How Visual Information Affects a Spatial Task

Peter Khooshabeh (khooshabeh@psych.ucsb.edu)
Mary Hegarty (hegarty@psych.ucsb.edu)
Department of Psychology, University of California, Santa Barbara, CA 93106 USA

Abstract
To examine the information content of mental representations of three-dimensional objects, 21 participants performed a mental rotation task with the classic Shepard and Metzler (1971) figures. We compared conditions in which the figures were monochromatic, a condition in which they were colored consistently, and a condition in which they were colored inconsistently. Color affected performance of participants with low spatial ability but not those with high spatial ability. This suggests that individuals with high spatial ability construct more schematic mental spatial representations for this task, whereas individuals with low spatial ability construct representations that include both visual and spatial information.

Keywords: color; mental rotation; individual differences.

Introduction
Understanding the format of mental images is one of the most fundamental questions in cognitive science. In particular, the depictive (Kosslyn, 1994) vs. propositional (Pylyshyn, 2002) nature of mental imagery has been hotly debated. Regardless of what is actually represented in the mind, the phenomenology is of a picture in the mind. But how much information is represented in that picture? Does it consist of abstracted spatial forms or does it also include visual details? This research is concerned with studying the information content of mental representations of three-dimensional (3-D) objects and the strategies people use in performing mental transformations of these objects.

Although there has been more than a century of research in spatial cognition, there are many open questions regarding the nature of mental representations and processes, such as mental rotation. Kosslyn (1994) suggests that mental rotation is a visuospatial process in which the content of mental images is represented and transformed in a visual buffer. However, there is evidence that individuals with high spatial ability may have more schematic mental representations, without much visual detail. Kozhevnikov, Hegarty, & Mayer (2002) classified individuals as verbalizers or visualizers based on a cognitive style questionnaire and also measured spatial ability. They 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 visuo-spatial 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. Finally, they show that individuals from different professions (visual artists and scientists) can be classified as having detailed visual imagery or abstract spatial imagery, respectively.

In this study, we examine the extent to which high- and low-spatial individuals represent visual details in the classic Shepard and Metzler (1971) mental rotation task. In this task, participants see a pair of figures (Figure 1) and they are asked to decide whether the two figures are the same or mirror images of each other. Previous research suggests that although participants see 2-D perspective renderings depicting the 3-D objects, the mental spatial transformation process mimics a physical rotation such that rotation time is proportional to the angle of rotation and rotations in depth have similar latencies to rotations in the picture-plane. This suggests that participants construct and transform 3-D mental images of the objects. However, a remaining question is what actually gets represented. Specifically, do people represent visual features of the objects or do they merely represent their more abstract 3-D shapes?

In order to study the effects of visual information on this spatial task, we created versions of the Shepard-Metzler figures that had some colored cubes. Although shape, like color, can be sensed through vision and can therefore be thought of as a visual feature (Marr and Nishihara, (1978), it can also be sensed through non-visual modalities whereas color can only be sensed visually (barring rare phenomena like synesthesias). The shape of the three pairs of objects in Figure 1 is the same, but they differ in color (a visual property). The first set is monochromatic like the classic Shepard-Metzler objects (Figure 1a). In the second set (Figure 1b), the figures have consistent colors, that is, the cubes are colored the same in the two figures. In the third set of figures (Figure 1c), the shapes are identical but the locations of the colors are inconsistent. In our experiments, participants are asked to judge whether the two shapes are identical while ignoring the colors of the cubes. Thus the
relevant property is shape (a spatial property) and color (a visual property) is irrelevant to the task.

A theoretical information-processing model suggests the following cognitive processes in mental rotation (Carpenter & Just, 1978; Just & Carpenter, 1985). First, the participant has to encode the figures. Second, he or she has to match common “arms” of the figures. The third step is to rotate them to congruence. Fourth, the participant makes a judgment of whether the figures are the same or mirror images of each other and finally he or she makes a response. Colored figures might affect performance, especially during the matching process. If the underlying representation has visual information, then we hypothesize that consistent colors will facilitate and inconsistent colors will impede performance. If the mental representation is a more abstract spatial representation, then color should not affect performance in the task.

Other researchers have attempted to support the process of mental rotation by using color. Metzler and Shepard (1974) described several experiments that investigated mental rotation in both the depth and picture planes. Based on their own phenomenology and comments from participants, they inferred that the process of matching corresponding arms of the figures was particularly difficult, especially for picture-plane rotations. In one experiment, they added color dots to the arms of the figures in order to make it easier to match their corresponding arms. However, the response time function was very similar to a control condition, suggesting that adding color to the figures did not affect task performance.

