Title

The Perception of Disoriented Complex Objects

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 3(0)

Author

Shwartz, Steven P.

Publication Date

1981

Peer reviewed

THE PERCEPTION OF DISORIENTED COMPLEX OBJECTS

Steven P. Shwartz
Cognitive Science Program
Yale University
New Haven, CT 06520

Abstract

Subjects were presented with pictures of real-world objects at varying orientations and required to name them. Naming latencies increased with the angular departure of the pictured object from the object's canonical orientation. The results suggest that, at least when orientation-invariant features that uniquely identify the object are not present, subjects mentally rotate an internal representation of the object into its canonical orientation in order to complete the recognition process.

Introduction

Theories of visual perception are divided on the issue of how the human visual system recognizes objects at non-standard orientations. Some theorists argue that the visual system searches for orientation-invariant features and bases perceptual recognition on these features (e.g. Milner, 1974). The fact that certain common objects seem to be immediately recognizable at non-standard orientations is compatible with this view. Other theorists (e.g. Rock, 1973) argue that in order to recognize a disoriented object, the object must be first imagined in its canonical orientation. The fact that some complex objects (e.g. faces) are difficult to recognize when disoriented supports this position.

The classic experiments of Cooper and Shepard (1975) suggest that in order to imagine disoriented objects at their standard orientations subjects perform a mental rotation of an internal representation of the object. The Cooper and Shepard paradigm required subjects to judge whether or not a presented letter had a normal or mirror-reversed form. They found that the time to make this decision was a monotonic function of the angular distance of the orientation of the presented letter from the canonical orientation of the letter. This suggests that before it was possible to make the required judgement, it was first necessary to perform a mental rotation of an internal representation of the presented letter in order to imagine it in its canonical orientation.

However, as Cooper and Shepard noted, these results may have been due to a special strategy developed by subjects for the purpose of performing the normal vs. mirror-image judgement because normal and mirror-image letters can not be distinguished on the basis of distinctive features. For this reason, the Cooper and Shepard experiments are not strong evidence that mental rotation is normally used in the recognition of disoriented objects. In fact, Corballis, Zbrodoff, Shetzer, and Butler (1978) have shown that under certain circumstances (i.e. when a subject is required to determine whether a presented stimulus was a specified target letter or one of 5 distractor letters), the time to identify a letter is independent of the orientation of the presented letter.

Thus, it is clear that people are capable of performing mental rotations on internal representations of perceptual stimuli in order to normalize the orientation of the perceptual stimulus prior to recognition. However, it is equally clear that mental rotation is not always required for the recognition of disoriented forms. Under certain conditions, people are capable of utilizing distinctive features of perceptual stimuli in order to recognize disoriented perceptual stimuli. In the Cooper and Shepard study, subjects were forced to use mental rotation because of a lack of available distinctive features. In the Corballis et. al. study, it was possible that subjects were able to extract a small set of features that could be used for identification of the target letter and that these features were recognizable at all orientations.

The purpose of the present study is to determine the extent to which mental rotation is used in the perceptual recognition of disoriented objects that a person might encounter in the physical world. In the present study, subjects are presented with drawings of physical world objects at varying orientations and required to identify them by name. If mental rotation is used to identify disoriented stimuli, then the time required to identity each stimulus object should be a monotonic function of the angular distance of the presented stimulus object from the canonical orientation of the object. Subjects were tested on novel stimuli -- i.e. subjects had no foreknowledge of the indentities of the objects they would see -- in order to ensure that they could not use a strategy of searching only for a few particular features.

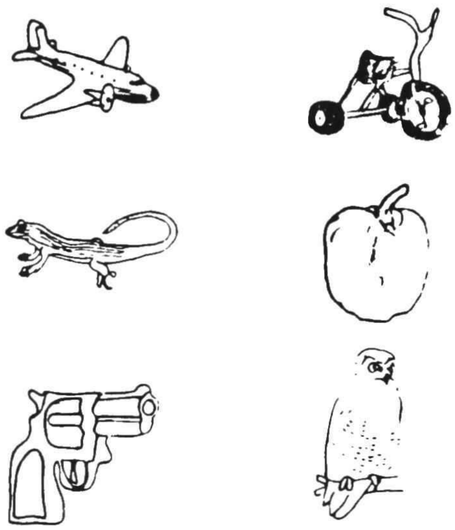

Figure 1. Examples of the stimulus objects used in the experiment.

Method

Subjects.

Twelve subjects from The Johns Hopkins University served as subjects in the experiment in partial fulfillment of a course requirement.

Materials and Apparatus.

The stimuli used in the experiment consisted of 48 xeroxed drawings of real-world objects, examples of which are shown in Figure 1. The stimulus set included 15 animals, 30 inanimate objects, and 3 well-known faces (Washington, Carter, and Reagan). Six versions of each of the 48 drawings were constructed, one at each of the following angular departures (in degrees) from the object's normal upright: 0, 60, 120, 180, 240, and 300. The stimuli were presented to subjects individually in a two-field tachistoscope, subtending an average horizontal and vertical angle of 3.2 degrees.

Procedure.

On each trial of the experiment, the experimenter simultaneously said the word "READY" and pressed a button initiating the trial. After a .5 sec warning interval, the stimulus was presented and remained visible until the subject verbally named the depicted object. The subject's verbal response triggered a voice activated relay that stopped a response timer and the subject's response latency was recorded. Before the start of the experiment, the experimenter explained the operation of the voice key and gave each subject practice at triggering the voice key only at the time of response. Each subject was instructed to verbally name each stimulus as rapidly as possible without making errors. In order to reduce the ambiguity of possible stimulus names, subjects were instructed that the response given should be the name that one would normally use in referring to the stimulus object and, as an example were told that the correct response to a picture of an animal is the name of the animal (and not "animal").

