Cognition, in press

Running heads: Reference Directions and Reference Objects

Reference Directions and Reference Objects in Spatial Memory of a Briefly-Viewed Layout

Weimin Mou[1,2], Chengli Xiao[1], Timothy P. McNamara[3]

[1]Chinese Academy of Sciences, [2]University of Alberta, [3]Vanderbilt University

Corresponding author:

Weimin Mou

Institute of Psychology

Chinese Academy of Sciences

10A Datun Road

Beijing, China 100101

Email: mouw@psych.ac.cn

Abstract

Two experiments investigated participants' spatial memory of a briefly-viewed layout. Participants saw an array of five objects on a table and, after a short delay, indicated whether the target object indicated by the experimenter had been moved. Experiment 1 showed that change detection was more accurate when non-target objects were stationary than when non-target objects were moved. This context effect was observed when participants were tested both at the original learning perspective and at a novel perspective. In Experiment 2, the arrays of five objects were presented on a rectangular table and two of the non-target objects were aligned with the longer axis of the table. Change detection was more accurate when the target object was presented with the two objects that were aligned with the longer axis of the table during learning than when the target object was presented with the two objects that were not aligned with the longer axis of the table during learning. These results indicated that the spatial memory of a briefly-viewed layout has interobject spatial relations represented and utilizes an allocentric reference direction.

Reference Directions and Reference Objects in Spatial Memory of a Briefly-Viewed Layout

When specifying a location in daily communication, people usually use a spatial reference direction and a reference location. For example, a statement such as, "the car is in front of me," uses an egocentric reference direction (front) and egocentric reference object (the speaker's body). By way of contrast, the statement, "the car is east of the house," uses an allocentric reference direction (geographic east) and an allocentric reference object (the house). Hence, both the spatial reference direction and the reference object can be either allocentric or egocentric.

In the past decade, one major goal of research on spatial memory has been to understand the reference systems that are used in memory to represent locations of objects in the environment; in particular, to determine the extent to which spatial memory is egocentric or allocentric. Two different criteria have been used in determining whether spatial memory is egocentric or allocentric. One criterion is the nature of the reference direction. If the reference direction is determined by the egocentric axis (body axis) of the observer, then spatial memory is classified as egocentric. Alternatively if the reference direction is determined by a direction independent of the observer, then spatial memory is allocentric. A second criterion is the nature of the reference object. If the reference object is the observer, such that self-to-object spatial relations are represented, then spatial memory is egocentric. Alternatively if the reference object is another object or set of objects, such that interobject spatial relations are represented, then spatial memory is allocentric. It is crucial to distinguish the reference direction from the reference object because memory may use a mixture of egocentric and allocentric reference systems (e.g., the car is east of me). Hence instead of asking whether spatial memory is egocentric or allocentric, we should ask two more specific questions: first whether spatial memory depends on reference directions and, if so, whether the reference directions are egocentric or allocentric; second whether self-to-object or object-to-object or both kinds of spatial relations are represented in memory.

Research in the past decade has shown clearly that long-term spatial memory utilizes reference directions. A large body of evidence has shown that long-term spatial memory is orientation dependent (e.g. Mou, McNamara, Valiquette, & Rump, 2004; Roskos-Ewoldsen, McNamara, Shelton, & Carr, 1998; Shelton & McNamara, 1997, 2001; Valiquette, McNamara, & Smith, 2003; for a review see McNamara, 2003). For example, Shelton and McNamara (1997) had participants learn the locations of several objects on the floor of a room from two orthogonal viewpoints. After memorizing the locations of the objects, participants moved to a different room and made judgments of relative directions ("Imagine you are standing at X, facing Y, please point to Z") using spatial memory. The results showed that judgments of relative direction were better at the imagined headings parallel to the learning views than at novel imagined headings. It is assumed that spatial relations consistent with the spatial reference direction used in memory can be retrieved more efficiently than spatial relations inconsistent with the spatial reference direction used in memory (e.g., Klatzky, 1998). Hence these results suggested that spatial memory has reference directions and the reference directions may be egocentric.

Subsequent findings reported by Shelton and McNamara (2001), however, indicated that reference directions in long-term spatial memory are allocentric. They reported that not all learning viewpoints were preferred in judgments of relative direction, raising the possibility that the reference direction in memory may be independent of the egocentric viewing direction. They reported that after people learned a layout of objects from two viewpoints, one aligned but the other misaligned with external structures (e.g. the walls of the surrounding room and the edges of a mat on which objects were placed), their performance in judgments of relative direction was better only at the imagined headings parallel to the learning viewpoints that were aligned with the external structures. The performance at the imagined heading parallel to the misaligned learning viewpoint was not better than the performance at novel imagined headings (see, also, Valiquette & McNamara, 2007; Valiquette, McNamara, & Labrecque, 2007). These findings provide strong evidence against egocentric models because they show that the spatial reference direction is not fixed with the egocentric front of the observer; otherwise participants should establish one spatial reference direction at each learning direction and judgments of relative direction should be equally good at all learning directions.

Mou, McNamara and their colleagues have published evidence that reference directions in long-term spatial memories are allocentric, and more specifically, intrinsic to the layout (e.g. Mou, Fan, McNamara, & Owen, 2008; Mou, Liu, & McNamara, 2008; Mou & McNamara, 2002; Mou, Zhao, McNamara, 2007). Mou and McNamara (2002) instructed participants to learn a layout along an intrinsic direction of the layout that was not parallel to their viewpoint and then required them to make judgments of relative direction using their memories. They reported that judgments of relative direction were best at the imagined heading that was parallel to the instructed direction and even better than at the imagined heading that was parallel to the learning viewpoint. Moreover Mou, Liu, and McNamara (2008) reported that when participants learned a desktop sized layout of objects with a symmetric axis and the symmetric axis of the layout was different from participants' viewpoint, participants moved their eyes more frequently along the directions aligned with (i.e., parallel or orthogonal to) the symmetric axis than along the directions aligned with their learning viewpoint. Consistent with the eye movement results, judgments of relative direction were better at the imagined headings aligned with the symmetric axis of the layout than at the imagined headings aligned with the learning viewpoint.

