Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Integrating Multiple Strategies Efficiently to Solve an Orientation Task

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 27(27)

ISSN

1069-7977

Authors

Anderson, John R.
Gunzelmann, Glenn

Publication Date

2005

Integrating Multiple Strategies Efficiently to Solve an Orientation Task

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

John R. Anderson (ja+@cmu.edu)
Department of Psychology
Carnegie Mellon University
Pittsburgh, PA 15213 USA

Abstract

This research compares the general strategy described by participants doing an orientation task to two strategies described in past research on a different kind of spatial task, perspective-taking (*array rotation* and *viewer rotation*). This evaluation indicated that participants were quite flexible and efficient in their approach to the task. The strategy described in participants' verbal reports made use of both of the perspective-taking strategies within individual trials. In addition, each alternative was applied in situations where previous research indicates that it holds an advantage over the other alternative. This research extends research on strategy use in spatial tasks by (1) showing how similar strategies can be applied to different kinds of spatial tasks and (2) illustrating how alternative strategies can be intermixed within a single task to produce efficient overall performance.

Introduction

Research on human performance in spatial orientation tasks has focused on the impact of misalignment on solution processes (e.g., Hintzman, O'Dell, & Arndt, 1981; Rieser, 1989; Shepard & Hurwitz, 1984). Other research has examined strategy differences in this area, showing that strategy variation can have important influences on performance (e.g., Gunzelmann & Anderson, 2004a; Huttenlocher & Presson, 1979; Just & Carpenter, 1985; Presson, 1982; Wraga, Creem, & Proffitt, 2000). This research typically uses instructional manipulations to encourage participants to use different strategies. Although this approach has been useful for uncovering differences in performance as a function of strategy use, it also leaves open the question of how strategies are selected by individuals to arrive at the solution. One motivation for this paper is to examine verbal reports of strategy use in an attempt to determine the extent to which efficiency (speed and accuracy) influences strategy selection in individuals solving spatial tasks.

Some research in the area of spatial cognition has attempted to identify the strategies individuals used. In many cases, this research has explored human performance on navigation tasks, using map-drawing or other tasks to infer how participants learn and represent routes through a space (e.g., Aginsky, Harris, Rensink, & Beusmans, 1997; Murakoshi & Kawai, 2000). Our study is somewhat different than those previous studies in that it does not involve moving through a space, either by real or virtual navigation. Instead, this research looks at performance on an orientation task where participants integrate information from different static representations of a space to make a spatial judgment. In this case, a visual scene and a map of the space were shown to participants. One of the objects in the visual scene was highlighted, and the task was to identify which of the objects on the map corresponded to that highlighted object. A sample trial is shown in Figure 1. The task is described in more detail below.

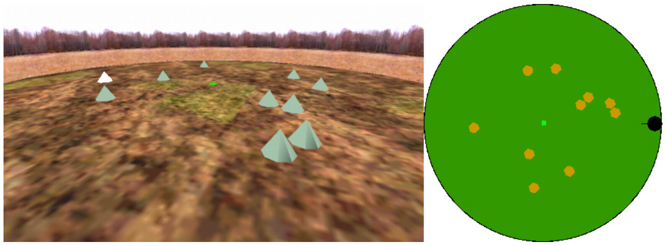

Figure 1: Sample trial. Participants click on the object on the map corresponding to the target.

To better understand human performance on this task, verbal reports were gathered from participants after they finished. In previous research, we have used these verbal reports to infer the general strategy that the participants were using to solve the task. This strategy is described below. It provides support for a theoretical explanation for participant performance on this kind of task (see Gunzelmann & Anderson, in press). In addition, the predictions of this strategy for performance have been validated against the human data using a computational cognitive model developed in ACT-R (Gunzelmann & Anderson, 2004b).

With a validated strategy for performing orientation tasks, it is possible to explore the relationship between it and strategies that have been described for other types of spatial tasks. This kind of comparison has not been performed in the past. In the next section, we briefly describe perspective-taking tasks and two strategies that have been described for doing them, *array rotation* and *viewer rotation* (Huttenlocher

& Presson, 1979; Presson, 1982; Wraga, et al., 2000). Then, after describing our experiment, we compare the perspective-taking strategies to the general strategies described by participants in our study for solving the orientation task. Although there are differences between orientation and perspective-taking tasks, they share important features as well. A careful analysis indicates that the strategy reported by participants for solving the orientation task consists of a combination of the perspective-taking strategies, executed in sequence.

