Gunzelmann, G. (2006). Understanding similarities in performance on different orientation tasks: strategy adaptation. In D. Fum, F. Del Missier & A. Stocco (Eds.), *Proceedings of the 7th International Conference on Cognitive Modeling* (pp. 124-129). Trieste, Italy: Edizioni Goliardiche.

Understanding Similarities in Performance on Different Orientation Tasks: Strategy Adaptation

Glenn Gunzelmann (glenn.gunzelmann@mesa.afmc.af.mil)
Air Force Research Laboratory
6030 South Kent Street; Mesa, AZ 85212 USA

Abstract

Spatial orientation is involved in a variety of common tasks, which are often difficult for individuals to perform. In this paper, a model is used to demonstrate that a general strategy may be adapted for use in different spatial orientation tasks. The predictions of the model, which was adapted from a model for another orientation task, closely match human performance on the target task. More importantly, the model provides insights into how individuals solve these sorts of tasks, and supports the conclusion that similar strategies can be adapted and used in different tasks.

Introduction

Navigation in unfamiliar environments is a common task performed by humans. In many cases, maps are used to guide decision making and minimize errors. The use of maps in these contexts requires that the information on the map be brought into correspondence with the visual world surrounding the individual. This is a potentially difficult process, and anyone who has gotten lost when visiting a new city can attest to the problems that can arise.

In situations where maps are used, it is often the case that the fundamental goal is to identify or locate some particular object, landmark, or place. There are two scenarios that may occur in this basic task. First, there may be something notable in the environment that the individual attempts to locate on the map. Visitors to theme parks are often drawn to visually impressive attractions in the park. To identify what a particular attraction is, however, requires locating it on the map, based upon the visual information available in the environment (a *find-on-map* task).

On the other hand, one can search in the environment for an item identified on the map (a *find-in-scene* task). Conference attendees may be given a map of the local area with the conference venue highlighted. In this situation, the features identified on the map must be used to locate the appropriate building (or room) in the environment. These two tasks are analogous, but not identical. They both require that information from a map be brought into correspondence with information available through visual perception (a visual scene). However, in the first scenario, the target's location is known relative to the egocentrically-based visual scene, while the target is identified within the allocentrically-based map in the latter scenario.

Research on spatial orientation has investigated a variety of tasks to explore how individuals represent and manipulate spatial information. Tversky and her colleagues (e.g., Franklin & Tversky, 1990; Tversky, 2003) have used orientation tasks to examine how individuals encode and remember spatial information. Her spatial framework theory emphasizes the role of the major axes of the body and the physics of the world in determining the ease with which this information can be accessed. Other research by Sholl has pursued a similar goal (e.g., Easton & Sholl, 1995; Sholl, 1995). Her research focuses on the role of alternative spatial coordinate systems, including body-centered and environment centered frames of reference.

Other research has used perspective taking as a paradigm to investigate the role of different strategies on performance (e.g., Huttenlocher & Presson, 1979; Presson, 1982; Wraga, Creem, & Proffitt, 2000). This research illustrates that the particular operations that are used to transform spatial information can have implications for performance. Gunzelmann and Anderson (2005) illustrate how the strategies identified for perspective taking can be applied to the kinds of spatial orientation tasks considered here. This provides initial evidence that general strategies may apply across a range of spatial tasks.

Finally, Hintzman, O'Dell, and Arndt (1981) provided a number of demonstrations of the results typically found in the sorts of orientation tasks described here. Across experiments, participants were asked to perform variations of the *find-in-scene* task, which required indicating the direction of a target relative to an indicated orientation (an egocentric judgment). Using a variety of stimuli and methods, they demonstrated a common pattern of results, with effects of both the relative direction of the target and the discrepancy between the egocentric orientation of the individual and the indicated orientation for the trial (misalignment).

Current Research

All of the research described above has produced similar patterns of results on tasks that require reasoning about spatial information presented in multiple reference frames. However, the investigations do not provide detailed explanations of the strategies or mechanisms that are involved in producing these results. Gunzelmann and Anderson (2004) present a model of human performance on a *find-on-map* task. This model made accurate quantitative predictions about human performance on that task using a strategy based on verbal reports from participants in the study (see Gunzelmann & Anderson, in press for a more detailed discussion of the strategy).

This paper explores the issue of whether similar cognitive mechanisms and strategies can be used to explain performance on *find-in-scene* tasks. While it seems reasonable to expect that performance would be similar, spatial information for the task must be manipulated differently depending on which sort of orientation task is performed. Thus, in this paper, the model from Gunzelmann and Anderson (2004) is extended to perform the *find-in-scene* task and to generate predictions of performance.