Furthermore Mou, Fan, McNamara, and Owen (2008) reported that scene recognition is also intrinsic orientation dependent. Participants were instructed to learn a layout of objects along an intrinsic axis that was different from their viewing direction. Participants were then given a scene recognition task in which they had to distinguish triplets of objects from the layout at different views from mirror images of those test scenes. The results showed that participants were able to recognize intrinsic triplets of objects, which contained two objects parallel to the instructed intrinsic direction, faster than non-intrinsic triplets of objects, which did not contain two objects parallel to the instructed intrinsic direction. This pattern occurred for all test views. Collectively, these findings indicate that spatial reference directions in long-term spatial memories are allocentric.

Research in the past decade has indicated that spatial reference objects can be either egocentric or allocentric in the spatial memory of a well-learned layout. Wang and Spelke (2000) concluded that self-to-object spatial relations were represented in memory after showing that disorientation disrupted the consistency in pointing to different objects. In a follow up study, Mou, McNamara, Rump and Xiao (2006, see also Holmes & Sholl, 2005) reported that there was no effect of disorientation after participants learned a layout with a good geometric structure. These results suggest that interobject spatial relations might also have been represented in memory at least when the layout has a good geometric structure. Thus the data so far suggest that transient self-to-object spatial relations and enduring object-to-object spatial relations may both be represented in memory (see also, Waller & Hodgson, 2006).

All of these studies examine memory of a well-learned layout. Participants' spatial memory was tested after participants studied the layout and could point to objects accurately. There are very few studies directly examining the spatial reference systems used in memory of a briefly-viewed layout in the environment, and the results of those studies do not provide clear evidence on whether the memories use allocentric reference directions and have interobject spatial relations represented.

In one relevant study, Jiang, Olson, and Chun (2000) reported that visual short-term memory of an array of squares presented on a CRT monitor stores relational information between the individual items in the array. In one experiment, participants remembered the locations of multiple briefly-presented display items, and were asked to detect changes in the position of a single item across a brief delay interval either when the non-target items were in the same position as they were originally or when the spatial configuration of objects was changed or eliminated at test. They discovered that change-detection performance was better when the spatial configuration of objects remained constant from memory display to test (context effect). This result suggests that inter-item spatial relations may be represented in visual working memory. If this context effect can also be shown in scene recognition based on spatial memory of a briefly-viewed array of objects in the environment, even after people physically move and adopt a novel perspective, it will suggest that the memory has interobject spatial relations represented.

In another relevant study, Simons and Wang (1998, see also Wang & Simons, 1999) reported that scene recognition based on the spatial memory of a briefly-viewed layout in the environment was viewpoint dependent when the novel view was caused by table rotation but became viewpoint independent when the novel view was caused by observer movement. They had participants detect the position change of one object in an array of five objects on a desktop either from the learning perspective or from a new perspective and either when the table was stationary or when it was rotated. They found that although display rotations produced viewpoint dependent recognition, observer movements apparently did not. Burgess, Spiers, and Paleologou (2004) dissociated two independent effects in Simons and Wang's experiments. One is the spatial updating effect, which is better recognition of a view expected from the updated spatial representation (e.g., the novel view one would expect to see after moving to a new viewing position) than of a view unexpected from the updated spatial representation (e.g., the same novel view after moving to a new viewing position and then back to the learning position); the other is the viewpoint effect, which is better recognition of the study view than of a novel view.

Both the updating effect and viewpoint effect have been explained by a model assuming that the spatial memory of a briefly-viewed layout is in the format of "visual-spatial snapshot" of the layout and the snapshot representations are updated during the physical movement of the observer (e.g. Wang & Spelke, 2002). A snapshot representation of a layout does not require any reference object and any reference direction to specify objects' locations. Scene recognition can be accomplished with a process of similarity matching between the input representation of the test scene and the snapshot representation of the study scene in memory. The input and snapshot representations matched best when participants were tested at the original learning view, so the viewpoint effect was expected. The updating effect can be explained if it is assumed that the snapshot representations can be updated with least effort during participants' physical locomotion.

In sum, several sources of evidence indicate the spatial memory of a well-learned layout in the environment has interobject spatial relations represented and utilizes an allocentric reference direction. In contrast there is no direct evidence that the spatial memory of a briefly-viewed layout in the environment has interobject spatial relations represented and utilizes an allocentric reference direction.

The current study sought direct evidence that the spatial memory of a briefly-viewed layout also has interobject spatial relations represented and utilizes an allocentric reference direction. Experiment 1 examined whether interobject spatial relations are represented in the spatial memory of a briefly-viewed layout. Participants briefly viewed a layout of five objects in the environment and after a short delay detected whether a probed object had been moved or not when all other objects stayed stationary (non-target stationary condition) or when all other objects changed their positions (non-target moved condition). The results showed that change detection was more accurate in the non-target stationary condition than in the non-target moved condition. Experiment 2 tested whether this pattern of results could be explained in a visual-snapshot model. Participants briefly learned the layout of five objects on a rectangular table, the longer axis of which was parallel to an intrinsic axis passing through two non-target objects in the layout. Participants detected whether a target object had been moved or not in two context conditions: In the intrinsic context condition, the two non-target objects fell on an imaginary line parallel to the longer axis of the table; in the non-intrinsic context condition, the two non-target objects fell on an imaginary line that was not parallel to the longer axis of the table. The results showed that change detection was more accurate in the intrinsic context condition than in the non-intrinsic context condition.