Design.

Each subject received 48 test trials, one for each of the 48 objects. Each subject was tested on 8 trials at each of the 6 orientations. Thus, each subject was tested on only one of the 6 orientations for each stimulus in order to ensure that a novel stimulus was presented on each trial. For each set of 6 subjects, each of the 48 objects were tested at each of the 6 orientations. Each subject received 12 practice trials with stimulus objects that were not included in the set of test stimulus objects.

Results.

Responses were considered correct when the name given was appropriate and was not the name of a superordinate category. The largest number of errors made by a subject was 4 and error trials were not included in the analyses.

The latency data were analyzed separately with both subjects and objects as random variables. The direction of disorientation -- clockwise (60 and 120 degrees) vs. counter-clockwise (300 and 240 degrees) -- had no effect on response latencies [$F < 1$ for both subjects and objects], nor was there a significant direction x orientation (60 vs. 120 degrees) interaction [$F < 1$ for both subjects and objects]. For this reason, the data for the 60 and 300 degree orientations as well as for the 120 and 240 degree orientations were not distinguished in the analyses.

The primary results of the experiment are shown in Figure 2. As can be in Figure 2, there is a monotonic effect of the angular departure from the canonical orientation of the object on response latencies [$\min F' = 3.519$, $p < .05$]. This analysis was also run without inclusion of the 3 faces and the pattern of results did not change as a result of this analysis. This result is compatible with the view that normalization of orientation occurs prior to recognition.

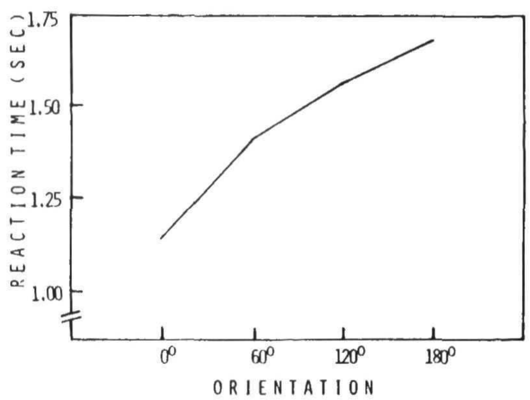

Figure 2. Reaction time as a function of the angle of departure from the canonical orientation of the object.

For each of the 48 stimulus objects, the best-fitting straight line relating the angular orientation of the object to response latency was computed using linear regression. The correlation coefficient of the regression provides a measure of how well the data are fitted by a linear function. The median correlation coefficient was .602, with 36 positive correlation coefficients and 12 negative coefficients. These data, along with the linear trend in evidence in Figure 2, constitute evidence that recognition difficulty increases as the degree of angular departure from the object's canonical orientation increases.

Discussion.

These results are consistent with the view that people use mental rotation to normalize their internal representations of visual stimuli in order to facilitate recognition. However, while a monotonic function of orientation was found for most of the stimuli used in the present experiment, there were stimulus objects for which such a trend was not found. Since orientation is manipulated between-subjects -- i.e. no subject sees a stimulus at more than a single orientation -- the non-monotonically-increasing trends could have been due to noisy data.

However, it is equally likely that for some of these stimuli, there do exist orientation-invariant features that are used to identify the object. For example, three objects for which a linear trend was not found were the owl, the alligator, and the gun. It is easy to imagine how the big eyes of the owl, the long teeth and curvy tail of the alligator, and the barrel and trigger of the gun can serve to uniquely identify each without mental rotation. These features are easy to spot at arbitrary orientations and are unique to these objects.

Corballis et. al. (1978) were similarly able to obtain identification functions for letters that were independent of orientation when the sets of positive and negative features were sufficiently restricted (by restricting the sizes of the sets of positive and negative letters). It is hypothesized that in the Corballis et. al. experiment these restrictions had the effect of increasing the cue validity of certain features of the letters that could be identified at arbitrary orientations to the point where these features could uniquely identify the letter. That is, Corballis et. al. were able to create a context in which the available orientation-invariant features were sufficient to uniquely identify the letter.

The results of the present experiment suggest the following algorithm for the identification of disoriented perceptual stimuli: When orientation-invariant features are available and are sufficient to uniquely identify an object, they are used. However, when such features are not available, or when these features do not uniquely identify the object, a mental rotation of the internal representation of the object is performed in order to extract features that are not orientation-invariant.

The author wishes to thank Pierre Jolicoeur for a helpful discussion of some of the issues involved in this research, Nancy Tang for collecting the data, and Ray Gibbs for comments on the manuscript.

References

Cooper, L. A. Shepard,R.N., Chronometric studies of rotation of mental images, in W.G. Chase (Ed.), _Visual Information Processing_, New York: Academic, 1973.

Corballis, M. C., Zbrodoff, N. J., Shetzer, L. I. & Butler, P. B. Decisions about identity and orientation of rotated letters and digits, _Memory and Cognition_, 1974, _6_(2), 98 107.

Milner, P. A. A model for visual shape recognition, _Psychological Review_, 1974, _81_, 521 - 535.

Rock, I. _Orientation and form_. New York: Academic Press, 1973.