More recently, Ruddle and Jones (2001) compared manual rotations of Shepard-Metzler figures in either a real physical condition or a virtual computer mediated condition. Participants performed real physical manual rotations much faster than virtual rotations. Ruddle and Jones attributed this to the difficulty in perceiving the figures’ orientation. In one experiment, they colored long and short arms of the Shepard-Metzler figures to facilitate the rotation process. Participants were faster at the virtual rotation with this manipulation than in the control condition (no color). Our research examines whether coloring parts of the figures facilitates mental as opposed to manual rotation.

Amorim, Isableu, and Jarraya (2006) also manipulated figures to make rotation easier. They made stimuli in the form of bodies in poses similar to Shepard-Metzler figures, with one arm of the figure showing the head and other showing arms and legs. Amorim et al. specifically predicted that rotation using a body-based coordinate system would be more holistic, but that it would still exhibit a linear increase with angular disparity. Rotation was easier with the body stimuli and as predicted, response times still showed a linear increase.

Our experiment provides visual information that might help or hurt performance, as illustrated by the sample stimuli in Figure 1. If people represent the visual information in the stimulus and not just the shape (spatial information), performance should be facilitated when the colors are consistent and impaired when the colors are inconsistent. Based on earlier work suggesting that low-spatial individuals represent visual details whereas high-spatial individuals form schematic spatial representations (Kozhevnikov et al., 2002; 2005), we predicted that low-spatial individuals would be affected by color in this way, whereas high-spatial participants would perform equivalently with the monochromatic and colored cubes.

**Method**

**Participants, Materials, Apparatus, and Design**

We recruited 21 participants (13 males and 8 females) from a summer research program at the University of California, Santa Barbara. All individuals participated voluntarily.

The stimuli (see Figure 1) were perspective drawings of 3-D abstract figures used by Shepard and Metzler (1971), but modified for the purposes of this experiment. Five different figures and the mirror images of these figures were used. Figures were paired such that they differed in angular disparity in 20° increments. By rotating the “same” figures along their elongated or vertical axis in depth, one could bring them into congruence, but this was not possible for the mirror images. There were 100 basic pairs for each of the three color conditions (monochromatic, consistent and inconsistent), i.e., 5 “same” pairs and 5 “different” pairs for 10 different angles varying from 0-180.

![Figure 1](image.png)

Figure 1: Three types of figures used in the study. All have the same shape and only differ with respect to color.

In addition to the monochromatic figures, the component cube faces of two other types of figures were colored. Consistently colored figures (see Figure 1b) were designed so that the position and order of the colors was the same for
the two figures to be compared. Inconsistently colored figures were designed such that the color order was different for the juxtaposed figures in the pair (see Figure 1c). We chose to color individual cubes within the figures as opposed to just putting color on the arms of the figures (Metzler & Shepard, 1974) or coloring whole arms of the figures (Ruddle & Jones, 2001). The 300 stimuli were presented on a computer monitor using Superlab, a commercial program from the Cedrus Corporation, which collected response times for the mental rotation trials. The Card Rotation test (Ekstrom, French, & Harman, 1976) was used to measure spatial ability.

A 3x10 within-subjects experimental design was used; each of the three types of figures (monochromatic, consistent color, inconsistent color) was presented at ten 20° increments between 0-180° in depth. The color conditions were blocked and presentation order of the three blocks was counterbalanced across participants.

**Procedure**

On arrival at the laboratory, participants were given thorough task instructions on paper, which stressed that they should make same-different judgments based on the shape and not the color of the figures and demonstrated the difference between the appearance of a figure and its mirror image. Once the participant understood the verbal instructions, he or she was administered practice trials with feedback. This condition familiarized participants with the computer display of stimuli. In order to show that color information was irrelevant to the task, the practice trials included pairs of figures in which the figures were mirror images but were colored consistently. Participants were instructed that the correct response to such a trial was 'different.' In another practice trial, the shape of the two juxtaposed figures was the same but they were colored inconsistently (as in Figure 1c). Participants were instructed that correct response for such a trial would be 'same.' After the practice trials, participants were given an opportunity to ask questions before they began the mental rotation task which they performed in three blocks of 100 trials. Finally, participants were administered Card Rotation test (Ekstrom et al., 1976).