Experiment 1

In this experiment, participants in a darkroom saw an array of five phosphorescent objects on a circular table and, after a short delay, indicated whether the object (target) specified by the experimenter had been moved both at the learning perspective and at a new perspective while the table stayed put. In half of the trials, the four objects (non-target) other than the target object were moved and placed in random positions near the center of the table before test (non-target moved); in the other half of the trials, the other four objects remained in their original positions (non-target stationary).

Method

Participants

Thirty two university students (16 men and 16 women) participated in this study in return for monetary compensation.

Materials and Design

The experiment was conducted in a room (4.0 by 2.8 m) with walls covered in black curtains. As illustrated in Figure 1, the room contained a circular table covered by a grey mat (80 cm in diameter, 48 cm above the floor), two chairs (seated 42 cm high), and 5 common objects (eraser, locker, battery, bottle, and a plastic banana sized around 5 cm) coated with phosphorescent paint but still recognizable in dark. The objects were placed on five of nine possible positions in an irregular array on the circular table. The distance between any two of the nine positions varied from 18 to 29 cm. The irregularity of the array ensured that no more than two objects were aligned with the observer throughout the experiment. The distance of the chairs to the middle of the table was 1.2 m. The viewing angle between the chairs was 98°. A bar stool was placed halfway between the two chairs to mark the half-way point between them. Participants wore a blindfold and a wireless earphone which was connected to a computer outside of the curtain. The lights were always off during the experiment, and the experimenter used a flashlight when she arrayed the layout. With the blindfold removed, participants were only able to see the five objects. In this way all irrelevant environmental cues were removed. The earphone was used to present white noise and instructions (e.g., to remove the blindfold and view the layout, to put the blindfold on). White noise was presented to mask auditory cues to object location. The experimenter indicated the target object by pointing to it with a laser-pointer.

Forty irregular configurations of object locations were created. In each configuration, one of the five occupied locations was selected randomly to be the location of the moved object. On those trials when an object moved, the object was placed on one of the other four unoccupied locations. This new location of the object was usually the open location closest to the original location and had a similar distance to the center of the table so that this cue could not be used to determine whether the target object had moved.

The configurations were randomly divided into four groups of ten. Four sets of 40 trials were created by assigning the four groups of configurations to four combinations of target movement (moved or stationary) and non-target movement (moved or stationary) with a Latin Square design. In that way, in each set of trials, different sets of configurations were assigned to different combinations of target movement and non-target movement, and across the four sets of trials each set of configurations was assigned to every combination of target movement and non-target movement once. Participants were assigned to one of the four sets of trials randomly so that across participants, any possible differences among the configurations were counter-balanced.

The primary independent variables were movement of non-target objects (moved vs. stationary) and testing perspective (0° vs. 98° away from the original viewing perspective). Non-target movement was manipulated within participants and testing perspective was manipulated between participants. The original viewing position (each of the chairs in Figure 1) was counter-balanced across participants. Eight participants were randomly assigned to each of the four combinations of testing perspective and learning position with the restriction that each group had four men and four women; in each group, a woman and a man were randomly assigned to each of the four sets of trials. The dependent variable was the percentage of the correct judgments in deciding whether or not the target object changed positions.[1]

Procedure.

Wearing a blindfold, participants walked into the testing room and sat on the viewing chair assisted by the experimenter. Each trial was initiated by a mouse click of the experimenter and started with a verbal instruction via earphone ("please remove the blindfold, and try to remember the locations of the objects you are going to see."). After three seconds, participants were instructed to walk, while blindfolded, to the other viewing position ("please wear the blindfold, walk to the other chair") or to the bar stool and then back to the original viewing position ("please wear the blindfold, walk to the middle point between the two chairs and then walk back to this chair"). Thirteen seconds after participants were instructed to stop viewing the layout, they were instructed to determine whether the object

---

[1] Results are the same when d-prime is used as the dependent variable.

pointed to by the experimenter with the laser pointer changed its position ("please take off the blindfold and make judgment"). The participant was instructed to respond as accurately as possible; speedy response was discouraged ("please respond as accurately as possible, respond only after you are sure this is your decision"). After the response, the trial was ended by another mouse click of the experimenter and the participant was instructed to be ready for the next trial ("please put on the blindfold and sit on the original viewing chair."). The experimenter wrote down the response on a sheet of paper.

Before the 40 experimental trials, four extra trials (one in each of the four combinations of non-target movement and target movement) were used as practice to make sure participants got used to the procedure and were able to walk to the other chair or bar stool and back to the original viewing position while blindfolded.

Results

Mean percentage of correct judgment as a function of non-target movement and testing perspective is plotted in Figure 2. The primary findings were as follows: First, participants were more accurate when the non-target objects stayed stationary than when they were moved to new positions. Second, participants who made judgments from the original viewing perspective were more accurate than those who made judgments from a new perspective. Third, the effect of non-target movement was comparable at different test perspectives.

Percentage of correct judgments was computed for each participant and each non-target movement condition, and analyzed in mixed analyses of variance (ANOVAs), with variables corresponding to non-target movement, testing position, and gender. Non-target movement was within participants. All others were between participants. The main effect of non-target movement was significant, $F(1, 28) = 15.13$, $p < .001$, $MSE = .008$. The main effect of perspective was significant, $F(1, 28) = 19.56$, $p < .001$, $MSE = .024$. The interaction between the two effects was not evident, $F(1, 28) = 0.04$, $p = .84$. No other effects were significant.

