SPATIAL COGNITION AND COMPUTATION, 4(3), 207–253

Orientation Tasks with Multiple Views of Space: Strategies and Performance

Glenn Gunzelmann
National Research Council
Air Force Research Laboratory

John R. Anderson
Carnegie Mellon University

Scott Douglass
Carnegie Mellon University

Two experiments examine how participants vary in their approach to solving an orientation task. Verbal reports from untrained participants in a pilot study revealed that some participants used a strategy based on mental imagery, while others used verbal descriptions to do the task. The two experiments presented here involved training participants to perform the orientation task using one of these strategies. Participants' performance, measured by response time and eye movements, differed as a function of strategy. An ACT-R model of the task that uses the strategies provides a validation of the proposed mechanisms, producing a close fit to both the response time and eye movement data. The model's success is achieved, in part, by performing all aspects of the task, from processing the information on the screen to making responses. Overall, the results indicate that strategic variability is an important feature of human performance on such tasks.

Keywords: Orientation, spatial reasoning, strategy, computational model, ACT-R, eye movements.

Correspondence concerning this article should be addressed to Glenn Gunzelmann, Air Force Research Laboratory, 6030 South Kent Street, Mesa, AZ 85212-6061; email: glenn.gunzelmann@mesa.afmc.af.mil.

Maintaining one's orientation within the environment is a basic component of successful navigation. In general, this involves integrating egocentric visual signals with spatial knowledge about the environment. This knowledge can be represented in a variety of forms. For instance, it may be based on information stored in memory or it may be contained in a physical map. Maps use stable, external identifiers to define the origin and the orientation, and are referred to as allocentric or exocentric representations (Klatzky, 1998; Tversky, 2000). If such a representation of space is oriented differently than the egocentric visual scene, then one must bring them into alignment somehow to make informed navigational decisions. An everyday application of this process involves determining which way to turn at an intersection by using a map. The visual scene presents an egocentrically-based representation of the space, while the map provides an allocentric representation. When the viewer is oriented differently from the map as the intersection is approached, some sort of transformation (e.g., rotation) must be done to bring the two perspectives into alignment. Of course, with a physical map it may be possible to actually rotate it to align it with the egocentric orientation. In other situations (e.g., cognitive maps or computer displays), mental transformations are needed to coordinate the views to make accurate decisions.

On a continuum of reasoning about orientation within a space, deciding whether to turn left or right at an intersection is a fairly straightforward task. Still, research has shown that this kind of judgment becomes more difficult as a function of the difference in orientations (misalignment) between the two views of space (Shepard and Hurwitz, 1984). The phenomenon bears a strong resemblance to findings in the mental rotation literature where the time needed to determine that two objects are identical increases linearly as a function of the angular disparity between them (e.g., Shepard and Metzler, 1971). This similarity has been used to support the conclusion that performance in orientation tasks involves mental imagery and rotation. However, orientation tasks add a layer of complexity to the mental rotation process since the information is generally presented in two different formats (egocentric and allocentric). Thus, determining exactly how the visual scene corresponds to the information on the map may require additional reasoning beyond the image rotation.

Experiments have been conducted using numerous orientation tasks in a variety of different contexts. For instance, Hintzman, O'Dell, and Arndt (1981) did a series of 14 experiments in which they investigated human performance in orientation tasks of various sorts. These ranged from simple 2-dimensional displays to room-sized environments populated with a number of objects. Performance across these situations was similar in many ways, although there were some differences. Beyond this one series of studies, there have been investigations of human orientation on college campuses (Abu-Obeid, 1998; Kirasic, Allen, and Siegel, 1984; McNamara and Diwadkar, 1997), in urban environments (Boer, 1991; Dogu and Erkip, 2000; Glicksohn, 1994), and in various sorts of virtual environments (Richardson, Montello, and Hegarty, 1999;

Rossano, West, Robertson, Wayne, and Chase, 1999). These studies have addressed a number of issues relating to learning and performance in orientation. They have produced a number of consistent results, which has led researchers to emphasize imagery-based processes in their theoretical accounts.

There is considerable research to suggest that the cognitive representations of large-size spaces differ depending on how the information was learned (Abu-Obeid, 1998; Klatzky, Loomis, Beall, Chance, and Golledge, 1998; Richardson, et al., 1999; Rossano et al., 1999; Taylor, Naylor, and Chechile, 1999), but this is only a part of the issue. In addition to understanding how spatial information is represented, it is important to examine the processes we use to manipulate representations of spatial information to bring them into correspondence with our current view of the world. Shepard and Hurwitz (1984) found that the time needed to determine whether a turn (presented on a computer as a pair of line segments) was left or right increased as a function of how far the initial direction of travel deviated from straight ahead ("up" on the monitor). The findings were similar when participants had to make comparable judgments in a more realistic simulation program (Gugerty, deBoom, Jenkins, and Morley, 2000). Boer (1991) found a pattern of data similar to Hintzman et al. (1981) in response times for pointing to cities (relative to one's current location). In an even simpler case, identifying directions on a compass appears to require an analogous kind of reasoning, since the pattern of data is similar (Loftus, 1978).

As we learn more about coordinating different perspectives, it becomes clear that there are multiple ways to do it. The most well established strategic difference in orientation tasks is a distinction between array rotation and viewer rotation (Huttenlocher and Presson, 1979; Presson, 1982; Wraga, Creem, and Proffitt, 1999, 2000). In an orientation task, array rotation involves mentally rotating the map or the objects in the visual scene to imagine how the objects would be arranged after the transformation. In contrast, viewer rotation involves imagining the movement of oneself to a new location in the space, and then determining where the objects would be relative to the new vantage point. Curiously, although the actual computational demands of the two strategies are equivalent, they produce different results depending on both the complexity of the array (i.e., the map or physical display; Presson, 1982; Wraga, et al., 2000) as well as upon the question that is to be answered (Huttenlocher and Presson, 1979; Presson, 1982). Although viewer rotation has been found to be easier for participants in most situations, both processes involve mental rotation.

The potential for multiple strategies cuts across all levels of spatial information processing. Just and Carpenter (1985) described a number of alternative strategies for performing a mental rotation task. One of these strategies involves a piecemeal rotation of the figure, a strategy that bears some resemblance to the performance of the cognitive model in the research described below. Fenner, Heathcote and Jerrams-Smith (2000) measured the visuo-spatial and verbal ability of children (age 5–6 and 9–10) and their performance on a wayfinding task. The results indicated that children high in visuo-spatial ability outperformed children low in visuo-spatial ability, but that this effect diminished with age. While they discussed the findings in terms of representational differences between young and old children, it seems reasonable to suspect that children who are low in visuo-spatial ability would develop alternative strategies for wayfinding that would allow them to perform as well as other children. This idea is supported by the findings of Aginsky, Harris, Rensink, and Beusmans (1997) who found that participants in a driving simulator learned a route either using a "visually-dominated" or a "spatially-dominated" strategy. While performance on the route-driving task was equivalent between the two strategy groups, differences emerged on other tests like map-drawing and identifying transformations to buildings on the course. Lastly, Murakoshi and Kawai (2000) found that some participants who were unable to perform tasks indicative of their knowledge of a space (e.g. draw a map, point to locations) were nonetheless able to successfully return to their starting point after being led on a route through the space. The verbal reports of these individuals indicated that the strategies they used to find their way back were less spatial in nature.

Clearly, multiple mechanisms exist for determining spatial orientation, but theorists have generally not emphasized this potential for variation. As mentioned above, several theorists have explained performance in orientation tasks by relating their findings to mental imagery and rotation (e.g., Hintzman, et al., 1981; Shepard and Hurwitz, 1984). Others have attempted to provide accounts of human navigation as a whole (Gopal, Klatzky, and Smith, 1989; O'Neill, 1991). Most efforts have not incorporated strategic variation in the orientation process, but even those accounts that have tend to focus solely on imagery-based mechanisms (Huttenlocher and Presson, 1979; Presson, 1982; Wraga, et al., 1999; 2000). This is true even though the existence of verbal strategies for tasks like mental rotation has been noted in the literature previously (Bethell-Fox and Shepard, 1988).

The research presented here attempts to provide a more detailed account of the orientation process in humans. The orientation task we used required participants to coordinate information presented in two different views of a space (one egocentrically-oriented and one allocentrically oriented; see Figure 1). The empirical results replicate the phenomena found by previous researchers (e.g., Boer, 1991; Gugerty, et al., 2000; Hintzman, et al., 1981), who have found that the degree of misalignment between the two views of the space has a strong impact on difficulty. In addition, our results produce the "M-shaped profiles" for the impact of the target's position that is common in studies of spatial orientation, with responses being fastest when the target was in line with the viewpoint (this is illustrated with our data below).

Besides replicating the basic pattern of results, this research provides evidence concerning the influence of different strategies on performance. In a pilot study involving untrained participants, some reported using mental rotation, but others reported using a verbal-description strategy. While a similar division in strategic approaches has been noted in other spatial tasks (Bethell-Fox and Shepard, 1988), it has not been examined in detail. In this case, differences in performance are predicted between the two strategies as a

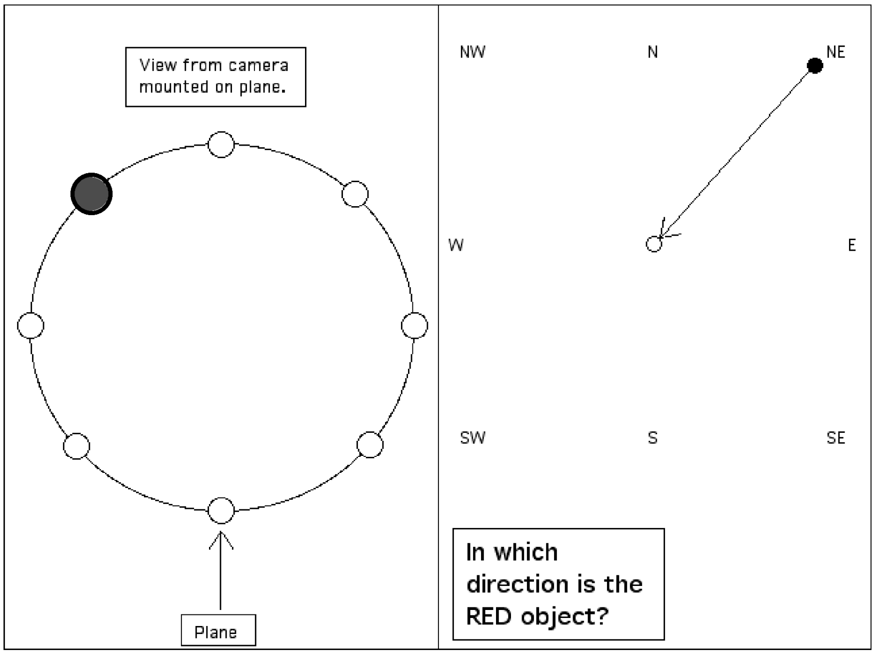

Figure 1. Sample trial for the orientation task used in the experiments presented here.

function of different manipulations in the task. The experiments presented here involve training participants to use one of the two strategies for the task to closely examine the differences between them, including eye movements in Experiment 2. Then, an ACT-R model of the two strategies is presented. This model serves two purposes. First, by constructing a model that includes perceptual, cognitive, and motor components, we demonstrate that the mechanisms of the proposed strategies are sufficient to explain the observed phenomena, including the eye movement data. Second, the model represents an attempt to address human performance in spatial tasks using a cognitive architecture, with general mechanisms that have been used to explain human performance in a variety of domains (see Anderson and Lebeire, 1998 for a review). This provides an opportunity to evaluate those mechanisms in the context of a domain that has not really been addressed by such general models of cognition in the past.

Figure 1 shows the orientation task used in the research presented here. In this figure, the left side represents the target field as viewed from a camera (attached to a plane flying above the area around the target field) and the darkened circle indicates the target. The right side shows an allocentric map with the target field at the center. The arrow on that side illustrates the camera's orientation for viewing the target field. Participants were asked to indicate the portion of the target field where the target is located (i.e., in which cardinal direction the target is located relative to the center of the target field). In the sample trial in Figure 1, the target is located in the southern portion of the target field. There are several typical findings for tasks of this sort (Boer, 1991; Gugerty, et al., 2000; Hintzman, et al., 1981). First, decisions for targets in line with the assumed orientation (i.e., targets that are horizontally in the center of the camera view) are made more rapidly. Second, response times for other targets increase as they occupy locations that are farther away from the bottom on the left or the right. These two results produce the "M-shaped profile" mentioned above. Finally, response times increase as the degree of misalignment between the two views increases. This ranges from 0 to 180 degrees, with 0 degrees being aligned (the plane is located in the south with the camera facing north) and 180 degrees being maximally misaligned (the plane is located in the north with the camera facing south).

The task in Figure 1 is loosely based on a simulator used for training unmanned air vehicle (UAV) pilots (see Gugerty, et al., 2000). This training system provides an egocentric view from a camera mounted on the UAV, as well as a map view that indicates the plane's location using GPS technology. One major simplification was to use a 2-D representation for the camera view. This representation made for a more straightforward relationship between the two views. In addition, this made our task similar in complexity to the one used by Hintzman et al. (1981; their Experiment 1), although there are some differences between them. Given the similarity, we expected that performance on our task should depend on many of the same cognitive processes as those used by the participants tested by Hintzman and his colleagues. Consequently, we expected this research to replicate the empirical findings they reported. They conclude that participants solve the task using mental imagery and rotation. However, there is evidence that mental rotation is difficult and that people often attempt to avoid it if possible. For instance, Kirsh and Maglio (1994) showed that individuals chose not to mentally rotate Tetris figures when they were able to actually rotate them on the screen. This suggests that individuals may choose other strategies when they are available. By gathering verbal reports from untrained participants in a pilot study, we were able to identify two distinct strategies that participants used to solve this task, one of which did not involve mental imagery or rotation.

Pilot Study

In a pilot study, 16 untrained participants completed the orientation task shown in Figure 1. After they finished, they were questioned about the methods they used to solve the task. These verbal reports provided two key pieces of information regarding the strategies that participants used. Firstly, participants reported that there were instances where they found the task particularly easy. Specifically, when the target was in line with the plane (the bottommost and topmost objects on the camera view), participants reported that they were able to quickly determine the answer. When the target was located at the bottom of the camera view, the correct response was the cardinal direction from which the plane was viewing the target field. If the target was located at the top of the camera view, the correct answer was the cardinal direction directly opposite the plane's location. Participants reported using quick verbal heuristics like "where I am" or "directly across from me" to complete these trials. It is interesting to note that this special-case strategy is not imagery-based, but uses a simple verbal description to indicate the target's position.

