Journal of Experimental Psychology:
Human Perception and Performance
1991, Vol. 17, No. 4, 1075-1089

0096-1523/91/$3.00

Shape Recognition Contributions to Figure–Ground Reversal: Which Route Counts?

Mary A. Peterson, Erin M. Harvey, and Hollis J. Weidenbacher
University of Arizona

Observers viewed upright and inverted versions of figure–ground stimuli, in which Gestalt variables specified that the center was figure. In upright versions, the surround was high in denotivity, in that most viewers agreed it depicted the same shape; in inverted versions, the surround was low in denotivity. The surround was maintained as figure longer and was more likely to be obtained as figure when the stimuli were upright rather than inverted. In four experiments, these effects reflected inputs to figure–ground computations from orientation-specific shape representations only. To account for these findings, a nonratiomorphic mechanism is proposed that enables shape recognition processes before figure–ground relationships are determined.

The ability to recognize a shape is coupled to figure–ground organization in that the region to which the contour is assigned has shape and meaning whereas the other region is shapeless and simply appears to continue behind the figural region, as illustrated by the Rubin vase-faces stimulus in Figure 1. When the figure–ground contour is assigned to the black region in Figure 1, the vase can be seen (recognized), but the faces cannot. The white region, lacking a contour, simply appears to continue behind the black region. Alternatively, when the figure–ground contour is assigned to the white region, the faces can now be seen (recognized), but the vase cannot. The black region now is lacking a contour, and consequently is seen as an undifferentiated background continuing behind the faces (Rubin, 1915/1958).

Despite the linkage between shape recognition and figure–ground organization, the Gestalt psychologists proposed that the assignment of a figure–ground contour was determined by "autochthonous" or stimulus variables; and that following this step, shape recognition routines were engaged for the figural region only. Support for this claim came from experiments that used displays devoid of meaningful shapes, in which figure–ground contours were shown to be assigned to those regions that were symmetric, enclosed, more convex, or smaller in area (e.g., Graham, 1929; Harrower, 1936; Kunnapas, 1957). Although these demonstrations clearly showed that figure–ground relationships *can* be determined without shape recognition input, they did not rule out the possibility that figure–ground computations can weigh inputs from shape recognition analyses as well.[1] Indeed, Rubin (1915/1958) observed that once regions depicting meaningful shapes are obtained as figure, they are maintained as figure for longer durations than meaningless regions; an observation that has, by and large, been disregarded (but see Rock, 1975, p. 358). In fact, since the time of the Gestalt demonstrations, the traditional view has been that the determination of figure and ground relationships precedes the operation of shape recognition processes (Biederman, 1987; Hebb, 1949; Kanisza & Luccio, 1987; Koffka, 1935; Kohler, 1929; Kosslyn, 1987; Marr, 1982; Neisser, 1967; Wallach, 1949; but see Lowe, 1985; Witkin & Tenenbaum, 1983).

The experiments described in this article examine whether shape recognition processes can contribute to those computations involved in the reversal of figure and ground relationships. Recent evidence suggests that two different processes, or routes, may be involved in shape recognition: One route may entail access to orientation-specific structural representations; another route may entail access to orientation-independent parts (e.g., Corballis, 1988; Humphreys, Riddoch, & Quinlan, 1988; Marr & Nishihara, 1978; see Ullman, 1989, for a review). In what follows, we first discuss the evidence for these two shape recognition routes, and we then present a series of experiments examining whether outputs from either or both of these shape recognition routes can serve as input to figure–ground reversal computations.

Orientation-Specific Structural Memory Representations

One process through which shapes are recognized entails matching a perceptual description of the structure of a stimulus shape to the best fitting *structural memory representation* (Marr, 1982; Marr & Nishihara, 1978). We use the term structural memory representation quite specifically to refer to memory representations that specify the parts (features) of a shape and their relative locations with respect to the object's canonical orientation (Hochberg, 1972; Hochberg, 1978; Pal-

This research was supported by National Science Foundation Grant BNS-8810997 to Mary A. Peterson, and by the U.S. Air Force Office of Scientific Research. We thank Julian Hochberg, Irvin Rock, Brad Gibson, Ken Forster, Walter Gerbino, and two anonymous reviewers for their comments on an earlier version of this article.

Correspondence concerning this article should be addressed to Mary A. Peterson, Department of Psychology, University of Arizona, Tucson, Arizona 85721.

[1] Gottschaldt's studies (cited in Koffka, 1935) conducted experiments pitting familiarity against other factors like good continuation, but the configurations he used were not figure–ground configurations.

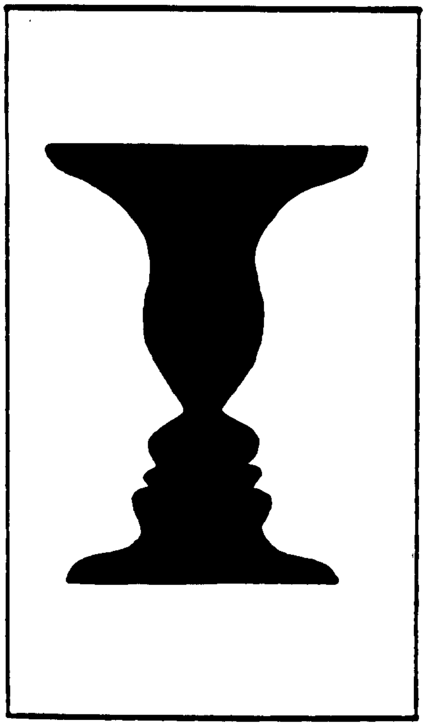

Figure 1. The classic Rubin vase–faces figure–ground stimulus.

mer, Rosch, & Chase, 1981). The perceptual description of a disoriented shape may have to be normalized before it is matched to the appropriate orientation-specific structural representation. A number of studies demonstrating that the time to name stimuli increases incrementally as they are increasingly disoriented from their canonical upright (e.g., Jolicoeur, 1985, 1988; Jolicoeur & Milliken, 1989; Maki, 1986; Rock, 1973; Tarr & Pinker, 1989) have been taken as evidence that this normalization process is time-consuming, with more time required to compensate for larger degrees of disorientation.

