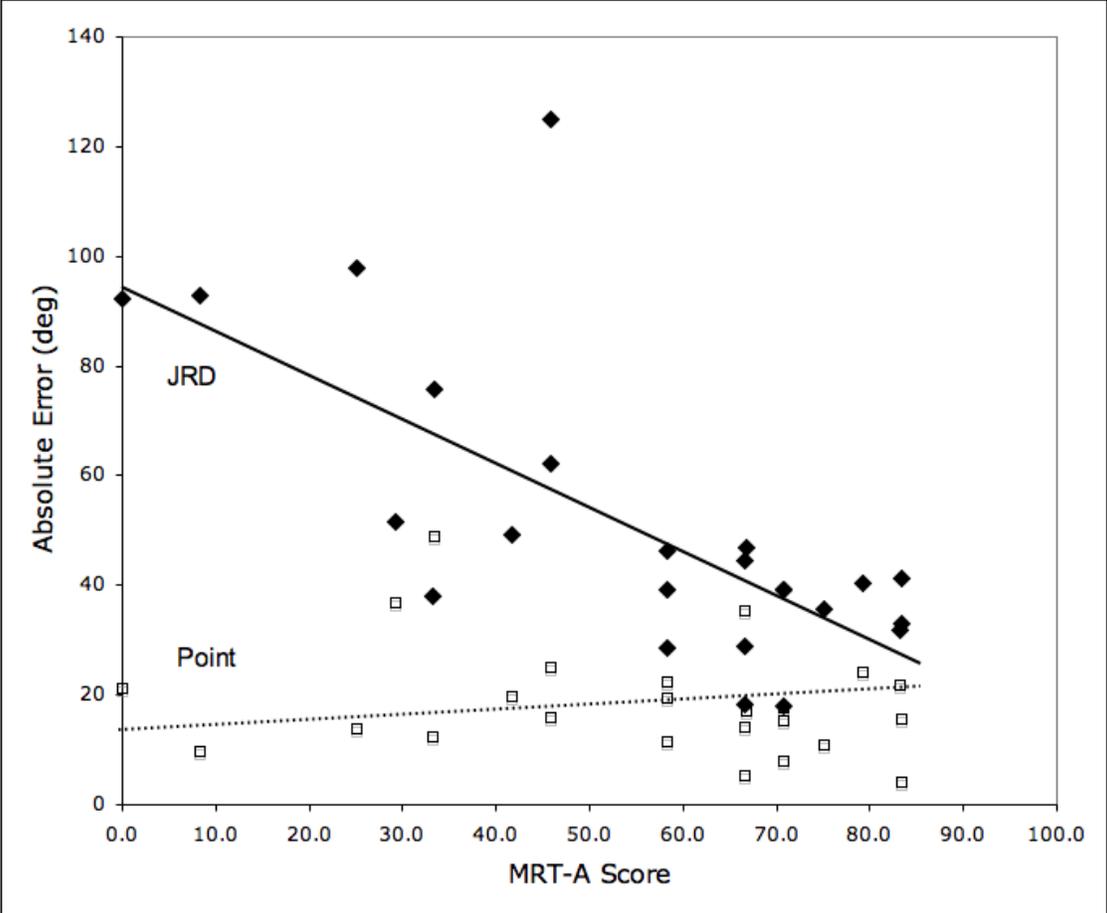


Figure 5

