

Representations of Shape during Mental Rotation

Peter Khooshabeh and Mary Hegarty

University of California Santa Barbara
Department of Psychology
Santa Barbara, CA 93106-9660
khooshabeh@psych.ucsb.edu

Abstract

How is shape represented during spatial tasks such as mental rotation? This research investigated the format of mental representations of 3-D shapes during mental rotation. Specifically, we tested the extent to which visual information, such as color, is represented during mental rotation using methods ranging from reaction time studies, verbal protocol analysis, and eye tracking. Another set of studies examined whether people use piecemeal or holistic strategies to rotate complex objects. Results show that individuals with good rotation ability do not represent color during mental rotation and rotate whole shapes; whereas poor rotators do represent color and rotate individual pieces of the shape using piecemeal strategies. This work contributes to theories about cognitive shape processing by showing that different information processing strategies may be one cause of individual differences in mental rotation performance.

Introduction

Mental rotation measures a component of spatial thinking that involves imagining the movement of objects external to our bodies. An example of this type of thinking in everyday life is imagining how to arrange luggage to fit into an automobile trunk without actually moving any bags. Physical objects that people imagine rotating have textures, colors, shading, and other properties. However, the crucial aspect of these objects for performing a mental rotation task is information about the shape of the objects.

Mental rotation has been a fundamental task in cognitive science research. The classic study by Shepard and Metzler (1971) presented participants with perspective drawings of 3-D shapes. They found that reaction time to rotate objects was linearly related to the size of the angle of rotation. Shepard and Metzler interpreted this result as a process of a mental simulation of the physical world. Just as it takes a

shorter amount of time to physically rotate an object 20° than it does 120°, the same applies for mental rotation. However, some questions still remain regarding the information processing during mental rotation. What is the format of representations that are mentally rotated? What strategies do individuals use to mentally rotate 3-D objects?

We have conducted several experiments that address these questions. One set of experiments investigates whether mental representations of objects incorporate either both visual features and spatial (shape) information or just spatial information. Another set of experiments asks whether individuals rotate components of shapes (using a piecemeal strategy) or whole shapes. In both experiments we examine differences between people with good and poor mental rotation performance with respect to both questions.

Previous Literature

Most research and ensuing discussion about mental imagery has been about whether images are similar in format to objects in the physical world or whether they are encoded in some other mental code that is not pictorial or depictive (Kosslyn, 1994; Kosslyn, Thompson, & Ganis, 2006; Pylyshyn, 2002). Kosslyn and colleagues suggest that mental rotation is a visuospatial process in which mental images are represented and transformed in a visual buffer. In this paper we assume that the representation is a 3-D mental image that includes metric information about the shapes of the objects (relative lengths of the arms and the angles between the arms of the figures) but question whether distinctive visual cues are also included in these mental images, during mental rotation. There is evidence that individuals with high spatial ability tend to have schematic mental images, without much visual detail, whereas individuals with low spatial ability encode more visual details in their mental images. Kozhevnikov, Hegarty, and Mayer (2002) classified individuals as verbalizers or visualizers based on a cognitive style questionnaire and also measured spatial ability. They

