Cognitive Science: A Multidisciplinary Journal
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t775653634

Strategy Generalization Across Orientation Tasks: Testing a Computational Cognitive Model

Glenn Gunzelmann [a]
[a] Air Force Research Laboratory, Mesa, Arizona

Online Publication Date: 01 July 2008

To cite this Article Gunzelmann, Glenn(2008)'Strategy Generalization Across Orientation Tasks: Testing a Computational Cognitive Model',Cognitive Science: A Multidisciplinary Journal,32:5,835 — 861
To link to this Article: DOI: 10.1080/03640210802221957
URL: http://dx.doi.org/10.1080/03640210802221957

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
JUL 2008	**Journal Article**	**01-01-2006 to 30-06-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Strategy Generalization across Orientation Tasks: Testing a Computational Cognitive Model	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER **61102F**

6. AUTHOR(S)	5d. PROJECT NUMBER
Glenn Gunzelmann	**2313**
	5e. TASK NUMBER **HA**
	5f. WORK UNIT NUMBER **2313HA09**

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)	8. PERFORMING ORGANIZATION REPORT NUMBER
Air Force Research Laboratory/RHA,Warfighter Readiness Research Division,6030 South Kent Street,Mesa,AZ,85212-6061	**AFRL; AFRL/RHA**

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
Air Force Research Laboratory/RHA, Warfighter Readiness Research Division, 6030 South Kent Street, Mesa, AZ, 85212-6061	**AFRL; AFRL/RHA**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) **AFRL-RH-AZ-JA-2008-0005**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Published in Cognitive Science: A Multidisciplinary Journal; 32(5), 835-861 (2008)

14. ABSTRACT

Humans use their spatial information processing abilities flexibly to facilitate problem solving and decision making in a variety of tasks. This article explores the question of whether a general strategy can be adapted for performing two different spatial orientation tasks by testing the predictions of a computational cognitive model. Human performance was measured on an orientation task requiring participants to identify the location of a target either on a map (find-on-map) or within an egocentric view of a space (find-in-scene). A general strategy instantiated in a computational cognitive model of the find-on-map task, based on the results from Gunzelmann and Anderson (2006), was adapted to perform both tasks and used to generate performance predictions for a new study. The qualitative fit of the model to the human data supports the view that participants were able to tailor a general strategy to the requirements of particular spatial tasks. The quantitative differences between the predictions of the model and the performance of human participants in the new experiment expose individual differences in sample populations. The model provides a means of accounting for those differences and a framework for understanding how human spatial abilities are applied to naturalistic spatial tasks that involve reasoning with maps.

15. SUBJECT TERMS
Testing; Prediction; Spatial Ability; Information Processing; Generalizability Theory; Performance Technology; Coding; Maps

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Public Release**	**27**	

Cognitive Science 32 (2008) 835–861

ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1080/03640210802221957

Strategy Generalization Across Orientation Tasks: Testing a Computational Cognitive Model

Glenn Gunzelmann

Air Force Research Laboratory, Mesa, Arizona

Received 29 August 2006; received in revised form 22 May 2007; accepted 7 August 2007

Abstract

Humans use their spatial information processing abilities flexibly to facilitate problem solving and decision making in a variety of tasks. This article explores the question of whether a general strategy can be adapted for performing two different spatial orientation tasks by testing the predictions of a computational cognitive model. Human performance was measured on an orientation task requiring participants to identify the location of a target either on a map (find-on-map) or within an egocentric view of a space (find-in-scene). A general strategy instantiated in a computational cognitive model of the find-on-map task, based on the results from Gunzelmann and Anderson (2006), was adapted to perform both tasks and used to generate performance predictions for a new study. The qualitative fit of the model to the human data supports the view that participants were able to tailor a general strategy to the requirements of particular spatial tasks. The quantitative differences between the predictions of the model and the performance of human participants in the new experiment expose individual differences in sample populations. The model provides a means of accounting for those differences and a framework for understanding how human spatial abilities are applied to naturalistic spatial tasks that involve reasoning with maps.

Keywords: Spatial orientation; Spatial cognition; Cogntive architecture; Computational model; Prediction; Strategy; Experiment; Adaptive Control of Thought–Rational (ACT–R)

1. Introduction

Traveling through an unfamiliar environment can be a challenging experience. To facilitate appropriate decision making, this task is often supported by using a map of the space. Maps provide a representation of space that is not tied to the viewer's egocentric perspective (i.e., they use *exocentric* or *allocentric* reference frames). This provides the advantage of allowing individuals to gain knowledge about the spatial layout of the space without direct experience

Correspondence should be sent to Glenn Gunzelmann, Air Force Research Laboratory, 6030 South Kent Street, Mesa, AZ 85212-6061. E-mail: glenn.gunzelmann@us.af.mil

navigating through it. However, because information about a space is represented differently on a map than it is from a first-hand, egocentric perspective, there are challenges associated with using the map-based information effectively to guide decision making and action.

One specific challenge associated with map use relates to the need to account for any discrepancies in orientation, or misalignment, between the exocentric reference frame of the map and the egocentric reference frame of our visual experience. Much of the research conducted on the topic of using maps has addressed this issue (e.g., Gunzelmann & Anderson, 2006; Hintzman, O'Dell, & Arndt, 1981; Levine, 1982; Levine, Marchon, & Hanley, 1984). Levine, Jankovic, and Palij (1982) described the requirements for this process using the "two-point theorem." To establish correspondence between two views of a space requires that two links between the representations be made. First, a common point must be known to link the two views. Then, to align the orientations of the two views, a second point or a reference direction is required.[1] Establishing these links allows an individual to translate accurately a direction specified in one frame of reference into the corresponding direction in the other.

An everyday example of the kind of reasoning required in this situation is determining which way to go in a mall to find a particular store. Of course, a simple search of the map will reveal the store's location, and these maps almost always include a "you-are-here" indicator, which shows the map's location in the mall. The indicator can be used as the first reference point. However, the store's location cannot be used to aid in establishing correspondence with the environment because its location in the environment is not known. If you knew that information, you would not need the map! However, if the map is placed appropriately, it should be aligned with the space, such that up on the map is forward in the mall when you are facing it (see Levine et al., 1982). Such placement facilitates identifying relative directions and, if this relation is also indicated in some way on the map, it provides the additional benefit of serving as a reference direction that can be used in establishing the appropriate relationship. However, this is often not the case (Levine, 1982; Levine et al., 1984). When it is not, other strategies are needed to complete the process of establishing correspondence, like finding a nearby store or locating a distinctive feature in both reference frames (e.g., an exit or elevator). Once this is accomplished, it should be possible to determine which way to go to get to the intended store.

Establishing correspondence between two views of a space is the fundamental process assessed by orientation tasks. Orientation tasks involve making some judgment about a space, which requires integrating information across multiple reference frames. The particular task can take a variety of forms, such as locating objects in the environment or on a map (e.g., Gugerty & Brooks, 2004; Gunzelmann & Anderson, 2006; Gunzelmann, Anderson, & Douglass, 2004; Hintzman et al., 1981), indicating the relative direction of something in the environment given an assumed position in the space (e.g., Boer, 1991; Rieser, 1989; Wraga, Creem, & Proffitt, 2000), or orienteering and navigation tasks (Aginsky, Harris, Rensink, & Beusmans, 1997; Dogu & Erkip, 2000; Malinowski, 2001; Malinowski & Gillespie, 2001; Murakoshi & Kawai, 2000; Richardson, Montello, & Hegarty, 1999). Regardless of the particular task, all orientation tasks require the fundamental step of determining how the representations correspond. Often there are two spatial views, an egocentric visual scene and an allocentric map (e.g., Gunzelmann & Anderson, 2006; Malinowski & Gillespie, 2001). In other cases, familiar environments are used, so individuals are asked to rely on internal cognitive representations of the space, rather than a map (e.g., Aginsky et al., 1997; Murakoshi & Kawai, 2000).

Tasks requiring coordination of spatial information from multiple frames of reference arise whenever there is uncertainty about location within an environment. This can be a frequent occurrence, especially the first few times that a town, mall, park, or other new area is visited. Despite the frequency with which individuals are challenged to perform these tasks, they are still difficult, often requiring significant cognitive effort to solve correctly (e.g., Aginski et al., 1997; Hintzman et al., 1981; Rieser, 1989). The research presented here examines human performance on two different orientation tasks to better understand the sources of difficulty and expand on previous studies.

One factor that has been shown repeatedly to influence the difficulty of orientation tasks is misalignment (e.g., Gunzelmann et al., 2004; Hintzman et al., 1981; Rieser, 1989). Misalignment is the difference (in degrees) between the orientations of two views of a space. For instance, traditional maps are oriented with north at the top, but individuals may be facing in any direction when using the map. If users happen to be facing north, then their view of the world is aligned with the map. However, the view of the world becomes increasingly misaligned as the direction being faced deviates more from north. In the worst case, the view of the world is misaligned by 180°, which happens when individuals are facing south while attempting to use a map oriented with north at the top (for a thorough discussion, see Levine et al., 1982). Resolving this discrepancy is a major source of difficulty in this kind of task. The impact of misalignment unites not only tasks of spatial orientation (e.g., Hintzman et al., 1981; Rieser, 1989), but also studies of mental rotation (e.g., Shepard & Hurwitz, 1984; Shepard & Metzler, 1971) and vision (e.g., Tarr & Pinker, 1989).

Although orientation tasks are a class of problem that share important features, there are distinct differences among them as well. For instance, to use a map to locate a store in a mall, you would locate the store on the map, along with your position and other features, and use that information to find the store in the actual space (find-in-scene). In contrast, the process goes in the opposite direction if the feature of interest is observed in the environment. In this case, the challenge is to locate it on a map (perhaps to identify what it is—a find-on-map task). Of course, other variations in the task result in more dramatic differences in task demands. Still, even in the two orientation tasks just described, the differences between them mean that they cannot be solved using an identical strategy. Moreover, it may be that the information processing demands of the tasks tax the human cognitive system in different ways. This research explores human performance on these two types of task, building on past research and evaluating the ability to account for performance on both of them using a general strategy and a common set of mechanisms.

