Orientation and Complexity Effects: Implications for Computational Models of Visual Analogical Reasoning

Lori Buchanan
buchanan@watserv1.uwaterloo.ca
Department of Psychology
University of Waterloo

Steve Joordens
joordens@watserv1.uwaterloo.ca
Department of Psychology
University of Waterloo

Roy Fleck
refleck@descartes.uwaterloo.ca
Department of Computer Science
University of Waterloo

Paul Thagard
prthagar@logos.uwaterloo.ca
Department of Philosophy
University of Waterloo

Abstract

Several computational models have recently been proposed to define or describe visual representations. Is it reasonable to accept these models as plausible explanations of human visual processing? One way to address this question is to examine whether the models are affected by variables that have been shown to affect human visual analogical reasoning. Two such variables are stimulus complexity and differences in orientation of stimuli that must be compared. Unfortunately, the experiments that have been used to uncover these effects typically use stimuli that are too complex to be easily defined within the structure of computational models. In the present paper this problem is resolved by producing the standard set of results for complexity and orientation with a set of easily defined stimuli. We therefore see this work as a preliminary step in the comparison of human and computational models of visual processing. We report results of a human experiment investigating mental rotation and complexity effects as well as an attempt to mimic these data with an implementation of one computational model.

Introduction

How are visual images represented in the brain? Are similar representations found in AI models of visual analogical reasoning? These questions may both have answers in work which compares the performance of humans to that of computational models on standard tasks of visual analogical reasoning. Such a comparison depends on a set of stimuli which can easily be translated for use in computational models. Experiments using such stimuli have yet to be reported for human subjects. This paper addresses this concern by testing an easily translated stimuli set for a problem involving visual analogical mapping (i.e. finding correspondences between two pictorial representations in a mental rotation context).

Human Visual Processing

Orientation Effects

The time required for subjects to judge whether two visually presented stimuli are the same or different is typically a linear function of the angular disparity between the two displays (e.g.Cooper, 1975; Shepard & Metzler, 1971; Tapley & Bryden, 1977). This finding has been taken to indicate that the task involves a "mental rotation" of the stimuli such that the representation passes through the intermediate representation that would correspond to the trajectory of the item if it were actually moved in space and leads to the general claim that the representation has qualities that are similar to the physical objects they represent.

Complexity Effects

A second manipulation that affects human visual analogical reasoning is complexity of the stimulus. Increases in complexity are typically associated with an increase in response times on matching tasks. Complexity also interacts with orientation such that as complexity increases the response time is more affected by offsets of orientation (Metzler and Shepard, 1974; Bauer, 1988; Bauer & Jolicoeur, 1993; Bethell-Fox and Shepard, 1988; Jolicoeur, Regehr & Smith, 1985). Our interest centers on whether these effects can be produced by computational models. In order to investigate this question we require a stimulus set that can be used to produce typical human orientation and complexity data and that can be easily translated for use in an implementation of the computational models.

Psychological Experiment

Our goal was to produce a very simple set of stimuli that can be used as a standard against which a number of models of visual representation may be tested. A variation of the Bethell-Fox and Sheppard (1988) manipulation in which the number of filled squares in a grid determined the complexity of the item is ideal except that this manipulation results in a very small set of possible complexities given that one filled square can not be differentiated from another. For that reason we chose to use colored areas rather than restrict ourselves to black and white. In addition, we used circles with pie shaped sections rather than square arrays so that the entire object can be seen as rotated rather than the individual pieces, and so that external information, such as straight edge versus corners can be eliminated.

Method

Subjects. Eight students from the University of Waterloo served as subjects. Five of these subjects were undergraduates who received course credit for their participation while the remaining three were graduate students who participated out of interest. All subjects had normal or corrected to normal vision.

Procedure. The experiment consisted of eight blocks of 32 trials each. Each trial proceeded as follows; 1) a white fixation cross was presented at the center of the black screen until the subject initiated the trial by pressing the spacebar, 2) a blank field was presented for 200 ms, and 3) two pie-stimuli were presented to the left and right of fixation and remained on the screen until the subject responded. The pie-stimuli were circles with colored sectors (see the Stimuli and Apparatus section for an exact description of the stimuli). Subjects were required to decide as quickly and accurately as possible whether the pie-stimuli were the same or different in all respects other than orientation.