Secondly, participants reported two distinct strategies for the remaining problems. Most of the participants (12) reported using mental rotation to complete the task. Though there was some variability in the exact nature of this transformation, all of the reports had the same character. In general, the participants reported rotating the relative positions of the camera (plane) and the target on the camera view to align them with the plane's location on the map view. This can be thought of as forming an angle that connects the plane to the target on the camera view, with the vertex at the center of the target field (a 135 degree angle in Figure 1). Then, this angle can be mentally moved to the map view and rotated to align the leg of the angle corresponding to the plane with the plane's location on the map (a rotation of 135 degrees counterclockwise in the trial shown in Figure 1). At this point, the answer can be determined by finding the target end of the angle.

The other strategy reported by participants avoids the need to mentally rotate the information presented on the screen. Four participants reported using a "counting" strategy to complete the task. For this strategy, participants counted the objects from the plane's location to the target position (3 in Figure 1) and noted the direction in which the target was located (on the left in Figure 1). Then, by counting the same number of steps (i.e., cardinal direction indicators) around the map view in the appropriate direction from the plane, participants were able to accurately complete the trials. This strategy does not involve mental imagery, and yet it is highly effective for doing the task. Performance in the pilot study did not differ significantly between participants who reported using the rotation strategy and those who reported using the counting strategy. It is likely that participants were using a mixture of strategies during the experiment and only reported one at the end. On the other hand, it might mean that the two strategies actually result in equivalent performance. The experiments presented below address this issue. There is, however, clearly some validity to the verbal reports, since the special-case strategy that was reported corresponded to unusually fast response times for those trials for every participant. In any case, a critical discovery from this pilot study was the counting strategy, an alternative that has not been reported in the literature. It indicates that imagery-based mechanisms are not necessary for solving this task.

Experiment 1

This experiment involves training participants to use one of the two strategies described above. We hope this will cause participants to adopt the instructed strategy from the start and use it throughout the experiment to allow us to investigate the performance differences between the strategies. This experiment will also give us further insight into why participants used the special case strategy when the target was in line with the plane. In the pilot study, it may have been the case that participants discovered this strategy while they were unsure of how to do the task. However, in this experiment participants began the task with knowledge of a strategy that was effective for all of the trials. If participants still used the special-case strategy, it suggests that it is a readily available shortcut that participants will apply even if they are instructed otherwise.

The counting and rotation strategies make different predictions about how and why misalignment (i.e., the plane's position) and the target's position relative to the plane should affect response times. In the counting strategy, counting is done from the location of the plane to the location of the target. The extent of this counting depends on the target's location relative to the plane, so response times should increase linearly as the target is positioned farther from the plane on the left or right. In contrast, the plane's position defines where counting begins on the map view. This is a potential source of difficulty for participants if they encode the target's location as *left* or *right* relative to the plane. This is because left and right of the plane do not always correspond to left and right on the map. In fact, when the plane is located in northerly positions (NW, N, and NE), left and right are actually reversed, meaning that when the target is located on the left side of the camera view, on the map this initially involves counting to the right. To execute this strategy correctly, extra cognitive steps would be required in these situations. Therefore, some effect of misalignment should appear in this strategy.

The rotation strategy leads to somewhat different predictions. With the rotation strategy, the degree of misalignment between the two views determines how much mental rotation of the angle needs to be done. Therefore, response times in the rotation strategy should increase linearly as a function of misalignment. The potential for left-right confusion arises for the impact of the target's location in the camera view. This possibility is discussed in more detail below, after the results of Experiment 1 are presented. Briefly, this confusion can arise during rotation, depending on the target's location, if the image that is rotated is not a coherent angle. Note that these predictions are opposite those from the counting strategy.

Based on the data from the pilot study, overall performance for the two strategies is likely to be similar. However, the predictions just described suggest that there should be interactions between the strategies for both the target's location and the degree of misalignment. Finally, if participants in this experiment still use the special case strategy mentioned above, then those trials should follow a different pattern. Namely, response times should be faster than the strategies predict and there should be little effect of misalignment, regardless of which strategy they were trained to use.

Method

Participants. The participants were 32 Carnegie Mellon undergraduates enrolled in an introductory level psychology course. There were 19 males and 13 females in the study (mean age 19.8 years). All were given course credit for their participation in the one-hour study.

Materials. The experiment and all instructions were presented on a Macintosh computer. Each trial consisted of a camera view and a map view (see Figure 1), and participants made their responses using the number pad on the right portion of the keyboard, which was spatially mapped to the cardinal directions on the map view of the space. So, if the correct response was "South" (as it is in the sample trial shown in Figure 1), participants responded by pressing the "2" on the number pad. In each trial, a single item was highlighted on the camera view and participants were asked to indicate in which portion of the target field that item was located given the orientation on the map view (see Figure 1). Also, the experiment incorporated a dropout procedure, which was explained to participants before they began. If a participant made an error on any of the trials during a block, it was repeated later in that block until he or she got it correct. Finally, participants completed 4 blocks of 64 trials (in random order) in the experiment. The 64 trials comprise all of the possible trials within the task (8 target locations crossed with 8 plane locations).

Procedure. Participants read the instructions for the task on the computer, and then they were given paper-based training on one of the two strategies described above (half of the participants were trained to use each strategy). The training consisted of (1) an illustrated description of the strategy and the steps involved in implementing it, (2) a detailed description of how the strategy applied to a sample problem, and (3) a set of 16 practice problems (see the Appendix for the steps participants were instructed to use to implement the strategies). Note that the instructions encourage participants to use clockwise and counterclockwise instead of left and right to solve the task. Since these terms do not need to be updated regardless of where the critical items are in the trial, using them would eliminate the left-right confusion issue described above. Sohn and Carlson (2003) have shown how left-right conflict can impact response times. In addition, the tendency to encode the locations of objects in space using left and right fits with research by Sholl (1987; Easton and Sholl, 1995) and Tversky (2003) who have explored how individuals encode information as a function of the body's natural axes (top/bottom, front/back, and left/right). Their research suggests that participants tend to use left and right to encode locational information and that left-right confusion is likely to appear when they do.

After reading the strategy description and its application to the sample problem, the participants completed the 16 paper-based practice problems by explicitly labeling them according to the strategy they were taught. In the counting strategy, this involved writing in the appropriate numbers on both the camera view and the map view. In the rotation strategy, this involved drawing in the appropriate angles on both views. The experimenter gave participants feedback on each of the practice trials and helped them to correct mistakes if needed. By the end of training, all participants reported understanding the strategies and were proficient at solving the practice problems correctly. Once the training was completed, participants completed four blocks of trials on the computer. After finishing, they were asked if they had used the strategy they were taught and if there were any instances where they did something differently. All participants reported using the strategy they were taught, and also reported using the special-case strategy reported by participants in the pilot study.

Results

Errors. The error data are presented in Figure 2. Overall accuracy was quite high (95%), but Figure 2 shows that participants using both strategies tended to make more errors when the target was not in line with the viewpoint and as the misalignment between the two views of the space increased. The exception to this was when the two views were maximally misaligned, where there were actually relatively few errors made by participants. Although this last trend does not match the response time data presented below, previous research has documented this result in terms of accuracy (Gugerty, et al., 2000; Hintzman, et al., 1981). Importantly, it should be noted that the pattern of errors was similar on the whole to the response time data (r = .77), indicating that the results were not due to a speed-accuracy trade-off.

Response Times. The data for the two strategy conditions are presented in Figure 3. Overall mean response times were similar, although response times for participants using the rotation strategy were a little longer than for participants using the counting strategy on average (2.60 seconds versus 2.47 seconds respectively). This difference in overall mean response times was not significant, $F(1, 30) = .233$, $p > .50$. However, there were differences in the pattern of data produced by participants using the two strategies as they performed the task (Figure 3). Participants using the counting strategy showed a larger and more consistent effect of the target's position relative to the plane (Figure 3a), while the participants using the rotation strategy showed a larger effect of misalignment (Figure 3b). Both of these interactions were significant, $F(7, 210) = 3.53$, $p < .02$ for the strategy by target position interaction, and $F(7, 210) = 3.81$, $p < .01$ for the strategy by misalignment interaction. These data were then compared using the slopes of these effects from the participants in the two conditions[1] (see Table 1). In this analysis, there was a significant difference between the two conditions in terms of the target's location, $F(1, 30) = 7.91$, $p < .01$, and in terms of misalignment, $F(1, 30) = 5.95$, $p < .03$. These results indicate that the impact of target location was greater for

---

[1]Slopes were calculated by averaging participants' data over left and right misalignments and target locations. In addition, target locations at the "bottom" and "top" of the camera view were excluded for target location. Then, a best fitting line was estimated for each of the conditions for each participant. The results of this process were used in the analyses.

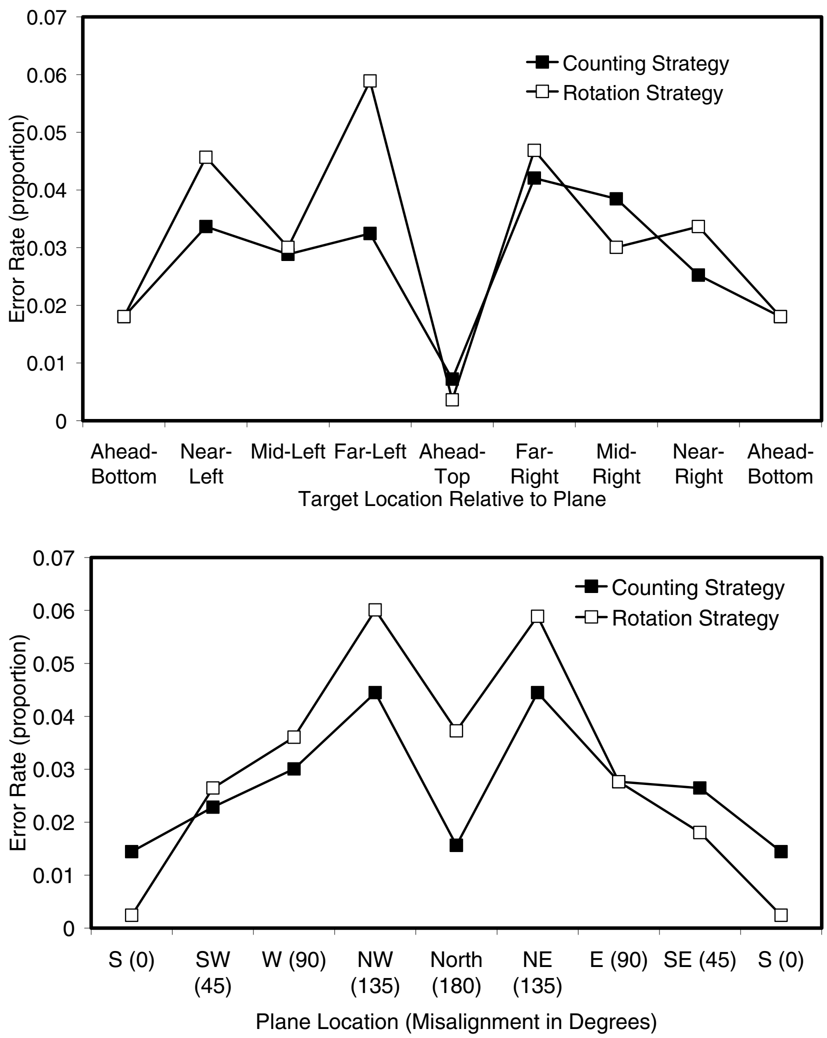

Figure 2. Error data for Experiment 1.

participants using the counting strategy, while the impact of misalignment was larger for participants using the rotation strategy.

In addition to the separate effects of misalignment and target location, the data showed that there was a more complex relationship between them (Figure 4). For participants using both strategies, there was a significant interaction between misalignment and the target's location relative to the viewer,

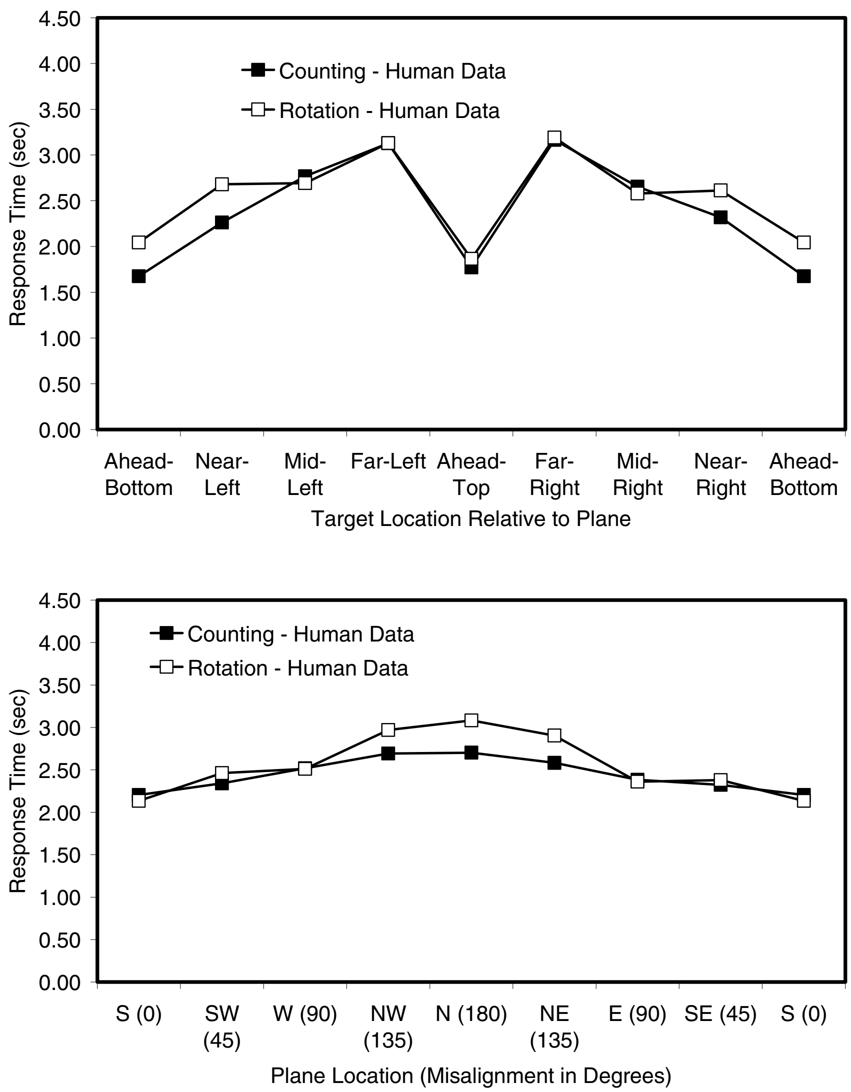

Figure 3. Response time data for Experiment 2, showing performance on the task based on strategy training condition.