In Gunzelmann and Anderson (2006), research is described that examines human performance on an orientation task where the target was identified in the visual scene and had to be located on the map (a find-on-map task). The present study extends this work in two important ways. First, the experiment described below replicates the experimental design for locating the target on a map from Experiment 2 in Gunzelmann and Anderson (2006), with one exception. In this experiment a completely within-subjects design was used, which was not the case in the earlier research (see the Experimental Methodology section). In addition to this methodological change, this study also extends the earlier research, where participants completed only a find-on-map task. In the current study, participants also did a corresponding find-in-scene task, where the target was highlighted on the map and the corresponding item

had to be identified in the visual scene. The human data collected here is used to test the predictions of a computational cognitive model, which was developed based on the findings from Gunzelmann and Anderson (2006). Thus, this research offers a means of both validating the earlier empirical results, and evaluating the generalizability of the computational account instantiated in a model to account for human performance in a different orientation task (find-in-scene).

The remainder of the article begins with an introduction to the experiment paradigm, followed by a description of the computational model for both tasks. The model was developed in ACT–R, and is based on a solution strategy reported by nearly all of the participants in the original empirical study using the find-on-map task (Gunzelmann & Anderson, 2006, Experiment 2). Following the detailed description of the ACT–R model, the performance predictions for both tasks are presented. The current empirical study is then described and the human data are compared to the predictions of the model. The article concludes by

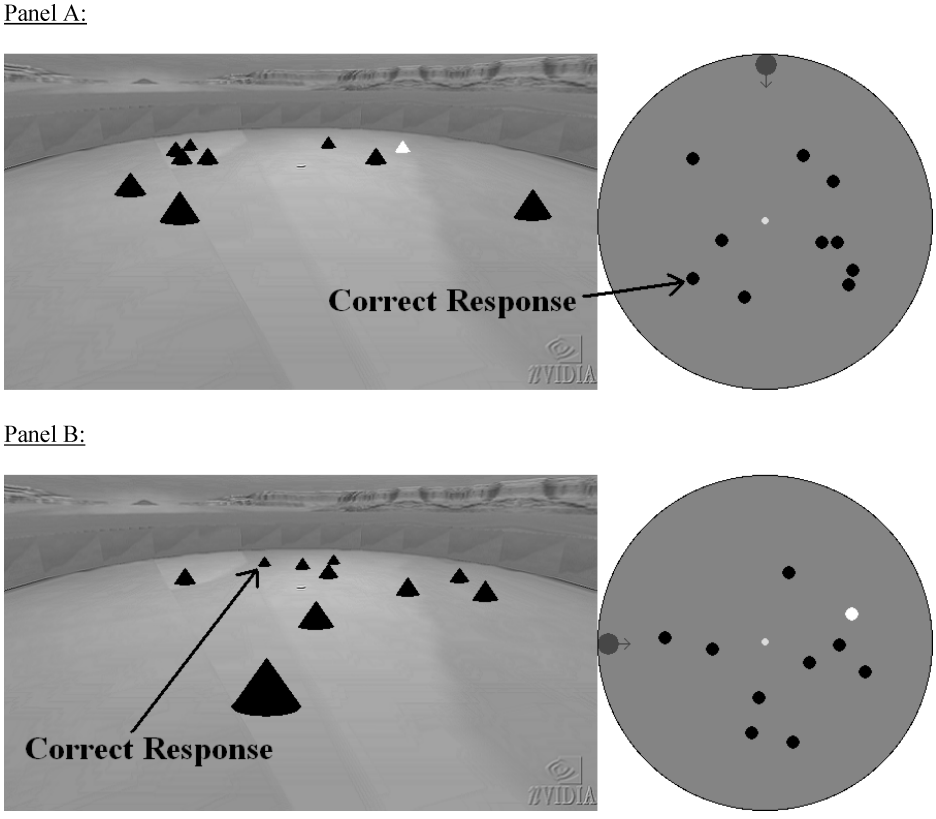

Fig. 1. Sample trials for the tasks used in this research. *Note:* Panel A shows the find-on-map task, with the target highlighted (in white) in the visual scene. Panel B shows the find-in-scene task, with the target highlighted (in white) on the map. In both tasks, the viewer's position is identified on the map as a circle on the edge with an arrow pointing to the center. Participants respond by clicking on the object in one view that corresponds to the target that is highlighted in the other. The correct response is indicated in each image.

discussing the implications of the modeling work and exploring possible extensions to the research.

1.1. Task paradigm

The research presented here involves an orientation task where a target is highlighted in one view of a space, and participants are asked to identify that target in the other view. One of the views is an egocentric visual scene, whereas the other is an allocentric map. A sample trial is shown for each version of this task in Fig. 1. In Panel A, the find-on-map task is illustrated, with the target highlighted in the egocentric view of the space. In Panel B, the find-on-scene task is shown, with the target highlighted on the map. In these examples, the target is shown in white, though it was red in the actual experiment. In both tasks, participants were asked to respond by clicking on the object corresponding to the target in the other view. The trial shown in Fig. 1a illustrates the task presented to participants in Gunzelmann and Anderson (2006).

The exact manner in which the stimuli were designed is described below, but there is one point that is influential with regard to the model. This relates to how objects were positioned within the space. There were 10 objects in the space in each trial, and they were arranged in groups (clusters). The space, itself, was divided into four quadrants, which had 1, 2, 3, and 4 objects, respectively. The arrangement of these quadrants relative to each other and relative to the viewer's position in the space was counterbalanced. The result is a set of spaces that did not have any obvious regularities to participants in the current study, none of whom were able to describe the constraints imposed on how the objects were positioned in the space. The model described next organizes the space by grouping objects according to the quadrant divisions. Groups, or clusters, of objects play an important role in the model's performance, by providing some context for narrowing the search for the target to a portion of the space in the other view. Evidence that participants used groups, and that they generally corresponded to the quadrant structure of the space, is provided below (see also Gunzelmann & Anderson, 2006). Ongoing research is directed at incorporating mechanisms into ACT–R to perform perceptual grouping (e.g., Best & Gunzelmann, 2005; Gunzelmann & Lyon, 2006).

2. Model overview

The model described in this section is able to complete both the find-on-map task and the find-in-scene task. The implementation of the model was based upon retrospective verbal reports of participants in Gunzelmann and Anderson (2006), who reported a common, general strategy for performing the find-on-map task. In addition, the data from that study were used to derive values for the parameters identified in this section. The mechanisms in this model are described in detail for the find-on-map task, based directly on the verbal reports, followed by a discussion of how the strategy was generalized to perform the find-in-scene task shown in Fig. 1b. The resulting model was used to produce *a priori* predictions of performance on the find-in-scene task, which are presented at the end of this section.

2.1. ACT–R cognitive architecture

The model was developed in the ACT–R cognitive architecture (Anderson et al., 2004), and it relies heavily on several fundamental properties of the architecture, such as the distinction between declarative knowledge and procedural knowledge. The strategy reported by participants in the original study was used to generate procedural knowledge to allow the model to complete the task. Declarative knowledge was incorporated into the model to support the solution process. For instance, knowledge about what the items on the display represent and concepts like *right* and *left* are represented in the model as declarative chunks. Besides the distinction between declarative and procedural knowledge, the perceptual and motor components of ACT–R, which give it the ability to interact with software implementations of tasks, are critical for generating performance predictions. The model's performance relies on its ability to encode information from the screen and to generate responses by making virtual mouse movements and clicks, with timing mechanisms that are based on existing psychophysical research. In tasks such as those used here, these aspects of human performance are critical components of a complete computational account.

2.2. Model for the find-on-map task

The performance of the model is driven primarily at the symbolic level, with the subsymbolic mechanisms influencing the latencies of events such as mouse movements, attention shifts, and retrievals of declarative chunks from memory. The model uses the default ACT–R parameter values for all of these mechanisms, except for retrieval latencies. The time required by the model to retrieve a chunk from declarative memory was set to .11 sec. However, even this value was based on previous research (Gunzelmann et al., 2004).

The basic task for the model is to encode the location of the target in enough detail so that it can be identified from among the objects shown in the other view. The model uses a hierarchical approach to accomplish this. First, the model identifies a group or cluster of objects that contains the target. The number of objects in this group and its location relative to the viewer (left, right, or straight ahead) are encoded, which provide sufficient information to allow the same cluster to be located on the map. In Fig. 1a, for instance, the target is within a cluster of three, located on the right side of the visual scene. In Fig. 1b, it is a cluster of four located straight ahead of the viewer. The second level of encoding is to identify the position of the target within the particular cluster. The model encodes this position verbally, as the *n*th object from one side or the other, and as the *n*th closest object in the group. The combined information about the location of the cluster and the position of the object within the cluster provides sufficient detail to accurately identify the correct object on the map. Note, however, that if the object is isolated in the space (a cluster of one), then the second level of encoding can be skipped. In these cases, finding the cluster is equivalent to finding the target.

2.2.1. Locating the correct cluster

Once the location of the target is encoded, the model shifts its attention to the map to identify the appropriate response. Finding the target on the map involves applying the description of the target's location to the perspective shown on the map. The viewer's location is identified

on the map, which facilitates this process (see Fig. 1). In fact, this provides all the information necessary to align the two views of the space because both the viewer's location and orientation are identified (see the Introduction section). Consequently, the model starts processing the map by finding the indicator showing the viewer's location. To find the cluster, then, the model searches the map for a cluster that is the correct size, and which is positioned correctly relative to the viewer.

To perform the search for the cluster efficiently, the model restricts its search to the appropriate region of the map based on the cluster's location relative to the viewer (to the left, to the right, or straight ahead). If the cluster is not straight ahead of the viewer in the visual scene, some spatial updating is required to determine the corresponding portion of the map when the viewer is not positioned at the bottom of the map (i.e., when the two views are misaligned). For instance, in Fig. 1a, the "right" side of the visual scene corresponds to the "left" half of the map. Note that in the trial shown in Fig. 1b, this update is not required. In that trial, the cluster is positioned straight ahead of the viewer, which corresponds to the same portion of the map, with or without the spatial updating step. Thus, the model skips the update in the situation depicted in Fig. 1b. For this experiment, the strategy of finding *any* cluster of the correct size in the correct qualitative location is sufficient because each quadrant contains a different number of objects. This process could be more complex if, say, there were multiple groups of the same size within the same space. The visuospatial reasoning required to disambiguate groups in such a situation is not addressed in this model, but it is an area of emphasis in current research.