Two aspects of the pie-stimuli were manipulated across trials; orientation and complexity. Orientation was varied according to the number of sectors by which the coloring of the left and right pie-stimuli were offset. For example, when the angular disparity was 0, the right pie-stimulus was an exact copy of the left pie-stimulus. When the angular disparity was 2, the right pie-stimulus was the same as the left pie-stimulus rotated 90 degrees (i.e., 2 sectors) in a clockwise direction. Disparities of 4 and 6 resulted in rotations of 180 degrees and 270 degrees respectively.

Complexity was manipulated by varying the number of colored sectors; three colored sectors (low complexity) versus four colored sectors (high complexity). The low complexity stimuli had blue, green and red sectors. The high complexity had the three colors of the low complexity display and a violet sector. For both high and low complexity trials the red and green sectors were always adjacent to each other and never adjacent to the blue sector. The blue sector was always presented in the uppermost left sector of the left pie-stimulus (see Figure 1 for an example of the stimuli).

In addition to the two critical manipulations, there was also a same / different manipulation that occurred across trials. On same trials, the right pie-stimulus was identical to the left pie-stimulus in all dimensions except orientation. On different trials the red and green sectors of the right pie-stimulus were reversals of those in the left pie-stimulus.

Stimuli and Apparatus. The pie-stimuli used in the current experiment were circles of 8 cm diameter which were divided into eight sectors. The sectors were created by drawing four lines that connected two points of the circumference and crossed the midpoint of the circle. The lines were drawn at angles of 0, 45, 90 and 135 degrees (see Figure 1).

On an experimental trial, two such stimuli were presented on the same horizontal plane with 4.5 cm separating the two stimuli. The screen background was always black. The colored sectors were colored in accordance with the complexity and orientation manipulations of the specific trial. The uncolored sectors were always light grey. Subjects sat approximately 50 cm from the display during the experiment.

All stimuli were displayed on a Zenith flat-screen true-black monitor (Model ZCM-1490) driven by a Zenith IBM compatible processor (Model ZDF-2236-BK). The necessary software was programmed using the Micro-Experimental Laboratory software package. The numeric keypad of an IBM extended keyboard was used by the subjects to indicate their responses ('1' for same, '2' for different).

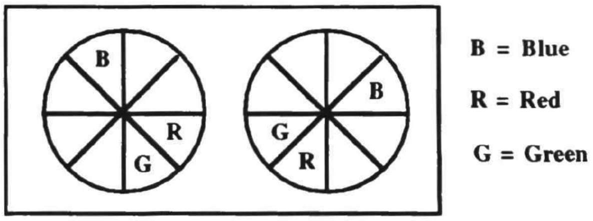

Figure 1. An example of a same trial consisting of a low complexity stimulus rotated 90 degrees.

Results

The mean judgment latencies for the correct responses on the same trials are depicted in Figure 2. Consistent with past studies of mental rotation, stimuli rotated 270 degrees (i.e., 90 degrees in a counter-clockwise direction) were associated with latencies similar to those of stimuli rotated 90 degrees in a clockwise direction, $t(7) = 0.3$, $p > 0.77$. Such a pattern implies that the stimuli were not rotated in a single direction across trials, but instead, were rotated in the most efficient direction for the current trial. Since this was the case we followed convention and collapsed across the two conditions prior to the analysis of variance.

The data were statistically analyzed using a 3 (rotation) by 2 (complexity) repeated-measures analysis of variance. This analysis revealed a significant main effect of both rotation, $F(2,14) = 18.98$, $p < .001$, and complexity, $F(1,4) = 30.09$, $p < .001$. The interaction between these factors, was also significant, $F(2,14) = 9.53$, $p < .002$. Thus, the statistical analyses confirm the implications of Figure 2. Decision times are a function of both the angular disparity and complexity such that complexity effects increase with increased angular disparity.