Feature Extraction

A second route through which shapes might be recognized entails the identification of orientation-independent distinctive features or parts (Corballis, 1988; Humphreys et al., 1988; Selfridge & Neisser, 1960). A number of lines of research, investigating both imagery and perception, have been taken as evidence that a normalization process is not always required prior to recognition (e.g., Cooper & Shepard, 1973; Corballis, 1988). For example, in some situations, reaction times to name letters and letterlike forms are unaffected by disorientation (Corballis, Zbrodoff, & Roldan, 1976; Eley, 1982; Jolicoeur & Milliken, 1989; cf. Takano, 1989). Additional evidence that shape recognition routes entailing access to distinctive features or parts differ from those entailing access to structural representations comes from neuropsychological experiments demonstrating that brain damage can interfere selectively with one or the other of these routes (Humphreys & Riddoch, 1984).

In the experiments described here, we examined whether reversals of figure–ground organization can be influenced by contributions from either of the two shape recognition routes described above. The stimuli we used are shown in Figure 2.[2] These stimuli were biased by Gestalt variables of symmetry, convexity, and/or enclosure to favor the center-as-figure interpretation, yet were drawn so that the surrounding region denoted a more meaningful shape than the center: The surround of Figure 2A denotes two half-silhouettes of standing women, and the surround of Figure 2B denotes two profiles of people with hooked noses.

In Experiment 1, we showed that despite the stimulus bias toward the center as figure, the surrounding region of these stimuli was maintained as figure for longer durations when it was upright rather than inverted. In Experiments 2–4, we tried to pinpoint the effects as reflecting contributions made to figure–ground reversal computations from orientation-specific structural memory representations.

General Method

In all of the experiments reported in this article, observers followed opposed-set instructions while viewing reversible figure–ground stimuli. Accordingly, we describe the opposed-set procedure in some detail and discuss previous research demonstrating its value.

The *opposed-set procedure* was introduced by Peterson and Hochberg (1983; Hochberg & Peterson, 1987; Peterson, 1986) to test whether viewers could intentionally influence the perceived organization of ambiguous displays. In the opposed-set procedure, observers' fixation is controlled and their intentions are manipulated by overt instructions to try to hold one of the potential interpretations of a reversible figure. Observers follow these hold instructions for 30-s trials and report continually about the perceived organization of the reversible figure.

In previous experiments, Hochberg and Peterson (1987; Peterson, 1986) recorded both direct responses about the variable to which the hold instructions referred and responses about a variable that was "perceptually coupled" to the instructed variable as an indirect measure of the perceptual status of the instructed variable. Perceptually coupled variables are variables that normally covary in the physical world (e.g., parallax and the relative distance to the two faces of a three-dimensional cube) and that also covary in perception (Epstein, 1982; Hochberg, 1974). A number of experiments and demonstrations imply that responses about perceptually coupled variables are unaffected by response bias (e.g., Gogel & Tietz, 1974; McCracken, Gogel, & Blum, 1980; Peterson & Shyi, 1988). The use of responses about perceptually coupled variables allowed Hochberg and Peterson (1987; Peterson, 1986) to localize their intention effects as occurring in perceptual processes rather than in memory or response proc-

[2] These figures were jointly devised by Julian Hochberg and Mary Peterson (see Peterson & Hochberg, 1989).