$F(49, 735) = 2.26$, $p < .01$ for the counting strategy and $F(49, 735) = 3.42$, $p < .001$ for the rotation strategy. In both conditions, this seems to be the result

Table 1
Slopes (ms/45º) of the effects found in Experiment 1

Factor	Condition	
	Rotation	Counting
Target Location	218	423
Misalignment	198	110

of the special case strategy that the participants reported using despite their training. This strategy applied when the target was in line with the plane (the first and last points on each line in Figure 4), and seems to have resulted in a diminished effect of misalignment when it applied. In fact, for the rotation strategy, the slope of the misalignment effect is 22ms/45º when the target is in line with the plane and 314ms/45º in other cases. This difference is significant, $F(1, 15) = 28.83$, $p < .001$. For the counting strategy the slope of the misalignment effect is 28ms/45º when the target is in line and 164ms/45º otherwise, which is also a significant difference, $F(1, 15) = 21.70$, $p < .001$.

Along with the verbal reports, the interactions just described provide evidence that the special-case strategies were used by the participants in this experiment. When the target was in line with the plane, response times were fastest and the impact of misalignment was greatly reduced. Since the counting and rotation strategies seem to have not been used in these situations, the trend analyses that are performed to examine how the target's location impacts performance within each strategy ignore those target locations. In addition, in all of the trend analyses, the data are averaged over left and right target locations and misalignments, since the data are basically symmetrical.

The effects in the data correspond to the hypotheses. For the counting strategy, it was expected that participants would show a linear increase in response time as a function of the target's position relative to the plane, as well as some impact of misalignment if left-right confusion were an issue for participants. The first of these expectations was clearly supported by the data, as there was a strong effect of the target's position on response time, $F(2, 30) = 53.17$, $p < .001$, and this effect showed a linear trend, $F(1, 30) = 106.32$, $p < .001$, and no evidence for a quadratic trend, $F(1, 30) = .01$, $p > .90$ (excludes "near" and "far" target locations). There was also an effect of misalignment in these data, $F(7, 105) = 11.84$, $p < .001$. Once again, this effect was reflected in a strong linear trend for these data, $F(1, 60) = 59.76$, p < .001, but no quadratic trend, $F(1, 60) = .07$, $p > .70$. The results suggests that participants were using left and right to encode the target's location, resulting in increasing left-right confusion as the two views were more misaligned.

We expected the effects to have different magnitudes for participants using the rotation strategy. Firstly, there should have been a linear effect of misalignment on response time. The data indicate that there was an effect of misalignment, $F(7, 105) = 14.844$, $p < .001$, and that this effect involved a strong linear trend, $F(1, 60) = 118.03$, $p < .001$, and not a quadratic trend,

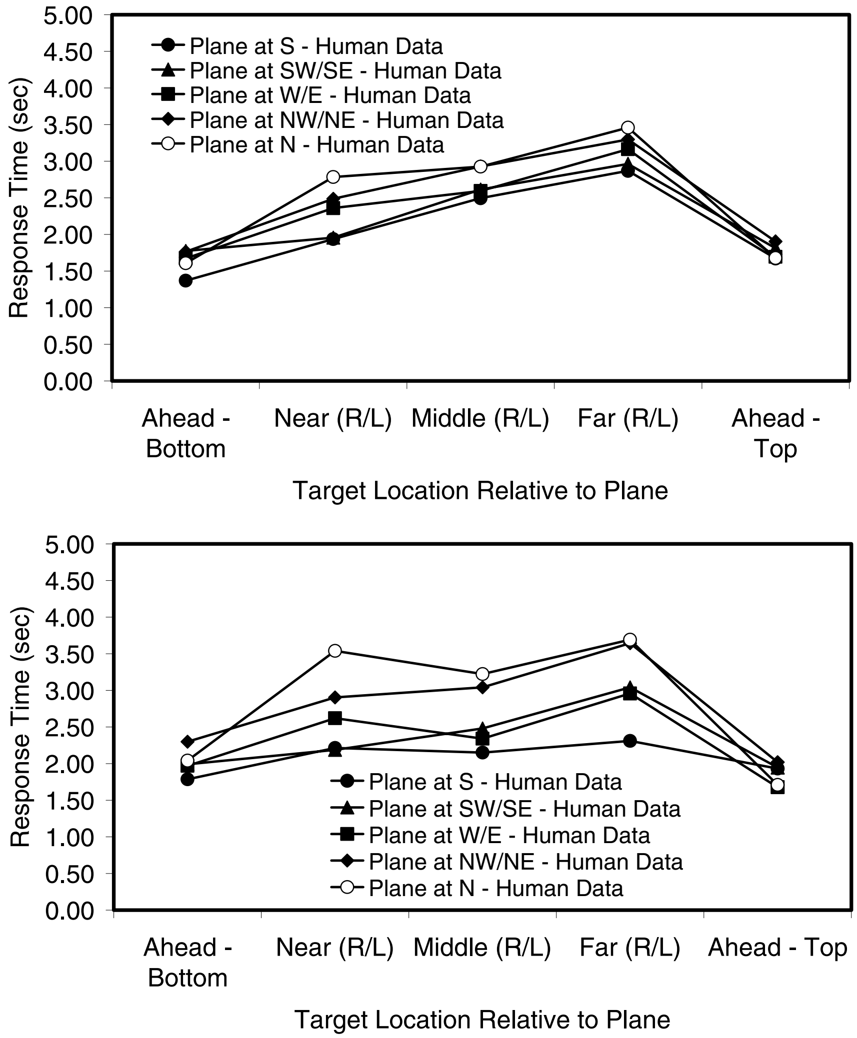

Figure 4. Interaction between misalignment and target location in Experiment 1. Values are averaged over left and right misalignments and target locations for clarity. Top panel is the data for the counting strategy, bottom panel is for rotation.

$F(1, 60) = .61$, $p > .40$. Secondly, if left-right confusion was an issue, there should have been an effect relating to the target's location in the camera view. The data show this effect, $F(2, 30) = 23.95$, $p < .001$, even excluding the special cases of targets at the top and bottom of the camera view. This effect contained both a linear linear component, $F(1, 30) = 35.24$, $p < .001$, and a quadratic component, $F(1, 30) = 12.67$, $p < .01$. As Figure 3a shows, when the "bottom" and "top" positions are excluded, the effect is monotonically increasing, but it does not appear to be strictly linear. The slopes of the effects for both strategies are shown in Table 1.

Discussion

The data collected in Experiment 1 show that there were differences in performance depending on which strategy was used to complete the task. The verbal reports and response time data also indicate that the special case strategies used by untrained participants were used by individuals trained to do something else. This finding supports other research on strategy use that finds lower-cost, highly effective strategies tend to be used in place of more effortful alternatives (Gunzelmann and Anderson, 2003; Lovett and Anderson, 1996; Reder and Schunn, 1999; Siegler, 1987). In this case, the shortcut strategy was so low-cost and effective that every participant reported using it when it was applicable. Beyond the special cases, the fact that there was no overall difference in performance on the task between individuals using the two strategies suggests that they are both just as effective for solving the task.

Both the effect of target location in the rotation strategy and the effect of misalignment in the counting strategy support the conclusion that participants encountered left-right confusion as they solved the trials. This conclusion is supported by the data on errors as well. Although overall accuracy was very high (96% for participants using the counting strategy), 40.4% of the errors in the counting condition were "left-right confusions". These are instances where the response given was the correct number of steps from the plane, but in the wrong direction. In the sample trial shown in Figure 1, a response of "West" would be considered a left-right confusion error. Note that "West" is 3 steps from the plane, but it is in the incorrect direction. Far more of these errors were made than would be expected by chance, $X^2(1) = 89.92$, $p < .001$, which would only be 14.3%. That such errors were the most common type of error in the data supports the conclusion that participants using the counting strategy were encoding the target positions as "left" and "right". As they began their search on the map, these labels would have to be updated to count in the appropriate direction when the views were misaligned. The model of the data, presented below, provides a more thorough explanation and demonstration of how this process could impact performance.

For the rotation strategy, the strong linear trend for the effect of misalignment supports the idea that mental rotation was being done as proposed by previous researchers (e.g., Hintzman, et al., 1981; Rieser, 1989; Shepard and Hurwitz, 1984). There was also an effect of the target's position on performance in the data. In addition, for participants using this strategy, 55.9% of their errors were left-right confusions (overall accuracy was 94%). Again, this is more than would be expected by chance, $X^2(1) = 735.37$, $p < .001$. In Figure 1, an angle pointing to West is the correct size (135 degrees), but opens in the wrong direction. There are at least two possible ways that these errors could be produced. First, it is possible that participants rotate the angle as a coherent unit, but that it somehow gets *flipped*, which would result in a left-right confusion error. The other possibility is that the image participants were rotating was not a rigid, coherent angle. Instead, it may be the case that participants only rotate the leg of the angle associated with the target until it forms an angle of the correct size with the line indicating the plane's orientation on the map.

There are reasons to suspect the latter explanation. In that case, left-right confusion errors would arise when the target leg was rotated too far, or in the wrong direction. That is, individuals would succeed in recreating an angle of the appropriate size. However, it would open in the wrong direction. In this experiment, the line indicating the plane's orientation shows, in a very real sense, the plane leg in its rotated position on the map view, which may encourage participants to not bother rotating that leg of the angle. In addition, rotating only a portion of the angle should be easier than mentally rotating the entire angle, since less information needs to be maintained and manipulated (Bethell-Fox and Shepard, 1988). Finally, this explanation fits with the "piecemeal" strategy that has been described for mentally rotating block figures (Just and Carpenter, 1985). While the second experiment does not resolve this issue, the ACT-R model described below illustrates how mentally rotating only a single leg of the angle can produce increased response times due to left-right confusion.

Based on the results, left-right confusion becomes more likely in the counting strategy as the orientations of the two views of the space become increasingly misaligned and the direction of counting differs on the map view and the camera view. Likewise, such confusion becomes more likely in the rotation strategy as the target's location gets nearer to the top of the camera view, and the direction that the target leg of the angle must be rotated differs from the direction the plane leg must be rotated. In essence, left-right confusion increases in either display as the critical elements are located closer to the top. The strategy (counting or rotation) is an algorithmic way of dealing with the confusion in one of the displays, but it does not eliminate confusion in the other. However, it is also the case that the pattern of errors is similar for both strategy conditions (Figure 2), suggesting that the strategies do not completely eliminate the difficulty associated with one display or the other. This provides an indication of why past research has found effects of both factors.

Experiment 2

In Experiment 2, we extend the results of Experiment 1 by collecting eye data to more closely examine how participants were performing the task. These data make it possible to identify what information participants were attending to, and in what order. As a result, it is possible to better understand the differences and similarities between the two strategies. At a more general level, it also is possible to examine what aspects of the task were important for the participants as they solved it.

Assuming that participants look at the information on the screen that they are thinking about (Just and Carpenter, 1976), the strategies that they were taught are explicit enough to allow us to make predictions about where on the screen they should look while they do the task. According to the instructions, participants should look at the camera view before the map view using both strategies. Participants using the counting strategy should spend their time looking at the target, the plane, and the items between the target and the plane in both the camera and map view (during the counting process), while participants using the rotation strategy should look at regions associated with the angle (the plane, target, and center of each view). Based on this, participants using the counting strategy should spend a greater proportion of their time looking at areas of the screen between the target and the plane while those using the rotation strategy should spend more time looking at the center of the two views. Participants in both conditions should spend some time looking at the plane and the target in both views. The eye data collected in this experiment are used to address these predictions.

Method

Participants. The participants in this experiment were 20 individuals recruited from an email b-board. This excludes three individuals who were unable to participate for technical reasons (see below). Each of the 20 participants was paid $15 for participating in the 1.5 hour experiment. There were 9 males and 11 females in the study (mean age 21.4 years).

Materials and Equipment. The eye tracker used to determine participants' point-of-regard (POR) was an ETL-500 manufactured by ISCAN, Inc. The ETL-500 is a high-resolution pupil/corneal reflection dark pupil tracker that uses a Polhemus FASTRAK magnetic head tracker to compensate for head movements. Because of this, participants did not use a chin rest, or other immobilizing device, during the experiment. The ETL-500 typically estimates a participant's POR accurately to within one degree of visual angle. However, accuracy does diminish somewhat as distance of the POR from the center of the field of view increases because of the geometry of the eye and ISCAN's proprietary algorithm for determining POR. The tracker was configured with ultra light head-mounted optics to minimize discomfort. This amounted to having participants wear a baseball cap, with the eye-tracking device mounted to the brim. Participants' POR was recorded every 16.7ms by the experiment delivery software, and keypress time was recorded and synchronized with the POR data. In this way, the experiment delivery software was able to record trial data files that could be analyzed and replayed after the experiment.

To make sense of such a large amount of data, it is necessary to classify the POR samples according to where on the screen each one falls. To do this, the screen was first segmented into 18 areas; one for each target position, one for each cardinal direction, and two others for the center of each view (Figure 5).

