Orientation-Dependent Mechanisms in Shape Recognition: Further Issues

Michael J. Tarr and Steven Pinker

Psychological Science 1991 2: 207
DOI: 10.1111/j.1467-9280.1991.tb00135.x

Commentary

ORIENTATION-DEPENDENT MECHANISMS IN SHAPE RECOGNITION: Further Issues

Michael J. Tarr[1] and Steven Pinker[2]

[1]*Department of Psychology, Yale University;* [2]*Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology*

Tarr and Pinker (1989, 1990) delineated the conditions in which the orientation of a shape affects the time subjects take to recognize it. Relatively large effects of orientation were found (1) when misoriented asymmetrical shapes were unfamiliar, and (2) when familiar asymmetrical shapes appeared in unfamiliar orientations. In contrast, relatively small effects or no effect of orientation were found (1) when familiar asymmetrical shapes appeared at familiar orientations, (2) when unfamiliar mirror-reversals of familiar asymmetrical shapes appeared at unfamiliar orientations, (3) for misoriented symmetrical shapes, whether they were unfamiliar or familiar, and (4) for misoriented bilaterally redundant shapes, whether they were unfamiliar or familiar. To account for these data, two questions must be addressed: What mental processes underlie large effects of orientation or their absence; and why do these two patterns of data depend on manipulations of familiarity, handedness, and shape geometry? In Tarr and Pinker (1989, 1990) we proposed a theory that simultaneously answered both of these questions (*what* processes are used when very small effects of orientation are observed, and *when* are these processes used).

The theory, *multiple-views-plus-transformation,* suggests that large effects of orientation in shape recognition are due to mental rotation. Crucially, this conclusion is based not only on the presence of an orientation effect, but on the similarity between our rotation rates and the rates observed in experiments by Shepard and Cooper (1982). These experiments used independent evidence to demonstrate the existence of an incremental mental rotation transformation, *not confined to the effects of orientation,* but depending on converging manipulations such as response time to probes at intermediate orientations and presentation points, effects of advance information and preparation time, and other techniques.

We suggested that the absence of such large effects of orientation may result from orientation-invariant mechanisms of three kinds: orientation-specific representations of familiar shapes at familiar orientations; a 180° rotation in depth to align unfamiliar mirror-reversals with their familiar standards; and 1½ D orientation-independent descriptions for symmetrical and bilaterally redundant shapes. Together these hypotheses not only explain the causes of the absence of mental-rotation-size orientation effects, but predict the reasons why this slope diminution occurs in particular conditions. For instance, for asymmetrical shapes, diminished effects of orientation occur only after extensive practice at familiar orientations and do not transfer to the same shapes in unfamiliar orientations (Tarr & Pinker, 1989; Tarr, submitted). Furthermore, response times at unfamiliar orientations generally increase with distance from the nearest familiar orientation, and at a rate comparable to the rate of rotation observed in the initial phase of the experiment, indicating that the same orientation-dependent process is used in both cases. The hypothesis not only accounts for this pattern by positing rotation to orientation-specific descriptions, but helps explain its onset, occurrence, or absence by positing that such descriptions develop only with experience. In a similar fashion, the near-flat—and much higher—response time functions for mirror-reversed versions of familiar shapes are explained by the geometric fact that a constant amount of time is required to rotate unfamiliar handedness versions into congruence with familiar versions by a "flip" in depth. This is confirmed by several additional, independent manipulations: Large orientation effects return when two-dimensional shapes are learned as mirror-image pairs, when handedness discriminations rather than shape recognition is the required task, and when three-dimensional, rather than two-dimensional, objects are learned (see Tarr, submit-

Send correspondence to: Michael J. Tarr, Department of Psychology, Yale University, P.O. Box 11A Yale Station, New Haven, CT 06520. e-mail: tarr@cs.yale.edu.

ted; Tarr & Pinker, 1989; for discussion of how flips in depth are not possible strategies in exactly these circumstances). Finally, positing orientation-independent descriptions for classes of bilaterally redundant shapes not only accounts for the absence of large effects of orientation (Tarr & Pinker, 1990), but predicts their occurrence for both symmetrical and bilaterally redundant shapes, even when they are unfamiliar.

Johnson (1991) asks: To what do we attribute small, but persistent, effects of orientation on the time to identify rotated shapes, in those conditions where large effects (comparable to those obtained by Shepard and Cooper, 1982) are absent? Johnson argues that such effects indicate an extremely fast rotation mechanism rather than a multiple-views theory of shape recognition or orientation-invariant theories (which suggest that near-flat response time functions are the result of object-centered or feature-based mechanisms). We feel that there is a problem with Johnson's hypothesis: Although he has raised a logically possible alternative explanation for some subsets of data, he does not provide an account for the overall pattern of results found within and across relevant experiments.

Johnson is correct that, in its simplest form, our explanation for these phenomena assumes a distinction between large effects of orientation (comparable to that found in "classic" mental rotation studies) and no effects of orientation, not the large/small distinction we actually find. There appear to be at least three plausible accounts for why we obtain small effects of orientation rather than perfectly flat response time functions: the existence of orientation-dependent mechanisms other than rotation; a mixture of rotation and nonrotation strategies; and, Johnson's suggestion, the existence of very rapid rotation.