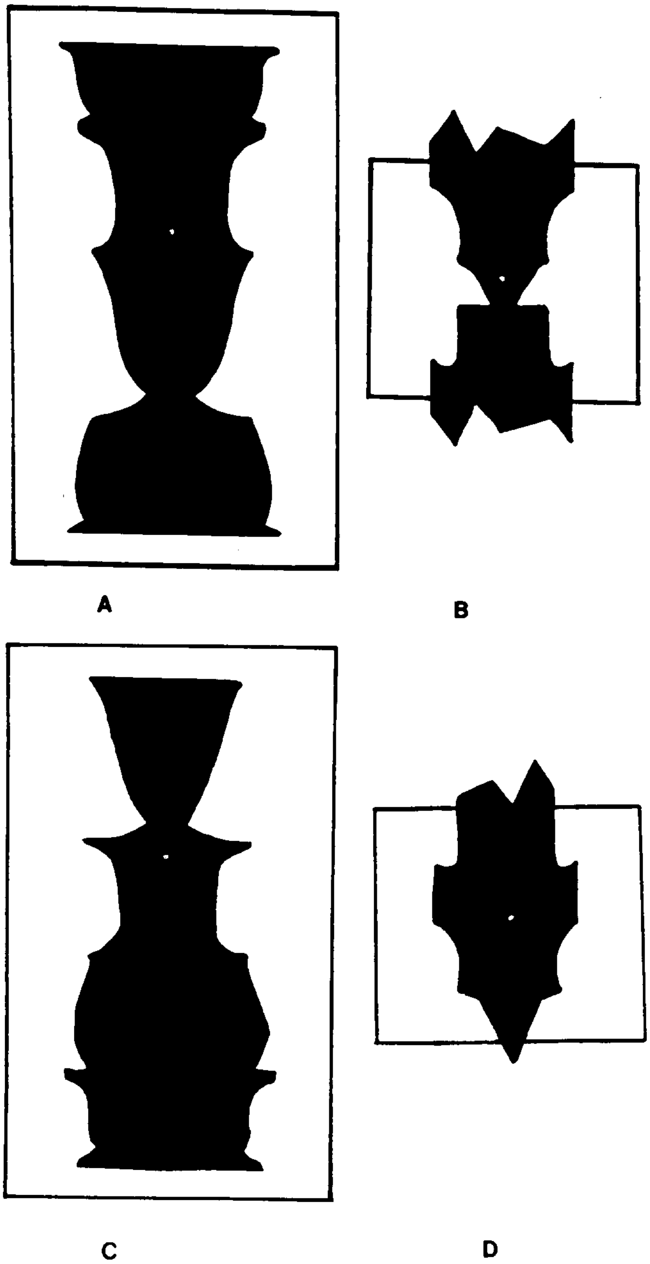

Figure 2. Panels A and B are the figure–ground stimuli used in Experiments 1, 3, and 4. (The surrounds in Panel A resemble silhouettes of standing women, and the surrounds in Panel B resemble profiles of faces.) Panels C and D are scrambled versions of Panels A and B, respectively.

esses. This was an important step because the locus of previous attempts to demonstrate that cognitive variables contribute to perceptual organization was questionable (e.g., Bruner & Goodman, 1947; Rock & Gutman, 1983; Schafer & Murphy, 1943).

One implication of Hochberg and Peterson's finding that observers' intentions can influence perceived organization is that hold instructions must be used in situations in which a large subset of the observers might have a preference to see one of the potential alternatives for a reversible figure. In such situations, it may be impossible to obtain baseline measures of reversal patterns before the introduction of the hold instructions. Instead, the use of the opposed hold instructions is intended to place upper and lower limits on observers' intentions and on their responses to the demand character of the experiment as well. Variables that influence perceived organization regardless of observers' intentions will be reflected in durations of both intended and unintended interpretations (see Hochberg & Peterson, 1987; Peterson, 1986; Peterson & Hochberg, 1983, 1989). Moreover, variables that influence the effectiveness of viewers' intentions can provide some indication of the mechanisms through which intention operates.

In the experiments reported in this article, we examined whether shape recognition inputs contribute to computations regarding figure–ground reversal, regardless of the viewers' intentions. We also examined differences in the effectiveness of the hold instructions from one condition to another for evidence regarding the mechanisms of intention. Previous research has shown that neither eye movements nor changes of convergence angle are necessary mediators of the intention effects (Peterson, 1984, 1986; Peterson & Gibson, 1991a), thereby leaving open the possibility of central mediation. One possibility is that intention can operate through one of the shape recognition routes examined here.

Experiment 1

In Experiment 1, we examined whether shape recognition inputs contribute to figure–ground reversal by comparing the durations that observers reported seeing the surround as figure in upright and inverted versions of our stimuli. We chose to compare performance with upright versus inverted stimuli on the basis of a preliminary experiment, in which we asked other observers to identify the shapes resembled by both the center and the surround when they were presented alone in either upright or inverted orientations. Consider first the results obtained for the upright and inverted surround: Of those observers who viewed the upright stimuli, 97% reported seeing the interpretations we intended for the surround, and most reported only those interpretations. Satisfying this joint criterion qualified the surrounding regions of the upright stimuli as highly *denotative* regions (See Peterson, Rose, Gibson, & Vezey, 1991).[3] Of those observers who viewed the inverted version of the surround, however, only 16% agreed on a single interpretation for the surround as specified in viewer-centered coordinates, and most gave more than one

[3] To "denote" means to "refer to explicitly" or to "mean." Hence, the *denotivity* of the pictorial image is the meaning of the image. We use the term *denotivity* rather than *meaning* to refer specifically to the structural meaning of a shape and not to the functional meaning or semantic associations; the latter are "connoted" by a shape rather than denoted. See Clowes (1971) for a similar use of the term denote: "To say that a situation in a domain (e.g., a picture or sentence in English) means something is to identify the situation in a second domain which it denotes" (p. 112).

interpretation. Hence, the inverted surround was classified as a *less denotative* shape than the upright surround. By these same criteria, the center region of both figures was found to be a low denotative region in both the upright (16% agreement) and the inverted (12% agreement) orientations.

We assume that the strength of the output from shape recognition routes is greater for regions of greater denotivity. Therefore, a finding that observers see the surround as figure for significantly longer durations when they view upright rather than inverted versions of the stimuli would be consistent with the proposal that outputs from shape recognition processes can serve as inputs to figure–ground reversal computations.

We chose to compare the durations of seeing the surround as figure across upright and inverted orientations rather than to compare performance across center and surround regions at any one orientation because the center and surround differ in area, symmetry, convexity, and/or enclosure, all of which favor the center as figure. None of these stimulus variables is changed by inversion, however, whereas the denotivity of the surround is. Consequently, a comparison of the durations of perceiving the surround as figure across upright and inverted orientations can reveal shape recognition inputs while stimulus inputs are held constant.

Method

Subjects. The subjects were 16 students at the State University of New York at Stony Brook who took part in this experiment to fulfill a research participation requirement for an introductory psychology course. All had normal or corrected-to-normal vision.

