INVESTIGATING STRATEGY IN A MENTAL ROTATION TASK USING EYE-TRACKING HEAT MAP ANALYSIS

Kabyashree Khanikar and Ritayan Mitra
Centre for Educational Technology
IIT Bombay, India

ABSTRACT

This study investigates the cognitive strategies employed during a mental rotation task through the integration of interview data and eye-tracking heat map analysis. A total of 20 interviews between 4 participants were analyzed independently by two coders to identify holistic and piecemeal rotation strategies and eye-tracking heat maps were examined in conjunction with interview data to reveal patterns of visual attention corresponding to these strategies. Preliminary results indicate that heat maps can effectively distinguish between holistic and piecemeal strategies, characterized by distinct visual attention patterns. Holistic strategies displayed uneven attention distribution with complementary focal points on the two orientations, while piecemeal strategies exhibited multiple focal points that were mirror images of each other. These findings suggest that heat maps could offer a robust method for identifying cognitive strategies in mental rotation tasks, paving the way for a visual analytical framework of spatial thinking studies.

KEYWORDS

Mental Rotation, Eye-Tracking Heat Maps, Holistic Strategy, Piecemeal Strategy, Visual Attention

1. INTRODUCTION

Spatial ability is a fundamental element of human cognition, playing a pivotal role in learning, problem-solving, and academic achievement, particularly in STEM fields (Uttal & Cohen, 2012). Mental rotation ability, the most extensively studied type of spatial ability, involves mentally manipulating two-dimensional (2D) or three-dimensional (3D) objects to envision their appearance after a specific angular rotation around a defined axis (Shepard & Metzler, 1971). This ability, which is intrinsic to spatial cognition, has been widely studied as an important dimension of intelligence (Carroll, 2009) and has been measured effectively using the Vandenberg Mental Rotation Test (Vandenberg & Kuse, 1978). Shepard and Metzler's work established the foundational understanding of how individuals manipulate spatial information, demonstrating a linear relationship between rotation degree and reaction time.

Shepard and Metzler's (1971) work on mental rotation established a link between response time (RT) and angular disparity, showing that larger rotations take longer. This led to two key models: holistic, where objects are rotated as whole units, and piecemeal, where objects are rotated part-by-part (Zhao & Della Sala, 2017). Folk and Luce (1987) expanded on piecemeal rotation, describing it as sequentially rotating image components. Research on individual differences in strategy use placed holistic and piecemeal approaches on a continuum (Paivio, 1971; Gluck & Fitting, 2003). Hegarty (2018) expanded on these strategies, noting that mental rotation tasks (MRTs) can be solved with both spatial and analytic strategies, such as cube counting and local shape comparisons. Differences in the strategies used in mental rotation tasks are crucial for examining individual variations in spatial thinking (Hegarty, 2010), as these strategies are amenable to training, supporting the idea that spatial intelligence is malleable (Uttal and Newcombe, 2013).

Various strategies, such as holistic rotation, piecemeal rotation, perspective-taking, local turns, counting cubes and global shape have been suggested as strategies that individuals employ to rotate an object mentally (Hegarty, 2018). Among these, past research has predominantly focused on a dichotomous framework contrasting holistic and piecemeal approaches. Shepard and Metzler (1971) introduced the concept of holistic mental rotation, emphasizing the mental manipulation of entire objects. This perspective was challenged by Just and Carpenter (1976) and later by Pylyshyn (1979), who proposed a piecemeal approach involving the

analysis of individual object components. Heil and Jansen-Osmann (2008) argued that men perform better in mental rotation tasks due to a preference for holistic strategies. Further research by Khooshabeh, Hegarty, and Shipley (2010, 2012) indicated that proficient imagers adapt their strategies based on task demands, while less proficient imagers rigidly adhere to piecemeal strategies. Nazareth (2018) advanced this understanding by emphasizing that strategy flexibility, rather than the superiority of any single strategy, leads to better performance.