Perspective-Taking Tasks

Perspective-taking tasks require the participant to identify what a display (e.g., an array of objects) would look like from a different viewpoint, or after it was rotated. The most recognizable example of such a task in psychological research is the Piagetian 3-mountains task (e.g., Piaget & Inhelder, 1956). In this task, the participant is asked to select an image that represents what a display (consisting of 3 mountains with identifiable characteristics) would look like from the perspective of another viewer. This requires the viewer to imagine how the components of the display would be arranged when viewed from that other perspective. Research has used a variety of variations of this basic problem to examine different strategies.

Strategies in Perspective-Taking As noted above, the two strategies that have received the most research attention in the perspective-taking literature are array rotation and viewer rotation (e.g., Huttenlocher & Presson, 1979; Presson, 1982; Wraga, et al., 2000). In array rotation, the participant makes a judgment by imagining the objects in the array rotating relative to its own axis, whereas viewer rotation involves imagining the rotation of the viewpoint around the array. In the first case, the question might be, What would this display look like from where you are now if the display were rotated by 180 degrees? In the latter case the question becomes, What would this display look like if you were standing on the opposite side of it?

Both the array rotation and viewer rotation strategies involve mental rotation, and should have the same computational complexity. However, Huttenlocher & Presson (1979; Presson, 1982) demonstrated that participants do not treat these two situations equivalently. That is, in some situations performance was better when participants were instructed to use the array rotation strategy, whereas in other cases instructions to use the viewer rotation strategy produced superior performance.

A key factor that influenced the relative difficulty of the two strategies was the type of question given to participants. In some cases, participants were asked *item questions*. These questions require participants to indicate the location of one of the items in the array after the transformation. For instance, which item is on the left? Or, where is the book? Other times, participants were asked *appearance questions*. These are questions which require knowledge of all of the items in the array. For instance, the question typically posed in the 3-mountain task is an appearance question (e.g., What would this scene look like from “over there”?). The important finding from this research for current purposes is that participants were more accurate on item questions when they were instructed to use the viewer rotation strategy, and more accurate on appearance questions when they were asked to use the array rotation strategy.

Experiment

The experiment described here is presented in more detail in Gunzelmann and Anderson (2004b; in press). Participants were asked to solve an orientation task, which involves integrating an egocentric visual scene with a map of the space. A target was identified as one of 10 objects in the visual scene, and participants indicated which of the objects on a map of the space corresponded to the target. Figure 1 shows a sample trial from this experiment. The exact design of the experiment is described below. However, an important feature of the stimuli is that the objects in the space were arranged in a somewhat irregular manner, making it difficult to use strategies that are based on regular distributions of objects in the space (e.g., Gunzelmann & Anderson, 2004a).

Orientation tasks differ in several ways from perspective-taking tasks. Most importantly, in perspective-taking participants are asked to imagine that the relationship between them and the display changes, producing a different situation. In orientation tasks, the relationship between the viewer and the space remains constant. However, participants must take information from one frame of reference and apply it to a different frame of reference. Despite the differences, both of these tasks require mental transformations to determine the appropriate response. Also, the transformations frequently involve mental rotation. It is this aspect of the solution process where the closest similarities lie.

Method

In each trial, participants were shown a visual scene and a map. There were 10 objects in the space on each trial, and all 10 objects were visible in both views. In the visual scene, one of the objects was highlighted in red to identify it as the target (it is white in Figure 1). To facilitate performance, the viewer’s position was indicated on the map as a black dot. The viewer was always positioned at the edge of the space, looking toward the center. The task was to identify the target on the map of the space. Participants made their responses by clicking on their answer on the map. Response times and accuracy were recorded. Twenty individuals participated, ranging in age from 17 to 31 (mean age = 21.9). Each participant was paid $10, and the experiment lasted no more than 1.5 hours.

The stimuli were designed by placing the objects into quadrants in the space. The quadrants were defined relative to the viewer, using either the main axes (horizontal and vertical) or the oblique axes (diagonal; this is the case in Figure 1) to divide them. Objects were placed randomly around a central point in the quadrants under the constraint that the four quadrants contain 1, 2, 3, and 4 objects respectively in each trial. The configuration of those quadrants relative to each other and relative to the viewer was counterbalanced (24 different maps).