Of these three explanations, there is good evidence only for the first. As elegantly shown in Carpenter and Just's (1978) study of eye movements, the mental rotation process is composed of at least *two* orientation-dependent mechanisms: a rapid search for the landmark features of a misoriented shape that determine its likely axis of rotation, top, and sides, followed by the actual, much slower, rotation. There is also evidence that the orientation-dependent effects of feature search may be removed from mental rotation data: By clearly marking the ends of three-dimensional objects with colored dots, Metzler (reported in Shepard & Cooper, 1982) obtained slight reductions in the rate of mental rotation, suggesting that a fast orientation-dependent process had been bypassed. Such findings are compatible with current computational theories of recognition by alignment. For instance, Ullman's (1989) model contains two major steps: the use of local landmarks or prominent axes in the input to compute an appropriate transformation, and the subsequent execution of the transformation. When memory contains a shape description that is stored in the same orientation as the input, the landmark or axis search will be used to match the input to it; in this case, the amount of rotation that is needed just happens to be zero. Therefore, response times for familiar orientations would still exhibit small effects of orientation—the empirical signature of a rapid orientation-dependent axis/feature-finding mechanism.[1]

The second possible explanation for why subjects' response times might display small orientation effects, even if they were generally recognizing shapes appearing at familiar orientations by direct matches to orientation-specific descriptions, is that they occasionally revert to a rotation strategy (to a "canonical" orientation). The slopes of the functions relating response time to orientation would be nonzero because they would be the weighted averages of data from a large number of trials where rotation was not used (with a slope of zero) and a small number of trials where rotation was used (with a standard rotation slope). This hypothesis has been suggested by Corballis (1988) and by Corballis, Zbrodoff, Shetzer, and Butler (1978).

Unlike these two explanations, it is unclear how Johnson's third, the "rapid rotation" account for small effects of orientation, can accommodate the much more dramatic effects of orientation on familiar shapes at unfamiliar orientations (shown in Tarr & Pinker, 1989). Clearly, subjects have learned something that is orientation-specific—a rapid rotation account would have to make the ad hoc proposal that what they learned is to rotate rapidly only at familiar orientations, but are incapable of the identical process, but at a different rate, for identical shapes at unfamiliar orientations.[2] In addition, the only result that Johnson cites in favor of a rapid rotation process, Jolicoeur and Landau's (1984) study, is inconclusive. A fast, orientation-dependent feature

[1] The existence of an orientation-dependent axis-finding mechanism within the mental rotation process also provides an explanation for the "M"-shaped response time functions sometimes obtained for naming misoriented shapes (see for instance, Jolicoeur, 1985). The "dip" in such functions occurs at 180°—the only orientation other than the upright where most shapes' vertical midline axis is already aligned with the egocentric and the environmental upright. Consequently, at these two orientations, the axis-finding mechanism may be bypassed. Therefore, while all other response times will have at least two components, the time to locate the axis and the time to execute the rotation, response times for shapes appearing at 0° and 180° may be missing the search-for-axis component, rendering them smaller than response times for the same shapes at nearby, nonvertically aligned orientations.

[2] Because some of the unfamiliar orientations (where slow rates of rotation were observed) fell *between* the familiar orientations and the upright, it is unlikely that subjects simply learned the appearance of shapes in intermediate orientations, thereby facilitating faster transformations to the upright.

search can also account for the small effect of orientation on error rates: The tachistoscopic presentation of the stimuli may have disrupted the successful completion of such a search. As stimuli were oriented further from the upright, the major vertical axis and other features may have taken longer to locate, exceeding the time available before the stimulus item was erased. Thus, Jolicoeur and Landau's results are consistent with current theories of shape recognition, all of which rely on the location of appropriate local features (see for example, Biederman, 1987; Marr & Nishihara, 1978; Ullman, 1989). Furthermore, a rapid feature-search mechanism does predict the small effect of orientation on response times: As shape misorientation increases, more time will be required to locate the features underlying recognition by alignment, direct match, or other mechanisms.

In sum, Johnson's proposal has the following problem: In attributing residual effects of orientation on recognition time to a mechanism that is identical to the one that yields large effects, only operating at a faster speed, he cannot explain why these differences in slopes can reliably be produced by *qualitative* manipulations such as handedness, symmetry, dimensionality, task, and familiarity. Furthermore, the explanation does not appear to have independent motivation, whereas at least one alternative, rapid feature/axis search, has both independent experimental and computational support. For these reasons, we suggest that Johnson's hypothesis, while logically compatible with subsets of our data, is not the preferred explanation for the phenomena as a whole.

REFERENCES

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115–147.

Carpenter, P.A., & Just, M.A. (1978). Eye fixations during mental rotation. In J.W. Senders, D.F. Fisher, & R.A. Monty (Eds.), *Eye movements and the higher psychological functions*. Hillsdale, NJ: Erlbaum.

Corballis, M.C. (1988). Recognition of disoriented shapes. *Psychological Review, 95*(*1*), 115–123.

Corballis, M.C., Zbrodoff, N.J., Shetzer, L.I., & Butler, P.B. (1978). Decisions about identity and orientation of rotated letters and digits. *Memory & Cognition, 6*, 98–107.

Johnson, S.H. (1991). Commentary on Tarr and Pinker. *Psychological Science, 2*(2), 206–207.

Jolicoeur, P. (1985). The time to name disoriented natural objects. *Memory & Cognition, 13*(*4*), 289–303.

Jolicoeur, P., & Landau, M.J. (1984). Effects of orientation on the identification of simple visual patterns. *Canadian Journal of Psychology, 38*(*1*), 80–93.

Marr, D., & Nishihara, H.K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. *Proceedings of the Royal Society of London, Ser. B, 200*, 269–294.

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

Tarr, M.J. (1990). Rotating objects to recognize them: The effect of orientation on the recognition of three-dimensional objects. Submitted.

Tarr, M.J., & Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology, 21*(2), 233–282.

Tarr, M.J., & Pinker, S. (1990). When does human object recognition use a viewer-centered reference frame? *Psychological Science, 1*(*4*), 253–256.

Ullman, S. (1989). Aligning pictorial descriptions: An approach to object recognition. *Cognition, 32*, 193–254.