Discussion

The purpose of this study was to develop a stimulus set that lends itself to testing visual representational schemes in AI programs. The critical finding was that this stimulus set was able to produce the pattern of data typically obtained in mental rotation studies using much more complex stimuli.

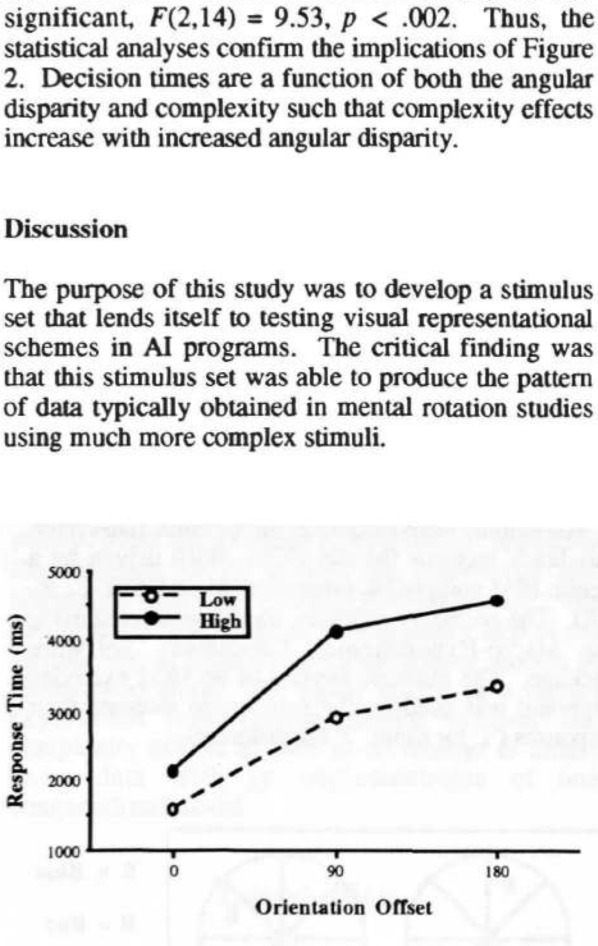

Figure 2. Mean reaction times as a function of orientation and complexity.

Computational Example

The data presented in Experiment 1 can be used both to assess current computational models of visual representation and to provide information regarding relevant characteristics in human performance. Our immediate interest concerns the degree to which the stimulus set lends itself to a test of a computational model. To this end we require computational models of visual representation which are capable of comparing two visual displays. Thagard, Gochfeld and Hardy (1992) provide us with such a model.

VAMP

Thagard, Gochfeld, and Hardy (1992) describe two Visual Analogical Mapping Programs (VAMP.1 and VAMP.2). VAMP.1 follows the array representational scheme described by Glasgow and Papadias (1992). VAMP.2 replaces the arrays with a network of interacting agents similar to those described by Minsky (1986). Thagard et al. prefer the representational scheme of VAMP. 2 because "arrays are too *boxy* to capture more complex spatial arrangements than *left, right, above, below*". Furthermore, this model lends itself to an examination of its performance in a mental rotation experiment in that its design includes a process by which one representation can be mapped to a second.

VAMP.2

Visual information is represented in VAMP.2 as a network of interconnected agents. These agents have knowledge about their immediate neighbors and can communicate this information to other agents. For example, when VAMP.2 compares two visual scenes it sets up two sub networks, one for each scene. The agents within each sub network represent some feature of the scene and spatial information is captured in terms of the relative positions of the agents to each other. The program then uses a parallel, constraint-satisfaction algorithm to pair up analogous agents across the two sub networks (cf. Holyoak and Thagard, 1989). The output of the program is a mapping of features in one scene to features in the other scene (for a more detailed description of VAMP.2 see Thagard et. al, 1992)

Thagard et al. (1992) provide several examples of VAMP.2's success in solving visual analogical reasoning. They have not, however, provided evidence that the tasks are solved in a way that could be considered similar to human visual analogical reasoning. A test of this would be to examine whether VAMP.2 produces data that mimics the human data from Experiment 1.