Research on the strategy debate in mental rotation has primarily utilized eye-tracking data, retrospective interviews, and think-aloud protocols. Eye-tracking metrics such as fixation count and saccades have lately been used to define a "strategy ratio" (SR), calculated as the number of fixations within the figure divided by the number of switches between figures. An SR of 1 indicates a holistic strategy, while a ratio greater than 1 indicates a piecemeal strategy (Khooshabeh and Hegarty, 2010). This assumes that good rotators using a holistic strategy look at the reference object once, encode and rotate it, then confirm with the other figure. This is grounded in the fundamental definitions of holistic and piecemeal strategies. In a holistic strategy, an individual rotates the entire mental image of an object. Conversely, a piecemeal strategy entails breaking down the image into segments, rotating one part to align with the comparison figure, and then applying the same rotation to the remaining parts to determine if they correspond (Just & Carpenter, 1985). Nazareth (2018) challenged the SR approach, noting that an SR of 1 could reflect different eye-tracking patterns and suggested including fixation duration as an additional metric to better differentiate strategies. Overall, there seems to be poor consensus about the usability of this ratio to differentiate piecemeal from holistic. Therefore, other than asking participants how they solved a mental rotation problem there is currently no good way to find out.

A rather visual and somewhat less utilized eye tracking analytics are heat maps, which shows the density gradient of fixations over the duration of problem solving for either an individual or a group. Heat maps are 2D graphical representation of data where the values of a variable are shown as colors, and the colors red, yellow and green are indicative of high, medium and low density visual attention, respectively (Bojko, A. 2009). Used extensively in consumer research, heat maps are more visual and less quantitative, hence ignored in many domains such as cognitive science. However, heat maps, when used correctly, can also reveal detailed patterns of visual attention and information processing, thereby offering a visually vivid summary of strategies and/or cognitive processes (Negi and Mitra, 2020).

In this ongoing study we propose to investigate whether heat maps can be used to differentiate between holistic and piecemeal strategies. Participants were asked to solve spatial thinking problems. They were then interviewed and the heatmaps were carefully examined to reveal patterns indicating either of the two strategies. Using these interviews and heatmaps we propose a tentative set of rules that can be used on heat maps to reveal underlying spatial strategies, namely, holistic and piecemeal. A more thorough and formal analysis is pending. In this study, we would like to report our preliminary results.

2. METHODOLOGY

This study uses data collected in a larger study on spatial thinking and multimodal analytics. The details of the study can be found in the paper by Ashwin T.S. et al. (under revision). The questions in the study were modeled after Peter and Battista library (2008). For the purpose of this study we have considered a subset of the main study and we have used 10 questions for our analysis, a sample question of which is shown in Figure 1.

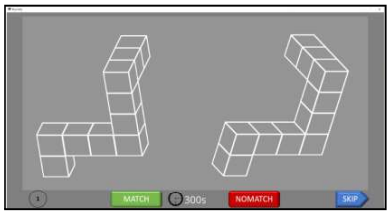

Figure 1. Screenshot of the sample question used for the study

After solving the questions the participants were interviewed and asked to reflect on what strategies they employed to answer the questions. A total of 20 interviews between 4 participants were analyzed for this study.

2.1 Interview Analysis

The interview data were analyzed independently by two coders, achieving an inter-rater reliability greater than 90%. Any discrepancies were resolved by a third rater and a coding scheme based on previous research by Hegarty (2018), was developed to identify holistic and piecemeal rotation strategies. For holistic rotation, we looked for keywords such as "rotate," "flip," "tilt," and "turn," and considered descriptions indicating that the participant mentally rotated the entire image while remaining stationary (Chu & Kita, 2011). Piecemeal rotation was identified through participants describing the process of breaking the image into parts to perform the rotation, without relying on a specific set of keywords. The interview analysis process is explained through Figure 2.

Although other strategies like local turn and cube counting were also identified, this study focused exclusively on holistic and piecemeal strategies. Only interviews explicitly stating the use of either holistic or piecemeal strategies were included in the analysis, ensuring a clear distinction between the two primary approaches.

