Orientation Specificity in Long-Term-Memory for Environmental Spaces

Tobias Meilinger (tobias.meilinger@tuebingen.mpg.de)
Bernhard E. Riecke (bernhard.riecke@tuebingen.mpg.de)
Heinrich H. Bülthoff (heinrich.buelthoff@tuebingen.mpg.de)
Max-Planck-Institute for Biological Cybernetics
Spemannstr. 38, 72076 Tübingen, Germany

Abstract

This study examined orientation specificity in long-term human memory for environmental spaces. Twenty participants learned an immersive virtual environment by walking a multi-segment route in one direction. The environment consisted of seven corridors within which target objects were located. In the testing phase, participants were teleported to different locations in the environment and were asked to identify their location and heading and then point towards previously learned targets. As predicted by view-dependent theory, participants pointed more accurately when oriented in the direction in which they originally learned each corridor. No support was found for a global reference direction underlying the memory of the whole layout or for an exclusive orientation-independent memory. We propose a "network of reference frames" theory to integrate elements of the different theoretical positions.

Keywords: Reference frame; environmental space; spatial memory; allocentric; egocentric; reference direction; view-dependent; self-localization; pointing; virtual environment; head-mounted display; navigation; spatial orientation

Introduction

Spatial memory is crucial for our lives as mobile organisms. Without having the capacity to orient oneself in space (which is largely reliant on spatial memory) we would have to search for our bathroom every morning and use aids to find the supermarket. Even when simply having to walk around a corner we would get lost, as is observed, for example in many patients suffering from Morbus Alzheimers. Of specific interest is how locations and spatial layouts are stored in memory.

Theories about the organization of spatial memory

There is an abundance of different and partially conflicting theories about the nature of spatial memory in humans and other animals. (e.g., Burgess, 2006; Mallot & Gillner, 2000; O'Keefe, 1991; Sholl, 2001; McNamara & Valiquette, 2004; Wang & Spelke, 2002). These different theories can roughly be categorized with respect to their assumptions regarding how people represent spatial information in long-term memory. More specifically, these theories assume that we store spatial information either: (1) in an orientation independent manner, (2) in an orientation dependent manner with respect to one or more reference directions, or (3) in an orientation dependent manner with respect to different experienced orientations.

Our goal for the current study was to distinguish between these three theories by designing an experiment in which each theory would independently predict different outcomes. A detailed description of each of these three theoretical positions will now be discussed.

Spatial memory is orientation independent. An orientation independent representation has mainly been argued for by Sholl and colleagues (e.g., Easton & Sholl, 1995; Holmes & Sholl, 2005; Sholl, 2001). They propose an allocentric organization of environmental knowledge. Essentially this means that object-to-object relations are stored in memory, as opposed to self-to-object relations. The defining characteristic of this theory is it assumes that memory content can be accessed equally well, independently of one's current position within the environment and/or facing direction. According to this theory, performance measures should not differ systematically when participants are asked to imagine a previously-learned environment from different perspectives. As such, this theoretic position is thus referred to as orientation independent. According to this approach, additional egocentric reference systems are assumed to exist in which space is not represented in object-to-object relations, but in self-to-object relations. Orientation independence is thought to only occur in well learned environments.

Spatial memory is orientation dependent with respect to a reference direction. Reference direction theory also assumes an allocentric (i.e. object-to-object) memory for space. The objects however, are encoded with respect to one or two reference directions like "north" or the main axis of a room (e.g., Mou, McNamara, Valiquette & Rump, 2004; Rump & McNamara, in press; McNamara & Valiquette 2004). The axes of coordinate systems which define spatial locations might also be interpreted as reference directions (e.g., O'Keefe, 1991). According to this theory, retrieving information from memory should be easiest when aligned with one of the reference directions. For example, imagining a certain position and orientation within a previously-learned scene should be easiest when the to-be-imagined orientation is aligned with one of the reference directions. This facilitating effect is expected to be reflected in improved performance measures such as faster response times and/or decreased errors. The resultant representation is consequently said to be orientation-dependent with respect to one or more reference directions. Such a reference direction is proposed to originate either from the initial exposure to an environment (e.g., the first view of a room), or from the most salient orientation or intrinsic axis of an environment (e.g., the reference axis of a rectangular room would most likely be aligned with the longer walls of the room).