Discussion

The most important result of this experiment was that participants were more accurate deciding whether the target object had been moved when the non-target objects stayed stationary than when the non-target objects were moved. This context effect occurred and was the same magnitude regardless of whether participants were tested in the original viewing position or in a new viewing position. This finding indicates that the spatial memory of a briefly-viewed layout in the environment might have interobject spatial relations represented, parallel to the findings in the study of visual working memory (Jiang, Olson, & Chun, 2000). The second important result was that participants were more accurate deciding whether the target object has been moved at the original viewing position than at the new viewing position, consistent with the finding of Burgess, Spiers, and Paleologou (2004). This finding indicates that scene recognition based on the spatial memory of a briefly viewed layout is viewpoint dependent even when the novel view is caused by observer locomotion.

However, a model based on snapshot representations may also be able to explain the context effect and its independence of the test perspective with the following assumptions. First, accuracy of location change detection increases with the similarity between perceived and stored view-specific representations of the scene (e.g., Poggio & Edelman, 1990). Second, the visual-spatial snapshot of the original viewing perspective is more similar to a view-specific representation of the test perspective when all non-target objects are in their original positions (non-target stationary condition) than to a view-specific representation of the test perspective when all non-target objects are in new positions (non-target moved condition). Third the difference in similarity between the two context conditions at the novel perspective is as high as that at the learning perspective. Hence a snapshot model can also explain the context effect and its independence of the test perspective. In addition, as we discussed in the Introduction, a model based on snapshot representations can also readily explain the viewpoint effect observed in Experiment 1 if we assume the familiar test view is more similar to the snapshot representation than is a novel test view.

Experiment 2

The purpose of Experiment 2 was to test a visual-spatial snapshot model's account of these findings and determine whether spatial memory of a briefly-viewed layout utilizes an allocentric spatial reference direction. A rectangular table was presented before and together with the presentation of each configuration of five objects used in Experiment 1 during learning (as shown in Figure 3). The table was presented such that its longer axis was parallel to an imaginary line that passed through two non-target objects in the array. For example, in the layout illustrated in Figure 3, Candle was the designated target object; Glue and Clip were along the longer axis of the table. In this way an allocentric spatial reference direction (e.g., the direction parallel to the line passing through Glue and Clip) was primed by the shape of the table for each learning configuration. During test, participants detected whether the target object (e.g. Candle) had been moved in the presence of two non-target objects on a round table (see Figure 4). In the intrinsic context condition, the two non-target objects were those two that had been placed along the primed spatial reference direction (e.g. Glue and Clip). In the non-intrinsic context, the two non-target objects were two that had not been placed along the primed spatial reference direction (e.g. Lock and Wood). In both intrinsic and non-intrinsic context conditions, test triplets of objects (one target and two context objects) were tested at the learning perspective.

According to the findings of Mou, Fan, McNamara, and Owen (2008), change detection should be better in the intrinsic context condition than in the non-intrinsic context condition if change detection is based on the spatial memory that utilizes an allocentric reference direction. In contrast, if change detection relies on a snapshot representation, no performance superiority in the intrinsic context should be expected as participants were tested at the same perspective for both intrinsic and non-intrinsic context conditions.

In order to easily implement the above design, the layouts and the tables were presented through an immersive head mounted display, the location and orientation of which were tracked.

Method

Participants

Twelve university students (6 male and 6 female) participated in return for monetary compensation.

Materials and design

The virtual environment with layouts and tables was displayed in stereo with a light-weight (about 7 oz) glasses-like I-glasses PC/SVGA Pro 3D head-mounted display (HMD, I-O Display Systems, Inc. California). Participants' head motion was tracked with an InterSense IS-900 motion tracking system (InterSense, Inc., Massachusetts). The apparatus was placed in a 6 m × 6 m laboratory with each wall covered by homogeneous black curtains. As illustrated in Figure 3, an 80 cm × 120 cm virtual table was presented on the floor in the middle of the room, and its orientation (determined by the longer axis) changed across the learning layouts. A chair was placed 1.9 m away from the center of the table, which participants were sitting on during both learning and test. A bar stool was placed 49 degrees away from the chair as in Experiment 1.

The same forty configurations and the same target locations in Experiment 1 were used. For each configuration, the four non-target locations were divided into two exclusive pairs (e.g. Glue-Clip, Lock-Wood in Figure 3). One pair (e.g. Glue-Clip) was randomly assigned as the context pair, objects on which would be presented with the target to consist of a test triplet of locations (e.g. Glue-Clip-Candle). The 40 learning configurations were then divided into two sets of configurations (20 in each set) with the restriction that the target would be moved in half of the configurations (10) in both sets. The two sets of the configurations were assigned to the two context conditions (intrinsic and non-intrinsic) with a Latin Square design. In the intrinsic context condition, the rectangular table would be presented so that its longer axis would be parallel to the intrinsic axis passing through the context objects in the test triplet. In the non-intrinsic condition, the rectangular table would be presented so that its longer axis would be parallel to the intrinsic axis passing through the two non-target objects other than the context objects in the test triplet. For example, suppose that Glue and Clip were the context objects and Candle was the target. In the intrinsic context condition, the rectangular table was placed as in Figure 3 such that the allocentric reference direction passed through Glue and Clip. In the non-intrinsic context condition, the rectangular table was placed so that its longer axis was parallel to the axis through Lock and Wood.

With the above design, across participants, the intrinsic and nonintrinsic context conditions were the same in terms of the learning configurations and the test triplets, and only differed in terms of whether the context locations were parallel to the allocentric reference direction that was primed by the rectangular table. During detection, a small red arrow was presented above the target object to indicate which object participants should detect its movement, and a gray circular table instead of the rectangular table was also presented to provide ground plane information (Figure 4).

Procedure.

In the preparation room, participants were instructed that they were going to view a layout of five objects on a rectangular table and judge whether one of the objects was moved. They were explicitly instructed that two objects in the layout were placed aligned with the longer axis of the table and were shown pictures of four extra configurations to help them understand.