Stimuli and apparatus. The stimuli were black and white drawings of the figure–ground stimuli shown in Figure 2. The center regions of the figures were black and were smaller in area than the surrounds: For Figures 2A and 2B, respectively, the areas of the center regions were .66 and .77 that of the surrounds.[4] The center of Figure 2A was symmetric; the center of Figure 2B was not, although it could be considered partially symmetric (see Palmer, 1989, for work on partial symmetry). The drawings were centered on a white background 21.5 × 27.9 cm at a distance of 105 cm from the subjects. At this viewing distance, Figure 2A subtended a visual angle of 5.2° × 3.0°; Figure 2B subtended a visual angle of 3.0° × 2.5°. A white fixation point was located in the middle of the black region of each figure, as shown in Figure 2.

Reports about figure–ground organization were recorded on two keys that were connected to an Apple IIe microcomputer with a Rogers A6 Timer/Driver card. The computer was used to time both the trial duration and the duration of each individual key press. A chin rest was used to keep viewers' heads steady throughout the trials.

Procedure. As an introduction to the experiment, observers were shown a black-and-white version of the Maltese Cross, and the linkage between shape recognition and figure–ground organization was pointed out to them (i.e., they were told that when the black region was figure, they would be unable to see the shape in the white region; likewise, when the white region was figure, they would be unable to see the shape in the black region); they were instructed to use this criterion to decide which region to report as figure.

Observers' intentions were manipulated with the opposed-set procedure (Peterson & Hochberg, 1983) described above. Before each trial, observers were instructed to maintain fixation for the duration of the trial on a point approximately centered on the stimulus. They were also instructed to try to hold either the black region or the white region as figure. Observers were asked to follow the intention instructions solely by concentrating, and not by blinking or by moving their eyes from the fixation point, and to report accurately about which region appeared to be figure at any moment, regardless of how successful they were at following the hold instructions. They were also asked to report all reversals, even extremely rapid ones. Observers reported what they were seeing by pressing one response key whenever (and for as long as) the surround appeared to be figure and pressing the other response key whenever (and for as long as) the center appeared to be figure, alternating keys over the course of the 30-s trial.

Observers participated in three practice trials with the Maltese Cross, with "hold white" and "hold black" trials counterbalanced across subjects. Practice trials were used to acquaint observers with the task and to reiterate the importance of fixation and of accurate reporting.

Experiment 1 was divided into two halves consisting of four trials each, one trial with each hold instruction with each figure. During the first half, all observers viewed inverted versions of Figures 2A and 2B. Between the first and second halves of the experiment, observers were assigned in counterbalanced order to one of two groups. One group (control group) continued to view inverted stimuli in the second half and were not informed of the shape denoted by the surround in the upright version. The other group (experimental group) viewed upright stimuli in the second half and were informed of the shape denoted by the surround. (Because 97% of the observers in the preliminary experiment had identified the surrounds of Figures 2A and 2B as standing women and face profiles, respectively, we can be reasonably confident that the subjects in the experimental group readily perceived the interpretation for the surround that was pointed out to them and that they would have perceived the same interpretation even without the verbal description provided by the experimenter.) During the first half of this and the following two experiments, the experimenter was unaware of which group the observer would be in during the second half of the experiment.

Figure order was counterbalanced in an ABBA order. In each experiment, each figure was seen first by half of the observers. Opposed-hold instructions were given on consecutive trials for a given stimulus figure. Hold instructions were counterbalanced within and across subjects. The left and right mapping of response button to center and surround response was counterbalanced across subjects.

Data analysis and predictions. For each 30-s trial, we calculated (a) the mean duration that an observer managed to maintain as figure that region which he or she was trying to hold as figure (intended region, *I*), and (b) the mean duration that the observer maintained the other (unintended, *U*) region as figure. Comparisons of the mean duration that a given region was seen as figure when it was intended (*I*) and unintended (*U*) can be made across trials. For example, a comparison of the mean duration that a surround-as-figure percept was maintained when it was intended versus when it was unintended can be made across the *I* surround-as-figure responses obtained on a "hold surround" trial and the *U* surround-as-figure responses obtained on a "hold center" trial.

---

[4] In a second preliminary experiment, observers were shown two inverted versions of each stimulus, one with a black center and one with a white center, and asked to report which region first appeared to be figure. For both stimuli, a larger proportion of observers reported seeing the center as figure for the black–center version than for the white–center version. Consequently, we used black-center figures in all the experiments reported here so that we could measure how variations in the denotivity of the surround affected the perception of stimuli that were originally biased toward the center as figure.

If denotivity influences figure–ground reversals regardless of viewer's intentions, both *I* and *U* surround-as-figure durations should be longer when the surround is upright (second half) as opposed to inverted (first half). Because of the use of a practice control group, this effect would be revealed in an analysis of variance (ANOVA) by a Half × Region × Group interaction in both the *I* and *U* measures.

Of course, if intentions are effective, then *I* durations should be longer than their corresponding *U* durations. Consequently, we measured the effectiveness of viewers' intentions by assessing the magnitude of the difference between *I* and *U* durations, $(I - U)$. If $(I - U)$ is significantly greater than zero, that will be taken to indicate that intention can lengthen the duration of seeing the intended region as figure. We will examine whether viewers' intentions are more effective when the surround is upright and high in denotivity (second half, experimental group) as opposed to inverted and low in denotivity (as it was in the first half for the experimental group and in both halves for the control group). Because of the nature of the experimental design we have used, this denotivity-dependent intention effect would be revealed by a Half × Region × Group interaction in the $(I - U)$ responses.