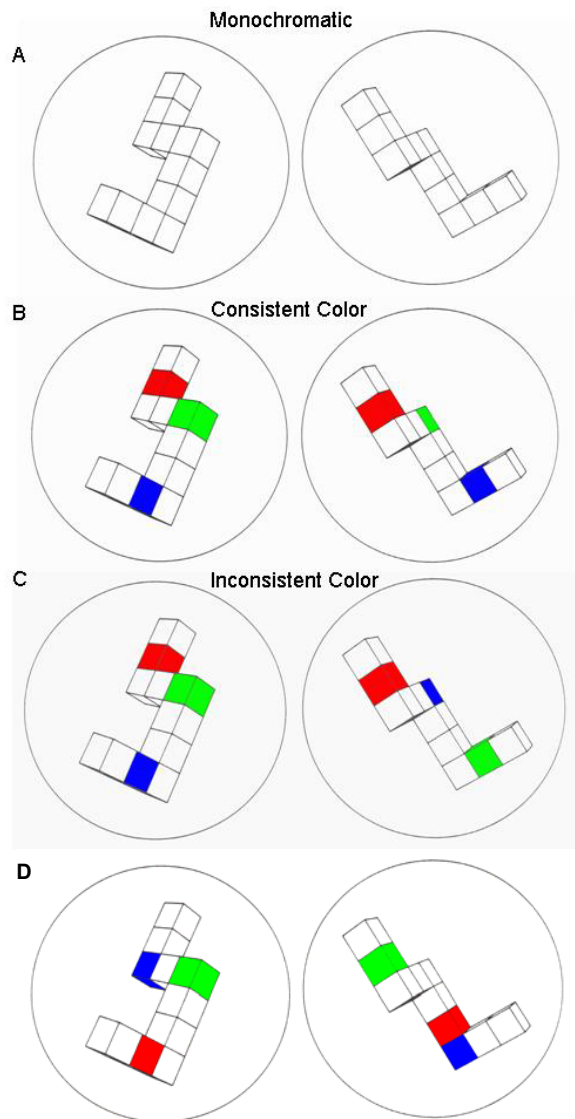


Fig. 1. All figures have the same shape and only differ with respect to color. Both the order and position of the Inconsistent figures in D are different.

identified two types of visualizers, those with high and those with low spatial ability. In studies of comprehension of kinematics graphs, low-spatial visualizers interpreted the graphs as pictures, whereas high-spatial visualizers correctly interpreted them as abstract spatial representations. In a related study, Hegarty and Kozhevnikov (1999) found that when solving mathematical word problems, low-spatial visualizers represented irrelevant visual details in the problems and performed poorly, whereas high-spatial visualizers constructed schematic spatial representations of the problem-relevant information and had superior performance.

Kozhevnikov, Kosslyn, and Shephard (2005) extended the dissociation between detailed visuospatial imagery and schematic spatial imagery. They showed that low-spatial

visualizers performed better than high-spatial visualizers on a task that involved focusing on detailed visual aspects of a stimulus, whereas the opposite was true for a mental rotation task. The difference in performance was not explained by differences in the ability to process abstract information, as the low- and high-spatial visualizers did not differ on a test of general intelligence.

The research of Kozhevnikov and colleagues suggests that one difference between those with good and poor performance on mental rotation tasks might be that good rotators construct schematic spatial representations of the figures that include metric information about their shapes but not visual details whereas poor rotators construct more detailed representations of these figures that include both metric shape information and distinct visual cues. In other research, Just and Carpenter (1985) modeled individual differences in strategies used by high and low spatial ability individuals to perform tasks such as cube-comparison and mental rotation. They found that high spatial individuals were more efficient at the tasks. Just and Carpenter (1985) account for this by saying that, "...high-spatial subjects use a more economical code to represent the figure that permits faster execution of the rotation and comparison" (p. 158).

An alternate but not incompatible explanation of individual differences in mental rotation is that high-spatial and low-spatial individuals may use different information processing strategies to perform the rotation process. For example, one explanation of individual differences in spatial thinking is that individuals with high spatial ability have more spatial working memory resources than low-spatial individuals (Miyake & Shah, 1999). If this is true, high spatial individuals can store more spatial information in working memory or process it more efficiently (assuming a tradeoff between storage and processing capacity in working memory). Because of their smaller spatial working memory capacities, low spatial individuals might be limited to represent or process only a part of the object at a time, thus affecting their strategy of mental rotation. For example in one previous study (Just & Carpenter, 1985), low-spatial individuals reported that their mental image fell apart while they were rotating it.