Wearing a blindfold, participants were escorted into the testing room and sat on the viewing chair. Then they put on the HMD. Before each trial, participants fixed their eyes on a large red arrow pointing to the middle of the upcoming rectangular table. Each trial was initiated by a key press of the experimenter. The virtual rectangular table was presented for five seconds, and then participants were instructed to remember the scene ("please try to remember the locations of objects you are going to see."). Then an irregular configuration was presented on the table. After three seconds, both table and objects disappeared and participants were instructed to stand up and walk to the bar stool and then back to the viewing position as in Experiment 1 ("please walk to the bar stool and then walk back to this chair"). Thirteen seconds after the disappearance of the learning configuration, the test triplet was presented and participants were instructed to press the mouse button to indicate whether the target object changed its position. Participants were required to respond as accurately as possible and as quickly as possible. The latency between the onset of the test triplets and the key press of the participants was also recorded in this experiment.

Before the 40 experimental trials, eight extra trials were used as practice to make sure participants understand the whole procedure.

Results and discussion

Mean percentage of correct judgments as a function of context (intrinsic or non-intrinsic) is plotted in Figure 5. As shown in the figure, participants were more accurate in detecting whether the target had been moved in the intrinsic context condition than in the non-intrinsic context condition.

Percentage of correct judgments was computed for each participant and each triplet type, and analyzed in a repeated measure analyses of variance (ANOVA) with one variable corresponding to context (intrinsic vs. non-intrinsic). The main effect of triplet type was significant, F (1, 11) = 10.47, $p < .01$, MSE = 0.008.

The mean latencies for the correct detections across participants were 4.81 s and 5.39 s for the intrinsic and non-intrinsic context conditions respectively. The difference was significant, F (1, 11) = 14.25, $p < .01$, MSE =0.14. This result showed that change detection was quicker in the intrinsic context condition than in the non-intrinsic context condition and assured that there was no trade-off between accuracy and latency.

These results clearly indicated that change detection in this experiment relied on a spatial representation that utilizes an allocentric reference direction rather than a snapshot representation.

General Discussion

The aim of this project was to investigate whether the spatial memory of a briefly-viewed layout has interobject spatial relations represented and utilizes an allocentric spatial reference direction. Participants learned a layout of five objects for only 3 s with all environmental cues removed in a dark room or in a virtual environment and after a 13 s interval detected whether a probed object was moved or not using their memories. Two important findings were obtained. First, participants were more accurate in change detection with the same context of other four objects (non-target stationary) than with the different context of other four objects (non-target moved) even when participants were tested at a novel viewpoint. Second, participants were more accurate in change detection with the same context of two objects that had been placed along the longer axis of a rectangular table during learning than with the same context of two objects that had not been placed along the longer axis of the rectangular table during learning.

The first finding strongly indicated that spatial memory of a briefly-learned layout has interobject spatial relations represented. If people only use themselves as the reference object and dynamically update self-to-object spatial relations during locomotion (e.g. Wang & Spelke, 2000), the self-to-object spatial relation for the probed object is the only basis of the detection. Self-to-object spatial relations for non-target objects are irrelevant, and therefore, changing their positions should not affect performance, so there should be no difference between the same context condition and the different context condition. Hence, just as Jiang et al. (2000) argued that visual short-term memory of an array of squares presented on a CRT monitor stores relational information between the individual items in the array, in parallel, we argue that spatial memory of a briefly-learned layout has interobject spatial relations represented.

One may argue that a snapshot representation that does not represent interobject spatial relations may also be able to explain the context effect. It is possible that when participants learned the layout, they stored a snapshot representation of the layout (e.g. Wang & Spelke, 2002). Change detection could be based on a process of global matching between the input representation of the test scene and the stored snapshot representation (e.g., Poggio & Edelman, 1990). The higher the similarity between the test scene and the snapshot representation, the better will be performance in change detection. If we further assume that the test scene in the same context condition is more similar to the snapshot than the test scene in the different context condition both when the test scene is presented at the original learning perspective and at the novel perspective, then the context effect can be explained.

The snapshot explanation, however, is challenged by the second important finding of this project: Change detection was better in the intrinsic context condition than in the non-intrinsic context condition in Experiment 2. It is difficult to envision how a snapshot representation model could be implemented so that the intrinsic context condition is more similar than the non-intrinsic context condition to the snapshot representation. In Experiment 2, the same learning configuration and the same test triplets were used in both the intrinsic context condition and non-intrinsic context condition. The rectangular table should also not able to change the level of similarity because it was only presented with the learning configuration but not presented during change detection. Hence there is no apparent reason why the test triplet in the intrinsic context condition is more similar to the snapshot representation than is the test triplet in the non-intrinsic context condition.

In contrast the finding that performance in change detection was better in the intrinsic context condition than in the non-intrinsic context condition can be easily explained by a model arguing that change detection was based on a spatial memory utilizing allocentric reference directions that were primed by the rectangular table.

In the intrinsic context condition, the context objects were consistent with the spatial reference direction primed by the rectangular table whereas in the non-intrinsic context condition, the context objects were inconsistent with the spatial reference direction primed by the rectangular table. With the assumption that spatial relations consistent with the spatial reference directions are easier to judge than spatial relations inconsistent with the spatial reference directions (e.g., Klatzky, 1998), it is expected that performance in change detection is better in the intrinsic context condition than in the non-intrinsic context condition.

The results of the current project directly indicated that spatial memory of a briefly-viewed layout has interobject spatial relations represented and utilizes allocentric reference directions. As we reviewed in the Introduction, the spatial memory of a well-learned layout also has interobject spatial relations represented and utilizes allocentric reference directions. Hence, spatial memory appears to be allocentric in terms of both reference object and reference direction regardless of the scale of the learning time. These findings are theoretically important because they can be used to constrain the contemporary models of spatial memory.