The model is used to illustrate that the same high-level strategy can be used in both tasks. However, the details of how the strategy is executed are different. A comparison of those details indicates that the same steps are involved, but in a different order. As a result, the model predicts that the pattern of results in the *find-in-scene* task should be identical to the *find-on-map* task. To test that prediction, I conducted an experiment, where participants were asked to perform the *find-in-scene* task, using the same experimental design as was used in Gunzelmann and Anderson (2004).

Experiment

The results described in this paper represent the data from only a portion of the entire experiment. Other aspects of the experiment are beyond the scope of this paper. Here I focus on the *find-in-scene* task, and the results that are used to evaluate the model presented below.

Method

Participants The participants in this study were 16 individuals recruited from the local community surrounding the Air Force Research Laboratory in Mesa, AZ, which includes Arizona State University's Polytechnic Campus. There were 6 males and 10 females in the study, with an average age of 32 years. Participants were paid $10 per hour for their participation.

Materials The task used in this research is illustrated in Figure 1. The figure shows a typical trial, with two views of a space depicted. On the left is an egocentrically-oriented view, while the right shows an allocentric map. On the map, a single object is highlighted in white (in the experiment it was red) to indicate that it is the target. Participants' task was to identify which of the objects in the egocentric visual scene corresponded to the target highlighted on the map. Responses were made by clicking on the appropriate object.

For this experiment, six unique configurations of objects in the space were used. These configurations are identical to those described in Gunzelmann and Anderson (2004), and were designed by organizing the objects into groups, arranged according to quadrants in the space. In each configuration, there were groups of 1, 2, 3, and 4, which were distributed among the four quadrants of the space. The six configurations represent the six possible arrangements of the group sizes relative to each other. In the experiment, each configuration was presented in all 8 45-degree

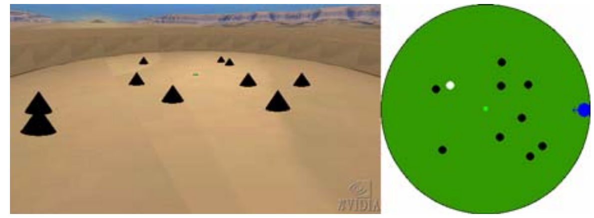

Figure 1: Sample Trial. The white dot on the map indicates the target, and the blue point (on the edge) indicates the viewer's location.

rotations on the map, with the quadrants defined by the main axes (+; Figure 1) or obliquely (X). This resulted in 48 different maps being shown to participants during the study.

The target in each trial could appear in any of the four quadrants. Any other objects in the quadrant served as local distractors. Finally, the viewer in each trial was positioned at the bottom, right, left, or top of the map, always looking straight across to the opposite side. This manipulation created 4 levels of misalignment between the orientation of the egocentric view (forward = up) and the orientation of the map.

Each participant completed all of the possible trials defined by the factors just described. This results in 768 trials (6 quadrant configurations in 8 rotations, with 4 possible target locations and 4 possible levels of misalignment for each). These were presented in a random order, using a drop-out procedure. If a participant made an error on a particular trial, that trial was repeated at some point later in the experiment, with the restriction that no trial was presented twice in a row (unless it was the last remaining trial in the experiment).

Procedure The experiment involved 2 sessions, however all of the trials for the task presented here were completed in one of those sessions, which lasted approximately 2 hours.[1] At the beginning of the session, participants were provided with instructions on the task, including a sample trial to complete. These instructions were computer-based, and each participant was required to respond to the sample trial correctly before beginning the experiment. The experimenter answered any questions participants had before they began.

[1] In the other session, participants completed the original *find-on-map* task from Gunzelmann and Anderson (2004; in press), with the order of tasks counterbalanced. Whereas the order in which these tasks were completed had an overall impact on performance, with the second task being completed more rapidly, $F(1,14)=18.12$, $p<.001$, none of the interactions involving task order were significant. While consideration of the *find-on-map* task is beyond the scope of this paper, it is worthwhile to note that there was no significant difference in overall performance on the two tasks, $F(1,14)=0.39$, $p>.50$, and the pattern of results was quite similar ($r=.931$). This provides further support for the model described below.

Trials were divided into blocks of 20, and participants were given the opportunity to take a break after each block. Response times (ms) and accuracy were recorded for each trial by the experiment software.