Spatial memory is orientation dependent with respect to experienced views. The third theory is typically referred to as view dependent. It assumes that the environment is stored in the local orientation in which it was experienced (e.g., Christou & Bülthoff, 1999; Mallot & Gillner, 2000; Wang & Spelke, 2002). The defining characteristic of view-dependent theory is that performance is assumed to be highest when one is aligned with the originally experienced orientation. Note that view-dependent representations are not modified or updated when one moves around. According to the definition by Klatzky et al. (1998), such a representation is classified as allocentric, as it is not dependent on the current position and orientation of the navigator. Similar to the two previously described theories, the current theory is based on location-to-location (object-to-object) information or allocentric representations. Alternatively, view-dependent theory can also be conceptualized as an egocentric representation (e.g., Burgess, 2006; Rump & McNamara, in press; Wang & Spelke, 2002).

Memory for environmental spaces

All three theoretical positions have found support from a series of experimental findings. The supporting evidence however, depends critically on the type of space used for testing. One basic distinction can be made between vista spaces and environmental spaces (Montello, 1993): Vista spaces are defined as spaces that are bigger than humans and that are visible from a single point of view. Typical examples for vista spaces include most rooms, open squares, or even small valleys. On the other hand, environmental spaces are defined as spaces where one has to move around and integrate different views to experience the entire space. Examples include buildings or towns (for a similar distinction see Tversky, 2005). The distinction between vista and environmental spaces is independent of the overall size of the spaces. Environmental spaces like buildings can, in fact, be much smaller than vista spaces like valleys or open squares. It is instead the extent to which a particular space can be accessed from one vantage point that remains the central issue.

In the context of vista spaces, all three theories have been supported (e.g., Diwadkar & McNamara, 1997; Holmes & Sholl, 2005; Mou, et al., 2004; McNamara, Rump & Werner, 2003). For environmental spaces, only orientation independent and view dependent theories have been tested (e.g. Christou & Bülthoff, 1999; Easton & Sholl, 1995). Reference direction theory has, however, hardly been investigated for environmental spaces. The purpose of this study therefore, was to test the predictions of these three theories for environmental spaces within one experiment – something that has not been done before.

Methods

For the experiment we used an immersive virtual environment presented via a head-mounted display (HMD). In the learning phase, participants experienced the virtual environment by walking through it. In the testing phase, participants were teleported to different locations in the environment. They were then asked to identify their location and heading and afterwards were instructed to point towards particular targets.

Figure 1: The virtual reality setup. The left image depicts a participant during the learning phase, equipped with a tracking helmet, head-mounted display (HMD), and notebook mounted on a backpack. The right image shows a participant pointing to a target during the testing phase.

Participants

Ten females and ten males between the ages of 19 and 36 (*M* = 25 years, *SD* = 3.8 years) participated in the experiment. They were recruited via a subject database and were paid for their participation.

Material

In the learning phase, participants were asked to learn the layout of the virtual environment and seven target objects located within the environment by walking through it several times. Participants' head position was tracked by 16 high-speed motion capture cameras at 120 Hz (Vicon® MX 13) while they walked freely in a large tracking space 15m×12m (see Figure 1). The participants' head coordinates were transmitted wirelessly (using WLAN) to a high-end notebook computer (Dell XPS M170) which was mounted on a backpack worn by the subject. This notebook rendered an egocentric view of a virtual environment in real-time using a NVIDIA GO 6800 Ultra graphics card with 256 MB RAM. Participants viewed the scene in stereo using a light-weight head-mounted display (eMagin Z800 3D Visor) that provided a field of view of 32×24 degrees at a resolution of 800×600 pixels for each eye. The overall setup provided important depth cues such as stereo vision and motion parallax, as well as all bodily cues important for orientation such as efference copy, vestibular and proprioceptive information.

Using this setup, the participants walked through a virtual environment that consisted of seven connected straight corridors of different colors and wall textures (see Figure 2). The corridors formed one closed loop without any junctions. Seven distinct target objects were placed at a height of 1.3 m, one in each corridor within a circular room. The seven target objects (a brush, telephone, shoe, watch, scissors, banana, and book) were selected to be similar to the objects used in earlier studies investigating the reference direction theory (e.g., Mou et al., 2004). To ensure that participants experienced the corridors only from one direction, they always walked through the corridor in a clockwise direction, without ever turning around. The structure of the environment and its initial exposure was arranged to establish a salient reference axis as predicted by the reference direction theory (see up/forward in the snapshot of Figure 2). This direction was parallel to the view that was first experienced as well as the longest straight path segment of the corridor and the overall orientation of the whole layout. Note that this reference direction was not experienced more often than the other directions represented by the six other corridors. This reference direction is a global orientation, much like a compass direction, as it is the same for all locations. Initial experience and main orientation of the physical lab space result in an identical reference direction in order to prevent interference from multiple reference frames of the physical hall and the virtual environment (e.g., May, 2004).