The target could appear in any of the four quadrants (among 0, 1, 2, or 3 other objects in the same quadrant), and the degree of misalignment between the two views was varied in 90° increments (0, 90, 180, and 270; determined by the viewer's location, at the bottom, left, top, or right of the map respectively). The resulting design contained 768 trials (2 quadrant alignment conditions, 24 different maps, 4 target locations, and 4 misalignments). For the data presented below, the 8 target locations represent the four quadrants relative to the viewer, crossed with the two quadrant alignment conditions.

In this study, participants completed half of the possible trials. Each participant was assigned to one of the quadrant alignment conditions and completed all 384 of the trials in that condition (10 participants were randomly assigned to each condition). Though it is not critical for this paper, the participants from the two conditions were ranked based on their scores on an assessment of spatial ability (Vandenberg & Kuse, 1978). These rankings were used as the basis for combining the data from participants in the two groups to create "meta-participants". The statistics presented in the results below are based on these meta-participants (Gunzelmann & Anderson, in press describes this in more detail).

Participants completed all 384 trials using a drop-out procedure. If an error was made on a particular trial, it was repeated later in the experiment under the constraint that the same trial was never presented twice in a row. The experiment was broken into blocks of 20 trials, and participants were permitted to take a short break between blocks. Once they finished, participants were questioned about how they solved the task. As part of this interview, they were given sample problems, and were asked to describe their solution process aloud. The experimenter followed up on vague responses to get a clear sense of the strategy being used. The verbal report data described here consist of summaries of participants' responses during this interview.

Results

In this section, the accuracy and response time data are discussed only briefly (see Gunzelmann & Anderson, in press for more detailed analyses of these data), focusing instead on the verbal reports from participants. These data were used to identify the strategy that participants were using. In conjunction with previous research on strategies in perspective-taking tasks, this effort leads to a better understanding of why participants performed the task in the manner they did.

Participants generally performed quite well in this experiment. Overall accuracy was 96%. And, the error data generally followed the same trends as the response time data (r=.83), supporting the conclusion that the outcomes were not the result of a speed-accuracy trade-off. The response times presented here include only correct responses. They indicated that there were several factors that influenced participants' performance on the task.

ANOVAs indicated that the target's location, misalignment between the two views, and the number of other objects in the quadrant all had significant influences on participants' response times in this study, $\underline{F}(7,63)=11.39$, $p<.01$ (MSE=4.45sec^2), $\underline{F}(3,27)=38.62$, $p<.001$ (MSE=6.47sec^2), and $\underline{F}(3,27)=60.67$, $p<.001$, (MSE=2.41sec^2), respectively. Participants took longer to respond when the target was off to the side of the viewpoint or farther from the viewer, when misalignment was greater, and when more objects were located in the same quadrant as the target. The data are presented in Table 1 and Figure 2. There was also an important interaction in the data. The magnitude of the effect of misalignment was larger when more objects were in the same quadrant as the target, $\underline{F}(9,81)=8.79$, $p<.001$. This effect is shown in Figure 2 as well.

Table 1: Mean response times (seconds) as a function of the location of the target relative to the viewer.

Target Location Relative to Viewer	Mean RT
Close in front	2.72 sec
Close to the side	3.90 sec
Middle to the side	3.90 sec
Far to the side	4.34 sec
Far in front	3.22 sec

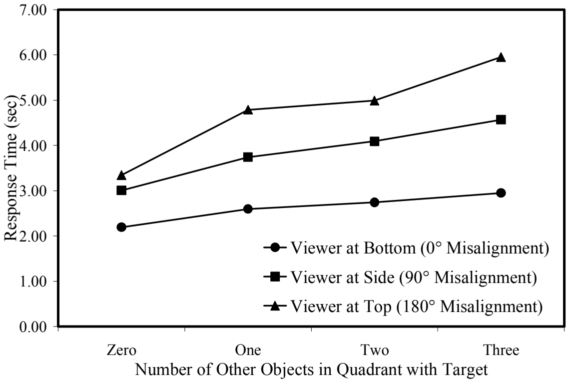

Figure 2: Mean response times (seconds) as a function of misalignment and the number of objects in the same quadrant as the target.

The results just mentioned provide clues as to how participants were performing the task. The impact of misalignment indicates that updating the frame of reference was a significant source of difficulty in this task, just as it is in perspective-taking. Also, the effect of the number of objects in the quadrant shows that nearby objects influenced how easily participants could do the task. The interaction between these two factors provides details about the relationship between the two sources of difficulty. These findings were very consistent across participants, suggesting that a common strategy may have been adopted by the participants. Meanwhile, their verbal reports provided further support for a common strategy, with additional clues to aid in the identification of that strategy.