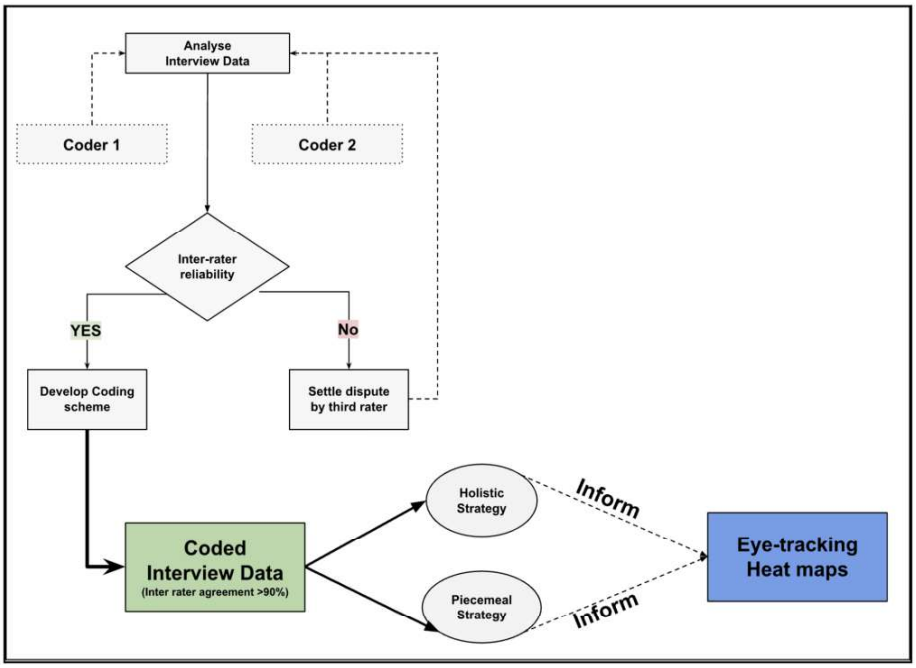

Figure 2. The Interview Analysis Process

2.2 Heat Map Analysis

The Tobii X3 sensor at a sampling rate of 120 Hz integrated with iMotions was used to collect the eye-tracking data. The data were analyzed using the I-VT (Interval Velocity Threshold) filter. The I-VT filter identifies fixations and saccades based on velocity thresholds, allowing for precise analysis of gaze patterns. Heat maps were then generated from the eye-tracking data, utilizing fixation counts to visualize areas of interest. The heat maps were then examined to see if any pattern emerged for holistic or piecemeal, as coded during the interviews.

ISBN: 978-989-8704-61-0 © 2024

3. RESULTS

3.1 Interview Analysis

We analyzed interviews from 4 participants for a total of 30 items, although there were initially 40 items (4 participants with 10 items each), as some participants could not recall the strategies they had used. Out of these 30 item-participant combinations, we discarded 10 items due to incorrect responses, as we surmised that analyzing incorrect responses could lead to misleading patterns of strategy use. Consequently, the final dataset for this study included 20 item-participant combinations. From these, we identified 10 holistic strategies, 9 piecemeal strategies, and 1 counting cubes strategy based on the interview coding.

Holistic rotations were evident from excerpts such as the following:

> *Interviewer: "How did you solve this problem? What strategies did you use?"*
> *Participant 1: "....I rotated it…"*
> *Interviewer: "Can you show me how you rotated it?"*
> *Participant 1: "...I started mapping the right figure to the left figure..I noticed that if I flip the right image upward it will take the shape of the left figure…"*
> *Interviewer: "Did you rotate the entire figure or did you rotate it in parts?"*
> *Participant 1: "Yes…I rotated the entire figure…I realized that the figures are the same only in different directions…so if I rotate the right image by some degrees I will be able to get the left figure…"*

In contrast to the above, piecemeal strategies were articulated quite differently in the interviews.