Procedure

In the *learning phase* participants were asked to walk eight times clockwise through the corridors. Their task was to learn where in the layout the objects were located. That is, participants had to learn in which corridor and where in the whole layout an object was located. Participants were asked not to turn around or look back into the corridor they were coming from. A learning criterion ensured comparable knowledge levels for all participants: At the end of the eighth passage, participants were shown the wall texture of a corridor and were then asked to name the object that is in the corridor of that texture. Participants who did not name all objects correctly could walk two extra rounds through the corridors before being asked again.

In the following *test phase*, participants were seated on a chair in front of a custom-built pointing device (see Figure 1, right). Through the HMD, they were presented with a view of one of the seven circular rooms at the location where an object had been situated during the learning phase. Contrary to the learning phase, all seven target objects were removed and the doors of the circular rooms were closed now in order to block the view to the rest of the corridor. The seven rooms were circular in order to avoid directional biases.

Participants were tested on eight different orientations in steps of 45° within each of the seven rooms resulting in 56 trials altogether. These test directions included the experienced orientation (i.e., along the corridor), and the reference direction as predicted by the reference direction theory (i.e., upwards in the snapshot of Figure 2). To ensure that participants had sufficient visual information to be able to determine their current location and heading even without having to turn their heads, the entrance doors had a wooden texture and the exit door on the opposite side had a metallic texture. Additional small objects (e.g., small rectangular plates that had a wooden and metallic texture on the side facing the wooden and metallic door, respectively) positioned in every circular room at ±45°, ±90° and ±135° indicated the other directions.

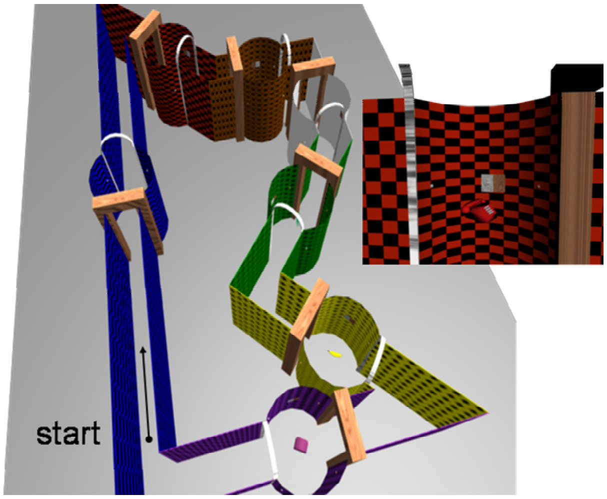

Figure 2: Perspective view of the virtual environment and of the interior of one room in detail (top right). Participants always walked around the environment clockwise, starting with the blue corridor. For the test phase, the doors were closed and the objects removed. From inside a room the participants had to first identify their location and heading and, second, point to the location of another object.

The participants were asked to identify their location and heading and afterwards point towards an instructed target. The time for self-localization was recorded as the time between the initial presentation of a new view and the time when participants indicated via button press that they had localized themselves in the environment (i.e., when they knew the depicted room and their orientation in the room).

Immediately afterwards, participants were asked to use the pointing device (see Figure 1, right) to point as accurately and quickly as possible to a goal target which was indicated by a text on the screen. During pointing, but not during self-localization, participants were asked not to turn their heads. If they did so during pointing against the instructions, the display turned black. During the entire testing phase, participants were physically seated facing the direction that corresponded to the reference direction during the learning phase. The direction displayed in the HMD during the test phase differed, however, for orientations other than the reference direction.

The pointing device consisted of a pointing handle which was connected to a fixed base by a buckling resistant flexible hose. This allowed participants to indicate any direction by moving the pointing handle in that direction. A two-axis acceleration sensor in the pointing handle recorded static and dynamic accelerations including gravitational acceleration, from which the pointing direction was reconstructed with an accuracy of about 1°. We measured pointing accuracy and pointing time (i.e., the time between presenting the goal object and the end of the pointing motion). The goal objects participants had to point to were chosen randomly as was the order of trials.