Results

The data were analyzed to explore the effects of three main factors, the position of the target relative to the viewer, the number of objects in the cluster containing the target, and the misalignment between the egocentric visual scene and the map. All three of these factors had an influence on performance. As targets were positioned off to one side or the other, and as they were farther from the viewer, response times increased, $F(7,98)=11.55$, $p<.001$. Response times also increased as more objects were positioned nearby to the target, $F(3,42)=92.47$, $p<.001$. Finally, as the misalignment increased, response times increased as well, $F(3,42)=45.65$, $p<.001$. Notably, error rates were low in this experiment, with overall accuracy of about 92%. These errors tended to follow a pattern similar to the response time data ($r=.64$), so they are not considered further here.

In addition to the main effects, there were interactions in the data. The impact of the size of the cluster was influenced by the extent of the misalignment between the two views. As the degree of misalignment between the two views increased, the impact of additional distractors near to the target increased (Figure 2). This interaction was highly significant, $F(9,126)=10.83$, $p<.001$. There was also an interaction between the location of the target and the number of nearby distractors, $F(21,294)=3.65$, $p<.01$. In this case, the pattern of results is less clear. The number of nearby distractors has an impact on performance, regardless of where the target is located relative to the viewer. The strongest effect appears to be that the location of the target had little impact when there were no nearby distractors (Figure 3).

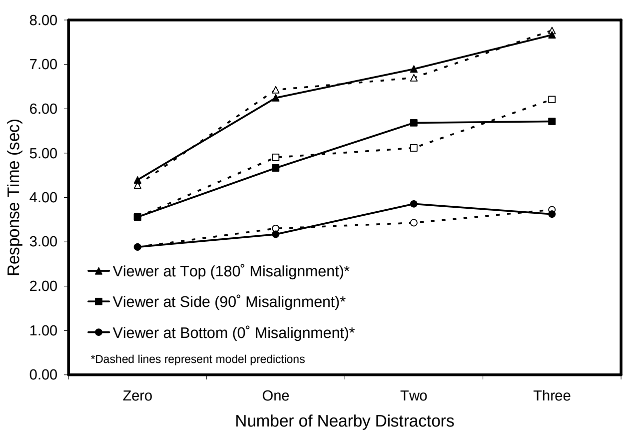

Figure 2: Empirical results and model data for the interaction between the number of nearby distractors and misalignment.

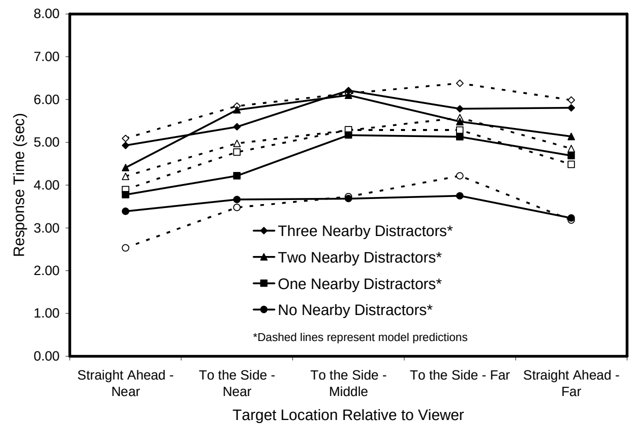

Figure 3: Empirical results and model data for the interaction between the number of nearby distractors and target location.

Finally, the location of the target influenced the impact of misalignment on performance (Figure 4). The effect of misalignment was reduced in cases when the target was located near the horizontal center of the field of view (the first and last points on each line). This effect was also significant, $F(21,294)=2.53$, $p<.02$.

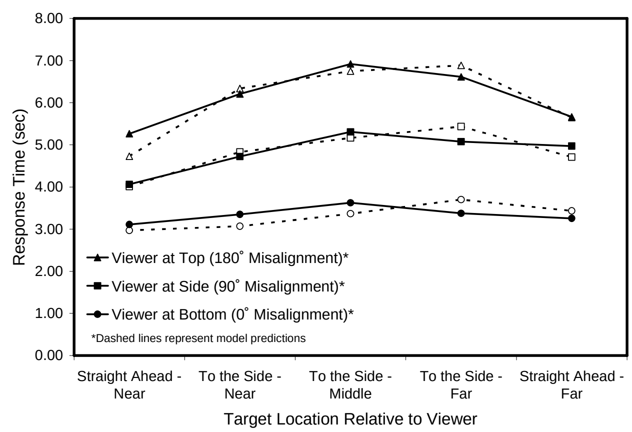

Figure 4: Empirical results and model data for the interaction between the location of the target and misalignment.