Results

The results of Experiment 1, shown in Table 1, imply that shape recognition routines contribute to figure–ground reversals. Furthermore, the results show that experimental group observers' intentions to hold the surround as figure were more effective when the surround was upright, and hence more denotative, than when it was inverted, and hence less denotative. In what follows, we first discuss the effects obtained in the separate ANOVAs on the *I* and *U* measures. Then we discuss the effects obtained in the ANOVA on the intention effects, $(I - U)$.

Intended and unintended responses. As can be seen in Table 1A and 1B, observers in the experimental group maintained the surround as figure longer in the second half of the experiment when they viewed upright versions of the stimuli than they had in the first half of the experiment when they viewed inverted versions of the stimuli. This effect was obtained in both intended and unintended responses. Observers in the control group showed no such change: The durations of their surround-as-figure percepts were approximately equal in both the first and second half of the experiment. Because upright and inverted versions of the stimuli differed only in the denotivity of the surround, these results suggest that a region can be maintained as figure longer when it is more denotative.

Table 1
Average Durations of Intended and Unintended Reports and Intention Effects in Experiment 1 (in Seconds)

Experiment half	Control group		Experimental group	
	Center	Surround	Center	Surround
	A. Intended (*I*)			
1	11.53	5.80	18.82	7.15
SE	2.48	0.90	2.87	1.44
2	12.48	5.71	6.81	13.85
SE	2.78	1.38	1.68	2.35
	B. Unintended (*U*)			
1	3.68	1.30	9.84	2.58
SE	1.01	0.28	3.54	1.08
2	2.62	1.12	3.88	4.81
SE	0.50	0.22	1.71	1.15
	C. Intention effects $(I - U)$			
1	7.85	4.50	8.98	4.57
SE	1.78	0.89	2.87	1.71
2	9.85	4.59	2.93	9.04
SE	2.74	1.36	1.16	2.21

The bias toward center as figure was evident in both the *I* and *U* responses for all observers in the first half of the experiment and for the control group observers in the second half of the experiment: In these conditions, the center region was maintained as figure longer than the surround.[5] The bias toward center as figure was not evident in the second-half responses of the observers in the experimental group, in which the surround was maintained as figure longer than the center. Indeed, observers in the experimental group maintained the center as figure for substantially shorter durations in the second half of the experiment than in the first half of the experiment, both when they tried to hold the center as figure (*I* center-as-figure responses, Table 1A) and when they tried to hold the surround as figure (*U* center-as-figure responses, Table 1B). This result suggests that when the surround is more denotative, as it is in the upright orientation, it is more likely to be obtained as figure by reversal out of the center-as-figure interpretation, even when the center is the intended region.

These results were reflected in ANOVAs by main effects of region, $F(1, 14) = 10.06$, $p < .007$, for *I* responses, and $F(1, 14) = 9.26$, $p < .009$, for *U* responses; two-way interactions between half and region, $F(1, 14) = 13.08$, $p < .003$, for *I* responses, and $F(1, 14) = 5.93$, $p < .03$, for *U* responses; and three-way interactions among half, region, and group, $F(1, 14) = 16.32$, $p < .002$, for *I* responses, and $F(1, 14) = 3.84$, $p < .07$, for *U* responses. We take the existence of the three-way interaction in both the *I* and *U* responses as evidence that regardless of viewers' intentions, regions of figure–ground stimuli are more likely to be seen as figure when they are high in denotivity than when they are low in denotivity.

In addition, the results suggest that the bias toward center-as-figure responses was greater in Figure 2A than in Figure 2B, which probably reflects the operation of the Gestalt laws of symmetry and relative area. This effect was apparent in an interaction between figure type and region, $F(1, 14) = 5.34$, $p < .04$, and $F(1, 14) = 7.57$, $p < .02$, in the *I* and *U* measures, respectively. Results for the two figures are not shown separately because figure type did not interact with group or orientation, the variables of interest in this study.

[5] The mean duration of seeing the center as figure in the first half of the experiment was longer in the experimental group than in the control group. We are reasonably certain that the difference between the first- and second-half center-as-figure durations in the experimental group is not an artifact of this difference because the experimenter was unaware of the group to which a subject would be assigned until after the first half of the experiment had been completed. Furthermore, the center-as-figure effects were replicated in a within-subjects design in Experiment 4.

Intention effects. The mere existence of denotivity effects in the I responses does not necessarily imply that intentions to hold a given region as figure were more effective when that region was high in denotivity. Intention effects can be measured only by assessing the magnitude of the difference between I and U responses. Were the effects in I and U responses found to be comparable, we would have evidence that shape recognition processes contribute to figure–ground reversal processes, but we would have no evidence regarding the relationship between shape recognition analyses and mechanisms of intentional control, a second theoretical question addressed by these experiments.

As can be seen in Table 1C, the $(I - U)$ measure showed that experimental group observers' intentions to hold the surround as figure were more effective in the second half of the experiment, when the surround was upright, than they had been in the first half of the experiment, when the surround was inverted. There was no change in the effectiveness of the control group observers' intentions to hold the surround as figure from the first to the second half of the experiment.

Intentions to hold the center as figure were more effective than intentions to hold the surround as figure in all conditions except the experimental group, second half. There, intentions to hold the center as figure were less effective than they had been in the first half, and were less effective than intentions to hold the surround as figure.

These effects were reflected in two-way interactions between half and region, $F(1, 14) = 11.64$, $p < .005$; between region and group, $F(1, 14) = 4.92$, $p < .05$; and a three-way interaction among half, region, and group, $F(1, 14) = 24.37$, $p < .0002$. Observers' intentions were effective in all conditions (all ps $< .025$).