Hypotheses

The experiment was designed such that the three above-mentioned theories about spatial memory would predict different patterns of performance:

(1) According to the *orientation independent theory*, participants should perform equally well for the different directions they faced in the test phase.

(2) The *reference direction theory* predicts better performance when the current view of the scene is aligned with the global reference direction. According to the theory, this reference direction should correspond to the "upward/forward" direction in the snapshot of Figure 2. Furthermore, participants' performance would be expected to vary depending on their orientation with respect to the global reference direction.

(3) *View-dependent theory* predicts best performance when participants are aligned with the viewing direction in which they experienced the environment. This orientation is locally defined by the orientation of the corridor. According to this, participants' performance should vary depending on their orientation with respect to the experienced orientation.

Results

The pointing accuracy was quantified as the mean absolute pointing error. It differed significantly from the chance level of 90° ($t(19) = 8.10$, $p < .001$). That is, participants did indeed acquire knowledge of the layout.

Participants' pointing accuracy varied as a function of local (experienced) orientation (see Figure 3; ANOVA within subjects; $F(7, 133) = 3.11$, $p = .005$, $\eta^2 = .14$).[1] As predicted by the view-dependent theory, they pointed more accurately when oriented in the direction in which they had experienced the corridor (0°) than when oriented in another direction ($t(19) = 3.99$, $p = .001$, $d = 0.89$). An alternative explanation of the results might be a speed-accuracy trade-off. However, the differences in pointing accuracy due to differences in pointing time, could be ruled out, as there was no effect of local orientation on pointing time ($F(7, 133) = 1.02$, $p = .419$, $\eta^2 = .05$). No effect was found in the time for self localization ($F(7, 133) = 0.71$, $p = .664$, $\eta^2 = .04$).

Participants' performance did not depend on the global orientation, neither in terms of the absolute pointing error (see Figure 4; $F(7, 133) = 1.43$, $p = .199$, $\eta^2 = .07$) or for the pointing time ($F(7, 133) = 1.01$, $p = .430$, $\eta^2 = .05$), or for the time for self-localization ($F(7, 133) = 0.23$, $p = .980$, $\eta^2 = .01$). The reference direction theory was, therefore, not supported by the current data.

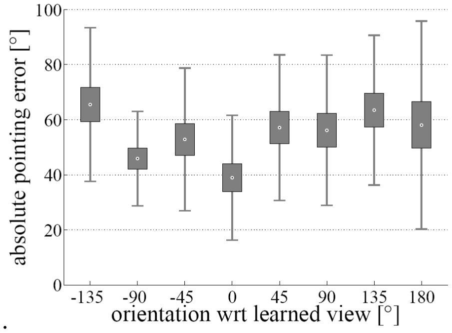

Figure 3: Pointing accuracy as a function of participants' local orientation in each corridor during pointing; that is, their heading relative to the experienced orientation (0°). Means, standard errors (boxes) and standard deviations (whiskers) are displayed.

We directly compared the reference direction theory and the view-dependent theory by comparing performance for the two conditions that are predicted to be the best by the two theories: Participants pointed more accurately when facing the experienced direction than when facing the reference direction ($t(19) = 4.38$, $p < .001$, $d = 0.98$).[2] However, no significant differences were found for the pointing time ($t(19) = 0.29$, $p = .773$, $d = 0.07$) or the time for self-localization ($t(19) = 1.78$, $p = .093$, $d = 0.40$).

Females and males did not differ in terms of their pointing time ($t(18) = 0.88$, $p = .388$, $d = 0.40$) or the time required for self-localization ($t(18) = 1.56$, $p = .137$, $d = 0.70$). Men pointed, however, more accurately ($t(18) = 4.34$, $p < .001$, $d = 1.94$).[3]

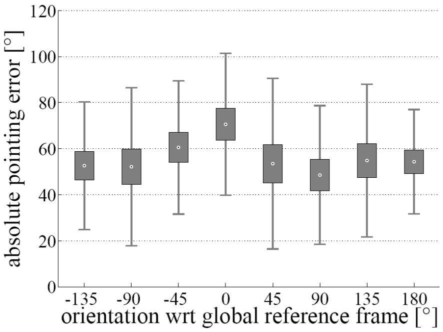

Figure 4: Pointing accuracy as a function of global orientation; i.e., heading relative to the reference direction (0°).