Discussion

The results of this experiment show that, although participants were able to successfully solve the problems, several factors that were varied made the task more or less challenging. The influence of misalignment reflects a commonly investigated aspect of human performance on this kind of task. When two views of a space become increasingly misaligned, it is more difficult to establish correspondence between them (e.g., Hintzman et al., 1981;

Rieser, 1989; Wraga et al., 2000). In addition, the results relating to the impact of the target's location and the number of nearby distractors present are consistent with results from the *find-on-map* version of this task reported previously (Gunzelmann & Anderson, 2004; in press).

All of the results support the conclusion, described in the introduction, that performance on these two types of orientation tasks is similar in many important ways. In the next section, this conclusion is examined in more detail, by evaluating how the strategy reported by participants completing the *find-on-map* task can be adapted and used to complete the *find-in-scene* task used here. To quantitatively assess this relationship, a computational cognitive model is used to evaluate how successful the adapted strategy is at accounting for the data reported for this experiment.

Computational Cognitive Model

The model described here was created in the ACT-R cognitive architecture (Anderson et al., 2004). ACT-R is a hybrid cognitive architecture, which posits a distinction between declarative and procedural knowledge at the symbolic level. Subsymbolic mechanisms influence the operation of the architecture by impacting which knowledge is used in particular situations, as well as controlling latencies in accessing or using that knowledge. The production cycle in ACT-R is serial, with a single production rule being executed (fired) on each cycle.

The production system component of the architecture operates by matching against the contents of buffers, which are the interface between the productions and other modules. Modules perform distinct processing tasks. For instance, there is a motor module, a vision module, and a declarative memory module. Modules operate in parallel, and subsymbolic evaluations within the modules are carried out in parallel as well. The results of the processes within each module are deposited in the module's buffer(s), which can then be accessed by the production system. The result of a production firing is to modify the contents of buffers directly, or to make requests of particular modules, which eventually result in changes to the buffers.

Model Overview

For ACT-R to perform a particular task requires that the knowledge necessary for performing the task be added to the architecture. In Gunzelmann and Anderson (2004), the knowledge took the form of a particular strategy for doing the *find-on-map* task, which was based on verbal reports from the participants in the study. The strategy involves encoding the location of the target in the visual scene hierarchically. First, a group of objects containing the target is identified, and the position of the cluster is encoded relative to the viewer. Then, the location of the target within that cluster is encoded relative to the other objects in the group. This representation is updated relative to the viewer's location on the map, and then used to identify the appropriate object. The steps involved in this strategy are summarized in Table 1.

Table 1: Steps for the *find-on-map* task.

Step	Description
1	Locate target (in visual scene)
2	Encode location of cluster -Egocentric reference frame -Right/left/center
3	Encode target location (in visual scene) -Egocentric reference frame -Right-left/near-far position
4	Find viewer on map
5	Identify allocentric reference frame
6	Update cluster location -Convert to allocentric reference frame
7	Locate cluster (on map)
8	Update target postion -Convert to allocentric reference frame
9	Locate target (on map)
10	Respond

The *find-in-scene* task is quite similar to the *find-on-map* task used in Gunzelmann and Anderson (2004). To examine the implications of an analogous strategy on performance for the task used here, a model of the *find-in-scene* task was generated by adapting the model described in Gunzelmann and Anderson (2004). Essentially, this model was produced by re-ordering the steps involved in performing the task to correspond to the demands of the new task. Table 2 presents the steps in executing the solution strategy in the model for the *find-in-scene* task.

Table 2: Steps for the *find-in-scene* task. Step numbers in parentheses reflect the step number for the *find-on-map* task in Table 1.

Step	Description
1 (4)	Find viewer on map
2 (5)	Identify allocentric reference frame
3 (1)	Locate target (on map)
4 (2)	Encode location of cluster -Egocentric reference frame -Right/left/center
5 (6)	Update cluster location -Convert to egocentric reference frame
6 (3)	Encode target location (on map) -Egocentric reference frame -Right-left/near-far position
7 (8)	Update target postion -Convert to egocentric reference frame
8 (7)	Locate cluster (in visual scene)
9 (9)	Locate target (in visual scene)
10	Respond

There are several special cases, which are not represented in Tables 1 or 2. For instance, if the target is very near to the viewer or directly in front of the viewer, then some steps may be skipped (see Gunzelmann & Anderson, 2004 for details). However, these special cases are the same for both tasks. As a result the same steps are involved in executing the general solution strategy for both versions of the task. It is only their order that differs. The implication of this is that performance on the *find-in-scene* task should be similar to the performance on the *find-on-map* task. The data presented above support this, showing a pattern of results similar to those presented in Gunzelmann and Anderson (2004). In the next section, additional details about the model are presented, followed by a description of the model's performance for the experiment presented above.