Computational Experiment

The implemented version of VAMP.2 operates on LISP code descriptions of visual scenes. LISP code was written to capture both the orientation and the complexity manipulations of Experiment 1. In this code the sectors of the stimuli were described with reference to their neighbors and to an external point (e.g. the top of a scene). See Appendix A for an example of the LISP code used to define the stimuli presented in Figure 1. The VAMP.2 program was run on a Macintosh Quadra 950.

The parallel constraint satisfaction network takes a certain number of cycles to solve any given visual analogical problem. The purpose of the current experiment was to assess what effect manipulations of orientation and manipulations of complexity had on the number of cycles need by the model to solve the problem.

Results

The results at each orientation and each complexity can be seen in Figure 3. The results indicate that VAMP.2 produces neither a classic rotation effect nor the expected complexity effect. Although the model is sensitive to differences in orientation, it has the greatest difficulty with stimuli offset 90 degrees. With respect to complexity, the model shows a reversed complexity effect solving the high complexity stimuli faster than the low complexity stimuli. Furthermore, although there is an interaction between complexity and orientation, the interaction does not mimic the interaction observed in Experiment 1.

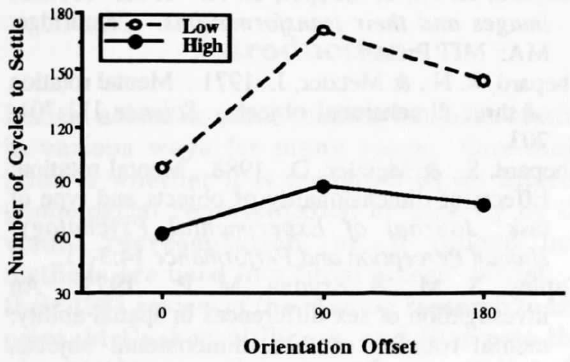

Figure 3. Mean cycles to settle as a function of orientation and complexity.

General Discussion

The rotation function in the human data is similar to past studies in that increased offsets of orientation result in increased reaction times. This function has typically been taken as evidence for some form of rotation transformation process. VAMP.2 does not produce the typical orientation function. It follows then that the failure to find the orientation function in VAMP.2 may be a consequence of VAMP.2 not including a separate rotation process.

In addition to VAMP.2's failure to produce human-like orientation effects, it also fails to produce a human-like complexity effect. Unlike human data VAMP.2 produces data which indicates an advantage for the more complex items. The failure to reproduce human-like complexity effects is most likely a due to the use of the parallel constraint satisfaction algorithm. This algorithm tries to find solutions by using constraints apparent in the definition of the visual scenes. The complexity manipulation results in an increased number of constraints. Thus, it should not be surprising that high complexity stimuli are associated with better performance.

Speculation on Future Models

Our long term goal in this project is to examine a variety of visual representational models. We expect the results of such an investigation to provide valuable information about the representational schemes required to produce an efficient computational model with human-like characteristics.

Given the data observed in Experiment 1, it is possible to make some speculations about the characteristics that a computational model should possess in order to produce human like patterns. First, the orientation function seen in the human data suggests that humans are 'mentally rotating' one of the pie stimuli prior to making a same/different response. This implies that computational models of visual processing should contain operators that are able to transform the visual representation of a given problem at various points in processing. Second, the complexity effect seen in the human data may suggest that humans simplify a visual problem down to the relevant features prior to performing any more detailed processing. This idea could be used to explain the interaction of complexity and orientation effects as well because it seems reasonable to assume that the more features that a problem has, the more difficult it would be to perform a transformation on the problem.

We are currently attempting to produce a computational model which has the characteristics discussed above. The model will be based on the graph-theoretic representation scheme proposed by

Ching, Wong and Thagard (1993). We believe that this scheme looks extremely promising for computational models of visual representation.

Acknowledgments

The authors wish to thank Ben Bauer, John Ching, Derik Hawley and Cameron Shelley for their discussion of relevant issues. This research was supported by a Natural Sciences and Engineering Reseach Council of Canada grant awarded to the fourth author.