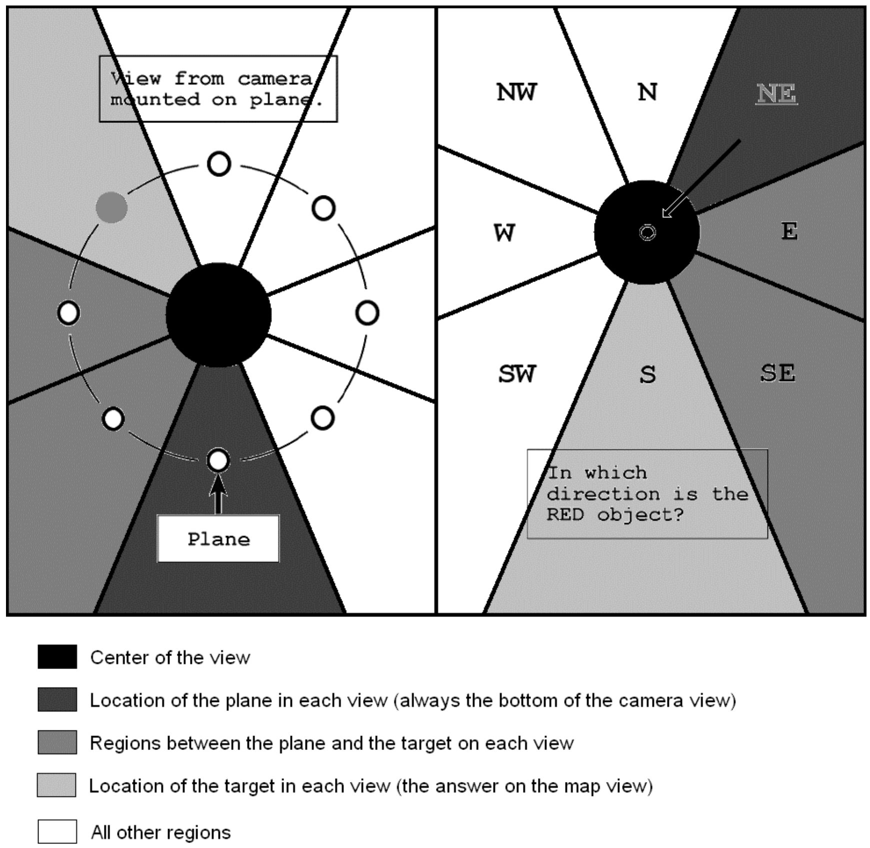

Figure 5. Defined regions for classifying the eye-tracking data in Experiment 2. The grid overlaid on the display was not seen by participants, but rather shows distances (in terms of visual angle) between objects on the screen.

The center areas were given a radius of approximately 1 degree of visual angle. Then, because the trials differed in terms of the location of the target and the location of the plane, these 18 areas were classified on each trial according to functional designations. Based on the predictions for the strategies described above, the nine areas in each view of the trial display can be classified as being in one of five functional regions for any given trial. The "functional regions" are illustrated for the sample trial shown in Figure 5, and the analyses presented below are based on those functional regions. Note that in trials where the target is located at the bottom of the camera view, the regions representing the target (the answer on the map view) and the plane are actually the same, meaning that samples in those regions for those trials are counted twice in the data presented below. Also, for trials where the target was located at the bottom, near-left, or near-right positions in the camera view there are no regions that fall between the plane and the target in either view.

For this experiment, we chose to calculate the number of samples (recorded every 16.7ms) that fell within each region. This method will allow for a more complete comparison of the experimental data to the model below. Finally, the experiment was reimplemented to work with the eye-tracking equipment, resulting in some minor changes to the display for this experiment (compare Figures 1 and 5). The most notable of these was that the plane's location is indicated as a highlighted cardinal direction here (grey and underlined in Figure 5), rather than as a black dot. This was mostly a coding convenience and did not seem to have an impact on participants' ability to do the task.

Procedure. Participants were tested in a single session that lasted no more than 1.5 hours. Before beginning the training portion of the experiment, they were asked to put on the eye-tracking equipment to make sure that their eye movements could be adequately tracked. Three individuals were unable to participate in the study for this reason. After this brief procedure, participants were presented with instructions for the task, followed by training on one of the strategies. This portion of the experiment was identical to Experiment 1, and half of the participants were trained to use each of the strategies. After completing the training, participants put on the eye-tracking equipment again. At this point, participants went through a procedure called calibration.[2] This procedure was repeated at the beginning of each block and opportunities for recalibration were presented every 16 trials during the block. Participants were also given the opportunity for a short break between blocks. As in Experiment 1, participants completed 4 blocks of trials, using the same drop-out procedure. After finishing, participants were questioned about whether or not they used the strategies they were trained to use, and what else they may have been doing. Once again, participants reported using the strategies they were taught and also the special-case strategy for targets in line with the viewpoint.

Results

Errors. The pattern of errors made by participants in this experiment closely matches the previous experiment (Figure 6). Accuracy remained high overall (93%) and right-left confusion errors were the most common errors in both strategies (50% for the counting strategy and 64% for the rotation strategy).

---

[2]Calibration is the process of identifying what the image of the eye looks like when the person is looking at different areas of the screen. In this process, participants are asked to look carefully at the cursor as it is programmatically moved to various locations on the screen. At each location, a snapshot is taken of the eye's position. Then, interpolation can be used to determine the participant's POR for other areas on the screen. The quality of the calibration was checked by having participants track the mouse as the experimenter manually moved it around the screen. In cases where the quality was poor, the procedure was repeated.

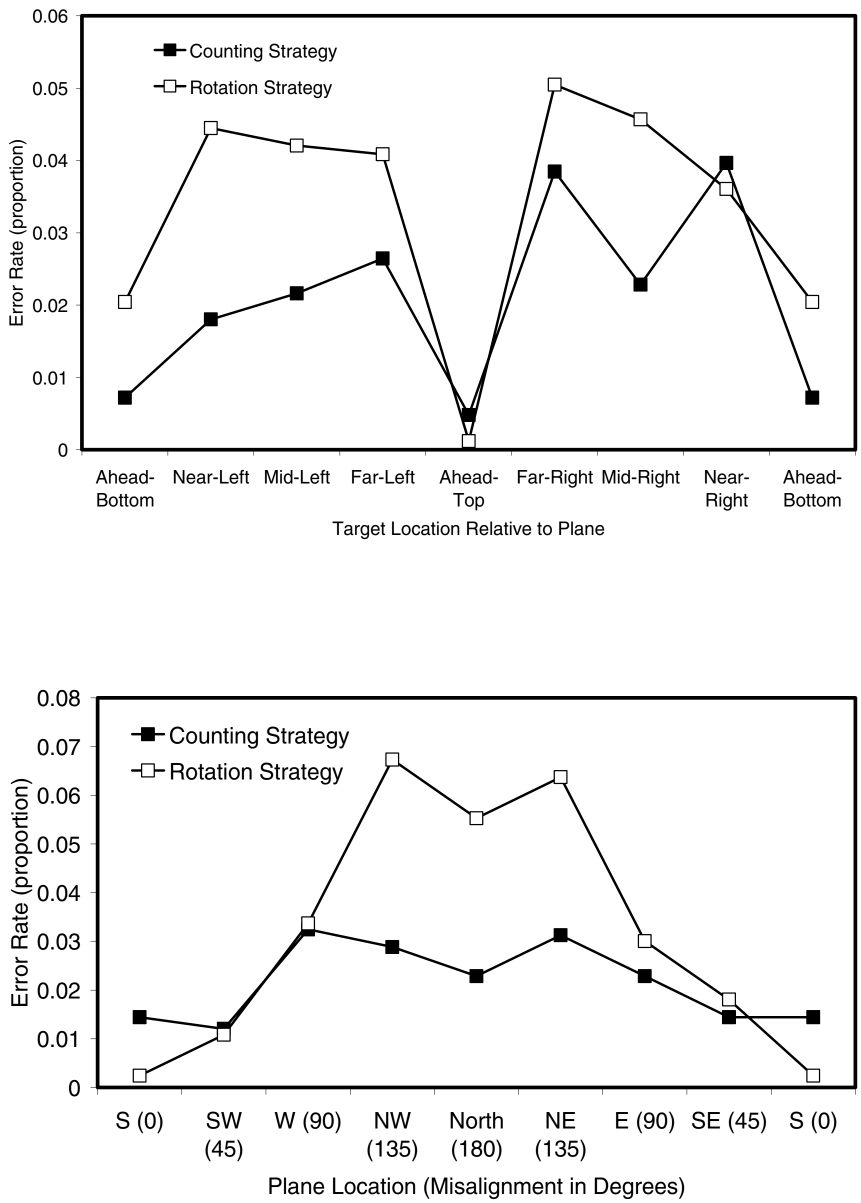

Figure 6. Error data for Experiment 2.

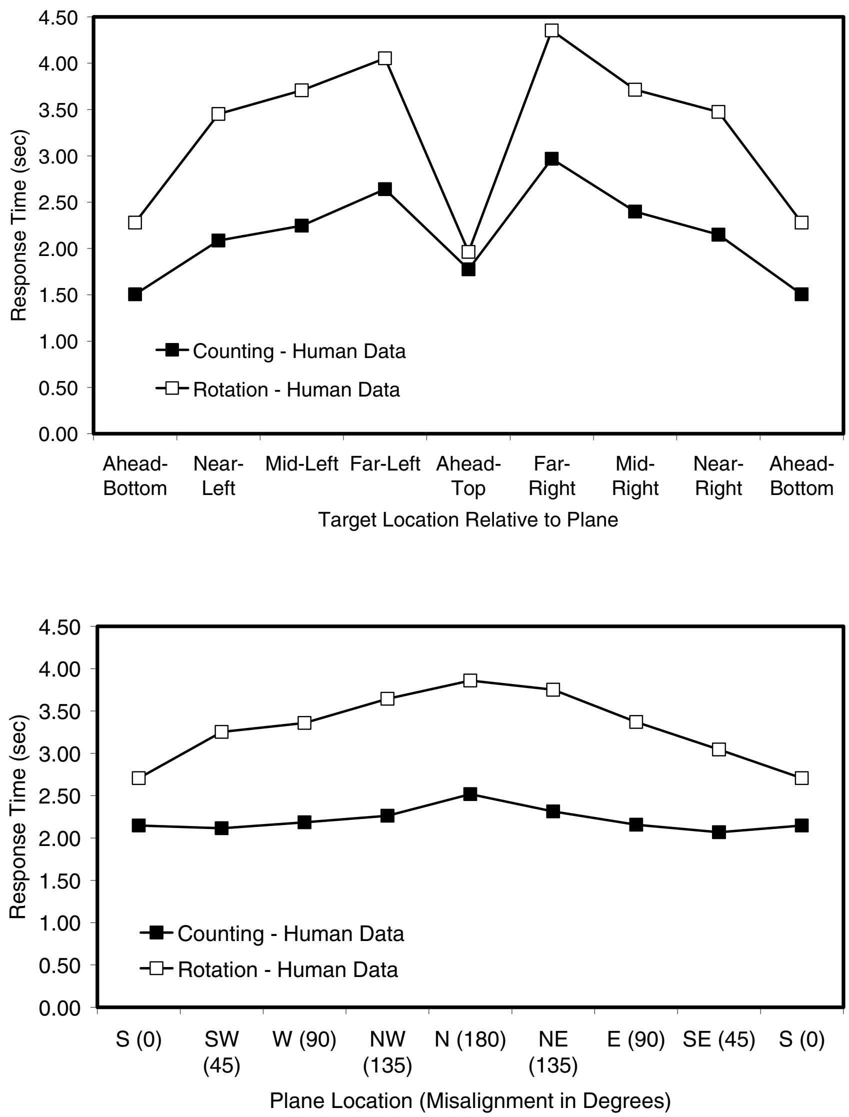

Figure 7. Response time data for Experiment 2, showing performance on the task based on strategy training condition.

When the target was in line with the plane and when the two views were aligned, error rates were very low. Error rates were higher when the target was located in one of the other positions, and also generally increased as the misalignment between the two views increased. This once again illustrates that the results are not due to a speed-accuracy trade-off ($r = .75$ with the response time data), and reinforces the conclusion that target location and misalignment both have an impact on task difficulty. Although it seems as though participants using the rotation strategy made more errors than those using the counting strategy, those data are skewed by one participant who made 131 errors (66% correct; this participant's performance was otherwise similar to the other participants, including average response time and the pattern of effects). When this participant's data are excluded, accuracy is similar between the two conditions (95% for the rotation strategy and 94% for the counting strategy).

Response Time Data. While the average response times from participants were generally in line with the data gathered in Experiment 1, performance was not equivalent between the two strategy conditions in this experiment (Figure 7). That is, participants using the rotation strategy took significantly longer to complete the trials than participants using the counting strategy, $F(1, 18) = 8.42$, $p < .01$. There was a slight tendency in this direction in Experiment 1, though the difference was not as large. This translated into a significant strategy by experiment interaction, $F(1, 48) = 4.82$, $p < .04$. Average response times were similar for participants using the counting strategy (2.47 versus 2.22 seconds for Experiments 1 and 2 respectively). In contrast, participants using the rotation strategy had faster response times in Experiment 1 than in Experiment 2 (2.60 versus 3.37 seconds respectively). Beyond this, the data seem similar for the two studies.

For the counting strategy, the data correspond well to the hypotheses and to the results of Experiment 1. Figure 7b illustrates that misalignment had an impact on performance, $F(7, 63) = 4.80$, $p < .03$, producing a significant linear trend in these data, $F(1, 36) = 16.58$, $p < .01$, with no quadratic trend, $F(1, 36) = 5.05$, $p > .05$. Response times increased as the misalignment between the two views increased, suggesting that left-right confusion was affecting participants' performance, as found in the last experiment. There was also a significant effect of target location, $F(7, 63) = 37.83$, $p < .001$ (Figure 7a). This effect had a significant linear component, $F(1, 18) = 82.25$, $p < .001$ and the quadratic trend was not significant, $F(1, 18) = 4.49$, $p > .05$ (trend analyses on target location exclude "near" and "far" target locations). This illustrates the expected effect that counting should have on response times.

For the rotation strategy, there was again an effect of misalignment, $F(7, 63) = 9.74$, $p < .001$, which was reflected in a linear trend, $F(1, 36) = 46.74$, $p < .001$, but no quadratic trend, $F(1, 36) = .81$, $p > .25$ (Figure 7b). In terms of the target's location there was a significant effect as well, $F(7, 63) = 32.03$, $p < .001$, but response times appear to increase more regularly as the target's position gets farther from the plane than they did in Experiment 1 (Figure 7a). Indeed, in this experiment, the linear trend for these data is significant,

$F(1, 18) = 32.40$, $p < .001$, while the quadratic component is not, $F(1, 18) = 1.20$, $p > .25$ (trend analyses on target location exclude “near” and “far” target locations). Despite this, it seems that the major trends in the data are comparable.

To compare the magnitudes of the effects in this study, the slope of the effects were calculated for each of the participants in the same way as Experiment 1 (Table 2). A comparison of those slopes reinforces the conclusion that the data from this experiment were similar to the results from Experiment 1. The magnitude of the misalignment effect was larger for participants using the rotation strategy, $F(1, 18) = 6.07$, $p < .03$, but there was no significant difference between conditions in terms of the impact of the target’s location, $F(1, 18) = 0.09$, $p > .75$.

The results again indicate that the impact of misalignment was diminished for targets directly in front of the viewer (Figure 8). However, the interactions were not significant, $F(49, 441) = 1.70$, $p > .15$ for the counting strategy and $F(49, 441) = 1.56$, $p > .19$ for the rotation strategy. However, by looking at the slopes of the misalignment effects for these different situations, the differences appear. For the counting strategy, this slope was 12ms/45º when the target was in line with the viewer and 121ms/45º in other situations, $F(1, 9) = 8.07$, $p > .02$. For the rotation strategy, the slopes were 42ms/45º and 366ms/45º for those respective conditions, $F(1, 9) = 19.03$, $p > .01$. Overall, the trends in the data correspond well with the hypotheses and to the results of Experiment 1. The next section describes the eye-tracking data, to examine how closely participants’ performance matches the predictions described above.