> *Interviewer: "How did you solve this problem? What strategies did you use?"*
> *Participant 2: "I was mostly looking at the figures block by block, point by point and then compare it between the two figures..I first picked one figure as my reference figure, here the left figure was the reference figure..then I looked for an easy point i.e an anchor point so that I can find the corresponding point in the right figure..so I try to find an anchor point in the left figure and I try to find the exact anchor point in the right figure..then I try to find a sub-anchor point and match it with the relative position in the other figure.."*

The participant clearly suggests that they investigated one of the figures 'block by block, point by point' and compared those with the other figure, indicating a piecemeal strategy.

3.2 Heat Map Analysis

The corresponding heat maps from those item-participant combinations (20) were subsequently investigated. Representative heat maps that typify holistic strategies as revealed by the interviews are shown in Figures 3-5. Similarly, piecemeal strategy heat maps are shown in Figures 6-8.

The heat maps were then examined to see if any pattern emerged for holistic or piecemeal, as coded during the interviews. Heat maps generated from participants employing a holistic strategy often displayed a pronounced focus in one figure. This is often evident as a single red area, most commonly not disjointed but may be of irregular shape (Figure 5), often on only one of the two figures (Figures 3 and 4) with some exceptions such as Figure 5 where we might get more than one such area between the two figures. More generally, the total attention between the two figures (all colors) often seem to be lopsided with one figure getting more attention than the other, as in Figure 4. Interestingly, however, we see that the areas of similar levels of attention (same color) between the two figures, mostly yellow and green (Figures 3 and 4) but also when red as in Figure 5, do not appear in corresponding parts of the figure. In fact, the areas seem to be complementary to each other. For example, in Figure 4, the middle two blocks seem to be covered more in the right side representation, whereas the two extreme parts of three and two blocks were in focus on the left. Similarly, in Figure 3 note the locations of the high attention areas in red in the both representations together make up the complete object. They appear on non-corresponding parts of the representations.

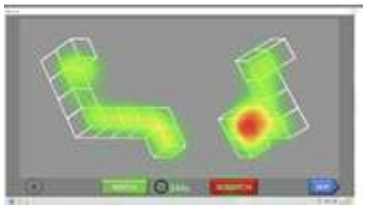

Figure 3. Heat map of participant 1

Figure 4. Heat map of participant 1

Figure 5. Heat map of participant 2

The heat maps that were coded as piecemeal through the interviews reveal starkly different patterns. Most importantly, unlike the holistic heat maps where we saw lopsided attention distribution between the two figures, the attention seems to be evenly distributed (Figure 7 and 8 but also Figure 6). Furthermore, the pattern in one figure seems to be a mirror-image of the other figure (Figures 5-8). Both of these are in stark contrast to what we observe for holistic items.

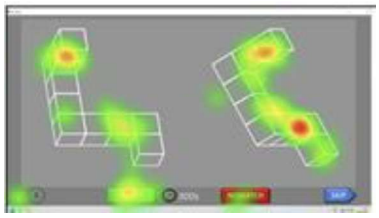

Figure 6. Heat map of participant 3

Figure 7. Heat map of participant 3

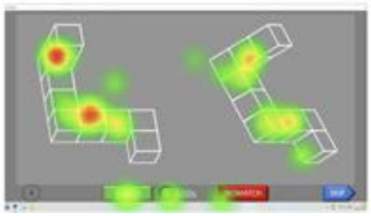

Figure 8. Heat map of participant 4

Out of the 20 item-participant combinations analyzed, we confirmed 8 as holistic and 7 as piecemeal strategies based on the heat map characteristics.

4. DISCUSSION

Eye-tracking heat maps are widely used in research to offer a visual representation of where users focus their gaze within an area of interest (AOI) (Kurzhals & Weiskopf, 2013). These maps provide a qualitative visualization by highlighting specific regions that attract the most attention during a task. A heat map

typically uses color coding—such as red for areas of highest concentration and green for lesser focus—to illustrate gaze distribution, allowing researchers to see not just where people look, but also how long they fixate on certain parts of an image. According to Wooding, fixation map analysis primarily seeks to answer the question, “Where in the image did people tend to look?” and provides an objective way to assess the similarity between different eye-movement patterns, though research into such comparisons is still relatively rare.