In summary, poor performance in mental rotation by low spatial individuals might be a consequence of the amount of irrelevant visual detail that they encode in their representations, in the use of piecemeal processes as a result of limited working memory capacity, or both. In other words, high- and low spatial individuals might differ in default data structures (schematic versus detailed mental representations), in algorithms that process those data structures (holistic or piecemeal rotation strategies), or both (e.g. Kozhevnikov, 2007).

Color and Mental Rotation

In order to examine whether visual (non-spatial) information is encoded during mental rotation, we studied a version of the Shepard and Metzler (1971) task, but

introduced distinctive visual cues (colors) to the 3-D figures that participants had to rotate (see Figure 1). Participants were instructed to judge whether the two objects had the same shape, and not to take color into account in making their judgments.

According to one model of mental rotation, participants accomplish this task by first finding corresponding arms of the two figures, rotating these arms first, and later examining whether the rest of the figure is also rotated to congruence. An important factor in finding corresponding arms is disambiguating the top and bottom of the figures (Carpenter & Just, 1978; Just & Carpenter, 1976; Metzler & Shepard, 1974). This is especially difficult for rotations with angular disparities larger than 90°.

The colors of the cubes are not relevant to the mental rotation tasks in our experiments so it is entirely possible to perform the match based on a spatial representation without any visual detail. If mental representations of objects during mental rotation possess only spatial (shape) information and not distinctive visual details, color should not affect performance on the task. If the representations of objects during mental rotation include both visual and spatial information, then participants should be affected by the color. Specifically, inconsistent colors on the two figures (as in Figure 1C and 1D) should hurt them and consistent colors should help them by allowing them to match common parts of each of the objects. Specifically, consistent colors are predicted to help individuals disambiguate the top/bottom of the stimuli. In a previous study, Metzler and Shepard (1974) attempted to disambiguate the ends of the objects by marking them with colored dots. This manipulation was not successful in improving performance. However, their dots were external to the figures (thereby adding another object to rotate) and it is possible that the dots became dissociated from their ends during the rotation process, accounting for why they did not find an effect of color. In our experiments, the colors were part of the objects, so that if they are rotated, they should be rotated with the rest of the objects.

Similarly, if individuals with low spatial ability are more likely to rotate objects piecemeal, due to limited working memory capacities (Just & Varma, 2007), consistent colors should help them decompose the large 3-D figures into more tractable pieces that are amenable to their piecemeal strategies of mental rotation and match common arms after the rotation processes.

We predicted that individuals with poor rotation ability would be affected by the addition of color on the figures either because they are more likely to represent details in their mental images or because they are more likely to use piecemeal strategies, due to low spatial working memory capacity. In contrast, high spatial ability should not be affected by color because they have schematic mental images that lack distinctive visual cues or because they have sufficient spatial working memory resources to rotate the object holistically.

Initial Reaction Time Experiments.

In initial studies, we measured errors and response times with monochromatic (control) figures, consistently colored figures or inconsistently colored figures (see Figure 1; Khooshabeh, 2009; Khooshabeh & Hegarty, 2008a, 2008b). As predicted, individuals with poor rotation ability were affected by distinctive colors in the images, whereas those with good rotation ability were not. Specifically, poor rotators were faster and more accurate on trials in which the shapes were consistently colored; there were no significant differences between performance on the inconsistently colored objects and the monochromatic objects. Therefore, it seems that the colors mainly enhance the performance of poor rotators. One interpretation of these results is that colors help differentiate the top from the bottom of the figures. Across all the experiments, participants with good rotation ability did not perform differently on the monochromatic trials compared to either type of colored trials. It is possible that good rotators do not need to explicitly distinguish the top and bottom because their larger spatial working memory enables them to represent and process each object as a unitary whole. We will explore this further in the section on Piecemeal Strategies in Mental Rotation.

Evidence from Eyetracking and Verbal Protocol Analysis.