Eye-Tracking Data. Each POR sample was identified as being located in one of the 10 functional regions (Figure 5). In addition to classifying the samples collected from participants, we also calculated the proportion of samples that would fall into each of the regions by chance. This was done by calculating the average area of the window that was occupied by each of the functional regions in the task across all trials. Figure 9 presents the proportion of samples that fall within each region for the two strategies and for chance. It is interesting to note that for both strategies, POR samples are more or less evenly divided between the camera view and the map view. In addition, POR samples for regions that are “off-task” are well below chance (the “other” categories).

Table 2
Slopes (ms/45º) of the effects found in Experiment 2

Factor	Condition	
	Rotation	Counting
Target Location	370	343
Misalignment	237	77

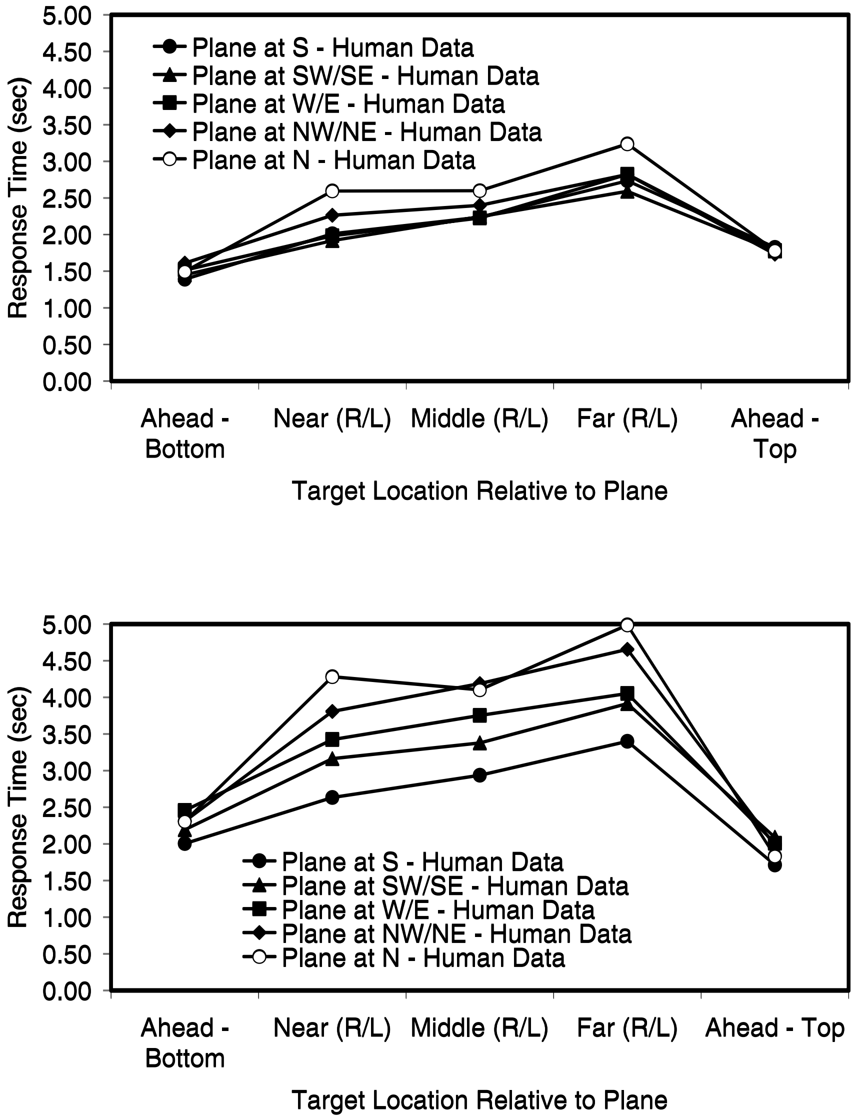

Figure 8. Interaction between misalignment and target location in Experiment 2. Values are averaged over left and right misalignments and target locations for clarity. Top panel is the data for the counting strategy, bottom panel is for rotation.

Using the data in Figure 9, comparisons can be made between the strategy conditions. For the most part, the proportions of samples in the regions are similar between conditions. Focusing on the areas for which differences were predicted, however, some variations appear. First, for the two "between" areas, the counting strategy shows a greater proportion of samples than expected by chance while the rotation strategy has fewer than expected by chance. This difference is significant in both views, $F(1, 18) = 5.62$, $p < .03$ for the camera view, and $F(1, 18) = 8.3$, $p = .01$ for the map view. Second, a greater proportion of samples fall in the center of the map view for the rotation strategy than for the counting strategy. This effect is also significant, $F(1, 18) = 10.11$, $p < .01$. While the trend is the same for the center of the camera view, this effect was not statistically significant, $F(1, 18)=2.99$, $p = .10$. These results all match the predictions described earlier. Participants using the rotation strategy spent a marginally greater proportion of time looking at the target on the camera view than those using the counting strategy, $F(1, 18)=4.35$, $p < .06$, which may reflect the time needed to encode the angle to be rotated. No other comparisons were significant.

The eye-tracking data can also be used to determine when participants were looking at different areas of the screen, which is informative about how they solved the problems. Figure 10 presents the average time at which participants were looking at the regions during the trials in the experiment. Regardless of the strategy they were taught to use, participants were taught to first encode the information from the camera view and then to use that information to determine the answer on the map view. The eye-tracking data support the conclusion that participants were performing the task in this order, as samples were recorded significantly earlier on the camera view than on the map view $F(1, 18) = 164.95$, $p < .001$. In addition, these data show how strategy impacted overall time, since participants using the rotation strategy looked at regions later overall than participants using the counting strategy, $F(1, 18) = 8.42$, $p < .01$. It is also the case that this effect was larger for the map view than for the camera view, which illustrates that participants using the rotation strategy were spending more time looking at both views, $F(1, 18) = 7.426$, $p < .02$ for the strategy by view interaction. These data suggest that participants using the rotation strategy had more difficulty both encoding the information from the visual scene and using that information to find the target on the map.

Discussion

The results of this experiment suggest that participants were using the strategies they were taught to complete the task. As hypothesized, participants using the counting strategy spent more of their time looking at regions between the plane and the target, and participants using the rotation strategy spent a greater proportion of their time looking at the center of the two views (though this effect was only significant for the center of the map view). In addition, both groups of participants looked at the plane and the target longer than would be expected by chance, and less time looking at strategy-irrelevant areas

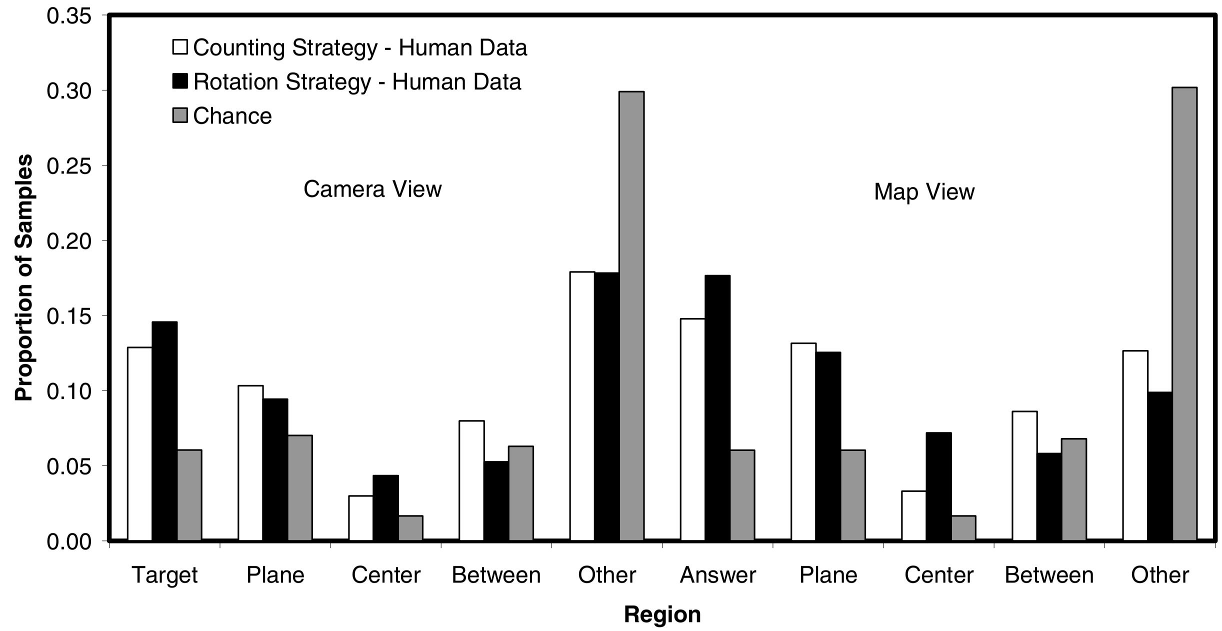

Figure 9. Eye movement data for Experiment 2, shown as the proportion of samples in each region for each condition, and for chance.

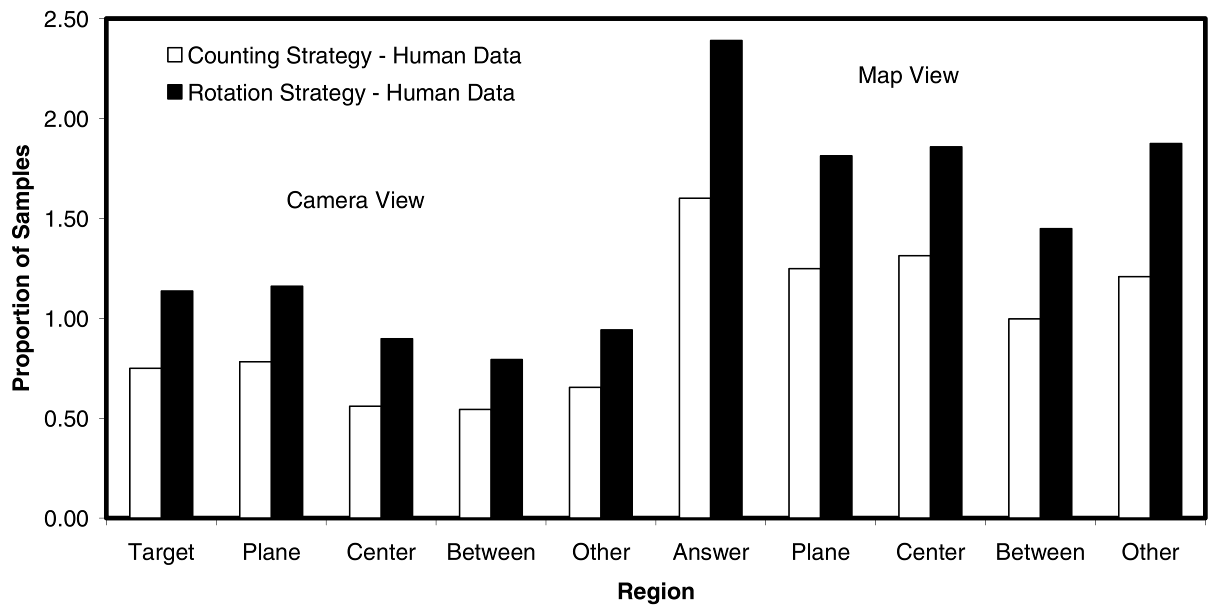

Figure 10. Eye movement data for Experiment 2, showing the average time that eye samples were recorded within each functional region.

ACT-R Model

The findings from the empirical studies provide evidence about how participants were performing the task in the two strategy conditions. The explanation is described in detail in this section, which presents the results from an ACT-R model of the task. The model incorporates perceptual, cognitive, and motor components, resulting in a simulation that is able to interact directly with the experiment, just as the human participants did. Thus, the model provides detailed quantitative predictions about both response times and eye-movements. Also, the model incorporates multiple strategies to more accurately portray how different participants are performing the task. We believe that this provides a comprehensive account of human performance in this kind of task. The model's source code can be downloaded at the ACT-R website (http://act-r.psy.cmu.edu).

ACT-R is a theory of human cognition that has been formalized as a running simulation (Anderson and Lebiere, 1998; Anderson, et al., 2004). At its most basic level, the cognitive component of ACT-R is a production system that operates in a loop, where the current state of the system is compared against possible actions, one is chosen and executed, and a new system state results. Within this system there is a division between declarative knowledge and procedural knowledge. Declarative knowledge is the repository for facts such as "three plus four equals seven" or "north is opposite to south". In contrast, procedural knowledge represents the information pertaining to how to transform one system state into another. This knowledge is represented as production rules, or condition-action pairs that describe these transformations. For instance, there may be a production for making a move in the Tower of Hanoi or one that encodes information about a visual stimulus. Within this basic framework, a wide variety of cognitive phenomena have been investigated (see Anderson & Lebiere, 1998 for a review).

In its latest instantiation (ACT-R 5.0), ACT-R contains cognitive, perceptual, and motor modules that work together to produce the system's behavior (Anderson, et al., 2004). These modules are able to operate in parallel. So, the motor module can be preparing and executing a mouse movement while the visual system is moving attention and the cognitive system is retrieving a piece of information from memory. Associated with each of the modules are "buffers" that hold information about the current state of the module. The buffers also serve as the interface between the modules and the production system. It is at the level of production execution that the system is serial. Productions match against the contents of the buffers, and only the contents of the buffers can be directly accessed or updated. In this sense, the complete set of buffers can also be thought of as the working memory storage for ACT-R.

In the model presented here, the buffers that are of primary interest are the retrieval buffer, the goal buffer, the visual-object buffer, and the visual-location buffer. The retrieval buffer holds declarative information that has been retrieved from long-term memory, while the goal buffer holds information about the current task situation. The two visual buffers hold information about “what” is being looked at and “where” something is on the screen respectively. As the model progresses through a trial, the information in these buffers is updated to keep track of the state of the solution process. When the answer has been found, buffers in the motor module are used in the planning and execution of the motor movements needed to make the appropriate keypress.