By integrating eye-tracking heat maps with self-reported cognitive strategies, researchers gain deeper insights into spatial thinking, especially in tasks like the Mental Rotation Test (MRT). Distinct heat map patterns often correspond to different mental strategies employed during these tasks, making heat maps a critical tool for understanding how participants process visual information. A thorough analysis of these maps, paired with qualitative data from participant interviews, can provide clues about which cognitive strategies are at play during mental rotation. A more structured, rule-based approach could be developed to systematically interpret eye-tracking data in relation to specific mental strategies. However, despite the potential, research into quantifying the patterns found in heat maps to reveal mental rotation strategies remains limited.

Heat maps generated by participants using a holistic strategy typically show a concentrated focal point on one of the figures presented in the MRT. This focal point, often marked by a red "hot spot", signifies where the participants are focusing most of their attention. The intensity of this focus suggests longer fixation durations in this area, which implies that participants might be mentally rotating the figure in its entirety during these moments. The eye remains fixed on one part of the figure while the participant performs the mental rotation task in their mind. This reflects the holistic model of mental rotation, where the object is rotated as a whole, similar to physically rotating an object without breaking it into parts.

Additionally, participants using a holistic strategy often exhibit complementary attentional distribution between the two figures presented in the task. This means that their attention is shared between the two objects, absorbing information from both orientations to construct a unified mental representation. Essentially, participants are gathering information from both objects to create a single mental model that can be rotated. This balanced focus indicates that the entire object is being mentally manipulated, not just individual parts, and the participant is processing the two orientations as parts of the same whole.

In contrast, heat maps from participants employing a piecemeal strategy show distinct focal points scattered across the figure. Unlike the holistic approach, where attention is concentrated on one area, piecemeal strategies involve breaking the figure into smaller parts and rotating each part independently. The heat maps reveal multiple focal points, typically aligned with specific components of the object. These areas of concentrated gaze indicate that the participant is analyzing each part of the figure separately, rotating it mentally before moving on to the next section. This process is characteristic of the piecemeal model, where the rotation of the object happens incrementally, part by part.

Interestingly, the focal points in piecemeal heat maps often appear as mirror images between the two figures being compared. This suggests that participants are focusing on corresponding parts of each figure, aligning them mentally to check for similarities or differences. The mirrored focal points indicate that participants are likely rotating individual parts of the object independently, comparing them across both figures before reassembling them into a cohesive whole. This method of mental rotation is more analytical and involves detailed attention to the object's smaller components.

Heat maps are not only useful for revealing visual attention patterns but also provide crucial insights into cognitive strategies. By visualizing where and how long participants focus on different parts of a figure, heat maps can differentiate between individuals who use holistic strategies and those who prefer piecemeal approaches. This ability to map out strategy use is particularly valuable for understanding inter-individual differences—how different people approach the same task—and intra-individual differences, where the same person might switch strategies across different trials. These maps provide a qualitative window into participants' thought processes, allowing researchers to infer cognitive approaches based on gaze patterns.

In the context of the MRT, heat maps can be particularly powerful for uncovering individual and group differences in spatial reasoning strategies. They allow researchers to visualize how participants' attention is distributed across different areas of the task, offering a visual summary of where and how participants engage with the mental rotation task. This can help researchers understand not only how different cognitive strategies manifest in visual attention but also how effective these strategies are in solving mental rotation problems.

In conclusion, eye-tracking heat maps offer a rich source of data for studying mental rotation and spatial reasoning. By providing both qualitative and quantitative insights into visual attention patterns, they enable researchers to explore how different cognitive strategies unfold during tasks like the MRT. Whether used to examine individual differences or broader group trends, heat maps have the potential to significantly enhance our understanding of the mental processes underlying spatial thinking.

5. CONCLUSION AND FUTURE WORK

A preliminary analysis of the interviews and heat maps have revealed the following rules:

Holistic strategy heat map signature:

1. Uneven distribution of attention between two figures. This may appear as a single dominant focus area (red spot) in one figure.
2. Complementary attentional distributions between both the figures.