The process of updating spatial references takes time in the model, and is controlled by a fixed cost spatial updating parameter. This parameter was estimated, based on the existing data from Gunzelmann and Anderson (2006), to be 0.6 sec. In addition to the costs of the update, when the viewer is located at the top of the map (the views are maximally misaligned, like Fig. 1a), the model also incurs a penalty for the resulting direct conflict between egocentric directional reference and the updated references for the map (right and left are reversed, and thus interfere with default egocentric references). For simplicity, the magnitude of this penalty is equal to the value of the spatial updating parameter. Note that when the cluster is straight ahead of the viewer, the model is able to save significant time by skipping this update. Finally, regardless of whether spatial updating is required, the model starts its search from the position of the viewer, and moves further away until it finds the appropriate group.

The costs associated with the spatial updating parameter cause the model to require longer to find solutions to trials when the two reference frames are more misaligned. This factor has been shown repeatedly to impact response times (RTs) in a variety of particular orientation tasks (Gunzelmann et al., 2004; Hintzman et al., 1981; Shepard & Hurwitz, 1984). Usually, the explanation provided for this effect involves a variant of mental rotation—that is, accounts of performance on this kind of task tend to make the claim that mental rotation is involved in transforming a representation of the information in one reference frame to match the orientation in the other reference frame. The updating process in the model does not explicitly instantiate that view. In part, this is due to architectural limitations of ACT–R, which do not support complex manipulations of spatial information like mental rotation. Thus, although participants often reported using verbal descriptions much like the model to describe the target's location, it is possible that mental rotation is utilized as a means of updating those

descriptions to apply appropriately on the map, and mental rotation is sometimes reported by participants as well (Gunzelmann & Anderson, 2006). With this in mind, the mechanisms described for updating those references can be seen as a simplification of the mental rotation processes that participants may be using.

2.2.2. Identifying the target in the cluster

Once the cluster is found, the model is faced with the challenge of identifying the appropriate item within the cluster. Recall that if there are no nearby distractors (a cluster of one), then this second step can be skipped. In those cases, the model responds as soon as the "cluster" is found by virtually moving the mouse and clicking on the object. In other cases, the model has to use the information encoded about the object's position to identify it. Once again, this requires that spatial references be updated based on the viewer's position on the map. This time, it is the references that were used to encode the object's position within the cluster. The implementation of the model assumes that these references are updated independently of the updates for the cluster location. Thus, spatial updating is required in this step as well when the two views are misaligned. The separability of these updates is discussed in more detail in Gunzelmann and Anderson (2006), where the empirical results were used to tease apart the updating costs for these two steps.

The process for updating the description of the object's position within the cluster operates as follows in the model. When the cluster contains 2 or 3 objects, updating the position description for the object is identical to the updating process for locating the cluster. In these cases, it is possible to encode the target's position using a simple qualitative reference, just like was possible for describing the cluster's position. Specifically, the target will be the leftmost or the rightmost, and also the nearest or farthest, in a group of 2. For a group of 3, the object may also be the one "in the middle" on each axis. Thus, the costs associated with the updating process in these cases are the same as the updates that were performed for the cluster, including the cost associated with conflicting references when the views are maximally misaligned. For a cluster of 4, however, a more sophisticated encoding will often be required. To reflect the additional complexity associated with transforming a representation of the location of the target in a cluster of 4, the updating costs for the model were doubled. Thus, the cost of making transformations becomes 1.2 sec in these cases. In addition, when the views are maximally misaligned, the cost rises to 1.8 sec, as a result of the penalty incurred for the directly conflicting references. The increased cost of these operations for clusters of 4 was estimated using the empirical data from Gunzelmann and Anderson (2006). As noted above, these costs reflect the effort required to update spatial references in the description of the target's location when the reference frames of the two views are misaligned. This process may involve mental rotation, which is not explicitly implemented here. The increased time required to make this transformation when more objects are in the group is related to findings that mental rotation is slowed by increased complexity in the stimulus (e.g., Bethell-Fox & Shepard, 1988).

Spatial updating is costly in the model, but once the process is completed, the model can apply the updated description of the target's location and identify the object on the map that corresponds to the target highlighted in the visual scene. Once the appropriate object is identified, the model executes a virtual mouse movement and click to make its response. In

addition to the parameters already noted (retrieval latency and spatial updating), the model has one additional parameter that was adjusted to match the quantitative results reported by Gunzelmann and Anderson (2006). This parameter is associated with moving between two-dimensional (2-D) and three-dimensional (3-D) coordinate systems represented by the map and the visual scene, respectively. This parameter was set to 0.25 sec to capture the cost of moving between these reference frames. The fit of this model to the data from Gunzelmann and Anderson (2006) is presented below. Next, however, the extension of the model to the find-in-scene task is described.

2.3. Model for find-in-scene task

The method used for generalizing the model described above to perform the find-in-scene task involved assuming that the same high-level strategy that participants reported for the find-on-map task would be employed. The implication of this approach is that the same steps will be executed to arrive at a response for the find-in-scene task, but those steps must be performed in a different order. This is illustrated in Table 1, which lists the steps involved in solving either task, with reference to the order of the steps in each. When the strategies are compared in this way, it obscures the fact that the knowledge state of the model will vary between the two tasks for any given step because they are performed in a different order. For example, in the find-on-map task, the allocentric reference frame is identified after the target's location has been encoded in egocentric terms. In contrast, this is the very first piece of information encoded by the model as it begins the solution process for the find-in-scene task. These differences reflect the variations in how the strategy is employed in the two tasks, and when relevant pieces of information about the trial are available and needed in each case.

There is an interesting difference in the processes that are required for completing the two tasks as well. In the find-on-map task, egocentrically encoded information must be converted to the allocentric coordinate system of the map. In the find-on-scene task, the process required is the opposite; information encoded relative to the map-based coordinate system must be

Table 1
Steps involved in executing the general strategy for both orientation task variants

	Step Number	
Viewer Position (Misalignment)	Find-on-Map	Find-in-Scene
Locate target	1	3
Encode location of cluster	2	4
Encode target location	3	6
Find viewer on map	4	1
Identify allocentric reference frame	5	2
Update cluster location	6	5
Locate cluster	7	8
Update target position	8	7
Locate target	9	9
Respond	10	10

converted to an egocentric reference frame to allow for appropriate search in the visual scene. This is the process that is reflected in the spatial updating parameter described above.

Differences in the timing of these processes could have important implications for the model's performance. However, in the model predictions presented next, the simplifying assumption is made that those two processes should be symmetric, and that the timing associated with them should be identical. By making this assumption, the model predicts that performance on the two tasks should be quite similar, qualitatively and quantitatively. In the next section, the model for the find-on-map task is compared directly to the data from Gunzelmann and Anderson (2006). In conjunction with these fits, the predictions of the model for the find-in-scene task are presented as *a priori* predictions about the match in performance across these two tasks.

2.4. Performance predictions

The data from Gunzelmann and Anderson (2006), in conjunction with the model predictions for both tasks, are shown in Figs. 2 through 4. Figure 2 illustrates the impact of misalignment on performance, with RTs increasing as the misalignment between the two views increases. The impact of misalignment is also apparent in Fig. 3. This effect in the model was discussed above, and stems from the spatial updating that occurs in cases where the two views are misaligned. Updates are needed whenever this situation exists, and when the views are maximally

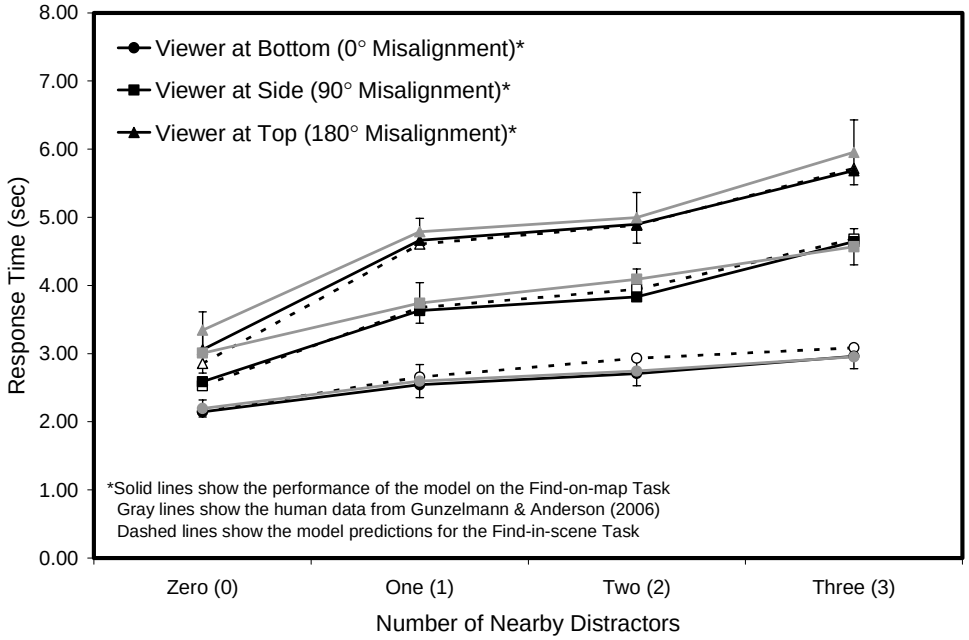

Fig. 2. Model predictions of performance in both tasks for the effects of misalignment and the number of nearby distractors. *Note:* Error bars for the human data show ±1 standard error.