The close coupling of the cognitive module of ACT-R to the perceptual and motor modules results in a system that is well-suited for investigating performance in the current task. The trials involve a fair amount of perceptual information that must be gathered from the screen and interpreted to find the solution (not to mention the motor actions needed to make responses). An explanation of participant performance that ignores such aspects of performance would not provide a an entirely accurate account of what participants are doing (Ritter, Baxter, Jones, and Young, 2000). In addition, since this account is implemented as a running simulation, it makes quantitative predictions that can be compared directly to the actual data from participants. While a verbal theory might provide predictions that are qualitatively like the data, it is a sterner test whether it can match the quantitative predictions. As ACT-R comes with strong constraints on its parameters, particularly the perceptual and motor processes, fitting quantitative data becomes a stronger test of the theory.

Model Design

In the following subsections, the implementation of the two strategies in ACT-R is described in detail. Since the model is an instantiation of our account of participant performance, these descriptions serve both as an explanation of the model and as a theoretical account of human performance on the task. Following these descriptions, some additional details of the simulation will be provided along with the model's fit to the empirical data. The steps in the execution of the model using both strategies are illustrated in Figures 11 and 12 for the counting and rotation strategies respectively. For both strategies, the location of ACT-R's visual attention is illustrated for each step in the solution process. For the rotation strategy, the current mental image of the angle is illustrated as well.

Counting Strategy. The counting strategy represents a solution to the orientation task that takes advantage of the regularity of the stimuli. Based on the data and the instructions that were given to participants, the model was constructed to complete the trials in the following way. First, the model counts the objects around the target field (starting at the bottom) to the target (steps 1 to 4 in Figure 11). Then, the count and the direction of the target (left or right) are encoded (“3 to the left” in Figure 11). Once the target information is encoded, the model finds the plane on the map view (step 5). At this point, the model needs to determine which way to count. In most cases, the direction of counting on the map view is the same as the direction on the camera view (left or right). However, problems arise when the plane is located in the north (NW, N, or NE). In these cases, the correct counting direction is actually opposite to the encoded direction, and extra operations are needed to correct the discrepancy (see "Model Details and Parameters" section below). Once the search direction is determined, the model counts around the cardinal directions on the map (steps 6 to 8). When it reaches the proper number in the count, it has found the answer (step 8). Then, the cardinal direction is mapped to the appropriate numerical response and the response is made by pressing the appropriate key. During each counting step for both counting sequences, the model goes through the three steps of finding the next location, moving visual attention to that location, and incrementing the count.

Rotation Strategy. In some respects the model of the rotation strategy is similar to the counting model, since a large portion of the task involves perceptual and motor actions. However, the rotation strategy involves mental imagery, and currently ACT-R does not have any built-in capacity for mental imagery. While this does not mean that the system is not suited for modeling tasks involving imagery, it does mean that the capacity to represent mental images had to be added to ACT-R to faithfully model the performance of participants using the rotation strategy. Recall from above that ACT-R is now composed of a number of modules, each with a corresponding set of buffers. The buffers are the repository for "current" information and can be thought of as a representation of the current state of the system or, alternatively, as the contents of working memory. Because of its implementation, ACT-R allows for the addition or subtraction of buffers. In this case, a buffer was added to hold imagery-based information (an imaginal buffer). While this buffer in its current form is quite limited, it does represent an initial attempt to expand the scope of ACT-R's competence into new areas of cognitive research.

The imaginal buffer at this point simply holds a single chunk of information that contributes to the current state of the system. That is, there are no mechanisms associated with it to allow for complex imagery-based manipulation of the information, like mental rotation. Rather, in this model, a representation of the angle is created in the imaginal buffer and updated in a step-wise manner as the model completes the trials. The representation consists of the locations of three points, the plane, the target, and the vertex. As the trial progresses, these slots are filled in and then updated iteratively to determine the correct answer.

In the instructions, participants were told to form an angle on the camera view and rotate it to align it with the plane's location on the map view. The model forms the angle by first looking at the bottom of the target field, then the center of the target field, and finally the target's location on the camera view (steps 1 to 3 in Figure 12). At each of these points, the visual-location is encoded (from the visual-location buffer) in the appropriate slot of the angle in the imaginal buffer (note the formation of a complete angle over steps 1 to 3). Once this is accomplished, the angle is "moved" to the map view by updating the visual-locations to be the corresponding locations on the map view (e.g., the vertex is updated to correspond to the center of the map view; step 4). Then, the model

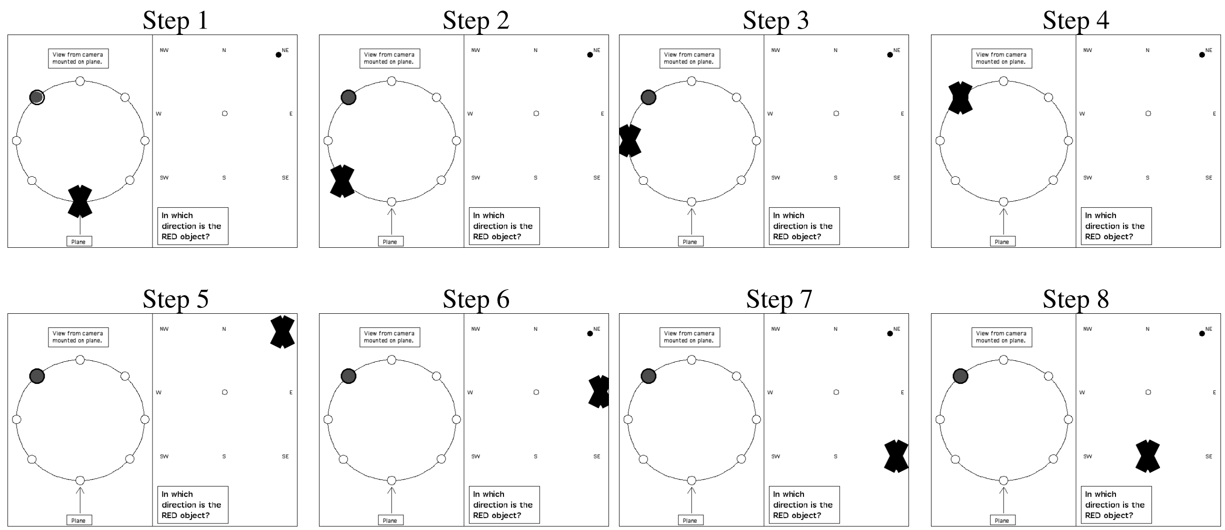

Figure 11. Illustration of the execution of the counting model. The images illustrate the process by which the model executes the counting strategy for the trial shown in Figure 1. The large, dark “X” represents the eye location. Note that the line showing the plane's orientation is deleted in this display to minimize confusion.

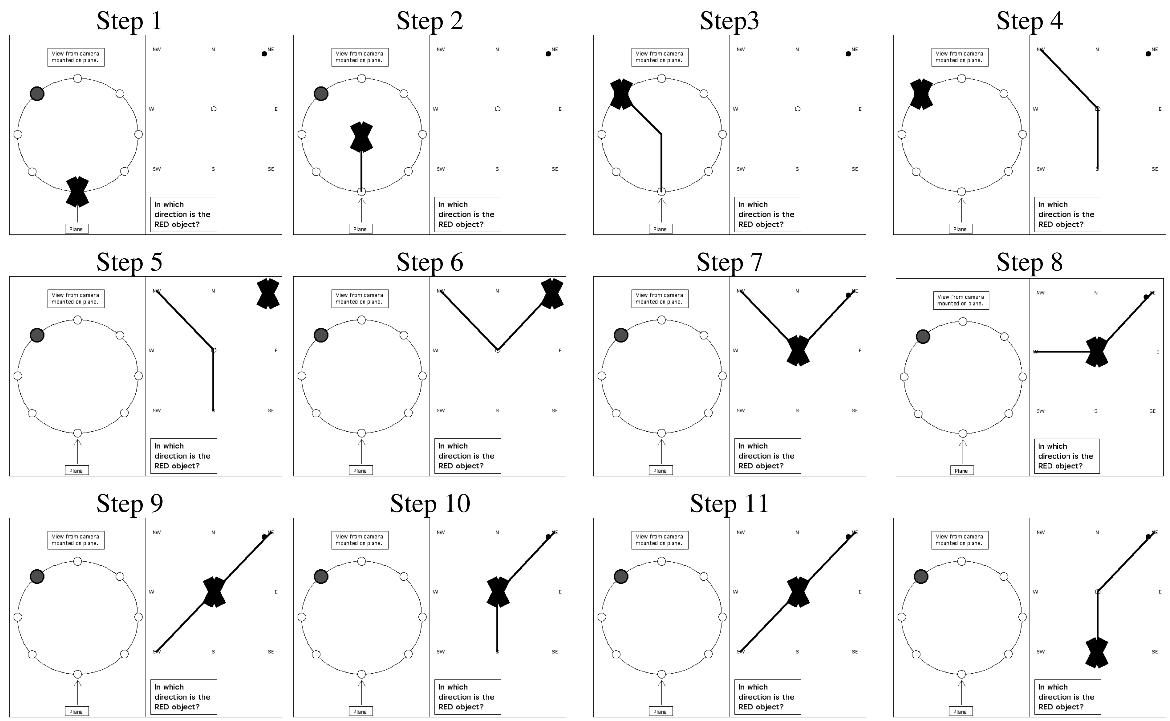

Figure 12. Illustration of the execution of the rotation model. The images illustrate the process by which the model executes the rotation strategy for the trial shown in Figure 1. The large, dark "X" represents the eye location, and the lines show the representation of the angle in the imaginal buffer. Note that the line showing the plane's orientation is deleted in this display to minimize confusion.

finds and moves its visual attention to the plane's actual location on the map view (step 5), and the plane's location is updated in the angle chunk in the imaginal buffer (step 6). Next, the model's visual attention is shifted to the center of the map view (step 7). Since the eye-tracking data showed that participants spent a relatively large amount of time looking at the center region of the map view when using the rotation strategy, the model executes the rotation while it is looking at the center of the map view (steps 8 to 10). This rotation only involves the target leg of the angle, since the location of the plane is updated in Step 6. Recall that this was one possible explanation for why the observed pattern of response times and errors were present in the empirical data. Once the rotation is finished, the model moves attention to the updated target location in the angle chunk in the imaginal buffer (step 11), encodes the cardinal direction, maps that direction to a response, and makes the appropriate keypress. After the target is identified, the process of responding is identical to that used by the model of the counting strategy.

The mental rotation in this model consists of iterative updates to the location of the target in the angle chunk in the imaginal buffer. This updating is based on the displacement of the plane from the bottom of the screen (135 degrees to the right in the sample trial in Figure 12). The location of the endpoint of the target segment of the angle is updated according to the extent of this disparity. The target location is updated by iteratively adjusting it to correspond to adjacent cardinal directions on the map. This process is illustrated in Figure 12 (steps 8 to 10). This mechanism is an approximation of analog mental rotation, which would involve continuous updates to this location.

In the rotation strategy, left-right confusion arises during the solution process when the *target* is located in the upper portion of the camera view (targets located in the far-left and far-right; recall that the shortcut strategy applies when the target is at the top). Since the model immediately updates the location of the plane in this angle when it locates it on the map (step 6), it only rotates the leg of the angle associated with the target location (steps 8 to 10). And, since the model rotates only the target leg and not a coherent angle, left-right confusions can arise. The situation in Figure 11 is an opportunity for such a confusion, since the target is in the far-left of the camera view. In this trial, the plane leg of the angle has been moved to the right, but the rotation of the target leg of the angle begins by moving it to the left (step 8). The confusions that can arise in such situations are identical in nature to those described for the counting strategy.

Model Details and Parameters. The two preceding sections have described the operation of the model for the two strategies taught to participants. What was left out of these descriptions was the special-case strategy that all participants reported using. This "shortcut" was also incorporated into the model. So, when the model encounters a target in line with the plane (at the top or bottom of the target field on the camera view), it does not bother to count or form an angle. Instead, the model immediately finds the plane on the map view. If the target was at the bottom of the camera view, the cardinal direction associated with the location of the plane is the answer, while the answer is the cardinal direction immediately opposite if the target was at the top of the camera view. Once the correct answer is determined, the process of responding occurs in the same way as described above.

In general, the performance of the model is a straightforward result of the execution of the strategies. By having the model use the reported strategies, the basic pattern of data is a parameter-free outcome. There are, however, two parameters that were estimated to fit the model to the response time data. A third parameter was estimated in fitting the eye data, but that parameter does not affect the model's behavior and will be described below. First, the time required to retrieve a piece of information (chunk) from declarative memory was set at 110ms for both strategies. The model depends on stored declarative knowledge to do the task. For instance, it has knowledge relating to cardinal directions, left and right, and what the information on the screen represents. When a piece of this knowledge is needed by the model to do the task, it executes a retrieval, placing the chunk in the retrieval buffer where it can be accessed. Each time the model retrieves one of these chunks from declarative memory, it takes 110ms. This has less of an effect on the trials that involve the short-cut strategies, as they require fewer retrievals to complete than the other trials. But, in general, this parameter affects the overall time needed to do the task.

The second parameter influences the time needed to change the search direction in both strategies in trials where left-right confusion is an issue (see above). In both models, the process is similar and requires extra steps, including an extra production. The production that actually executes the update to the search direction was given an execution time of 200ms. However, the whole impact of left-right confusion in the model also includes 2 additional retrievals, which results in a total effect of approximately 420ms.[3] All of the other parameters that influence the behavior of the model were given their default ACT-R values. ACT-R has a number of parameters associated with its performance, but the ACT-R theory has default values for these and they do not need to be estimated. For instance, ACT-R has default times for executing keypresses, shifting visual attention, and encoding visual information (many of these are derived from the EPIC theory; Kieras and Meyer, 1997). These values are based on vast amounts of psychophysical research that has looked at these issues in the past. None of these values were changed in the model presented here. The next section compares the model to the participant data, both in terms of response times and eye movements.

---

[3]To resolve the left-right confusion and switch the direction of counting/rotation, the model retrieves a declarative chunk relating to the encoded direction (left or right). In this chunk is the identity of the opposite direction, which is used to update the direction for counting or rotation. Then, a chunk from declarative memory is retrieved to reestablish a state where the searching can begin. As described, this whole operation adds one production (with a 200ms execution time) and 2 retrievals (110ms each) over situations where no change is necessary.

Comparison of Model to Data

Response Times. The model was compared to the combined data from the two studies described above. The average data for target location and misalignment are presented in Figure 13, along with the model's predictions. For both effects, the model does a fairly good job of predicting the response times of participants (r = .97, *RMSD* = 92ms for the effect of target location; r = .98, *RMSD* = 118ms for the effect of misalignment). The slopes of these effects across both experiments are shown in Table 3, along with the model's predictions. Overall, the model does a good job of predicting these data as well.