In Sholl's model (e.g., Easton & Sholl, 1995; Holmes & Sholl, 2005; Sholl, 2001; Sholl & Nolin, 1997), an egocentric self-reference system codes self-to-object spatial relations in body-centered coordinates, using the body axes of front-back, right-left, and up-down (e.g., Bryant & Tversky, 1999; Franklin & Tversky, 1990). This system provides a framework for spatially directed motor activity, such as walking, reaching, and grasping. Self-to-object spatial relations are continuously and efficiently updated as an observer moves through an environment. The allocentric object-to-object system codes the spatial relations among objects in environmental coordinates using an orientation independent reference system. A dominant reference direction in this system is established by egocentric front when participants are perceptually engaged with the environment (Sholl & Nolin, 1997, p. 1497).

Mou, McNamara, Valiquette, and Rump (2004) recently proposed a model of spatial memory and navigation. According to this model, the human navigation and spatial representation system comprises two subsystems: The egocentric subsystem computes and represents transient self-to-object spatial relations needed to control locomotion (e.g. obstacle avoidance). These spatial relations are represented at sensory-perceptual levels and decay relatively rapidly in the absence of perceptual support or deliberate rehearsal. The environmental subsystem is responsible for representing the enduring features of familiar environments. In this subsystem, the spatial structure of the environment is represented in an orientation dependent manner using an intrinsic reference system (e.g., Shelton & McNamara, 2001). Interobject spatial relations are specified with respect to a small number (typically 1 or 2) of intrinsic reference directions (e.g., Mou & McNamara, 2002). The learning locations and orientation are also represented with respect to the same intrinsic reference system. When people move in the environment they update their location and orientation with respect to the intrinsic frame of reference. Spatial updating does not change the spatial reference direction of the allocentric spatial representation. People can calculate self-to-object spatial relations using the representation of their own location with respect to the intrinsic reference frame and the representation of object-to-object spatial relations.

Wang and Spelke (2000, 2002) have proposed a model of spatial memory and navigation that consists of three interrelated systems. According to this model, an egocentric system computes and dynamically updates spatial relations between the navigator's body and important objects in the surrounding environment. This dynamic egocentric system supports path integration. A second system represents the appearances of familiar landmarks and scenes for recognition. These representations are viewpoint-dependent and can be conceived of as visual-spatial "snapshots" of the environment (e.g., Burgess, Spiers, & Paleologou, 2004; Diwadkar & McNamara, 1997; Wang & Simons, 1999). Finally, an allocentric system represents the geometric shape of the environment (e.g., the shape of a room). The allocentric system does not represent interobject spatial relations in the environment; its purpose is to support reorientation when the path integration system breaks down (e.g., as a result of disorientation).

These three models are similar as all of them claim that self-to-object spatial relations are represented in spatial memory and spatial updating occurs when people locomote in the environment. However they differ in whether spatial memory contains interobject spatial relations and whether spatial memory utilizes allocentric reference directions. With respect to the interobject spatial relations, Sholl’s model and Mou and McNamara’s model propose that interobject spatial relations are represented in spatial memories whereas Wang and Spelke’s model proposes that an allocentric system represents the geometric shape of the environment but does not represent interobject spatial relations in the environment.

Regarding the allocentric reference direction, Sholl's model and Wang and Spelke's model propose that spatial memory does not have an allocentric reference direction whereas Mou and McNamara's model proposed that spatial memory utilizes allocentric reference direction. Sholl's model and Mou and McNamara's model can readily explain the context effect of the current project and why disorientation does not disrupt spatial memories when participants learn a layout with good shape (Mou, McNamara, Rump, & Xiao, 2006). In contrast Wang and Spelke's model cannot easily explain these results. Mou and McNamara's model can readily explain the intrinsic orientation dependent results in the current project and in the projects where the spatial memory of a well-learned layout was examined using various measurements including eye movement (Mou, Liu, & McNamara, 2008), judgments of relative direction (Mou & McNamara, 2002), and scene recognition (Mou, Fan, McNamara, & Owen, 2008). In contrast neither Sholl's model nor Wang and Spelke's model can easily explain these intrinsic orientation dependent results.

In Experiment 2 of the current project, the allocentric reference direction was established by the cue of an external frame (the rectangular table). This raises the question of which cues were used in Experiment 1 to select an allocentric reference direction given that all environmental cues were eliminated by presenting objects in a dark room. We speculate that participants used their viewing perspective or the self-to-array axis (e.g. Waller, Lippa, & Richardson, 2008) to establish a spatial reference direction to specify interobject spatial relations. Note that such a spatial reference direction should not be interpreted as egocentric, rather it is still allocentric because the spatial reference direction is fixed in the mental representation and not fixed to the egocentric front of the observer when he or she moves and changes perspective. Mou, McNamara, Valiquette, and Rump (2004) reported that performance in judgments of relative direction was better at the imagined heading that was parallel to the original learning heading than at the imagined headings that were not parallel to the learning heading even when participants physically moved to different test perspectives so that their actual egocentric front was always parallel to the imagined heading. These results strongly suggested that participants established a spatial reference direction that was parallel to their learning heading, but when participants moved the spatial reference direction was fixed in the representation rather than fixed to participants' egocentric front. In this sense, the spatial reference direction is intrinsic to the layout. Hence participants' viewing perspective is just one of the various cues that people can use to establish an intrinsic reference direction. Future research should also be able to demonstrate that the spatial reference direction used in the spatial memory of a briefly-learned layout can also be established based on the cue of the viewing perspective and change detection is better with the context relevant to the viewing perspective than with the context relevant to a novel perspective even when participants physically move and adopt the novel perspective.