Verbal Reports Participants' verbal reports indicated that they performed the task in a hierarchical manner. First, they identified an area of the space that contained the target, and then they determined the target's position within that area. These two steps can be thought of as (1) identifying the position of a cluster of objects containing the target relative to the viewer and (2) determining the position of the target within the cluster. These steps, and evidence for them from the verbal reports, are discussed further here. Additional arguments supporting this strategy can be found in Gunzelmann and Anderson (2004; in press).

There was, of course, some variability in exactly how participants reported solving the task. However, there was a great deal of consistency in the general approach. The most critical aspect of the verbal reports provides evidence for the hierarchical decomposition just described. Of the twenty participants, 17 explicitly mentioned either "splitting" the view into sub-areas or using "clusters", "configurations", "groups", or "patterns" of objects to solve the task.

In the hierarchical solution process, the first step was to find the general area of the map where the target was located. Participants reported identifying a cluster of objects that contained the target and finding that cluster on the map of the space. By focusing in on a subset of the objects in the space, participants were able to reduce both the amount of area on the map that they needed to search as well as the number of objects that they needed to consider. They were able to accomplish this reduction by locating a single feature from the visual scene on the map (i.e., the cluster).

Step 2 of the solution strategy involves more detailed encoding and processing of the area of the space where the target was located. There was a little more variation in how participants described completing this step. Eight of the participants stated that they used mental rotation to line up the target in the visual scene with the object locations shown on the map. They did this by rotating the cluster of objects containing the target so that it lined up with the corresponding cluster on the map. Two of these participants did not explicitly mention using the first step. However, research has shown that rotating displays of greater complexity is more difficult (Bethell-Fox & Shepard, 1988). Thus, it is likely that these participants found some way to reduce the number of objects they needed to rotate to solve the problem.

The remaining participants reported a verbal strategy for the second step, which generally involved identifying the target's position in the cluster (e.g., "directly behind the closest object in the group" for Figure 1). These participants did not report using a mental image. While this process is somewhat different from rotating the cluster, it still requires that the internal relationships be maintained and updated during the transformation. Thus, although the details are somewhat different, the general approach remains consistent. One of these participants did not specifically mention an attempt to narrow the search area by focusing on a cluster or region.

Regardless of the particular methods, the second step requires that the target's location be encoded relative to the other objects in the cluster, and that the information be transformed so that it can be applied to the representation on the map. Using either mental rotation or a verbal approach, successful completion of this step requires that the participants maintain the internal relationships among the objects in the cluster as they make the transformation. The interaction between misalignment and the number of objects in the quadrant with the target reflects the difficulty of this process. It shows that this step became increasingly difficult as information about more objects had to be maintained and transformed. The consistency of this result indicates that this was true regardless of the details of how participants reported executing the second step.

The verbal reports suggest that the general strategy for solving the task was quite similar across participants, even if some of the specific methods differed somewhat. Participants seem to take a hierarchical approach to locating the target, by finding an area of the map to search and then focusing in on that area. This is the general strategy that served as the basis of the model described in Gunzelmann and Anderson (2004). That model used this strategy to solve the task and produced a close fit to the human response time data. In addition, Gunzelmann and Anderson (in press) describe in more detail the empirical evidence supporting this strategy. The remainder of this paper considers this general approach to solving the task further. The solution strategy is compared to the strategies reported for perspective-taking tasks, to explore the relationship between the two tasks and how people solve them.

Comparison of Strategies

There are a number of differences between the perspective-taking tasks used by Huttenlocher and Presson (1979; Presson, 1982) and the orientation task used in the experiment described above. The two most obvious differences between them are that the perspective-taking tasks require imagining a scene from a different position in the space, and that the orientation task requires a transformation of the current point of view to a different frame of reference.

Despite any differences between the tasks, they have important features in common. In both tasks, participants are presented with a representation of an array of objects and they must perform some transformation that allows them to identify the location of those objects in a different representation. In perspective-taking, the different representation is a new viewpoint relative to the array, whereas the orientation task used here requires a change from an egocentric frame of reference to an allocentric frame of reference. The following discussion illustrates how the steps participants reported for doing the orientation task relate to the perspective-taking strategies described above.

The verbal reports from participants indicated that they began each trial by attempting to identify a cluster of objects containing the target, which they could then locate on the map. What is crucial is how this step was achieved. Most of the participants mentioned that they used the left-right axis to help them divide the space. To find a cluster of objects "on the left", though, requires using a frame of reference based upon the viewer's location on the map. This process is analogous to the viewer rotation strategy described in perspective-taking tasks. Essentially, it involves imagining oneself on the map, allowing for a determination of which half of the map corresponds to the left half of the visual scene.