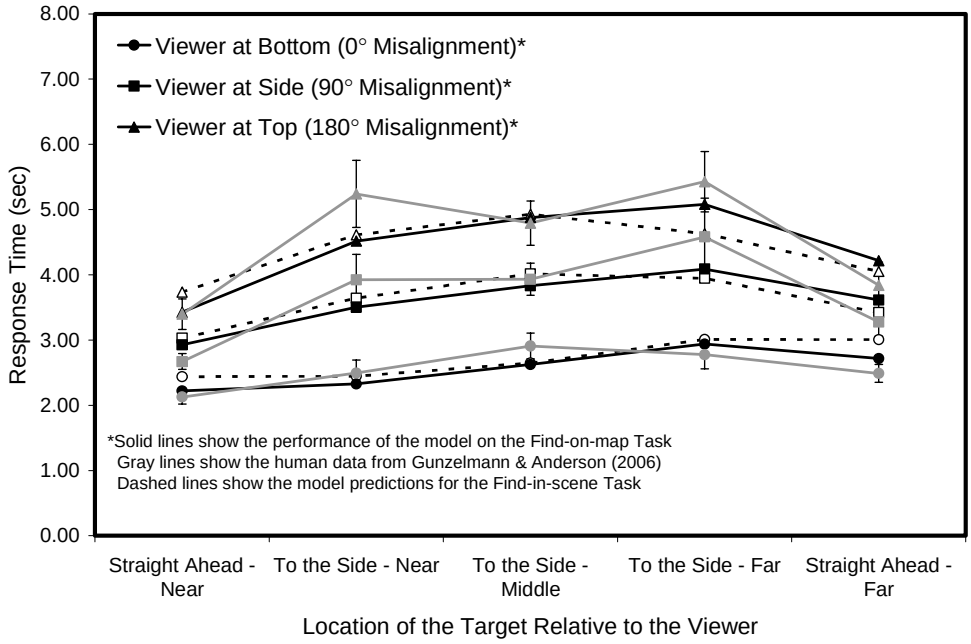

Fig. 3. Model predictions of performance in both tasks for the effects of misalignment and the location of the target relative to the viewer. *Note:* Error bars for the human data show ±1 standard error.

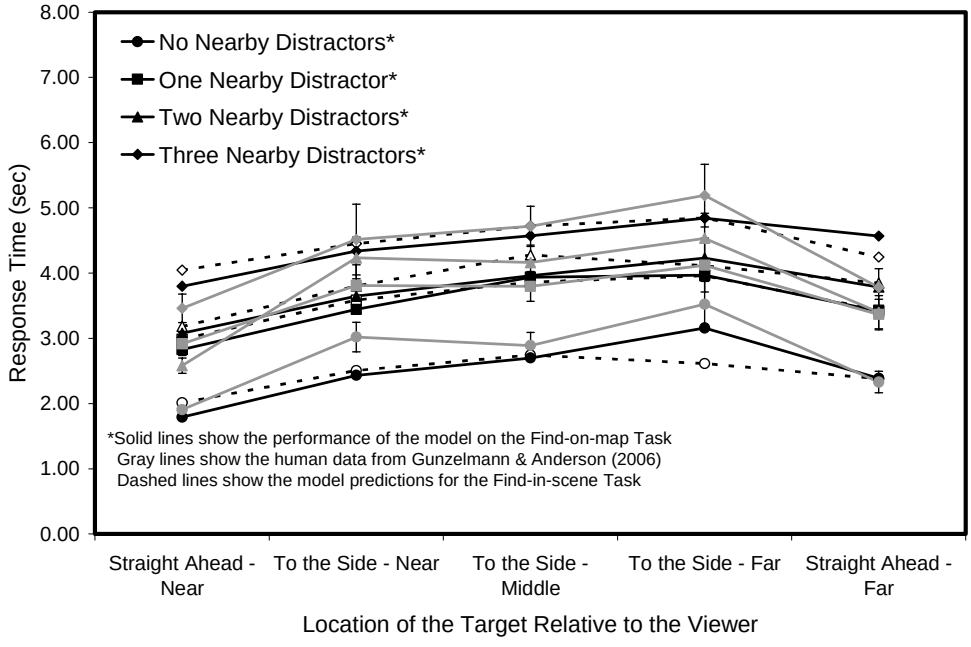

Fig. 4. Model predictions of performance in both tasks for the effects of the number of nearby distractors and the location of the target relative to the viewer. *Note:* Error bars for the human data show ±1 standard error.

misaligned the model also pays an additional penalty for direct conflict between the spatial terms.

Figure 2 also shows the influence of nearby distractors on performance in both the human participants and the model. Recall from the brief description of the task that clusters of objects were positioned within quadrants in this task. The other objects located in the same quadrant as the viewer are identified as nearby distractors. Because groups of 1 to 4 were used, the target was among 0 to 3 nearby distractors on any given trial. As the number of these distractors increased, so did RTs. The impact of nearby distractors is the result of the additional processing that is needed as the size of the cluster increases. Simply encoding a larger group takes longer due to the costs of shifting visual attention. In addition, it is more costly, on average, to encode the target's position in a larger cluster, and then to apply that encoding on the other view because more objects must be considered. This main effect can be seen in Fig. 4 as well.

The interaction between misalignment and nearby distractors, shown in Fig. 2, arises from the increased cost associated with making spatial updates to resolve misalignment as the number of nearby distractors increases. With no nearby distractors, the second encoding step is skipped by the model. For groups of 2 and 3, the target's position within the cluster must be encoded, but the costs of updating that information is less than for clusters of 4, where the target's position within the cluster is still needed, and the cost of updating that information is greater.

Figures 3 and 4 illustrate the impact of the target's location on performance. Because of the experimental design in Gunzelmann and Anderson (2006), one group of participants completed trials where the target was off to the side in the near and far positions while another group of participants did the other trials. This may explain the pattern of data in Figs. 3 and 4, which is not completely captured by the model. By using a completely within-subjects design in the experiment described below, this effect can be evaluated more systematically. In contrast, the predictions from the model reflect the manner in which the model searches for the cluster containing the target object. The model initiates this search close to the viewer and moves away until the cluster is found. The impact of this is that nearby clusters are found more quickly, resulting in faster RTs for those trials. In addition, no spatial updating is required when the cluster is approximately straight ahead of the viewer. Because the updating cost is avoided in those cases, RTs are faster for those conditions as well. The latter mechanism produces the interaction between the target's location and misalignment, shown in Fig. 3. The first and last points on each line illustrate conditions where the cluster is located straight ahead of the viewer. In these cases, the impact of misalignment is reduced relative to other target locations. Last, the model predicts no interaction between the location of the target and the number of nearby distractors (Fig. 4). The first factor influences the search for the cluster—that is, the location of the target defines the location of the cluster, regardless of how many objects are in the cluster. Meanwhile, the number of objects in the cluster influences the second step of identifying the target within the cluster. The location of the cluster in the space does not impact the solution process once the cluster has been identified. Because they impact different steps in the solution process, these two factors do not influence each other in the model. This prediction in the model is supported by the human data from Gunzelmann and Anderson (2006), as illustrated in Fig. 4.

The data shown in Figs. 2 through 4 illustrate that the model is accurate in capturing the performance of human participants on the find-on-map task. The model predicts all of the major trends in the data. The strategy implemented is responsible for the qualitative predictions of the model. The hierarchical approach of locating a cluster followed by identifying the target position within the cluster leads to the prediction that the size of the cluster should have an impact on performance. The search strategy utilized for locating the cluster causes the position of the cluster relative to the viewer in the space to influence RTs. Finally, resolving misalignment between the two views increases RTs on misaligned trials as a result of the processes required. Over all of the data, the model is in line with these qualitative trends ($r = .90$). The quantitative predictions of the model were fit by manipulating the two spatial parameters. The retrieval latency (0.11 sec) was based on previous research (Gunzelmann et al., 2004). These parameters indicate that spatial processes are central to performance on this task, and also suggest that there are substantial challenges associated with processing the spatial information presented in the display. With the parameter values identified above, the model for the find-on-map task captures human performance quite well at the quantitative level as well (Root Mean Squared Deviation, RMSD = 0.536 sec).

The model for the find-on-map task was generated based upon the verbal reports of participants in the Gunzelmann and Anderson (2006) study. In addition, the parameters were estimated to fit those data. Thus, perhaps it is unsurprising that the model provides a good fit to the empirical data. However, the model was also extended to perform the find-in-scene task, which was not performed by human participants in Gunzelmann and Anderson (2006). The predictions of that model are presented in Figs. 2 through 4 as well. What is most interesting about the predictions is the degree of correspondence they have to the predictions of the model for the find-on-map task. This is true not only at a qualitative level ($r = .97$), but also at a quantitative level (RMSD = 0.273 sec). These predictions offer an opportunity to evaluate the general account embodied by the model of human performance on spatial orientation tasks because the model for the find-in-scene tasks utilizes the same high-level approach to the task as the model for the find-on-map task. In the next section, an experiment is presented that provides a test of the model's predictions, and of the generalizability of the solution strategy reported by participants in Gunzelmann and Anderson (2006).

3. Experiment

This experiment provides a replication and extension of Experiment 2 in Gunzelmann and Anderson (2006), including an important procedural modification. Gunzelmann and Anderson divided participants into two groups, and individuals in each group completed one half of the set of trials that are possible given the stimulus design. Data were merged across these groups by pairing participants using an assessment of spatial ability based upon Vandenberg and Kuse's (1978) Mental Rotation Test. Although the conclusions based on those meta-participants were partially validated later in that article, those particular results have not been replicated using a within-subjects design. Consequently, each participant completed all 768 of the possible trials in this study to provide such a replication. In addition to this modification of the earlier methodology, this experiment extends the previous research by having the

participants complete a find-in-scene task, in addition to the find-on-map task. By modifying the stimulus materials so that the target was highlighted on the map, rather than in the visual scene, a direct comparison was made of performance on these two orientation tasks using otherwise identical stimuli. This also allowed for a direct evaluation of the predictions of the ACT–R model described above.

Demonstrating that the effects found previously occur in a within-subjects context is an important step in validating the account developed in Gunzelmann and Anderson (2006) and the predictions of the model presented above. In addition, direct comparisons between the two different orientation tasks explored here have not been conducted previously. The model makes the prediction that performance should be qualitatively and quantitatively very similar between them, despite differences in the processing demands and the kinds of transformations that are required. This experiment provided the data needed to evaluate the theoretical claims embodied by the model.