Was there any evidence that self-to-object spatial relations were formed at learning in the current project? The above-chance level performance in change detection in the non-target move condition in Experiment 1 may suggest that participants relied on self-to-object spatial relations. But it is not clear whether such self-to-object spatial relations were represented during learning and dynamically updated during locomotion as proposed by Wang and Spelke (2000, 2002) or were computed using updated location and orientation of the body with respect to the intrinsic frame of reference as proposed by Mou, McNamara, Valiquette and Rump (2004). Further research is required to distinguish these two possibilities.

In conclusion, this project demonstrated that detection of the position change of an object in a briefly viewed layout was context dependent and intrinsic orientation dependent. These results suggest that the spatial memory of a briefly-viewed layout has interobject spatial relations represented and utilizes an allocentric reference direction.

Acknowledgements

Preparation of this paper and the research reported in it were supported in part by the 973 Program of Chinese Ministry of Science and Technology (2006CB303101) and a grant from the National Natural Science Foundation of China (30770709) to WM and National Institute of Mental Health Grant 2-R01-MH57868 to TPM. We are grateful to Dr. Andrew Hollingworth and three anonymous reviewers for their helpful comments on a previous version of this manuscript. We are also grateful to the Virtual Reality Laboratory of The Institute of Psychology, Chinese Academy of Sciences for using its facility and space in collecting data.

Please address correspondence to Weimin Mou, Institute of Psychology, Chinese Academy of Sciences, 10A Datun Road, Beijing, China 100101.

References

Burgess, N., Spiers, H. J., & Paleologou, E. (2004). Orientational manoeuvres in the dark: dissociating allocentric and egocentric influences on spatial memory. *Cognition, 94,* 149-166.

Bryant, D. J., & Tversky, B. (1999). Mental representations of perspective and spatial relations from diagrams and models. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 137-156.

Diwadkar, V. A., & McNamara, T. P. (1997). Viewpoint dependence in scene recognition. *Psychological Science, 8,* 302-307.

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.

Franklin, N., & Tversky, B. (1990). Searching imagined environments. *Journal of Experimental Psychology: General,* 119, 63-76.

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.

Jiang, Y., Olson, I., & Chun, M. (2000). Organization of visual short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26,* 683-702.

Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel & K. F. Wender (Eds.), *Spatial cognition: An interdisciplinary approach to representing and processing* *spatial knowledge* (pp. 1-17). Berlin: Springer-Verlag.

McNamara, T. P. (2003). How are the locations of objects in the environment represented in memory? In C. Freksa, W. Brauer, C. Habel & K. Wender (Eds.), *Spatial cognition III, Routes and navigation, human memory and learning, spatial representation and spatial learning,* LNAI 2685 (pp. 174-191). Berlin: Springer-Verlag.

Mou, W., Fan, Y., McNamara, T. P., & Owen, C. (2008). Intrinsic frames of reference and egocentric viewpoints in scene recognition. *Cognition, 106*, 750-769.

Mou, W., Liu, X., & McNamara, T. P. (2008). Layout geometry in encoding and retrieval of spatial memory. *Journal of Experimental Psychology: Human Perception and Performance*, in press.

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., Rump, B., & Xiao, C. (2006). Roles of egocentric and allocentric spatial representations in locomotion and reorientation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32,* 1274–1290.

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.

Mou, W., Zhao, M., & McNamara, T. P. (2007). Layout geometry in the selection of intrinsic frames of reference from multiple viewpoints. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. *33,* 145-154.

Poggio, T., & Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature, 343*, 263-266.

Roskos-Ewoldsen, B., McNamara, T. P., Shelton, A., & Carr, W. (1998). Mental representations of large and small spatial layouts are orientation dependent. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24,* 215-226.

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

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

Sholl, M. J. (2001). The role of a self-reference system in spatial navigation. In D. Montello (Ed.), *Spatial information theory: Foundations of geographical information science* (pp. 217-232). Berlin: Springer-Verlag.

Sholl, M. J., & Nolin, T. L. (1997). Orientation specificity in representations of place. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23,* 1494-1507.

Simons, D. J., & Wang, R. F. (1998). Perceiving real-world viewpoint changes. *Psychological Science, 9*, 315-320.

Valiquette, C. M., & McNamara, T. P. (2007). Different mental representations for place recognition and goal localization. *Psychonomic Bulletin and Review, 14,* 676-680.

Valiquette, C. M., McNamara, T. P., & Labrecque, J. S. (2007). Biased representations of the spatial structure of navigable environments. *Psychological Research, 71,* 288-297.

Valiquette, C. M., McNamara, T. P., & Smith, K. (2003). Locomotion, incidental learning, and the selection of spatial reference systems. *Memory & Cognition, 31*, 479-489.

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

Waller, D., Lippa, Y., & Richardson, A. (2008). Isolating observer-based reference directions in human spatial memory: Head, body, and the self-to-array axis. *Cognition, 106,* 157-183.

Wang, R. F., & Simons, D. J. (1999). Active and passive scene recognition across views. *Cognition, 70*, 191-210.

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

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

Figure Captions

Figure 1. The experiment setup of Experiment 1

Figure 2. Correct percentage in detecting position change as the function of non-target movement and testing perspective in Experiment 1. *(Error bars are confidence intervals corresponding to ±1 standard error of the mean, as estimated from the analysis of variance.)*

Figure 3. An example of learning layout of Experiment 2

Figure 4. An example of test layout of Experiment 2

Figure 5. Correct percentage in detecting position change as the function of context (intrinsic vs. non-intrinsic) in Experiment 2