On the other hand, the second step of participants' solutions seemed to relate more closely to the array rotation strategy. Many of the participants reported that they used mental rotation to complete this step. This involved rotating the cluster of objects containing the target so that it lined up with the appropriate cluster on the map. Participants could then zero in on the correct object to identify the target. This corresponds exactly to the array rotation strategy described for perspective-taking. Here, the objects in the cluster are rotated to the new orientation. Even those participants who reported using a verbal strategy indicated that they incorporated information about the other objects in the cluster into their description. It is the relationships among those objects that allowed participants to identify the target. Thus, the transformations must have preserved this information. So, while the method may have been different, the nature of those transformations seems to have been the same.

There is overlap in the kinds of transformations that are required to arrive at the correct solution in orientation tasks and perspective taking tasks, which allows similar approaches to apply in both kinds of task. For instance, the first step of the strategy described here seems to be an example of an item question, which requires the individual to keep track of only a single item in the array. In the orientation task, participants grouped objects into clusters, which could then be treated as individual items. As a result, the first step requires finding the answer to an item question such as, Where would the cluster be if I were at "this location" (i.e., the black dot) on the map?

It is telling that participants answer the question of where the cluster is located by using viewer rotation. Huttenlocher and Presson (1979; Presson, 1982) found that participants were more accurate in their responses to item questions when they were instructed to use the viewer rotation strategy. In this case, participants opted to use viewer rotation to locate the cluster on the map. This provides some evidence that participants are able to effectively choose an appropriate strategy to complete this step.

In the second step, participants continue to show efficiency in their solutions. Identifying the location of the target within the cluster requires information about all of the objects in the cluster, since the spatial relationships among them are critical in determining which one is the target. Thus, the second step is an example of an appearance question. Participants' verbal reports suggest that they use the array rotation strategy here, which maintains the locational and relational information. Huttenlocher and Presson (1979; Presson, 1982) showed that using array rotation in this kind of situation will lead to greater accuracy than viewer rotation.

The comparison of the strategies from these different tasks illustrates ways in which they correspond. This demonstrates that individuals may have general strategies that they can apply flexibly in different circumstances. Further, the examination of when the different perspective-taking strategies were used here shows that participants were able to choose strategies that were locally efficient to serve a larger goal. By approaching the task hierarchically, they were able to choose a strategy at each step that provided the better opportunity for success.

Conclusion

The correspondence of the strategy for orientation to strategies from perspective-taking is not entirely surprising. The two types of tasks share important features, which require participants to perform analogous transformations to determine the correct response. What is somewhat surprising is that participants selected different approaches for difference steps in each trial. However, what is most interesting is that participants seem to choose the more effective strategy at the decision points in the solution process to efficiently solve the problem.

The efficiency of participants' strategy choices shows that they are actually quite sophisticated in how they approach spatial tasks. One reason for this may be the vast amount of experience that is gained with such tasks in naturalistic settings. Maps are common in navigation tasks, and we frequently have to interpret spatial information from perspective other than our own (e.g., in giving or following directions; Taylor & Tversky, 1996). The results described here indicate that these experiences have allowed us to learn efficient approaches to different spatial tasks.

Acknowledgments

The experiment described here was conducted as part of the first author's dissertation, and was supported by AFOSR Grant #F49620-99-1-0086 to Dr. John R. Anderson at Carnegie Mellon University and by NRSA training grant #5-T32-MH19983 from NIMH to Dr. Lynne Reder at Carnegie Mellon University. We would like to thank Don Lyon and Kevin Gluck for their helpful comments.

References

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

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

Gunzelmann, G., & Anderson, J. R. (in press). Location Matters: Why Target Location Impacts Performance in Orientation Tasks. *Memory & Cognition*.

Gunzelmann, G., & Anderson, J. R. (2004a). Orientation Tasks Involving Multiple Views of Space: Strategies and Performance. *Spatial Cognition and Computation, 4*, 209-256.

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

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

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

Piaget, J., & Inhelder, B. (1956). The child's conception of space. London, England: Routledge & Kegan Paul.

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

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

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

Taylor, H. A., & Tversky, B. (1996). Perspective in spatial descriptions. *Journal of Memory and Language, 35*, 371-391.

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

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