Discussion

Experimental group observers reported seeing the surround as figure for longer durations in the second half of Experiment 1, when they viewed upright stimuli and were aware of the meaningful shape it denoted, than they had in the first half of the experiment, when they viewed inverted stimuli and were unaware of the meaningful shape it denoted. These results were found both when observers tried to hold the surround as figure (i.e., I surround-as-figure reports) and when they tried to hold the center as figure (i.e., U surround-as-figure reports). This effect suggests that regardless of viewers' intentions, regions of reversible figure–ground stimuli are maintained as figure longer when they are highly denotative.

Note also that the center was seen as figure for shorter durations when the stimuli were upright than when they were inverted, both when observers intended to hold the center as figure (I center reports) and when they intended to hold the surround as figure (U center reports). This finding implies that regions of reversible figure–ground stimuli denoting upright meaningful shapes are likely to be obtained as figure by reversal out of the alternative interpretation even when observers are intentionally trying to perceive the alternative interpretation.

The concomitant decrease in the duration of center-as-figure percepts and increase in the duration of surround-as-figure percepts found in the second-half responses of the experimental group may indicate that the relative denotivity of the two regions competing for a figure–ground contour is assessed prior to figure–ground reversal (cf. Rock, 1975, p. 288).

The $(I - U)$ measure showed that observers' intentions to hold both the surround and the center as figure also depended on the orientation of the surround: Intentions to hold the surround as figure were more effective when the stimulus was upright rather than inverted, indicating that intentions can operate more effectively when the intended region denotes an upright recognizable shape. On the other hand, intentions to hold the center were less effective when the stimulus was upright than when it was inverted. Thus the effectiveness of observers' intentions varied with the denotivity of the unintended region as well as with the denotivity of the intended region.

Can these effects obtained in Experiment 1 be localized in one of the shape recognition routes introduced earlier? Both the maintain and the obtain effects obtained in Experiment 1 depended on the presence of an upright, as opposed to an inverted, shape. This orientation dependence might imply that the denotivity effects were mediated by access to an orientation-specific structural memory representation that was accomplished prior to figure–ground reversal for upright stimuli, but not for inverted stimuli. At this point, however, the effects cannot be pinpointed as contributions from the structural memory representation system, because observers viewing upright stimuli were informed of the identity of the shape denoted by the surround, whereas observers viewing inverted stimuli were not. Consequently, we cannot be sure whether to attribute the experimental group's change in performance to the change in stimulus orientation, which would implicate the structural representation route; to a knowledge-dependent search for distinctive features or parts; or to knowledge-dependent semantic mediation.

Another possibility is that the special knowledge about the surround imparted to the experimental group subjects between the two halves of the experiment carried with it an implicit demand to increase efforts to hold the surround as figure. In response to such an implicit demand, experimental group observers may have made different eye movements on "hold surround" trials during the second half of the experiment, or they may have engaged in covert behaviors that enhanced the effectiveness of their intentions to hold the surround as figure; for example, increasing their concentration or allocating their spatial attention to the surround more often. In at least some theories, fixation location or spatial attention location can contribute to figure–ground organization (Hochberg, 1971; Sejnowski & Hinton, 1987). In Experiments 2–4, we eliminate these and other alternative explanations of the results of Experiment 1 and identify the effects in Experiment 1 as reflecting contributions from the structural representation route.

Experiment 2

Many current theories of shape recognition (e.g., Biederman, 1987; Hoffman & Richards, 1985; Marr & Nishihara,

1978) propose that shape recognition processes entail an initial step in which the contour of the input shape is decomposed into parts at the minima of curvature. The parts are then matched to components of structural memory representations. The stimuli used in Experiment 2 were scrambled versions of the stimuli used in Experiment 1, created by partitioning the contours at minima of curvature defined from inside the surround. Consequently, if we take the parts of the shape to be those defined by contour minima of curvature, then the surrounds of both the original stimuli and the scrambled versions of the stimuli should access the same representational components, and therefore the same orientation-independent part representations. However, the parts of the scrambled stimuli were rearranged so that no two parts that were connected in the original stimuli remained connected in the scrambled versions, as can be seen in Figures 2C and 2D. Consequently, the orientation-specific structural memory representations accessed by the scrambled stimuli should be different from those accessed by the original stimuli.

The procedure of Experiment 2 was very similar to that of Experiment 1: All subjects viewed scrambled stimuli in the first half of the experiment. Between the first and second halves of the experiment, the experimenter showed the observers in the experimental group the upright *un*scrambled figures and pointed out the correspondence between the parts of the scrambled shape and the parts of the upright unscrambled shape, naming each part as she went along. In Experiment 2, unlike Experiment 1, observers in both the experimental and control groups continued viewing scrambled stimuli during the second half of the experiment.

If figure-ground computations are influenced by shape recognition procedures entailing local part or feature identification only, without regard for the relative locations of the parts, then Experiment 2 should replicate Experiment 1. Moreover, if the results of Experiment 1 were due to strategy changes induced by changes in the demand character of the experiment, then they should be replicated in Experiment 2, because such strategy changes should not be constrained by stimulus structure. Alternatively, if the obtain and maintain effects found in Experiment 1 reflected access to structural memory representations, they should not be replicated with scrambled stimuli.

Method

Subjects. The subjects were 16 students at the University of Arizona who took part in this experiment to fulfill a research participation requirement for an introductory psychology course. All had normal or corrected-to-normal vision.