Model Details

The processing steps involved in executing the general strategy for these tasks is responsible for the qualitative performance of the model. The quantitative performance, however, depends on the values for several parameters. The first, the retrieval latency, uses the same value (0.11s) as was used in Gunzelmann and Anderson (2004). This parameter controls the time required to retrieve a chunk from declarative memory. In this task, knowledge is needed about the objects on the screen, *right* versus *left*, etc. Each time one of those pieces of information is needed, it takes 110ms for the declarative module to place the appropriate chunk into the retrieval buffer, where it can be accessed by the central production system.

A second parameter reflects the costs associated with moving between 2-D and 3-D coordinate systems, which is required to align information in the two views. This parameter was set to 1.0 s, which is higher than where it was set in the previous model (0.25 s). This parameter helps to capture the substantially longer response times of participants in this study (4.93 s versus 3.77 s on average).

The final parameter that was manipulated reflects the speed of processing the spatial information in the images. These operations are involved in updating the descriptions of the cluster location and the location of the target within the cluster as the model solves the task. This occurs in the task during steps 5 and 7 in Table 2, which involve conversion to an egocentric reference frame. The model does this by updating the descriptions generated using the allocentric frame of reference of the map to match the egocentric frame of reference in the visual scene. The spatial updating parameter used in this model differs from the value used in Gunzelmann and Anderson (2004). In this model, it was set to 0.9s, whereas it was set to 0.6s in the earlier model.

In the model, each of the operations needed to update a piece of spatial information requires an amount of time equal to the value of the spatial updating parameter. The number of operations increases as the number of nearby distractors increases, and as the misalignment between the two views increases. For instance, no updating is necessary when the two views are aligned (i.e., when the viewer is located at the bottom of the map). The special cases, mentioned above, largely represent instances where these updates can be skipped. For instance, if a cluster is located straight ahead of the viewer, there is no need to encode information about *left* versus *right* to support searching for the cluster on the map; *straight ahead* suffices. Jacobsen and Waters (1985) provide an empirical demonstration of how coordinating left-right and near-far axes impacts difficulty, including the benefits associated with being "in the middle." Additional details concerning this mechanism are described in Gunzelmann & Anderson (2004).

All of the other parameters in the model were set to the same values as were used in Gunzelmann and Anderson (2004). The parameters that were different may be associated with individual differences in abilities that are relevant for this task. For instance, the parameter that is associated with extracting allocentric information from the visual scene may be associated with familiarity with the types of virtual environments used in this experiment. The second parameter is associated with spatial ability, relating to the speed with which individuals can update their frame of reference. Research has shown individual differences in mental rotation ability (Just & Carpenter, 1985), so it should not be surprising that different participant groups would differ with respect to this ability in this experiment. Familiarity with the kinds of virtual environments used here could also impact performance with respect to this factor. Mental transformations can be performed more quickly with familiar material (Bethell-Fox & Shepard, 1988).

Model Performance

The model presented here was adapted from the model reported in Gunzelmann & Anderson (2004). Thus, it predicts that performance on the *find-in-scene* task should be influenced by the same factors as the find-in-map task. Since the empirical data conform to this prediction, the qualitative predictions of the model are a good fit to the results from this experiment. In fact, the overall qualitative fit of this model to these data (r=.96) is as good as the fit of the original model (r=.95) to the data reported in Gunzelmann & Anderson (2004).

The model captures the trends illustrated in Figures 2-4. In Figure 2, the impact of a.dditional nearby distractors increases as misalignment increases (r=.983). In Figure 3, the model predicts that the impact of target location will arise, regardless of how many nearby distractors there are (r=.912). This figure illustrates the largest discrepancy between the model and human performance. In the empirical data, there is little or no effect of the target's location when there are no nearby distractors, whereas there is an effect in the model. Finally, Figure 4 illustrates that the model makes accurate predictions about how the target's location relative to the viewer influences the effect of misalignment on performance (r=.980). Overall, these results show that the model accurately captures the relative influence of these different factors on performance.