**Results**

**Accuracy**

A repeated measures analysis of variance with angular disparity and type of figure (monochromatic, consistent color, inconsistent color) as independent variables showed a linear decrease in accuracy as a function of angular disparity, $F(1,20)=14.8$, $p<.001$ (see Figure 2). There was no main effect of figure type, $F(2,40)=0.3$, $p=.7$, but there was an interaction between angular disparity and figure type, $F(16,320)=1.98$, $p<.05$.

**Response Times**

Response data were based on correct responses for true trials only.$^1$

The response time data are presented in Figure 3. A repeated measures analysis of variance with type of figure and angular disparity as independent variables showed a significant effect of angular disparity, $F(8,160)=20.1$, $p<.001$, and no significant effect of figure type, $F(2,40)=1.9$, $p>.10$. The effect of angular disparity conformed to a linear trend $F(1,20)=33.3$, $p<.001$. There was also a marginal significant interaction between type of figure and angle, $F(16,320)=1.6$, $p=.1$. As Figure 3 suggests, the response time on the colored figures is faster for angular disparities greater than 100°. A focused analysis of response times collapsed for angles between 120-160° indicated a significant effect of figure type, $F(2,40)=7.1$, $p<.01$.

---

$^1$ Due to experimenter error, not all participants completed the zero-degree disparity trials, so we omitted these from the analysis.
Individual Differences

We assigned participants to either a high or low spatial ability group based on a median split of scores on the Card Rotation test. The maximum score possible on this test is 160. Those in the high-spatial group (n=10) scored an average 137.1 (SD=13.2) and the low-spatials (n=11) scored an average 103.7 (SD=16.7). This difference was significant based on an independent samples t-test, t(19)=5.05, p<.001.

![Figure 4](image)

**Figure 4**: Response time as a function of angular disparity for the three different types of figures. This is for individuals with high spatial ability.

![Figure 5](image)

**Figure 5**: Response time as a function of angular disparity for the three different types of figures. This is for individuals with low spatial ability.

Using the dichotomous classification of spatial ability as a between-subject variable, we conducted a mixed model analysis of variance to test the effects of figure type and spatial ability for both accuracy and response time. There were no significant effects of spatial ability or figure type for the accuracy measure. The analysis of response times revealed a significant interaction between figure type and spatial ability, F(2,38)=9.6, p<.001 (see Table 1 for descriptive statistics). High spatial participants were faster in general F(1,19)=5.1, p<.05. Furthermore, analysis of simple effects indicated that their speed of rotation was not significantly affected by the type of figure (monochromatic, consistent color, inconsistent color), F(2,18)=2.5, p=.1 for high-spatial participants. In contrast there was a significant effect of type of figure for the low-spatial participants F(2,18)=6.8, p<.006. As Table 1 and Figure 5 show, these participants were faster on both the consistent and inconsistent colored figures than on the monochromatic figures.

<table>
<thead>
<tr>
<th>Spatial Ability</th>
<th>Color</th>
<th>Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>Monochromatic</td>
<td>7858 (773)</td>
</tr>
<tr>
<td></td>
<td>Consistent</td>
<td>4967 (383)</td>
</tr>
<tr>
<td></td>
<td>Inconsistent</td>
<td>4780 (586)</td>
</tr>
<tr>
<td>High</td>
<td>Monochromatic</td>
<td>3232 (814)</td>
</tr>
<tr>
<td></td>
<td>Consistent</td>
<td>3911 (404)</td>
</tr>
<tr>
<td></td>
<td>Inconsistent</td>
<td>4812 (618)</td>
</tr>
</tbody>
</table>

To ensure that the differences between high- and low-spatial individuals were not due to outliers, we computed the difference between the average response times on the Consistently-colored and Monochromatic figures and plotted this variable against spatial ability. As Figure 6 shows, there was a linear trend for the amount of speed-up in mental rotation due to colors to be related to an individual’s spatial ability with low spatial individuals being helped more than high-spatial individuals (r=.60, p<.01)

Discussion

In summary, this research replicated classic findings on mental rotation (Shepard & Metzler, 1971). Response times were linearly related to the angular disparity of the two figures to be compared. Accuracy also declined with angular disparity as is typical for non-practiced participants on this task (e.g., Amorim et al., 2006). The more novel results of this experiment concerned the effects that resulted from coloring some of the cubes that made up the figures. Color facilitated performance on trials with larger angular disparities. Analysis of individual differences indicated that as predicted, color did not significantly affect the accuracy or response times of high-spatial individuals. In contrast, low-spatial individuals were significantly faster with colored figures than with monochromatic figures.