Figure 1

Figure 2

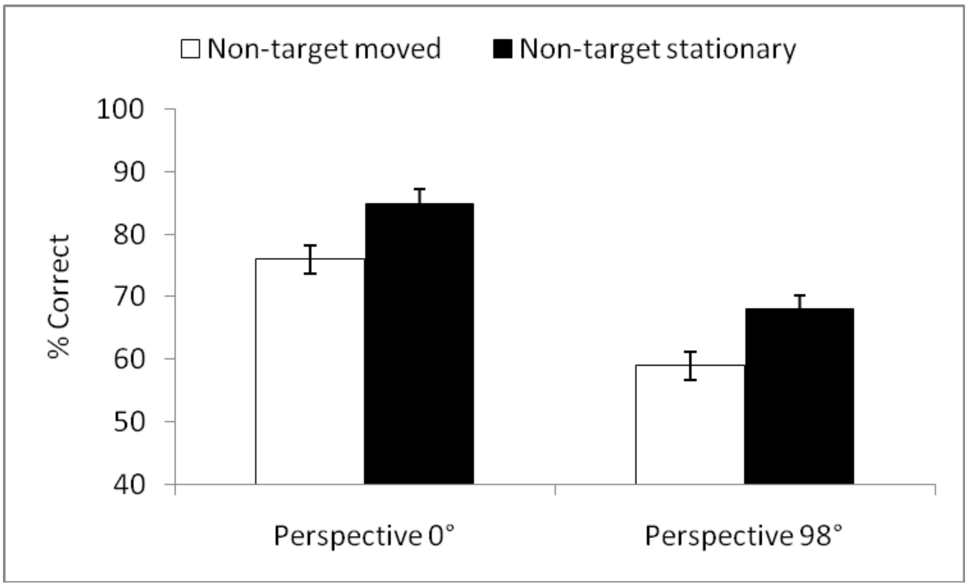

☐ Non-target moved
■ Non-target stationary
100
90
80
70
60
50
40
% Correct
Perspective 0°
Perspective 98°

Figure 3

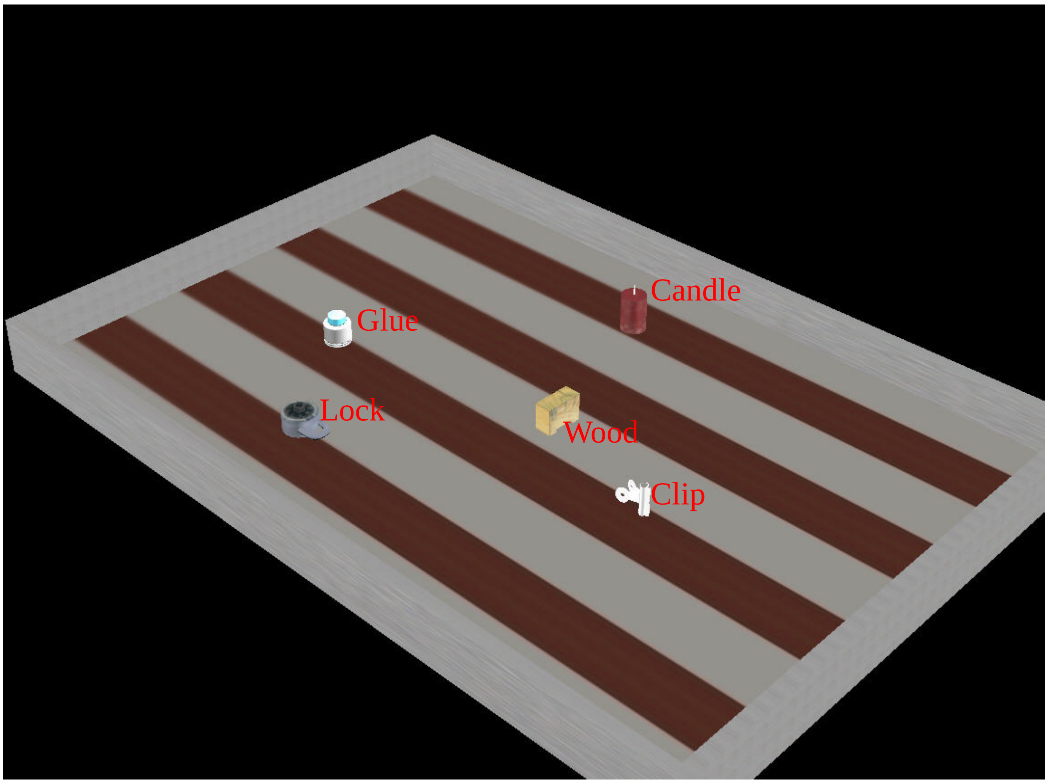
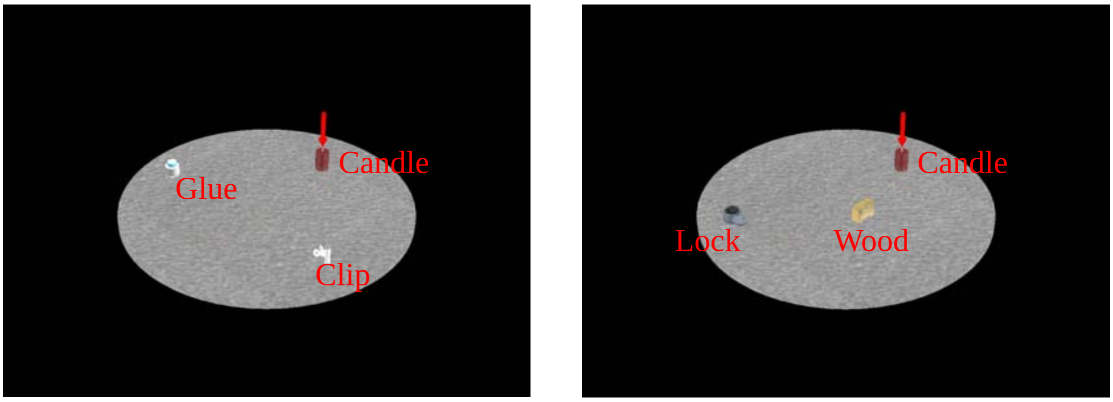


Figure 4



A: Intrinsic context

B: Non-intrinsic context

Figure 5