In another study, we used a combination of eye tracking and verbal protocol analysis. Participants' eyes were tracked while they completed a subset of mental rotation trials (similar to those in Figure 1A, B, and D) that included monochromatic, consistently colored and inconsistently colored objects; the same participants completed a set of similar trials while thinking aloud. These individuals were also pre-tested using a battery of spatial ability psychometric measures.

First, the behavioral results showed that consistent color made the low spatial ability faster and more accurate but had no effect on the high spatial ability, thereby replicating the previous experiments. The verbal protocol analysis showed that individuals with low spatial ability mentioned color more often than participants with high spatial ability. Specifically, on the consistent trials, the low spatial ability mentioned color on approximately one third of trials, on average. In contrast, participants with high spatial ability mentioned color on only 5% of trials. This is an indication that low spatial ability consciously thought about color more often while they performed the mental rotation task.

In the eye tracking analysis, we defined regions of interest that circumscribed the figures. We compared the number of consecutive fixations made within each figure to the number of saccades made between the two figures. If a holistic rotation strategy is being used, then the number of fixations within an object should be equal to the number of switches of fixation between objects. This is, during holistic rotation participants should look only once at the whole figure on each side to encode the whole object, and then make a saccade to the other figure. However, if a

piecemeal strategy is used, then individuals will make multiple fixations on one figure, in order to look at different pieces to rotate, before making a saccade to the other figure.

For high spatial, the number of fixations within an object was about equal to the number of saccades between objects, and this was true for monochromatic, consistently colored and inconsistently colored figures, indicating that they appeared to use the holistic strategy for all figures. For low spatial, the number of fixations within objects was equal to the number of saccades between objects only for the consistently colored trials. On monochromatic and inconsistently colored trials, these participants made more consecutive fixations within objects than saccades between objects. Therefore, consistent color appears to have helped individuals with poor rotation ability encode the object holistically, whereas they had to perform a piecemeal rotation on trials with monochromatic or inconsistently colored pieces. One speculation is that the consistent colors allow the poor rotators to construct a sparse representation of just the pattern of the three colored cubes.

What Happens When Color is Relevant?

One explanation of performance in our experiments is that good rotators do not represent color during mental rotation, but we do not know if this is strategic or automatic. If it is strategic, then good rotators might explicitly suppress the representation of color when it is irrelevant but encode it when it is task relevant. If it is automatic, then participants with good rotation ability should have difficulty if color is relevant to the task. In one experiment, we used stimuli similar to those in Figure 1 but the instructions were changed. Participants were now told that they should take color into account in addition to shape in judging the identity of the objects. Thus, in this experiment, the correct answer for Figure 1D was “different” because the colors were different, even though the shape of the objects was the same. The results showed that all participants, both good and poor rotators, performed accurately and quickly on the trials with inconsistent color; that is they were able to quickly judge that the objects were different based on the color information. This shows that good rotators were able to take color into account when it was relevant to the task, (although they were not affected by color in the previous experiments). This is not surprising because many tasks in daily life involve making decisions with respect to visual information like color, so good rotators are just as able to take color into account as poor rotators when it is task relevant.

Is Distinctiveness of Visual Cues Important?

Instead of using three different colors for the colored blocks in our earlier experiments, in later experiments we colored three blocks the same color (black in one experiment and red in another). For example, in one of these experiments the three colored blocks in Figure 1B were all colored black and in the other they were all