Piecemeal strategy heat map signature:

1. At least 2 disjoint focal points - indicating that the figure was broken up in at least two parts.
2. Attention to equivalent points in the two figures. This makes the heatmaps look like mirror images as opposed to complimentary in holistic.

We have come up with the above rules by comparing interviews that were coded separately with the heat maps. However, we need to undertake a thorough evaluation of the aforementioned rules by looking at the entire dataset covering all interviews and all heatmaps. Only then we would be able to establish valid and reliable ways of identifying mental rotation strategies with heat maps.

We have also conjectured certain things, such as mental rotation happening in areas of high attentional focus (red areas in lopsided holistic heat maps). However, we do not yet have any evidence for the same. Gaze patterns, where we can see the time evolution of attention, could indicate whether the focus is due to actual rotation or information gathering . For example, if the gaze pattern is stalling in the said area in the middle or later parts of the problem-solving process then it could be argued to be due to active mental rotation at that point in time. However, if the attention is largely happening in the beginning then that could be simply information processing or out of confusion. We believe gaze patterns that show time evolution of attention could play a significant role in building more nuanced and accurate models of visual signatures of mental rotations (RQRQ1Mitra and McNeal, 2017).

We also remain open to the possibility of other types of mental rotation strategies apart from holistic and piecemeal coming to light with a thorough and more detailed investigation of the interviews and heat maps together.

REFERENCES

Ashwin, T. S., Ranganath, S., Khanikar, K., Khan, K., Rajendran, R., & Mitra, R. (under review). A multisensor dataset of South Asian post-graduate students working on mental rotation tasks.

Bojko, A. (2009). Informative or Misleading? Heatmaps Deconstructed. In: Jacko, J.A. (eds) *Human-Computer Interaction. New Trends*. HCI 2009. Lecture Notes in Computer Science, vol 5610. Springer, Berlin, Heidelberg, 30-39.

Chu, M., & Kita, S. (2011). The nature of gestures' beneficial role in spatial problem solving. *Journal of Experimental Psychology: General, 140*(1), 102–116.

Hegarty, M. (2018). Ability and sex differences in spatial thinking: What does the mental rotation test really measure? *Psychonomic Bulletin & Review, 25*(4), 1212–1219.

Heil, M., & Jansen-Osmann, P. (2008). Sex differences in mental rotation with polygons of different complexity: Do men utilize holistic processes whereas women prefer piecemeal ones? *The Quarterly Journal of Experimental Psychology, 61*(5), 683–689.

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 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*(2), 137–172.

Khooshabeh, P., & Hegarty, M. (2010, March). Representations of shape during mental rotation. In *2010 AAAI spring symposium series*.

Khooshabeh, P., Hegarty, M., & Shipley, T. F. (2013). Individual differences in mental rotation. *Experimental Psychology*, *60*(3), 164–171.

Mitra, R., & Negi, S. (2020). Fixation duration and the learning process: An eye-tracking study with subtitled videos. *Journal of Eye Movement Research*, *13*(6), 1–11.

Nazareth, A., Killick, R., Dick, A. S., & Pruden, S. M. (2019). Strategy selection versus flexibility: Using eye-trackers to investigate strategy use during mental rotation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(2), 232–245.

Peters, M., & Battista, C. (2008). Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain and Cognition*, *66*(3), 260–264.

Pylyshyn, Z. W. (1981). The imagery debate: Analogue media versus tacit knowledge. *Psychological Review*, *88*(1), 16–45.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703.

Uttal, D. H., & Cohen, C. A. (2012). Spatial thinking and STEM education: When, why, and how? In *Psychology of Learning and Motivation,* vol 57, Academic Press, 147-181.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, 139(2), 352–402.

Wooding, D.S. (2002). Fixation maps: Quantifying eye-movement traces. *In Proceedings of the symposium on Eye tracking research & applications—ETRA* ’02, New Orleans, LA, USA, 25–27.