The model's performance depends on executing the strategies and actually doing the task. The special-case strategies that were implemented in the model (based upon the participants' verbal reports) account for the comparatively fast response times for targets located at the bottom and top of the camera view. In these trials, the model encodes the answer as being the same cardinal direction where the plane is or the one directly across from it. Then, the model finds the plane on the map and can quickly identify the correct answer.

For the counting strategy, the linear increase in response time as a function of target location results because the model does more counting when the target is located farther from the plane (Figure 13a). Here, the model counts to the target on the camera view and then to the answer on the map view. In both cases, the model has to find the next item on the screen, move visual attention to that object, and increment a count. This takes 185ms[4] for each iteration of the count on both views, producing a slope of approximately 370ms per counting step (Table 3). Once again, this effect does not show up for targets at the top of the camera view since the special-case strategy applies. For the effect of misalignment, the counting model takes longer for northerly plane locations because of left-right confusion (Figure 13b). The details of this are described above.

For the rotation strategy, there is a linear effect of the plane's location (misalignment) because the image formed on the camera view is rotated as a function of where the plane is on the map (Figure 13b). In this process, the model iteratively updates the endpoint of the target leg of the angle in the angle chunk in the imagery buffer. For each update, the model locates the nearest cardinal direction in the proper direction and replaces the initial endpoint with the location of the new cardinal direction. This effect reflects the time needed by the model to perform these actions (approximately 260ms;[5] Table 3). Again, this mechanism is an approximation of continuous mental rotation. Meanwhile, there is also an effect of the target's location (Figure 13a). These data can be

[4]This process requires 2 productions (50ms each) and a shift of attention (85ms). These execution times are the ACT-R default values.

[5]This operation involves 3 productions and a retrieval, so the slope is partially dependent on the parameter set to control the retrieval time for chunks from declarative memory. The 50ms production execution time is ACT-R's default value.

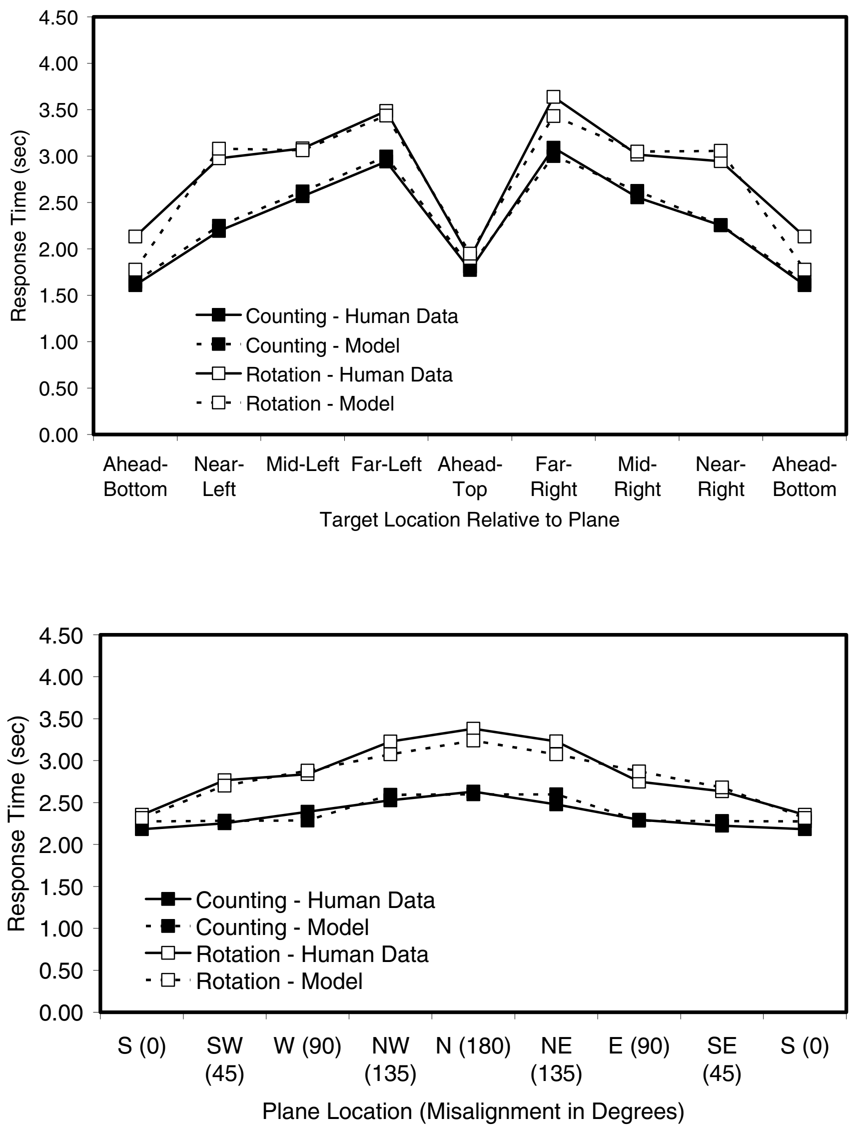

Figure 13. Comparison of the model's performance to the response time data produced by participants for Experiments 1 and 2 combined.

Table 3
Average Slopes (ms/45º) of the Effects Combined across Experiments 1 and 2. The Model's Predictions Are in Parentheses.

Factor	Condition	
	Rotation	Counting
Target Location	276 (183)	392 (372)
Misalignment	213 (223)	97 (96)

characterized as resulting from three situations relating to the target position on the camera view. First, for targets at the bottom or the top, the special-case strategy applies, making response times in these cases fast relative to others. Then, for near-left, near-right, mid-left, and mid-right targets, response times are somewhat longer. This relates to both the effort needed to form the angle and the difficulty in performing the mental manipulations to that angle on the map view. Finally, response times for targets in the far-left or far-right are longer still, owing to the left-right confusion issue described above. This complexity is in addition to the effort needed to form and manipulate the angle, which is why response times in these trials are the longest.

In addition to the main effects, there were interactions between the target's location and misalignment in both strategy conditions. These data and the model's predictions are shown in Figure 14a and 14b for the counting and rotation strategies respectively. In both cases, the interaction arises in the model because of the influence of the special-case strategy. The special-case strategy applies when the target is located at the bottom or the top of the camera view (the first and last points on each of the lines in Figure 14). For these trials, response times are the same in the model regardless of the orientation on the map view. For other target locations, however, misalignment does have an impact on performance. For the counting strategy, the model essentially produces two lines, one line for trials where left-right confusions do not arise, and one where it does (Figure 14a). When the data from participants are averaged over these two situations, the model provides a close fit to the human data ($r = .99$; RMSD = 135ms). The data from the participants without this averaging are illustrated for the individual experiments in Figures 4 and 8. In the rotation strategy, the effect is regular, with response times increasing linearly as a function of the misalignment between the two views when the special cases are ignored (the different lines in Figure 14b). However, note that when the plane is in the South (the two views are aligned), there is no increase in response time in the model for "far" targets. In these cases, no rotation is required, so the model does not need to resolve left-right confusion. Again, the model's performance does a good job of matching the human data ($r = .95$; RMSD = 279ms). Notice that the overall vertical separation between the lines is greater for the rotation strategy and the slope of the lines is steeper for the counting strategy. These features illustrate the larger effect of misalignment for the rotation strategy (vertical separation) and the larger effect of target location for the counting strategy (slopes). These predictions from the model correspond well to the data produced by participants in the experiment.

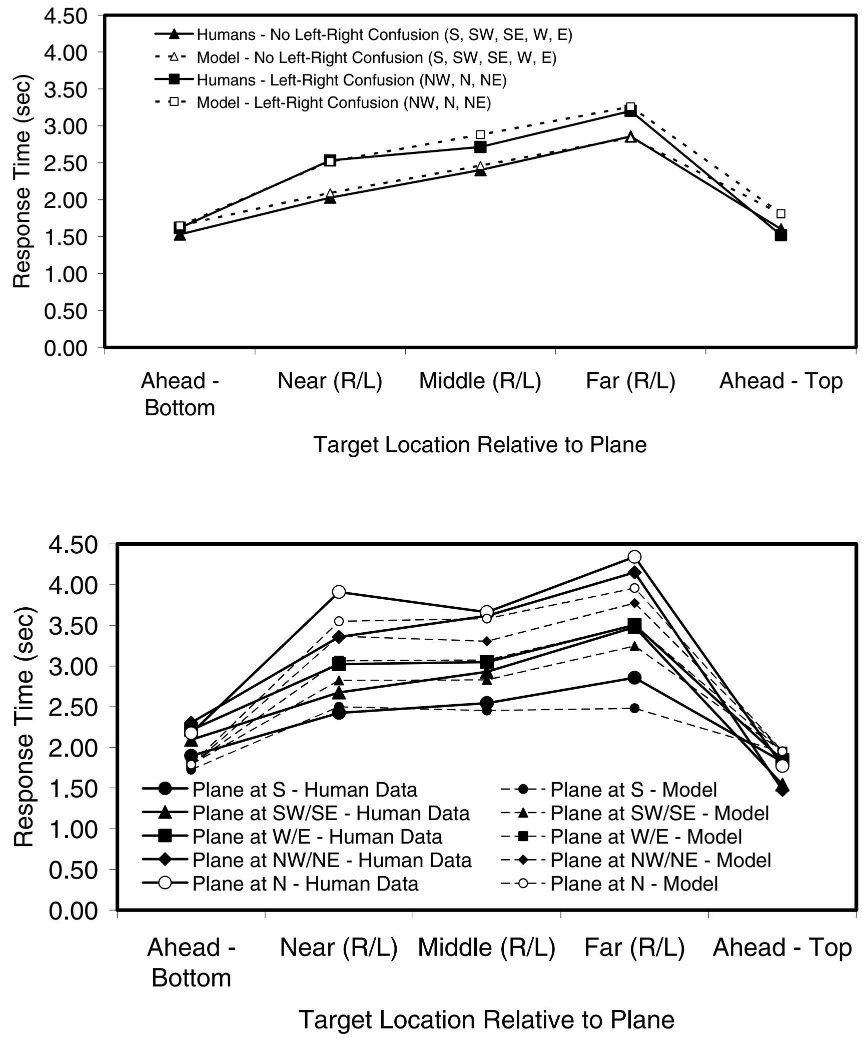

Figure 14. Response time data for Experiments 1 and 2 and the model, showing the interaction of target position and misalignment for the counting strategy (top) and for the rotation strategy (bottom), averaged across left and right deviations for both factors.

Eye Data. In terms of eye movements, the model attends to the information on the screen that it is processing at that point in its solution. So, in the counting strategy, the model attends the items on the screen as it performs the counting. In the rotation strategy, the model attends the appropriate regions of the screen as it encodes the points that form the angle, attends the location of the plane as it updates that leg of the angle, and attends the center of the map view as it rotates the other leg. Both models attend the answer as they encode and make their responses. In this sense, the eye movements the model produces are extremely indicative of what it is doing to solve the task. We feel that the same is true of participants in the study. However, ACT-R's visual attention is currently persistent. When ACT-R moves its visual attention to an item on the screen, it attends to it continuously until it is given the command to attend somewhere else. This is clearly not how human eye movements operate. Humans blink, look away, glance at various regions haphazardly, and take some time to execute a saccade (while ACT-R takes time to shift visual attention, it does not produce intermediate POR's during the shift). At this point, ACT-R itself cannot predict these random movements. In addition, the eye-tracking equipment is not perfect, meaning that the eye data have some biases (see above).

In order to have the ACT-R model account for the eye data in this study, a parameter was used to control the proportion of eye samples that were "on-task" versus "off-task". This value was set to .5. This means that the locations where ACT-R directs its persistent attention contributes 50 percent of the prediction for the eye data. The other 50 percent is randomly distributed across the entire screen. Using only this single parameter, the model makes very good predictions of what the eye data should look like for both strategy conditions (Figure 15). Overall, the correlation for these data is .88, with a RMSD of .028 (2.8%). This is a sizeable improvement over both pure chance (r = .52, *RMSD* = .087; Figure 9) and the model without any error added (r = .44, *RMSD* = .098). These data add credence to the theoretical account of participant performance in this task.

Looking more closely at the eye movement data for the model in the two conditions, the differences in how the model performs the task can be seen. First, the proportion of time spent by the model attending the "other" regions is entirely a function of the parameter set to control on-task versus off-task samples. In fact, without this parameter, the model (using either strategy) never attends these regions. In addition, using the counting strategy, the model never attends the center of either view and using the rotation strategy, the model never attends the regions between the target and plane on either view. So, the real differences that are predicted by the model correspond to the proportion of time spent attending the center of both views and the regions between the plane and the target in both views. When using the rotation strategy, the model spends more time attending the center of the camera view because this is where the model attends when it encodes the vertex of the angle. It attends the center of the map view longer because this is where it attends when it is executing the rotation. In contrast, when using the counting strategy, the model attends more at the regions between the plane and the target on both views because it attends to those areas as it counts to the target on the camera view and to the answer on the map view. These differences were found in the data although the model is sometimes more extreme than the participants. Combined with the response time data, the model corresponds quite closely with the behavioral data presented above.

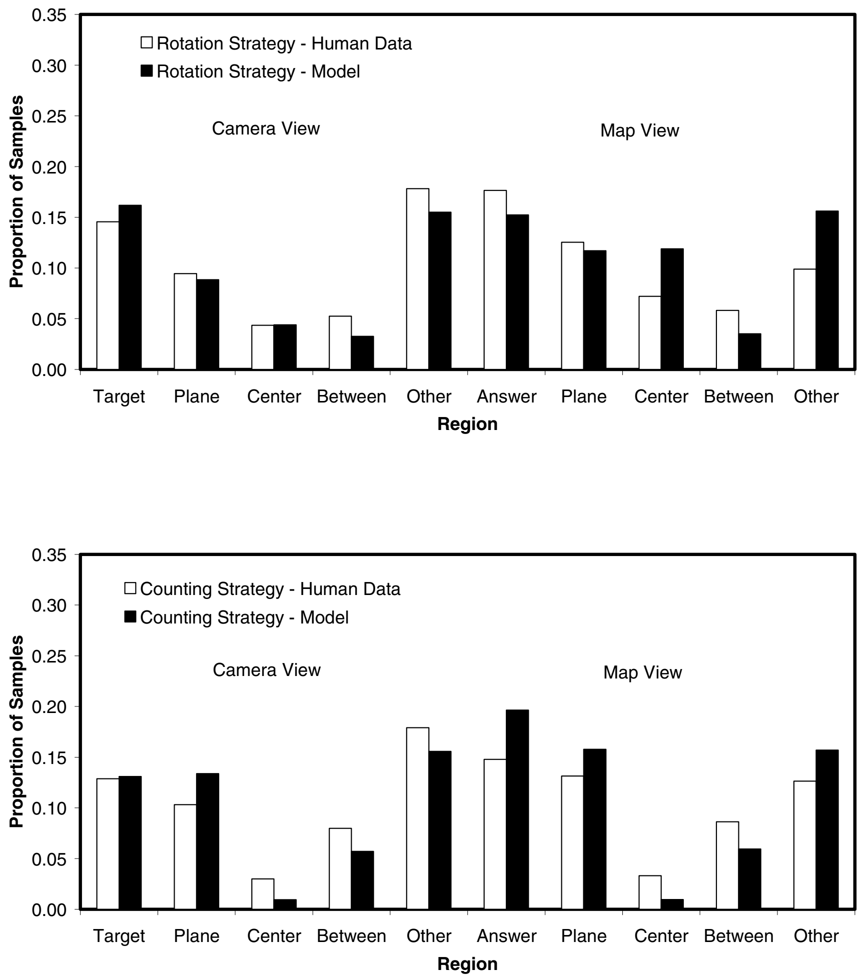

Figure 15. Comparison of the model to the data produced by participants in terms of average proportion of samples falling within each functional region.