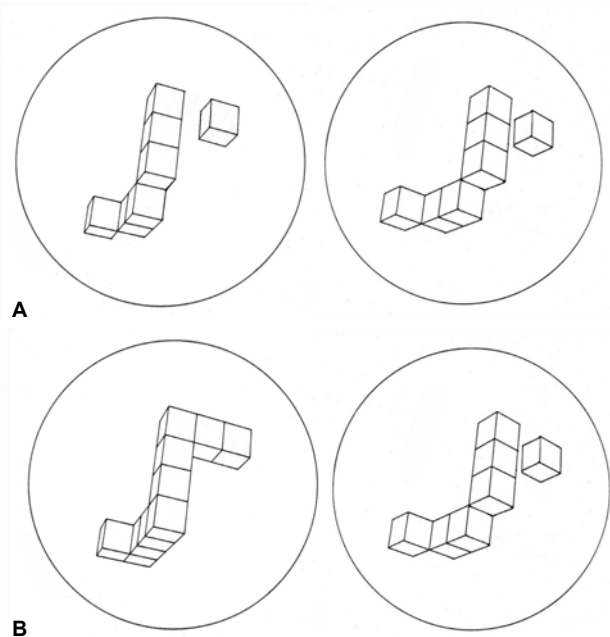


Fig. 2. The two types of incomplete trials rendered with missing cubes from the 3-D figures.

colored red. In these experiments, there were no significant differences between colored and monochromatic figures (Khooshabeh, 2009). These experiments indicated that it is the distinctiveness provided by three different colors that was important to performance in the earlier experiments, not just the fact that specific blocks were marked with a salient cue. This result is consistent with the idea that the advantage of colors to poor rotators is in allowing them to match corresponding parts of the objects by making them distinct. Although we used color as a visual cue in our experiments, it is also likely that other distinctive markers, such as 3 unique textures, could benefit individuals with poor rotation ability.

Piecemeal Strategies in Mental Rotation

Other related work in our laboratory explicitly tested whether poor rotators are more likely to use piecemeal strategies in mental rotation. In two experiments, 3-D Shepard and Metzler objects were manipulated to promote holistic and piecemeal strategies. In one condition of the experiments, participants had to rotate “fragmented” Shepard and Metzler figures in which three of the cubes were deleted, and in the other condition, they had to rotate complete objects (control condition, see Figure 2). We assumed that the fragmented objects would induce a piecemeal strategy, so that a difference between the fragmented objects and complete objects would indicate that a piecemeal strategy is not being used for the control (complete) objects. Across two experiments, the results showed that good rotators were slower and less accurate on

the trials with fragmented figures compared to complete figures (such as traditional 3-D Shepard and Metzler objects shown in Figure 1A). There were no differences in performance on fragmented versus complete objects for the poor rotators. This suggests that poor rotators use piecemeal rotation as their default strategy for complete objects, whereas high spatial individuals prefer holistic strategies for these objects.

General Discussion

In summary, individuals with good rotation ability are not affected by distinctive visual cues (colors) in mental rotation, when these cues are irrelevant to the task. These individuals also appear to use a holistic strategy when rotating both monochromatic and colored objects. In contrast, individuals with poor rotation ability are affected by color and appear to use a piecemeal strategy. It appears that color can enhance mental rotation performance by enabling participants to keep track of the pieces of an object during piecemeal rotation. In this account, good rotators are not affected by color because they do not use piecemeal strategies. Poor rotators appear to be able to use color strategically to disambiguate the different pieces of the object (e.g., distinguish the top from the bottom) during piecemeal rotation.

Although these results are based on mental rotation of the 3-D cube figures from Shepard and Metzler, it is likely that they would generalize to other complex 3-D objects. For example, Stull, Hegarty, and Mayer (2009) studied a situation in which students learned the structure of an anatomical object (a vertebra) by rotating a virtual model of the vertebra. In one condition of the experiment, they placed distinctive visual cues (colored poles) as orientation references on the 3-D virtual models. Low-spatial individuals in the orientation reference condition physically rotated the bone more efficiently and also demonstrated better anatomical learning of the various parts of the vertebrae bone compared to low-spatial individuals who did not receive these distinctive cues. Individuals with good spatial ability performed well in both conditions. One interpretation of these results is that the orientation references provide a distinct feature to disambiguate the symmetry of the bone, just as distinct colors help distinguish the top and bottom of the 3-D cube figures used in this study.