---

[1] For each condition and each participant the median values of each measure was computed in order to control for outliers.

[2] Performance in the first corridor was excluded from this analysis as both theories have identical predictions.

[3] Including gender in the analysis of pointing error produced identical results.

Discussion

The present study examined reference frames used to encode environmental spaces in long-term memory. As predicted by view-dependent theory, participants pointed more accurately when oriented in the direction that they had experienced each corridor. However, no support for a global reference direction underlying the memory of the whole layout could be found. When directly comparing the pointing accuracy between theories, participants performed better in the condition predicted to be best by the view-dependent theory than in the condition predicted to be best by the reference direction theory.

Orientation-independent theory would predict equal performance for all facing directions (e.g., Easton & Sholl, 1995; Holmes & Sholl, 2005; Sholl, 2001). The current data showed, however, clear orientation dependency with respect to the experienced view. This is inconsistent with orientation-independent theories of mental representations for environmental spaces. The time of exposure to the environment (participants walked on average 8.1 times through the environment) might, however, not have been sufficiently long to form a perspective-free memory of the environment. Using much longer learning times might eventually have lead to different results. Similarly, the pattern of results might have been different if participants were allowed to freely explore, thus experiencing the corridors in multiple orientations. Furthermore, our results cannot, exclude the possibility that an orientation independent representation exists in addition to an orientation dependent representation.

View-dependent theory predicts that environments are encoded in the orientation they were originally experienced (e.g., Christou & Bülthoff, 1999; Mallot & Gillner, 2000; Wang & Spelke, 2002). According to this theory, when experiencing an environment from an orientation that is different from the learned orientation, performance should decrease, which is exactly what was observed in the present study. In such misaligned situations, additional mental processes must compensate for the discrepancy between one's current orientation in the environment and the orientation or reference frame in which it was encoded. This compensation could be accomplished, for example, by a shift in perspective or a mental rotation (e.g., Iachini & Logie, 2003; Shepard & Metzler, 1971). Such an explanation is consistent with results from a second experiment that used a very similar setup and procedure as the one reported here. In that experiment, participants were not required to mentally shift their perspective. Instead, they could (and most did) align themselves during the test phase simply by turning their head in the experienced orientation, thus facing 0°. Conversely, in this case none of the participants showed a pattern of rotating their head to align it with any global reference direction.

Note that the explanation of encoding the environment in a view-dependent manner does not necessarily assume an egocentric representation in the sense that only self-to-object relations of the environment are stored. We think that long-term-storage of environmental information as in the case of our experiments always encompasses object-to-object relations, or more general, location-to-location relations. This information would not be updated while the participant moves around, and is in the sense of Klatzky (1998), therefore, an allocentric and not an egocentric representation. This is a mere terminological difference to other positions in order to distinguish updating form long-term memory (cf. Burgess, 2006; Rump & McNamara, in press; Wang & Spelke, 2002). Memory for the environmental space of this experiment would hence be classified as allocentric, because it is stored in long-term memory. Nevertheless, our results clearly show that it is view-dependent.

The results reported here were found in a virtual reality setup using a rather restricted field of view. We therefore cannot exclude the possibility that participants might encode the environment differently when provided with a larger field of view or more natural stimuli. During the pointing itself however, participants were visually oriented with respect to the simulated environment and had to rely on their memory to point to other objects not visible. Hence, no additional restrictions due to the field of view should be expected.

In summary, the current results suggest that spatial memory for environmental spaces is encoded with respect to the local orientation in which it was experienced. Conversely, we could not find support for a global reference direction underlying the spatial memory of all participants, even though the environment used was designed to provide a strong global reference direction. Individual participants might, of course, have used individual reference directions which are not necessarily identical to the direction predicted by the reference direction theory.