3.1. Method

3.1.1. Participants

The participants in this study were 16 volunteers solicited from the local community around the Air Force Research Laboratory in Mesa, AZ, which includes Arizona State University's Polytechnic Campus. Participants ranged in age from 18 to 50, with a mean age of 32 years. There were 6 men and 10 women in the sample. Participants were paid $10 per hour for their participation, which consisted of two sessions, each lasting approximately 2 hr.

3.1.2. Design and materials

The stimuli used in this study were nearly identical to those used in Gunzelmann and Anderson (2006). The only difference was in the background landscape that was used for the egocentric views of the space. This was modified for greater clarity and discriminability of the objects relative to the background. The size of the space and the positions of objects were identical. The stimuli were created using the *Unreal Tournament* (2001) game engine. In each trial, a space containing 10 objects was shown. On the left was a visual scene, showing the perspective of a viewer standing on the edge of the space. On the right was a map of the space, which included an indication of the viewer's position. All 10 objects were visible in both views on every trial. Fig. 1a shows a sample trial for the find-on-map task, and Fig. 1b shows a sample trial for the find-in-scene task. The objects, themselves, were placed into the space according to quadrants, which contained 1, 2, 3, and 4 objects, respectively. For the experiment, six unique spaces were created, which represent all of the possible arrangements of the quadrants relative to each other. Then, by presenting each of these maps in 8 possible 45° rotations, all of the possible arrangements of the quadrants, oriented at 45° intervals relative to the viewer, were included. These variations were incorporated to offset any influence that the particular arrangement of those quadrants might have on performance.

In each trial, the target was highlighted as a red object (it is white in the sample trials shown in Fig. 1). For the find-on-map task (Fig. 1a), the target was highlighted in the visual scene. For the find-in-scene task, the target was highlighted on the map. As in Fig. 1, the viewer's position was indicated on the map for each trial, regardless of which task was being performed.

The viewer was located in one of 4 positions at the edge of the map. The viewer was either at the bottom, one of the sides (left or right), or top of the space. In all cases, the viewer was facing toward the center of the space (visible as a light-colored dot in both views). This manipulation produces different degrees of misalignment: 0° when the viewer is at the bottom, 90° when the viewer was at either side, and 180° when the viewer was at the top of the map. These manipulations to the stimuli result in a large number of potential trials for each task (768). There are 8 possible locations of the target relative to the viewer[2], from zero to three distractors located in the same quadrant as the target[3], four different levels of misalignment, and 6 different configurations of quadrants relative to each other (6 different maps). All of these conditions were repeated for both tasks. The only difference in generating trials for the two tasks was which view contained the highlighted target, with the other view being the one where participants had to locate the target and make their response. The procedural details are described next.

3.1.3. Procedure

Participants completed all of the possible trials for both tasks in two sessions. One task was done in Session 1, and the other task was completed in Session 2. The order of tasks was counterbalanced across participants to offset learning and other possible order effects. Each session lasted no more than 2.5 hr (only 2 of 32 sessions lasted more than 2 hr). Participants began each session by reading a set of instructions for the task, including a sample trial. Participants were required to respond correctly to the sample trial before beginning the experiment. In addition to the instructions, the experimenter was available to answer any questions participants had before they began. Each session was divided into blocks of 20 trials, allowing participants to take a short break between them if desired. In addition, their progress through the experiment was indicated by providing information about how many trials they had completed and how many they had gotten correct. Participants made their responses by clicking on the object in one view that they believed corresponded to the object highlighted in red in the other view. Feedback was given on each trial regarding whether the response was correct or incorrect. Their RTs and click locations were recorded on each trial. Clicks that did not fall on one of the objects in the appropriate view were ignored.

This experiment also incorporated a drop-out procedure. If a participant made an error on any of the trials, that trial was repeated later in the experiment until the participant got it correct. The only constraint on this was that the same trial was never presented twice in a row, unless it was the last remaining trial in the experiment. The large number of trials made it virtually impossible for a participant to recognize when a previous trial was being presented again. This procedure motivated participants to respond both quickly and accurately, as both aspects of performance influenced overall time to complete the experiment. The same procedure was followed for both tasks. Once participants finished each task, they were asked to describe the approach they used to complete the trials. The experimenter asked questions, when necessary, to clarify these reports. In addition, each participant was asked to step through a couple of sample trials to illustrate the solution method.

3.2. Results

A complete discussion of the verbal reports is beyond the scope of this article, but a general account of them is relevant to the current focus. Recall that the model described above was based on the verbal reports of participants in the study conducted by Gunzelmann and Anderson (2006). It is important to note that the verbal reports from participants in this study are compatible with the model that has been described. Overall, participants reported strategies that were quite similar to the one that has been instantiated in the model. In fact, every participant used "patterns," "groups," or "clusters" to describe how they organized the space to find the solution. By matching corresponding groups of objects in the two views, they were able to bring the two views into correspondence and identify the target. This general technique is identical to the strategy described by participants in previous experiments, suggesting that the general approach taken in the modeling work is appropriate for this new group of participants. It is also interesting to note that all of the participants reported a similar overall approach to both tasks, which provides initial support for concluding that participants are using the same general strategy for both tasks.

3.2.1. Errors

Participants were very accurate in performing the tasks, completing over 90% of the trials correctly across both tasks (93.1%). There was a small difference in accuracy between the task conditions (94.4% correct for the find-on-map task vs. 91.8% on the find-in-scene task), but this difference was not significant, $t(15) = 0.92$, $p > .35$. Formal analyses are not conducted on the error data due to the sparseness of the data. Still, the data do reveal interesting patterns. Table 2 shows the errors rates as a function of misalignment for each of the tasks. In both tasks, increased misalignment resulted in a higher proportion of errors. For the location of the target, errors increased as the target was located farther off to one side or the other. This was true for both tasks (Table 3. Finally, errors increased substantially in both tasks as the number of nearby distractors increased (Table 4).

Although the errors made in these tasks are not modeled here, they do point to sources of difficulty, providing important information about how participants were solving the task. The fact that nearby distractors were an important influence on accuracy indicates that local features were being used by participants to locate the target. In addition, nearly one half of the errors made by participants across both tasks (44.3% overall) involved clicking on one of the other objects in the same quadrant. This is more often than would be expected by chance

Table 2
Errors as a proportion of responses as a function of misalignment

	Errors (Proportion)	
Viewer Position (Misalignment)	Find-on-Map	Find-in-Scene
Viewer at bottom (0°)	0.007	0.019
Viewer at side (90°)	0.069	0.081
Viewer at top (180°)	0.076	0.139

Table 3
Errors as a proportion of responses as a function of the target's location relative to the viewer

	Errors (Proportion)	
Target Location Relative to the Viewer	Find-on-Map	Find-in-Scene
Nearby, directly in front	0.020	0.029
Nearby, off to side	0.052	0.084
Intermediate distance, off to side	0.063	0.104
Far away, off to side	0.065	0.084
Far away, directly in front	0.064	0.080

given the presence of nine non-target objects in the space on each trial, $\chi^2(1, N = 1097$ total errors$) = 249.57$, $p < .001$. Therefore, it seems that much of the time participants were able to locate the correct area of the space, but made their error in determining which of the objects in that cluster or quadrant was the target. This kind of error fits quite nicely with the implementation of the model described above. The pattern of errors is similar to the RT data as well ($r = .547$), which supports the conclusion that the results were not a consequence of a speed–accuracy trade-off. The RT data for this study are described next, followed by an evaluation of the model's performance relative to participants in this study.

3.2.2. RTs

Because error rates were relatively low, the RTs provide a more sensitive measure of the sources of difficulty for the tasks presented in the experiment. First, it is important to note that the order in which the tasks were performed did not have a significant effect overall, $F(1, 14) = 0.32$, $p > .5$ (mean square error [*MSE*] = 1,264.63 sec). In addition, there was no overall difference in performance between the two tasks, $F(1, 14) = 0.39$, $p > .5$ (*MSE* = 256.95 sec). The interaction between these two factors, however, speaks to the learning that was occurring as participants worked through the experiment. Participants' RTs were quite a bit longer on the task they completed first (average RT was 5.37 sec) versus the task they completed second (average RT was 4.50 sec). This is reflected in a significant interaction between the task being performed and the order in which they were performed, $F(1, 14) = 18.1$, $p < .001$ (*MSE* = 256.95 sec).

Table 4
Errors as a proportion of responses as a function of the number of nearby distractors

	Errors (Proportion)	
Number of Nearby Distractors	Find-on-Map	Find-in-Scene
Zero	0.014	0.020
One	0.054	0.065
Two	0.083	0.112
Three	0.069	0.124

For the other effects in the experiment, it is useful to consider whether they may result from some subset of the maps that were used in the study. Recall that six unique spaces were generated to include all possible configurations of quadrants relative to each other. It is possible to test the effects of the other factors in the experiment, using the maps as the participants in the analyses. In the analyses described below, statistics are presented in this way (F_m), in addition to the standard statistical results, with the data analyzed over participants (F_p). If a result is statistically significant over participants, but not over maps, it suggests that characteristics of a subset of the maps may, in fact, be responsible for the effect, rather than some general information processing characteristic of the participants. More confidence can be placed in the robustness of effects that are significant according to both analyses. In addition to presenting these different statistics, Greenhouse–Geisser corrected p values are used where applicable.