Response times for the two sets of colored figures are almost identical for the low-spatial participants, but are higher for the monochromatic figures. In contrast, color has no significant effect on the average response times for high spatial ability participants. Thus color (a purely visual
between their stimuli and ours was that the dots were not consistent with those of Metzler and Shepard (1974) who found no effect when they added color dots to the arms of the figures. Mental rotation of the body figures showed a linear increase with angular disparity. However, the slope of the response time was attenuated for the body posture figures, and response times for these figures were faster for angles larger than 90°, similar to the effects of colored blocks in our experiment. When the rotation became more difficult, the additional spatial cues helped low-spatial individuals in our study just as the additional spatial information in the body figures helped participants perform large rotations in the study by Amorim et al.

While the preliminary results of our experiment are intriguing, more research is needed to understand the mechanisms by which coloring cubes facilitates performance for low-spatial individuals. A likely explanation is that the colored cubes facilitated the process of matching corresponding arms of the figures and this had greater effects on performance of low-spatial individuals than high-spatial individuals, because the latter are able to find corresponding arms based on shape alone. Another possibility is that colored cubes allow different strategies. While mental rotation most likely involves analogue processes, it is not clear whether individuals encode the whole object and rotate it in its entirety or perform a piecemeal component-by-component rotation. Stimulus complexity certainly has an effect on the type of strategy that an individual undertakes, with less complex stimuli being more amenable to holistic rotation strategies (Aretz & Wickens, 1992). It is possible that these figures are too complex to be rotated holistically and that coloring some of the cubes suggests a more successful piecemeal strategy to low-spatial individuals whereas high-spatial individuals adopt this strategy regardless of whether some of the cubes are colored.

Another remaining question is whether color is automatically represented in a mental image, and therefore has to be inhibited in the case of inconsistently colored figures, or whether people can view a colored object and construct an abstracted representation that does not contain color. In research on reading, the Stroop effect suggests that reading is automatic and conflicts with color naming. By analogy, in our mental imagery task, representing the colors

---

**Figure 6:** The difference in average response time between the Consistent and Monochromatic figures as a function of spatial ability.

![Response Time (RT) and Spatial Ability](image_url)
of the stimuli may be automatic, and may slow performance with a mental rotation task in which color is to be ignored.

Given previous literature on sex differences in mental rotation (Linn & Petersen, 1985), a note about sex differences is warranted. We observed no sex differences in the Card Rotation test or the mental rotation task. Most previous work that reports sex differences in mental rotation is based on the Vandenberg and Kuse (1978) mental rotations test, which is a speeded test in which participants have to compare a single 3-D Shepard-Metzler figure to four others in each trial. The measures in this study are simpler. The Card Rotations test is based on 2-D stimuli and the experimental task involves a comparison of only two figures in each trial. Finally, time pressure was not as great in our task. It is possible that sex differences are evident only in more complex and speeded mental rotation tests. It should also be noted that the number of participants in our study was small, so we had limited power for detecting a difference. It is important to study the extent to which sex differences in the Vandenberg and Kuse (1978) test and in the Shepard and Metzler task show consistent patterns.

Future Directions. A limitation of this experiment is that the inconsistently colored figures were not as different as they might have been. In current research we are examining performance with inconsistent figures in which the colors move to different cubes within each arm, as well as moving to different arms of the figures. We predict that high-spatial individuals will not be perturbed by this manipulation but that it will impair performance of low-spatials relative to a monochromatic control condition. In future research, we will also conduct verbal protocols and use eyetracking methodology to better understand the differences in strategies between high- and low-spatials on this task and how these are moderated by our color manipulations. For example, low-spatials might pay attention to the colors more than high-spatials, especially in the case of an inconsistency in both the hue and spatial dimensions.

This research addresses the underlying format of mental representations of 3-D shape. In particular, it is providing new evidence regarding the extent to which mental rotation relies on purely spatial representations or visuospatial images. Given that the brain processes visual information in a separate pathway compared to spatial information, another future direction will be to investigate the neural processing of performance in this task.

Acknowledgments

We thank David Herschorn and Amanda Baratz for collecting data, and Andrew Stull for comments on earlier drafts of this manuscript. 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. Michael Tarr and his laboratory provided the original figures from Roger Shepard.

References