Stimuli and apparatus. Figures 2C and 2D were our stimuli. Figure 2C subtended a visual angle of 5.2° × 2.9°, and Figure 2D subtended a visual angle of 2.8° × 2.3°. The area of the center regions was .65 and .79 that of the surrounds for Figures 2C and 2D, respectively. Responses were recorded on the two keys of a mouse attachment for a Compaq 386 microcomputer. In other respects, the stimuli and apparatus were the same as in Experiment 1.

Procedure. The procedure used in Experiment 2 was similar to that used in Experiment 1, with the following exceptions. The Rubin vase-faces stimulus was used during the instruction and practice phases of the experiment rather than the Maltese Cross. In this experiment, hold instructions referred to the "center" and the "surround" regions. All observers viewed scrambled versions of Figures 2A and 2B (i.e., Figures 2C and 2D) throughout the experiment. Between the two halves of the experiment, the experimenter showed the upright unscrambled stimuli (Figures 2A and 2B) to the observers in the experimental group, identified the shape depicted by the surround, and pointed out the correspondence between the parts of the scrambled stimulus and the parts of the upright unscrambled stimulus, naming each part of the denotative regions as she did so (e.g., for Figure 2B: "forehead," "nose," "chin," etc.). In Experiment 2, unlike Experiment 1, the stimuli were viewed in the same orientation in the first and second halves of the experiment.

Results

Intended and unintended responses. As can be seen in Table 2A and 2B, the pattern of results obtained in Experiment 2 was different from the pattern obtained in Experiment 1. Neither the intended nor the unintended surround-as-figure percepts reported by subjects in the experimental group were longer in the second half of the experiment than they had been in the first half. Thus, regions denoting scrambled versions of highly denotative shapes were not maintained as figure longer after they were identified (i.e., in the second half of the experiment) than they were before they were identified (i.e., in the first half of the experiment).

There was an overall bias toward seeing the center as figure. However, observers in both groups reported seeing the center as figure for shorter durations in the second half of the experiment than in the first half. Because this effect was found in both groups, it cannot reflect either knowledge-dependent behaviors or contributions from shape recognition routines; it might simply be a consequence of practice with the stimuli. Durations of surround-as-figure reports also tended to be shorter in the second half of the experiment.

The ANOVA revealed a main effect of region in both the I and U reports, reflecting the stimulus bias toward center as

Table 2
Average Durations of Intended and Unintended Reports and Intention Effects in Experiment 2 (in Seconds)

Experiment half	Control group		Experimental group	
	Center	Surround	Center	Surround
	A. Intended (I)			
1	10.77	5.63	8.85	6.27
SE	2.39	1.67	1.00	0.92
2	8.81	4.25	5.90	5.54
SE	2.92	0.99	0.83	0.88
	B. Unintended (U)			
1	4.57	2.36	3.17	0.96
SE	1.48	1.25	1.02	0.24
2	4.25	1.22	1.19	1.05
SE	2.35	0.25	0.27	0.33
	C. Intention effects ($I - U$)			
1	6.20	3.27	5.68	5.32
SE	1.74	0.70	1.21	1.06
2	4.55	3.03	4.71	4.48
SE	1.43	1.05	0.89	0.82

figure: $F(1, 14) = 12.26$, $p < .004$, and $F(1, 14) = 8.95$, $p < .01$, for I and U respectively. In addition, there was a main effect of half, $F(1, 14) = 13.09$, $p < .003$, and $F(1, 14) = 6.24$, $p < .03$, for I and U reports, respectively. No interactions involving half were significant, however (all $ps > .34$); in particular, the three-way interaction among half, region, and group failed to reach significance in either the I responses ($p > .64$) or the U responses ($p > .34$). The only other significant effect obtained in either the intended or the unintended responses was an interaction between figure type and region in the intended responses, reflecting a greater bias toward the center-as-figure percept in Figure 2D than in Figure 2C, perhaps because the center of Figure 2D overlapped the surround, $F(1, 14) = 7.09$, $p < .02$.

Intention effects: $(I - U)$. Experimental group observers' intentions to hold the surround as figure were not more effective in the second half of the experiment than they had been in the first half, suggesting that the intention effects obtained in Experiment 1 are not attributable to knowledge-dependent behaviors unconstrained by structure. Overall, viewers' intentions were less effective in the second half of the experiment than they had been in the first half, as reflected in a main effect of half, $F(1, 14) = 4.62$, $p < .05$. The responses of the two groups did not differ: the $(I - U)$ measure showed no main effect of group ($p > .57$) and no interactions involving group. In particular, the three-way interaction among half, region, and group failed to reach significance ($p > .44$).

The only other effect found to be significant in the ANOVA was a two-way interaction between figure type and region, $F(1, 14) = 5.05$, $p < .05$, reflecting the fact the viewers' intentions to hold the center of Figure 2D and the surround of Figure 2C were more effective than their intentions to hold the alternatives. t tests revealed that observers' intentions were effective in all conditions (all $ps < .025$).

Comparison of Experiments 1 and 2. ANOVAs comparing the results of Experiments 1 and 2 supported the hypothesis that different patterns of results were obtained for the center and surround regions across group and half in the two experiments. We looked for and obtained a four-way interaction among half, region, group, and experiment, $F(1, 28) = 9.24$, $p < .006$, and $F(1, 28) = 21.00$, $p < .001$, for the I and $(I - U)$ responses, respectively, which showed that the differences observed by comparing the tabled values were significant. We did not find a significant four-way interaction in the ANOVA on the U responses ($p > .36$), perhaps because of floor effects.

Discussion

The finding that the obtain and maintain effects of Experiment 1 were not found in Experiment 2 rules out the possibility that those results can be ascribed to contributions from a shape recognition route entailing access to orientation-independent representations that do not preserve the relative locations of shape parts. Other types of orientation-independent feature or part extraction routes are possible. For example, Corballis (1988) discusses an orientation-independent part extraction route in which connectivity relationships are specified.