One of the primary motivations for the experiment was to evaluate the similarity in performance between the two orientation tasks presented to participants. As predicted by the model, performance was similar. As noted above, there was no overall difference in RTs for the two tasks; this was also supported by the analysis over maps, $F_m(1, 5) = 1.50$, $p > .25$ ($MSE = 6.23$). Looking more closely at the data, there appears to be a similar effect in both tasks with regard to the effect of misalignment (Figs. 5 and 6). The impact of nearby distractors is also similar (Figs. 5 and 7, as is the influence of the target's location relative to the viewer on performance (Figs. 6 and 7). In these figures, the data are averaged over left and right for

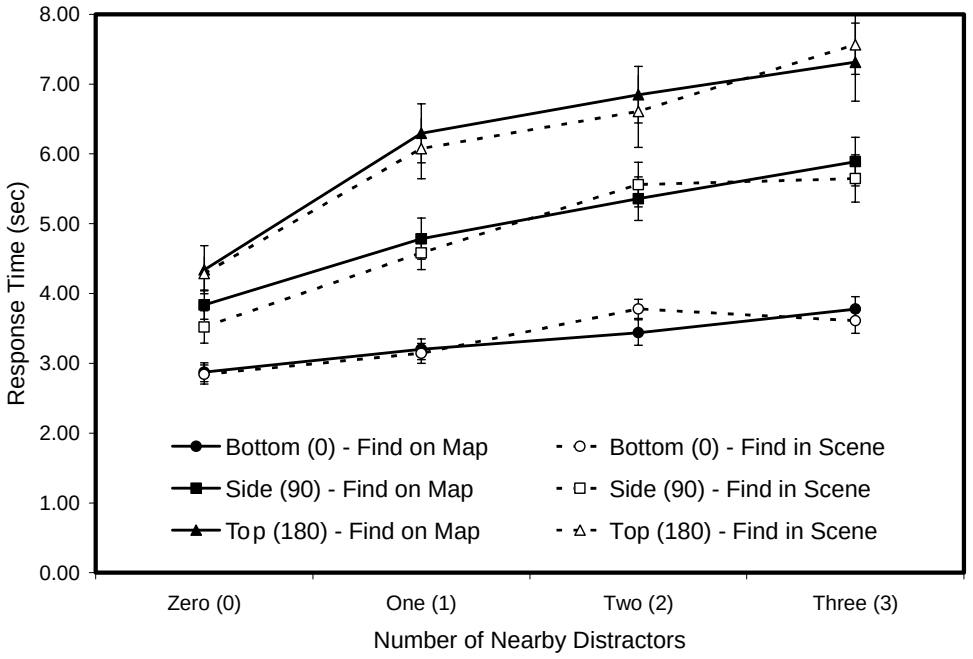

Fig. 5. Human performance data for both tasks showing the effects of misalignment and the number of nearby distractors. *Note:* Error bars for the human data show ±1 standard error.

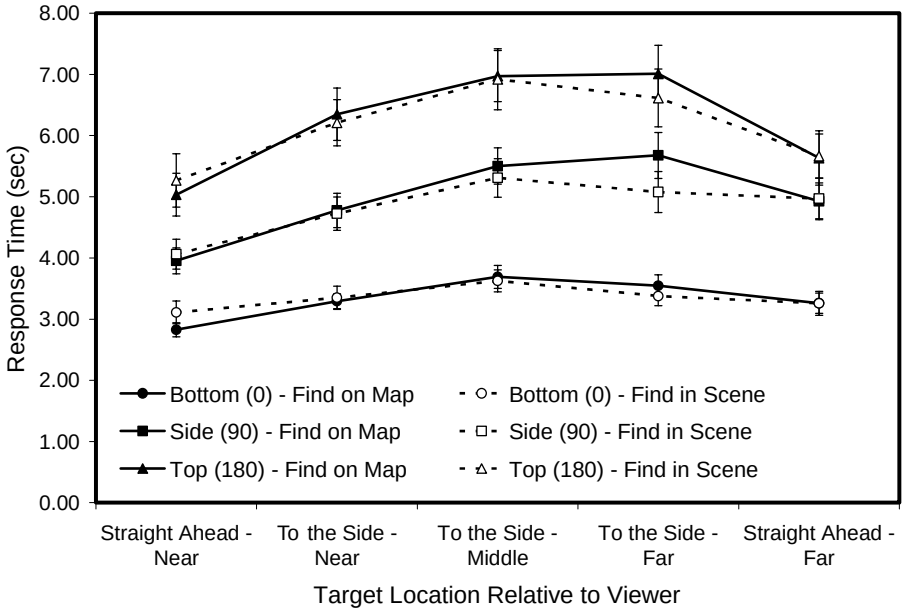

Fig. 6. Human performance data for both tasks showing the effects of misalignment and the location of the target relative to the viewer. *Note:* Error bars for the human data show ±1 standard error.

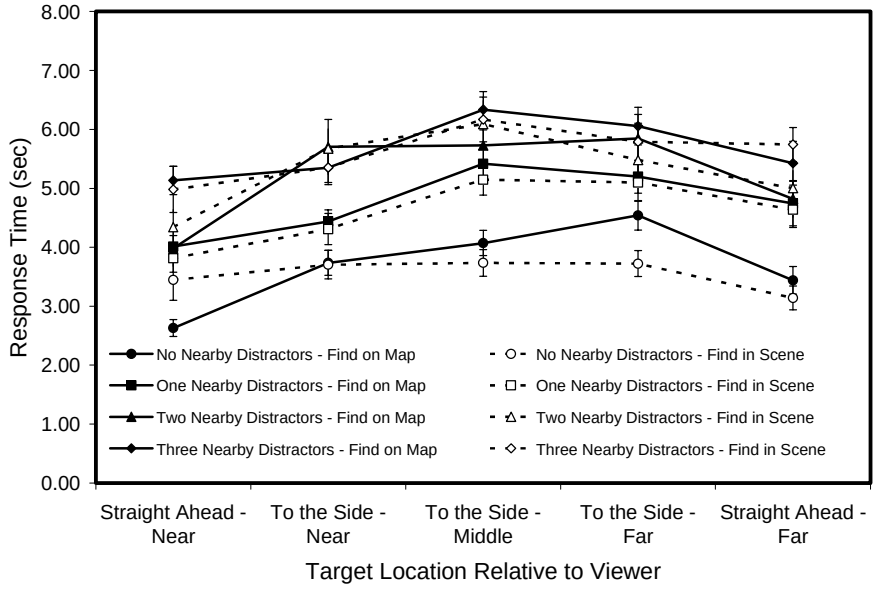

Fig. 7. Human performance data for both tasks showing the effects of the number of nearby distractors and the location of the target relative to the viewer. *Note:* Error bars for the human data show ±1 standard error.

Table 5
Summary of statistical results comparing performance between the two tasks

Analysis	df_{effect}	df_{error}	F[a]	MSE
T × M	3	42 (15)	0.34 (1.49)	45.89 (0.64)
T × L	7	98 (35)	2.45 (2.54)	15.77 (1.68)
T × D	3	42 (15)	2.20 (1.29)	15.47 (0.93)
T × M × L	21	294 (105)	0.65 (0.65)	12.63 (0.71)
T × M × D	9	126 (45)	1.24 (1.38)	11.19 (0.70)
T × L × D	21	294 (105)	1.48 (0.95)	11.41 (1.12)
T × M × L × D	63	882 (315)	1.32 (1.19)	9.80 (0.68)

Note. Values in parentheses represent results of the analyses over maps. MSE = mean square error; T = task; M = misalignment (denoted by the viewer's position on the map); L = location of the target relative to the viewer; D = nearby distractors.

[a]No effects significant at $p < .05$.

both misalignment and target location, to increase clarity and readability. The graphs indicate that each of the factors had similar effects on performance in both tasks. This is reflected in the statistical findings, which showed that none of the interactions between the task and these three factors were significant. Moreover, none of the higher level interactions involving task were significant either. This includes the three-way interactions and the four-way interaction of the task with all of these factors. Table 5 contains a summary of these statistical results. These findings show that there is little evidence from this experiment that performance on these two orientation tasks differs, either at the quantitative level of average RTs (RMSD = 0.567 sec) or in terms of the qualitative impact of the manipulated factors on performance ($r = .931$). These findings are in line with the predictions of the model described above.

Because the pattern of results is similar for both of the orientation tasks, the remaining analyses are reported by averaging the data over the two tasks. In this analysis, it would be surprising if misalignment failed to have an impact on performance given previous research in this area. As is shown in Figs. 5 and 6, this factor did impact participants in the expected direction. RTs increased as the degree of misalignment increased, $F_p(3, 42) = 64.61$, $p < .001$ ($MSE = 140.65$ sec) and $F_m(3, 15) = 240.21$, $p < .001$ ($MSE = 2.36$ sec). When the data are averaged over left or right, this effect is strongly linear, $F_p(1, 14) = 86.92$, $p < .001$ (MSE = 29.96 sec) and $F_m(1, 5) = 298.59$, $p < .001$ (MSE = 3.23 sec), with little evidence of a quadratic effect over participants, $F_p(1, 14) = 1.49$, $p > .20$ (MSE = 6.19 sec), but some indication of an effect over maps, $F_m(1, 5) = 7.218$, $p < .05$ (MSE = 0.88 sec). The inconsistent results for the quadratic effect suggest that it may not be reliable. Increasing numbers of nearby distractors also resulted in longer RTs for participants, as shown in Figs. 5 and 7, $F_p(3, 42) = 113.24$, $p < .001$ ($MSE = 43.84$) and $F_m(3, 15) = 21.12$, $p < .001$ ($MSE = 14.69$). This effect was also marked by a linear trend, $F_p(1, 14) = 152.04$, $p < .001$ ($MSE = 1{,}285.02$ sec) and $F_m(1, 5) = 55.38$, $p < .001$ ($MSE = 78.75$ sec), but also contained some evidence of a quadratic component, $F_p(1, 14) = 35.04$, $p < .001$ ($MSE = 373.32$ sec), although the effect did not reach significance in the analysis over maps, $F_m(1, 5) = 4.59$, $p = .085$ ($MSE = 63.59$ sec). Finally, the location of the target relative to the viewer impacted

participants' RTs, $F_p(7, 98) = 22.37$, $p < .001$ ($MSE = 30.92$ sec) and $F_m(7, 35) = 22.77$, $p < .001$ ($MSE = 1.90$ sec). This effect can be seen in Figs. 6 and 7.