The results of these studies suggest both algorithmic and representational differences between good and poor rotators in mental rotation. The difference between holistic and piecemeal rotation is an example of an algorithmic difference that differentiates good and poor rotators and which possibly reflects differences in spatial working memory capacity. Independently of this, good and poor rotators may represent different information in their mental images, with poor rotators being more likely to represent purely visual information (color) and not just spatial information. However, the representation of color by poor rotators may be strategic, and a way of compensating for

poor mental rotation abilities. Thus the default representation in mental rotation may be schematic, but individuals with poor rotation ability may use strategies to take advantage of other sources of information (such as color) in order to make rotation more tractable. Similarly, both good and poor rotators take color into account when it is task relevant. Thus, visual information can be used flexibly to solve spatial problems when the task demands it, or when the spatial information processing abilities of the individual are lacking.

Acknowledgements

Peter Khooshabeh performed this research while on appointment as a U.S. Department of Homeland Security (DHS) Fellow under DOE contract number DE-AC05-00OR22750. Support from the NSF IGERT grant #DGE-0221713 is also appreciated.

References

- Carpenter, P., & Just, M. (1978). Eye fixations during mental rotation. In *Eye movements and the higher psychological functions* (pp. 115–133). Hillsdale, NJ: Erlbaum.
- Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology, 91*, 684–689.
- Just, M., & Carpenter, P. (1985). Cognitive coordinate systems: Accounts of mental rotation and individual differences in spatial ability. *Psychological Review, 92*, 137 - 172.
- Just, M., & Carpenter, P. (1976). Eye fixations and cognitive processes. *Cognitive Psychology, 8*, 441–480.
- Just, M., & Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Cognitive, Affective, and Behavioral Neuroscience, 7*, 153–191.
- Khooshabeh, P. (2009). Is mental rotation color blind? Unpublished doctoral dissertation, University of California, Santa Barbara.
- Khooshabeh, P., & Hegarty, M. (2008). How visual information affects a spatial task. In *Proceedings of Cognitive Science Society*. Washington, DC: Cognitive Science Society.
- Khooshabeh, P., & Hegarty, M. (2008b). Differential effects of color on mental rotation as a function of spatial ability. In *International Spatial Cognition Conference*. Freiburg, Germany.
- Kosslyn, S. M. (1994). *Image and brain: The resolution of the imagery debate*. Cambridge, MA: MIT Press.
- Kosslyn, S. M., Thompson, W. L., & Ganis, G. (2006). *The case for mental imagery*. New York, NY: Oxford

University Press.

- Kozhevnikov, M. (2007). Cognitive styles in the context of modern psychology: Toward an integrated framework of cognitive style. *Psychological Bulletin*, *133*(3), 464-481.
- Kozhevnikov, M., Hegarty, M., & Mayer, R. (2002). Revising the visualizer/verbalizer dimension: Evidence for two types of visualizers. *Cognition and Instruction*, *20*, 47-77.
- Kozhevnikov, M., Kosslyn, S., & Shephard, J. (2005). Spatial versus object visualizers: a new characterization of visual cognitive style. *Memory and Cognition*, *33*(4), 710-26.
- Metzler, J., & Shepard, R. N. (1974). Transformational studies of the internal representation of three-dimensional objects. In *Theories in cognitive psychology: The Loyola Symposium* (pp. 147-201). Hillsdale, NJ: Erlbaum.
- Miyake, A., & Shah, P. (1999). *Models of working memory: Mechanisms of active maintenance and executive control*. New York: Cambridge University Press.
- Pylyshyn, Z. (2002). Mental imagery: In search of a theory. *Behavioral & Brain Sciences*, *25*(2), 157-238.
- Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*, 701-703.
- Stull, A. T., Hegarty, M., & Mayer, R. (2009). Getting a handle on learning anatomy with interactive three-dimensional graphics. *Journal of Educational Psychology*, *101*, 803-816.