General Discussion

The experiments presented here provide evidence that individuals vary in how they solve an orientation task. While many of the participants in the pilot study used mental rotation to complete the trials, others opted for a verbal approach that avoided the need for imagery-based cognitive operations. In addition, we found evidence that although these strategies produce somewhat similar results overall, there are reliable differences in performance, both in terms of response times and eye movements. An ACT-R model that uses these strategies reproduces all of the major trends in the data, and provides a close fit to the response times and eye movement data.

While previous researchers have focused on mental imagery and rotation in explaining performance on these tasks, the data collected here clearly demonstrate that such operations are not necessary to produce fast and accurate solutions. This finding adds to research that has demonstrated the influence of language on performance in spatial tasks (Bethell-Fox and Shepard, 1988; Gallistel, 2002; Li and Gleitman, 2002; Sholl and Egeth, 1981; Taylor and Tversky, 1996). In fact, participants using the verbal counting strategy performed significantly faster than those using the imagery-based rotation strategy overall, $F(1, 48) = 7.62$, $p < .01$. In addition, the fine-grained eye movement data exposed additional details about how participants were doing the task, helping to differentiate between alternative accounts. Still, the effects produced by participants using both strategies are similar in many ways and have the appearance of rotation functions. It would be easy to mistakenly attribute the results to mental rotation without knowing the instructions that were given to participants, even though only one of the functions for the model in Figure 13 is produced by mental rotation in the model (i.e., the effect of misalignment for the rotation strategy). Thus, it is not surprising that past research in this area has tended to ignore strategy differences.

The model provides good support for the theoretical account. It hinges on strategic variability and the possibility of completing orientation tasks without resorting to mental rotation. This research illustrates that explanations relying solely on mental rotation are missing an important aspect of variability and performance on such tasks. The model also provides a detailed explanation of human performance, producing a close correspondence to the data in terms of both response times and eye movements. And, the model performs the task by using primarily the same mechanisms that have been used to account for human performance in many other domains (see Anderson and Lebiere, 1998). The lone exception to this is in the model of the rotation strategy, which extends the ACT-R theory by incorporating an imaginal buffer to allow ACT-R to hold representations of images in memory. It is interesting to note that the operation of this buffer is analogous to others in ACT-R. The fact that so much of the model derives from basic ACT-R mechanisms supports the idea that participants tend to use well-practiced perceptual and memory-based processes when they can, to avoid using more complicated procedures (e.g., Kirsh and Maglio, 1994).

The model actually does the task, using the same interface as the participants, and so demonstrates that we have a complete account of task performance without significant holes that could be overlooked in a model that dealt with an abstraction of the actual task. Rather, the model acts just like a participant, gathering information from the screen, operating on it, and eliciting responses by making key presses. All of these aspects of performance are important in the task presented here. Using such an implementation allows the model to produce robust quantitative predictions of performance, and provides a more comprehensive account of human performance in this area than has been achieved previously. In this sense, the eye-tracking data provide crucial evidence about performance on this kind of task. They illustrate how participants distribute their attention during the trial, and show that participants using different strategies emphasize different parts of the screen as they do the task. The fact that the model is able to match these different distributions of attention supports the theoretical account that it embodies. In addition, the ability to use a cognitive architecture, like ACT-R, to evaluate such fine-grained data is an important step in the evolution of cognitive modeling.

The empirical data presented here closely resemble the findings presented by Hintzman et al. (1981), who proposed a theoretical account based on mental imagery. In the experiment they conducted that was most similar to these (their Experiment 1), they did not find a significant effect of the target's location on response time (when they ignored targets in line with the viewpoint), though the trend was in the same direction. So, their explanation did not consider the location of the target beyond the special cases. Interestingly, Hintzman et al. also conducted a number of other experiments involving cognitive maps of 8 different objects arranged in a circle (Experiments 3–5; 8–13). In these experiments, the target's location relative to the orientation did influence performance, as in our research. They concluded that there was a process of inference that followed the mental rotation, which was needed to identify the appropriate response item. They posit that the *position* of the object after the mental rotation was available, but that its *identity* was not. So, their account indicates that after the rotation participants had to execute a search of memory to determine which of the 8 objects occupied the location. In the experiments presented here, however, the position of the rotated target is all that is needed to make the correct response. Thus, the results cannot be attributed to a search of memory for the correct item for the given location.

It is also worthwhile to compare this research to studies that have examined different imagery-based strategies in orientation tasks (Carpenter and Proffitt, 2001; Huttenlocher and Presson, 1979; Presson, 1982; Wraga, et al., 2000). In those studies, *array rotation* and *viewer rotation* strategies were described. Array rotation involves imagining the objects rotating, while the viewer remains stationary. While the materials are certainly different in the research described here, the rotation strategy that some of our participants were instructed to use is comparable to array rotation. Participants were asked to imagine an angle rotating, which contains information about the locations of objects in the target field (or array).

On the other hand, a component of the counting strategy described here can be related to the viewer rotation strategy. In viewer rotation, individuals imagine their location in the space changing, and then locate objects in the space according to the new vantage point. This involves updating the viewer's spatial frame of reference to match the new location. The model embodies a verbal approach to updating the spatial frame of reference, but imagery-based mechanisms are also possible. In our task, participants could resolve left-right confusion by imagining self-rotation in the space. Once this process has been completed, the verbal encoding of the target's location could be applied on the map based on the updated frame of reference. Since the target's location is encoded as a position relative to the viewer in this strategy, that description can be applied directly once the update is completed, based on the new orientation. The relationship between the verbal spatial updating described here and the imagery-based viewer rotation strategy warrants consideration in future research.

In conclusion, the experiments and model presented here provide an alternative to theoretical accounts of orientation that rely exclusively on mental imagery and rotation. In the pilot study, many participants did use such strategies, but some developed a different approach that was equally effective for doing the task. In addition, the finding that both strategies seem to reduce the complexity of the task to something more like basic mental rotation is interesting. By ignoring some of the details of how the two views actually correspond, participants were able to significantly simplify their task. As a result, they were able to use straightforward strategies to determine the correct answers. It is likely that individuals do this in more realistic environments as well.

Acknowledgments

This research was conducted as part of the first author's dissertation while at Carnegie Mellon University. It was supported in part by grant number F49620-99-1-0086 from the Air Force Office of Scientific Research (AFOSR) to the second author and by National Research Service Award (NRSA) training grant number 5-T32-MH19983 from the National Institute of Mental Health (NIMH) to Dr. Lynne Reder. Some of the preparation of this manuscript was done while the first author was a National Research Council Posdoctoral Research Associate at the Air Force Research Laboratory in Mesa, AZ. Portions were presented at the 24th Annual Conference of the Cognitive Science Society in Fairfax Virginia (August, 2002). We would like to thank Roberta Klatzky, Myeong-Ho Sohn, and Marlene Behrmann for their comments on earlier drafts of this paper.

References

Abu-Obeid, N. (1998). Abstract and scenographic imagery: The effect of environmental form on wayfinding. *Journal of Environmental Psychology, 18,* 159–173.

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., & Lebiere, C. L. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum.

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*, 12–23.

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

Carpenter, M., & Proffitt, D. R. (2001). Comparing viewer and array mental rotations in different planes. *Memory & Cognition, 29*, 441–448.

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

Easton, R. D., & Sholl, M. J. (1995). Object-array structure, frames of reference, and retrieval of spatial knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21,* 483–500.

Fenner, J., Heathcote, D., & Jerrams-Smith, J. (2000). The development of wayfinding competency: Assymtrical effects of visuo-spatial and verbal ability. *Journal of Environmental Psychology, 20,* 165–175.

Gallistel, C. R. (2002). Language and spatial frames of reference in mind and brain. *TRENDS in Cognitive Sciences, 6*, 321–322.

Glicksohn, J. (1994). Rotation, orientation, and cognitive mapping. *American Journal of Psychology, 107,* 39–51.

Gopal, S., Klatzky, R. L., & Smith, T. R. (1989). NAVIGATOR: A psychologically based model of environmental learning through navigation. *Journal of Environmental Psychology, 9,* 309–331.

Gugerty, L., deBoom, D., Jenkins, J. C., & Morley, R. (2000). Keeping north in mind: How navigators reason about cardinal directions. In *Proceedings of the Human Factors and Ergonomics Society 2000 Congress* (pp. I148–I151). Santa Monica, CA: Human Factors and Ergonomics Society.

Gunzelmann, G., & Anderson, J. R. (2003). Problem solving: Increased planning with practice. *Cognitive Systems Research, 4*, 57–76.

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

Huttenlocher, J., & Presson, C. C. (1979). The coding and transformation of spatial information. *Cognitive Psychology, 11,* 375–394.
Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8,* 441–480.
Just, M. A., & Carpenter, P. A. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92,* 137–172.
Kieras, D. E., & Meyer, D. E. (1997). An overview of the EPIC architecture for cognition and performance with application to human-computer interaction. *Human-Computer Interaction, 12,* 391–438.
Kirasic, K. C., Allen, G. L., & Siegel, A. W. (1984). Expression of configurational knowledge of large-scale environments: Students performance of cognitive tasks. *Environment and Behavior, 16,* 687–712.
Kirsh, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science, 18,* 513–549.
Klatzky, R. L. (1998). Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel, and K. F. Wender (Eds.). *Spatial cognition: An interdisciplinary approach to representing and processing spatial knowledge* (pp. 1–17). New York: Springer-Verlag.
Klatzky, R. L., Loomis, J. M., Beall, A. C., Chance, S. S., & Golledge, R. G. (1998). Spatial updating of self-position and orientation during real, imagined, and virtual locomotion. *Psychological Science, 9,* 293–298.
Li, P., & Gleitman, L. (2002). Turning the tables: Language and spatial reasoning. *Cognition, 83*, 265–294.
Loftus, G. R. (1978). Comprehending compass directions. *Memory & Cognition, 6,* 416–422.
Lovett, M. C., & Anderson, J. R. (1996). History of success and current context in problem solving: Combined influences on operator selection. *Cognitive Psychology, 31,* 168–217.
McNamara, T. P., & Diwadkar, V. A. (1997). Symmetry and Asymmetry of human spatial memory. *Cognitive Psychology, 34,* 160–190.
Murakoshi, S., & Kawai, M. (2000). Use of knowledge and heuristics for wayfinding in an artificial environment. *Environment and Behavior, 32,* 756–774.
O'Neill, M. (1991). A biologically based model of spatial cognitive and wayfinding. *Journal of Environmental Psychology, 11,* 299–320.
Presson, C. C. (1982). Strategies in spatial reasoning. Journal of Experimental Psychology: Learning, Memory, and Cognition, 8, 243–251.
Reder, L. M. & Schunn, C. D. (1999). Bringing together the psychometric and strategy worlds: Predicting adaptivity in a dynamic task. In D. Gopher and A. Koriat (Eds). *Cognitive regulation of performance: Interaction of theory and application. Attention and Performance XVII* (pp. 315–342). Cambridge, MA: MIT Press.

Richardson, A. E., Montello, D. R., & 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.

Ritter, F. E., Baxter, G. D., Jones, G., & Young, R. M. (2000). Supporting cognitive models as users. *ACM Transactions on Computer-Human Interaction, 7*, 141–173.

Rossano, M. J., West, S. O., Robertson, T. J., Wayne, M. C., & Chase, R. B. (1999). The acquisition of route and survey knowledge from computer models. *Journal of Environmental Psychology, 19,* 101–115.

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

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

Sholl, M. J. (1987). Cognitive maps as orienting schema. Journal of Experimental Psychology: Learning, Memory, and Cognition, 13, 615–628.

Sholl, M. J., & Egeth, H. E. (1981). Right-left confusion in the adult: A verbal labeling effect. *Memory & Cognition, 9*, 339–350.

Siegler, R. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General, 116,* 250–264.

Sohn, M., & Carlson, R. A. (2003). Viewpoint alignment and response conflict during spatial judgment. *Psychonomic Bulletin & Review, 10,* 907–916.

Taylor, H. A., Naylor, S. J., & Chechile, N. A. (1999). Goal-specific influences on the representation of spatial perspective. *Memory & Cognition, 27,* 309–319.

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of memory and language, 35*, 371–391.

Tversky, B. (2000). Remembering spaces. In E. Tulving and F. I. M. Craik (Eds.). *The Oxford handbook of memory* (pp. 363–378). New York, NY: Oxford University Press.

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

Wraga, M., Creem, S. H., & Proffitt, D. R. (1999). The influence of spatial reference frames on imagined object and viewer rotations. *Acta Psychologica, 102,* 247–264.

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

APPENDIX

Directions for strategies given to participants as training in Experiments 2 and 3.

Steps in the Counting Strategy:

Locate the target on the camera view and calculate the number of steps (positive for clockwise or negative for counterclockwise) from the plane to the target.

Find the plane in the map view and count around the map view the number of steps (clockwise for positive or counterclockwise for negative) to the target.

Respond with that location.

Steps in the Rotation Strategy:

Locate the target on the camera view and form the angle joining the target to the plane with the vertex at the center of the target field. Remember to keep track of whether the angle opens clockwise or counterclockwise.

Mentally rotate that angle and overlay it on the map view so that the location of the plane matches the location of the plane in the map view.

Find the "target location" end of the angle. This is the correct answer.