The main effects provide evidence of how the factors were influencing performance. However, it is also the case that there were interesting interactions between the factors that modulated their influence on human performance. These interactions are presented in Figs. 5 through 7 as well. The interaction between misalignment and nearby distractors is evident in Fig. 5, which illustrates the finding that the impact of misalignment was larger when there were more nearby distractors. This interaction was significant, $F_p(9, 126) = 16.52$, $p < .001$ ($MSE = 14.91$ sec) and $F_m(9, 45) = 14.03$, $p < .001$ ($MSE = 1.01$ sec). The data in Fig. 6 illustrate the influence of misalignment and the location of the target relative to the viewer on performance. The analyses indicate that there was a significant interaction between these factors, $F_p(21, 294) = 4.45$, $p < .01$ ($MSE = 13.59$ sec) and $F_m(21, 105) = 6.04$, $p < .01$ ($MSE = 0.63$ sec). As in previous research and the model, the impact of misalignment was reduced when the target was in a cluster positioned approximately straight in front of the viewer. Finally, the interaction between the target's location and the number of nearby distractors is illustrated in Fig. 7. Although the data appear to be fairly regular, the analysis over participants provided evidence for a significant interaction, $F_p(21, 294) = 5.37$, $p < .001$ ($MSE = 12.28$ sec). However, the analysis over maps was not significant, $F_m(21, 105) = 2.01$, $p > .14$ ($MSE = 2.05$ sec), suggesting that the effect observed in the analysis over participants may be a spurious result. Note that these results are consistent across tasks, even though Fig. 7 may give the impression of some disparity between the two tasks (see Table 5, T × L × D interaction). Finally, the three-way interaction of these factors was not significant in either analysis, $F_p(63, 882) = 1.25$, $p > .25$ ($MSE = 10.70$ sec) and $F_m(63, 315) = 1.09$, $p > .30$ ($MSE = 0.76$ sec).

The results just described address one of the motivations for the experiment, which was to replicate the findings from Gunzelmann and Anderson (2006), using a purely within-subjects design. Success on this goal can be evaluated directly by comparing performance of participants in the earlier work to the performance of the current participants on the find-on-map task. The pattern of results was quite similar for this comparison ($r = .915$). The most important difference between the two datasets was that participants in this study took substantially longer to respond, on average, than participants in the previous work (5.00 sec in this experiment vs. 3.77 sec in Experiment 2 of Gunzelmann & Anderson, 2006). In fact, this difference in average RT was significant, $F(1, 24) = 9.27$, $p < .01$ ($MSE = 128.01$ sec). This probably reflects differences in the populations from which the participants were recruited for the two different studies.[4] This is addressed further in the conclusion. Despite the large quantitative differences, however, the high correlation of the data from the two experiments indicates that the factors that were manipulated had similar impacts on performance for both groups. In fact, none of the interactions were significant between these two datasets ($p > .10$ for all interactions).

3.3. Discussion

The model described above provides a means for understanding human performance on spatial orientation tasks requiring that multiple views of a space be brought into correspondence.

The original model was developed to perform the find-on-map task, and was then generalized to perform the find-in-scene task. Despite differences in the procedures required to solve these two tasks, the model predicted that performance on them should be quite similar. This prediction was borne out by the experimental results (compare Figs. 2–4 with Figs. 5–7).

The experiment provides a within-subjects validation of previous results. In this study, all of the participants completed all of the possible trials for both tasks. The data provide encouraging support for the pattern of results predicted by the model. In addition, the population from which the participants from this study were selected differs greatly from the population involved in the original research. The similarity in the pattern of results provides evidence that the factors shown to influence difficulty are general influences on human spatial performance, which are not limited to a particular subgroup. In the next section, the model's performance is compared in detail to the human data from this study.

4. Model performance

A close comparison of Figs. 2 through 4 against Figs. 5 through 7 produces two conclusions. First, the pattern of data is similar for the model and for the participants in this experiment. The qualitative similarity between them is quite good: $r = .92$. The other obvious conclusion is that the model is much faster than participants at completing the task. The average RT for the model was 3.64 sec, whereas it was 4.94 sec for the human participants, across both tasks. As a result, the quantitative fit is less impressive (RMSD = 1.45 sec). This discrepancy can be viewed as a consequence of estimating the spatial updating and translation parameters using the empirical results from Gunzelmann and Anderson (2006). Those participants responded much more quickly, on average, than participants in this study. Thus, the model's predictions fail to match these data at a quantitative level. As mentioned above, it may be that differences in the populations from which the samples were drawn may be responsible for this overall difference in performance. Differences in spatial ability, familiarity with the kinds of 3-D virtual environments portrayed in this experiment, or both may be contributing to those performance differences. These factors can be seen to relate, respectively, to the spatial updating parameter and the parameter associated with the transitions between the 2-D and 3-D perspectives, which are influential in determining the model's RT on any given trial. It may be that the spatial updating parameter can be associated, to an extent, with familiarity with the virtual environments used here as well, not just overall spatial ability. For instance, note that practice with one of these tasks resulted in a large speed-up on the second task, indicating that practice and experience are important contributors to performance.

If the spatial parameters in the model are allowed to vary to account for differences in ability or practice between the participants in the two studies, then the model's performance can come much closer to the performance of the individuals in the experiment described here. The performance of the model with revised parameter estimates is shown in Figs. 8 through 10. These data are based on the model using 0.9 sec as the spatial updating parameter and 1.00 sec for the 2-D/3-D transition parameter. These are changed from the values of 0.6 sec and 0.25 sec, respectively, which were used to account for the data from the earlier work. Not surprisingly, the correlation between the data from the model and human performance

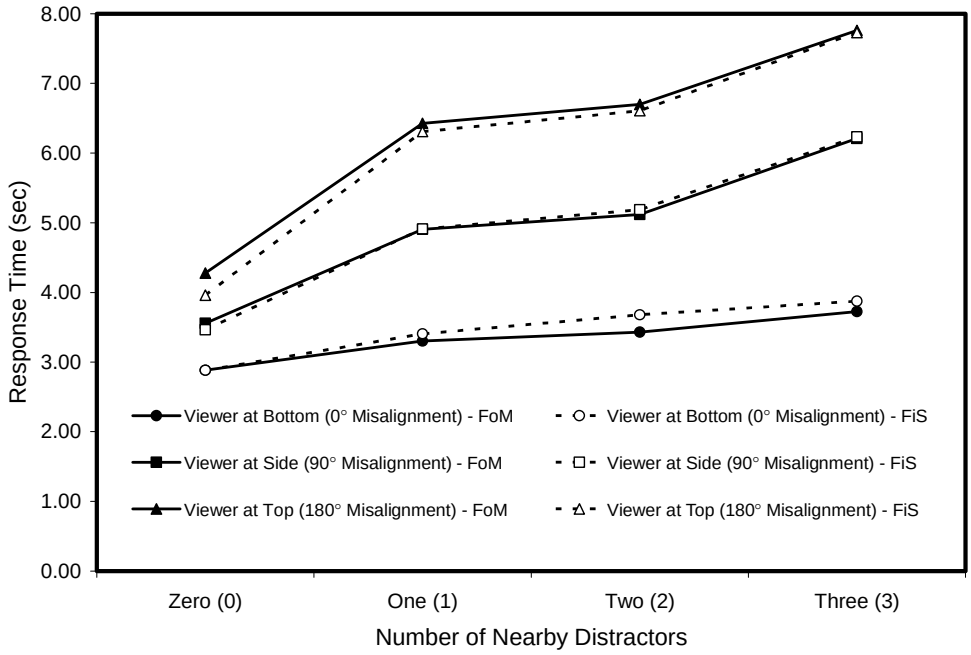

Fig. 8. Model performance, based on revised parameter values, for both tasks as a function of misalignment and the number of nearby distractors.

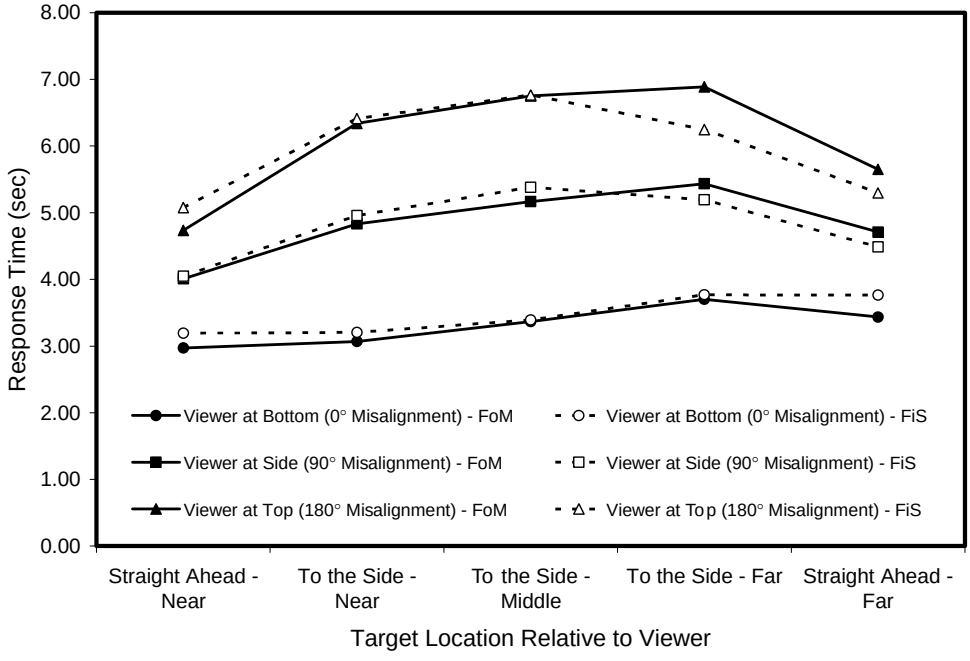

Fig. 9. Model performance, based on revised parameter values, for both tasks as a function of misalignment and the location of the target relative to the viewer.