The model captures the quantitative level of performance of the participants as well, using the values described above for the parameters (RMSD=0.278 s, 0.416 s, and 0.250 s for Figures 2, 3, and 4, respectively). The average RMSD is 0.315 s, which is comparable to the fit reported in Gunzelmann and Anderson (2004), where the average RMSD was 0.287 s.

General Discussion

The experiment presented here illustrates several phenomena associated with orientation tasks, including the well-established impact of misalignment on performance. In addition, this research adds additional evidence that the location of the target, as well as the location of the distractors, can have significant influences on performance in this kind of task.

The model extends earlier work to a somewhat different orientation task. The changes that were implemented illustrate how the strategy for the *find-on-map* task can be adapted to successfully apply to the *find-in-scene* task. The general strategy has been shown to make use of efficient strategies from the perspective-taking literature (Gunzelmann & Anderson, 2005), and careful analysis of the empirical data provides additional support for the strategy, in addition to the verbal reports from participants (Gunzelmann & Anderson, in press).

The model captures the trends in the human data, with about the same degree of accuracy as the model presented in Gunzelmann & Anderson (2004). This further supports the notion that similar high-level strategies may be adapted for other variants of this kind of spatial task. The hierarchical encoding that participants report using for these tasks can be applied much more generally to representing spatial information (McNamara, Hardy, & Hirtle, 1989; Stevens & Coupe, 1978). Combined with the capacity for making spatial transformations, like mental rotation, the general strategy reported here could be adapted for use in a variety of tasks that require reasoning about spatial information. Future research will be directed at understanding how the strategy adaptation process occurs, and what the limits are on the ability to adapt general strategies to novel tasks.

Acknowledgments

The experiment described here was supported by AFOSR Grant #02HE01COR, and by a National Research Council Research Associateship Award to the author. I would like to thank Sharon Lebeau for her help with data collection, and Don Lyon and Kevin Gluck for their helpful comments.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y . (2004). An integrated theory of the mind. *Psychological Review 111*, 1036-1060.

Bethell-Fox, C. E., & Shepard, R. N. (1988). Mental rotation: Effects of stimulus complexity and familiarity. *Journal of Experimental Psychology: Human Perception and Performance, 14(1)*, 12-23.

Easton, R. & Sholl, M. (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.

Gunzelmann, G., & Anderson, J. R. (2004). Spatial orientation using map displays: A model of the influence of target location. In K. Forbus, D. Gentner, and T. Regier (Eds.), *Proceedings of the Twenty-Sixth Annual Conference of the Cognitive Science Society* (pp. 517-522). Mahwah, NJ: Lawrence Erlbaum Associates.

Gunzelmann, G., & Anderson, J. R. (2005). Integrating multiple strategies efficiently to solve an orientation task. In B. G. Bara, L. Barsalou, and M. Bucciarelli (Eds.), *Proceedings of the Twenty-Seventh Annual Meeting of the Cognitive Science Society* (pp. 851-856). Mahwah, NJ: Lawrence Erlbaum Associates.

Gunzelmann, G., & Anderson, J. R. (in press). Location matters: Why target location impacts performance in orientation tasks. *Memory & Cognition*.

Hintzman, D., O'Dell, C. & Arndt, D. (1981). Orientation in cognitive maps. *Cognitive Psychology*, *13*, 149-206.

Huttenlocker, J. & Presson, C. (1979). The coding and transformation of spatial information. *Cognitive Psychology, 11*, 375-394.

Jacobsen, T. & Waters, H. (1985). Spatial perspective taking: Coordination of left-right and near-far spatial dimensions. *Journal of Experimental Child Psychology, 39*, 72-84.

Just, M. & Carpenter, P. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92*, 137-172.

McNamara, T., Hardy, J. & Hirtle, S. (1989). Subjective hierarchies in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 211-227.

Presson, C. (1982). Strategies in spatial reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 8*, 243-251.

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

Sholl, M. (1995). The representation and retrieval of map and environment knowledge. *Geographical Systems, 2*, 177-195.

Stevens, A. & Coupe, P. (1978). Distortions in judged spatial relations. *Cognitive Psychology, 10*, 422-437.

Tversky, B. (2003). Structures of mental spaces: How people think about space. *Environment and Behavior, 35*, 66-80.

Wraga, M., Creem, S. & Proffitt, D. (2000). Updating displays after imagined objects and viewer rotations. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 151-168.