Previous studies have shown that the preference of global vs. local orientation depends also on the specific task circumstances. When pointing while being positioned *within* an environment, participants tend to use a local orientation (e.g., Wang & Spelke, 2002). In contrast, when direction judgments are made while *outside* the environment (i.e., imagined pointing), they often rely on a reference direction (e.g., McNamara, Rump & Werner, 2003; Shelton & McNamara, 2001). Furthermore, local orientation appears to dominate in scene recognition, whereas a reference direction effect seems to occur more often in judgments made without visual cuing (e.g., imagined pointing; Valiquette & McNamara, in press). Consistent with these results participants in the current experiment oriented on the local orientation while being located within the environment, not outside and they located on the local orientation while having visual cues available. It is an open question whether orientation dependency with respect to the experienced local view will also be found in imagined pointing when participants are located outside of the environment and can also not see the environment.

The crucial difference between the current study and previous experiments is that in these experiments vista spaces were tested rather than environmental spaces. For memory of vista spaces like the individual corridors in our experiments, our findings do not contradict the predictions of the reference direction theory. Both the reference direction theory and the view dependent theory predict, in fact, the same performance advantage.

To integrate these two theories, we propose to extend the reference direction theory to environmental spaces by allowing for a network of multiple, local reference frames (cf. Meilinger, 2007). Such multiple reference frames are separable conceptual units connected with each other in a network (cf. Mallot & Gillner, 2000). Information from these separate reference frames can be integrated (e.g., during pointing; cf., Wang & Brockmole, 2003). The individual reference frames (e.g., one for each vista space), are not necessarily co-aligned. As for the reference direction theory by McNamara and colleagues, a single reference frame could be selected either based on the initial viewing direction or on salient environmental features like geometry, symmetry, slant, etc. Note that in this context, view-dependent theory is just a subset of our proposed "*network of reference frames*" theory, which could serve as a means to integrate the seemingly incompatible theoretical positions.

Acknowledgments

This research was supported by the EU grant "Wayfinding" (6th FP - NEST). The authors thank Naima Laharnar for help in data collection and processing, Michael Weyel, Gerald Franz and Hans-Günther Nusseck for support in programming and setting up the virtual reality, Daniel Berger for help in writing, Jenny Campos for proof reading, and three anonymous reviewers for their helpful comments.

References

Burgess, N. (2006). Spatial memory, how egocentric and allocentric combine. *Trends in Cognitive Science, 10,* 551-556.

Christou, C.G. & Bülthoff, H.H. (1999). View dependence in scene recognition after active learning. *Memory & Cognition, 27,* 996-1007.

Diwadkar, V.D. & McNamara, T.P. (1997). Viewpoint depence 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.

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,* 1069-1078.

Iachini, T. & Logie, R.H. (2003). The role of perspective in locating position in a real-world, unfamiliar environment. *Applied Cognitive Psychology, 17,* 715-723.

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 representation and processing of spatial knowledge* (pp. 1-17). Berlin: Springer.

Mallot, H.A. & Gillner, S. (2000). Route navigation without place recognition: What is recognized in recognition-triggered responses? *Perception, 29,* 43-55.

May, M. (2004). Imaginal perspective switches in remembered environments: transformation versus interference accounts. *Cognitive Psychology, 48,* 163-206.

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

McNamara, T.P. & Valiquette, C.M. (2004). Remembering Where Things Are. In G. L. Allen (Ed.), *Human spatial memory: Remembering where* (pp. 251-285). Mahwah, NJ: Lawrence Erlbaum Associates.

Meilinger, T. (2007). Orientation in environmental spaces: goals, mechanisms, knowledge, and representations. *Unpublished doctoral dissertation.*

Montello, D. R. (1993). Scale and multiple psychologies of space. In A.U. Frank & I. Campari (Eds.), *Spatial information theory: A theoretical basis for GIS* (pp. 312-321). Berlin: Springer.

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. (1991). An allocentric spatial model for the hippocampal cognitive map. *Hippocampus, 1,* 230-235.

Rump, B. & McNamara, T.P. (in press). Updating Models of Spatial Memory. *International Conference on Spatial Cognition 2006.*

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, 191,* 952–954.

Sholl, M.J. (2001). The Role of a Self-Reference System in Spatial Navigation. In D.R. Montello (Ed.), *COSIT 2001* (pp. 217-232). Berlin: Springer.

Tversky, B. (2005). Functional significance of visuospatial representations. In P. Shah & A. Miyake (Eds.), *The cambridge handbook of visuospatial thinking* (pp. 1-34). Cambridge: Cambridge University Press.

Valiquette, C. & McNamara, T.P. (in press). Different mental representations for place recognition and goal localization. *Psychonomic Bulletin & Review.*

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

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