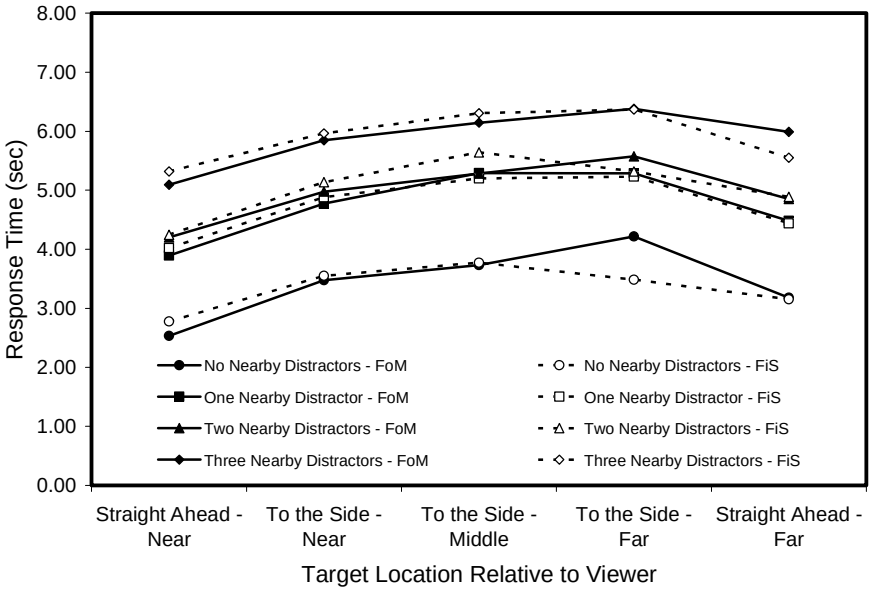

Fig. 10. Model performance, based on revised parameter values, for both tasks as a function of the number of nearby distractors and the location of the target relative to the viewer.

remains high ($r = .93$). In addition, with the revised parameter estimates, the model makes good quantitative predictions about performance as well. The RMSD between the model data shown in Figs. 8 through 10 and the human data in Figs. 5 through 7 is 0.55 sec. The average RT for the model with these new parameter values is 4.88 sec, which is more in line with the human data in this study. Thus, allowing for different parameter values to reflect the substantially different overall performance in the two studies, the model provides a very good account of human performance on these tasks. In fact, the parameters that were varied provide some clues about the source of individual differences in this task. This topic is addressed in the conclusion.

5. Conclusion

The research described in this article explores human performance on tasks involving spatial orientation with maps. The results support past research, using tasks similar to those presented here (Gunzelmann & Anderson, 2006), as well as studies using a wide array of variations on the general theme (e.g., Aginsky et al., 1997; Boer, 1991; Dogu & Erkip, 2000; Gugerty & Brooks, 2004; Gunzelmann & Anderson, 2006; Gunzelmann et al., 2004; Hintzman et al., 1981; Malinowski, 2001; Malinowski & Gillespie, 2001; Murakoshi & Kawai, 2000; Richardson et al., 1999; Rieser, 1989; Wraga et al., 2000). The model captures all of the trends in the data, providing evidence to support the account of performance that it embodies. Additional support for the model comes from the verbal reports of participants from the

original study. The strategy that the model uses for the find-on-map task is based upon the verbal reports from Gunzelmann and Anderson (2006), and the verbal reports for this study were similar. The strategy for the find-in-scene task is adapted from the original task, with the principal changes being a reordering of the steps (Table 1), with a corresponding change to the representations maintained and manipulated in the solution process. This model was applied in an a priori manner to the experimental paradigm used here, generating the predictions shown in Figs. 2 through 4. Human participants generated data that matched the trends predicted by the model, despite the finding that they were significantly slower than the model predicted (and also slower than the participants in the original study).

The predictions generated by the model represent a case of near transfer in producing *a priori* predictions using a computational cognitive model. Despite the differences in the processing that is required in the two situations, there are significant similarities in the demands for the two tasks. Still, it is important to note that even near transfer explorations of the generalizability of computational cognitive models are rarely done. In addition, the extension of the model was achieved using a principled adaptation of a strategy based on general principles of human spatial ability. Thus, this general approach can be adapted to provide a means of making performance predictions on other spatial tasks. This will provide a fruitful avenue for future research into understanding how spatial abilities are brought to bear in a range of tasks.

Besides illustrating the potential to generalize across tasks, the ACT–R model provides a foundation for understanding individual differences in performance on spatial tasks. The empirical results demonstrate that the participants in this study and in Gunzelmann and Anderson (2006) were affected by the same factors. Thus, despite being faster, the participants in the earlier study were still impacted by misalignment, nearby distractors, and the location of the target relative to the viewer. In addition, the pattern of results was quite similar in the two experiments ($r = .915$), indicating that the same factors were influencing performance in much the same way. In the model, the overall difference in performance was captured by varying parameters associated with performing spatial operations. Note that this explanation supports the empirical findings by suggesting that the same general strategy was being applied by both groups. Manipulating the spatial parameters in the model does not impact its qualitative performance, which is a function of the solution strategy. Not surprisingly, this account suggests that spatial ability and familiarity with the types of virtual environments portrayed should be important influences on people's ability to perform this task rapidly. The particular values that were used for the spatial parameters in the model appear to reflect a level of proficiency in these contexts. The point is made clearly in the data from the two studies because the results are substantially similar other than the discrepancy in overall RTs.

The ability of the model to capture the behavior of a different sample of participants, drawn from a different population, suggests common predispositions for how to process spatial information across individuals and across tasks. Central to the account developed here is the use of hierarchical encoding in the model, a tendency for which there is substantial evidence in the experimental literature (Hirtle & Jonides, 1985; McNamara, 1986; McNamara, Hardy, & Hirtle 1989; Stevens & Coupe, 1978). However, the operations that are performed on those representations, as well as the sources of difficulty that impact performance, also appear to be similar across individuals and tasks. The model provides an explanation for the similarities, as well as a way to conceptualize individual differences that were observed. The foundation on

which the general strategy is based provides a means for understanding human performance on other spatial tasks. Future research will be directed at extending the general mechanisms of this model to additional tasks. This process is already underway (see Gunzelmann & Lyon, 2006). Consideration of an increasingly broad range of tasks will contribute to refinement of the representations and processes in the model, resulting in a more complete account of human spatial competence.

Notes

1. Note that a second point also provides information about scale by representing the distance between two points in both reference frames.
2. Target locations will tend toward the center of the quadrant where it resides, such that 45° rotations produce eight approximate target locations relative to the viewer (4 quadrant center points in each of 2 possible alignments of the quadrants –, +, and ×)
3. This value was defined by the number of objects in the same quadrant as the target. Although the random placement of objects into each quadrant meant that the target could be closer to an object in a neighboring quadrant, this was rare; and the creation of multiple maps was partially intended to offset such random effects.
4. The participants in the experiment described in Gunzelmann and Anderson (2006) were recruited from Carnegie Mellon University, a population notable for their generally high aptitude at spatial tasks and perhaps also for their propensity to play computer games like *Unreal Tournament*, which was used to generate the stimuli for this research.

Acknowledgments

The author was supported as a National Research Council Postdoctoral Research Associate while this research was conducted. Additional support came from the Air Force Office of Scientific Research Grant #02HE01COR. Portions of this research were presented at the 25th annual meeting of the Cognitive Science Society (2004) and at the 7th international conference on Cognitive Modeling (2006). Particular thanks are in order for Sharon Lebeau for her help with data collection. The original model for the data in Gunzelmann and Anderson (2006) was developed while the author was a graduate student at Carnegie Mellon University. Thanks also to John R. Anderson, Jerry Ball, Kevin Gluck, and Don Lyon for their helpful comments on this work.

References

Aginsky, V., Harris, C., Rensink, R., & Beusmans, J. (1997). Two strategies for learning a route in a driving simulator. *Journal of Environmental Psychology, 17*, 317–331.

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.

Best, B. J., & Gunzelmann, G. (2005, July). *Hierarchical grouping in ACT–R.* Paper presented at the 12th annual ACT–R workshop, Trieste, Italy.

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*, 12–23.

Boer, L. C. (1991). Mental rotation in perspective problems. *Acta Psychologica, 76*, 1–9.

Dogu, U., & Erkip, F. (2000). Spatial factors affecting wayfinding and orientation: A case study in a shopping mall. *Environment and Behavior, 32*, 731–755.

Gugerty, L., & Brooks, J. (2004). Reference-frame misalignment and cardinal direction judgments: Group differences and strategies. *Journal of Experimental Psychology: Applied, 10*, 75–88.

Gunzelmann, G., & Anderson, J. R. (2006). Location matters: Why target location impacts performance in orientation tasks. *Memory & Cognition, 34*, 41–59.

Gunzelmann, G., Anderson, J. R., & Douglass, S. (2004). Orientation tasks with multiple views of space: Strategies and performance. *Spatial Cognition and Computation, 4*, 207–253.

Gunzelmann, G., & Lyon, D. R. (2006). Qualitative and quantitative reasoning and instance-based learning in spatial orientation. In R. Sun & N. Miyake (Eds.), *Proceedings of the 28th annual meeting of the Cognitive Science Society* (pp. 303–308). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

Hirtle, S. C., & Jonides, J. (1985). Evidence of hierarchies in cognitive maps. *Memory & Cognition, 13*, 208–217.

Levine, M. (1982). You-are-here maps: Psychological considerations. *Environment & Behavior, 14*, 221–237.

Levine, M., Jankovic, I. N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology: General, 111*, 157–175.

Levine, M., Marchon, I., & Hanley, G. L. (1984). The placement and misplacement of you-are-here maps. *Environment & Behavior, 16*, 139–157.

Malinowski, J. C. (2001). Mental rotation and real-world wayfinding. *Perceptual and Motor Skills, 92*, 19–30.

Malinowski, J. C., & Gillespie, W. T. (2001). Individual differences in performance on a large-scale, real-world wayfinding task. *Journal of Environmental Psychology, 21*, 73–82.

McNamara, T. (1986). Mental representations of spatial relations. *Cognitive Psychology, 18*, 87–121.

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

Murakoshi, S., & Kawai, M. (2000). Use of knowledge and heuristics for wayfinding in an artificial environment. *Environment and Behavior, 32*, 756–774.

Richardson, A., Montello, D., & Hegarty, M. (1999). Spatial knowledge acquisition from maps, and from navigation in real and virtual environments. *Memory & Cognition, 27*, 741–750.

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

Shepard, R., & Hurwitz, S. (1984). Upward direction, mental rotation, and discrimination of left and right turns in maps. *Cognition, 18*, 161–193.

Shepard, R., & Metzler, J. (1971). Mental rotation of three dimensional objects. *Science, 171*, 701–703.

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

Tarr, M. J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*, 233–282.

Unreal Tournament. (2001). [Computer software]. Raleigh, NC: Epic Games, Inc.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations: A group test of three-dimensional spatial visualization. *Perceptual & Motor Skills, 47*, 599–604.

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.