Appendix A

```
 (make-copies-thing 'l-slices' (l-slice1 l-slice2
 l-slice3 l-slice4 l-slice5 l-slice6 l-slice7 l-slice8)
 (put-things-adjacent '(
   (right l-slice1 l-slice2)
   (right l-slice2 l-slice3)
   (right l-slice3 l-slice4)
   (right l-slice4 l-slice5)
   (right l-slice5 l-slice6)
   (right l-slice6 l-slice7)
   (right l-slice7 l-slice8)
   (right l-slice8 l-slice1) ) )
 (part-of top l-slice2)
   (set-shape 'blue l-slice1)
   (set-shape 'grey l-slice2)
   (set-shape 'grey l-slice3)
   (set-shape 'red l-slice4)
   (set-shape 'green l-slice5)
   (set-shape 'grey l-slice6)
   (set-shape 'grey l-slice7)
   (set-shape 'grey l-slice8)
(setq left-stimulus (make-scene 'left-stimulus
   l-slice1))
(make-copies-thing 'r-slices' (r-slice1 r-slice2
r-slice3 r-slice4 r-slice5 r-slice6 r-slice7 r-slice8)
(put-things-adjacent '(
  (right r-slice1 r-slice2)
  (right r-slice2 r-slice3)
  (right r-slice3 r-slice4)
  (right r-slice4 r-slice5)
  (right r-slice5 r-slice6)
  (right r-slice6 r-slice7)
  (right r-slice7 r-slice8)
  (right r-slice8 r-slice1) ) )
(part-of top r-slice2)
  (set-shape 'grey r-slice1)
  (set-shape 'grey r-slice2)
  (set-shape 'blue r-slice3)
  (set-shape 'grey r-slice4)
  (set-shape 'grey r-slice5)
  (set-shape 'red r-slice6)
  (set-shape 'green r-slice7)
  (set-shape 'grey r-slice8)
  (setq right-stimulus (make-scene 'right-stimulus
    r-slice1)) )
(defun map-left-to-right ()
  (make-left-stimulus)
  (make-right-stimulus)
  (run-vis-mapping left-stimulus right-stimulus) )
```

References

Bauer, B. 1988. Mental rotation of 2-dimensional and 3-dimensional objects. Unpublished Honours Thesis, University of Waterloo.

Bauer, B. & Jolicoeur, P. 1993. Stimulus dimensionality effects in mental rotation. In preparation.

Ching, J., Wong, A., & Thagard P. 1993. Computational imagery: An alternative framework. In preparation.

Cooper, L. A. 1975. Mental rotation of random two-dimensional shapes. *Cognitive Psychology* 7:20-43.

Glasgow, J., & Papadias, D. 1992. Computational Imagery. *Cognitive Science* 16:355-394.

Holyoak, K., & Thagard, P. 1989. Analogical mapping by constraint satisfaction. *Cognitive Science* 13:295-355.

Jolicoeur, P., Regehr, S., Smith, L. B. J. P., & Smith, G. N. 1985. Mental rotation of representations of two-dimensional and three-dimensional objects. *Canadian Journal of Psychology* 39:100-129.

Metzler, J., & Shepard, R. N. 1974. Transformational studies of the internal representation of three-dimensional objects. In R. Solso (Ed.) *Theories in cognitive psychology: The Loyola Symposium.* (pp. 147-201) Hillsdale, NJ:Erlbaum.

Minsky, M. 1986. *The society of mind.* NY:Simon and Schuster.

Shepard, R. N., & Cooper, L. A. 1982. *Mental images and their transformations.* Cambridge, MA: MIT Press.

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

Shepard, S., & Metzler, D. 1988. Mental rotation: Effects of dimensionality of objects and type of task. *Journal of Experimental Psychology: Human Perception and Performance* 14:3-11.

Tapley, S. M., & Bryden, M. P. 1977. An investigation of sex differences in spatial ability: mental rotation of three-dimensional objects. *Canadian Journal of Psychology* 31:122-130.

Thagard, P., Gochfeld, D. & Hardy, S. 1992. Visual analogical mapping. *Proceedings of the Fourteenth Annual Conference of the Cognitive Science Society.*