Experiment 2 can also serve to rule out one class of strategic behavioral changes observers might have made in response to the demand character of the experimental situation—the class of strategic changes that should be unaffected by the relative locations of the parts, such as changes in motivation, or in fixation or attention location.

It remains possible that the effects shown by the experimental group observers in Experiment 1 were due to knowledge-dependent strategies that operate with unscrambled versions of the stimuli, but not with scrambled versions. One such strategy depends on the fact that when the figure–ground organization reverses from surround as figure into center as figure, the shape seen in the surround disappears. Suppose that once having seen an identifiable shape in the surround, experimental group observers in Experiment 1 were tempted to try to recover it on those occasions when it disappeared. Such "recovery attempts" (as we will refer to them) could lead both to shorter intended center-as-figure percepts because they would undermine intentions to hold the center as figure and to shorter unintended center-as-figure percepts because they could supplement other intentional strategies used in service of the "hold surround" instructions. In Experiment 3, we show that this "recovery attempts" explanation cannot account for the effects obtained in Experiment 1. In addition, we begin to isolate the single shape recognition route responsible for the Experiment 1 effects.

Experiment 3

In Experiment 3, observers in both the experimental and control groups viewed inverted versions of Figures 2A and 2B in both halves of the experiment. As before, the experimenter showed upright versions of the stimuli to the observers in the experimental group between the first and second halves of the experiment, and made certain that observers could identify the highly denotative shape depicted by the surround in both the upright and inverted versions.

The parts of inverted shapes are presented in their proper relative locations with respect to an orientation-independent representation; parts that should be connected are connected. Therefore, the figure–ground effects obtained in Experiment 1 should be replicated here if they can be mediated by outputs from an orientation-independent recognition route in which connectivity relations are specified (perhaps supplemented by semantic processing). Moreover, observers should be just as tempted to recover inverted denotative shapes as upright denotative shapes, provided they are aware of the inverted denotative shape depicted by the inverted surround. Therefore, if the effects of Experiment 1 were due to recovery attempts, they should be replicated in Experiment 3.

If the results obtained in Experiment 1 are not replicated in Experiment 3, that could imply that they reflect contributions from structural memory representations and that the normalization process preceding the match between the inverted surround and the orientation-specific memory representation was not completed prior to figure–ground reversal.

Method

Subjects. The subjects were 16 students at the University of Arizona who participated in this experiment to fulfill a research requirement for an introductory psychology course. All had normal or corrected-to-normal vision.

Stimuli, apparatus, and procedure. The stimuli were inverted versions of Figures 2A and 2B. The procedure was similar to that used in Experiment 2 with the following exceptions. All observers viewed inverted shapes throughout the experiment. Between the first and second halves of the experiment, observers were assigned in a counterbalanced order to either a practice control group or an experimental group. Observers in the experimental group were shown the upright figures between the two halves of the experiment. The figures were identified, and the correspondence between the upright and inverted figures was demonstrated. Observers in the control group received no new information between the two halves of the experiment. Experiment 3 differed from Experiment 1 in that observers in both the control and experimental groups continued to view inverted stimuli during the second half of the experiment.

Results

Intended and unintended responses. The results of Experiment 3 are quite different from the results of Experiment 1, as can be seen in Table 3. Consider first the *I* and *U* responses. Observers in the experimental group did not report significantly longer surround-as-figure percepts in the second half of the experiment. Therefore, observers could not maintain the inverted surrounds as figure any longer after they had seen the upright versions than they could before they had seen the upright versions. Observers in both groups reported shorter (intended and unintended) center-as-figure percepts in the second half of the experiment. Because this effect was found in both groups, it cannot reflect knowledge-dependent behaviors; it is probably a consequence of practice.

Table 3
Average Durations of Intended and Unintended Reports and Intention Effects in Experiment 3 (in Seconds)

Experiment half	Control group		Experimental group	
	Center	Surround	Center	Surround
	A. Intended (*I*)			
1	7.91	4.69	8.12	5.72
SE	2.21	0.81	1.58	0.64
2	4.87	3.98	4.89	6.75
SE	0.92	0.42	0.54	1.40
	B. Unintended (*U*)			
1	4.20	1.83	3.01	2.12
SE	1.39	0.50	1.03	1.03
2	2.49	2.08	1.39	2.05
SE	0.77	0.55	0.43	0.74
	C. Intention effects ($I - U$)			
1	3.71	2.86	5.11	3.60
SE	0.90	0.42	1.06	0.61
2	2.38	1.90	3.50	4.70
SE	0.97	0.51	0.36	1.15

The ANOVAs performed on the intended and unintended responses showed no effect of group ($p > .43$ and $p > .61$) and no interactions involving group. In particular, there was no three-way interaction among half, region, and group in either the intended responses ($p > .51$) or the unintended responses ($p > .81$). Both intended responses and unintended responses revealed a two-way interaction between half and region, $F(1, 14) = 5.11$, $p < .04$, and $F(1, 14) = 4.64$, $p < .05$, reflecting the finding that for both groups the center-as-figure percepts were shorter in the second half of the experiment. The intended responses also showed a main effect of figure type, $F(1, 14) = 6.73$, $p < .02$, showing that the mean durations of both center-as-figure and surround-as-figure reports were longer for Figure 2B ($M = 6.18$) than for Figure 2A ($M = 5.55$).

Intention effects. The results obtained with the ($I - U$) measure were similar to those obtained with the *I* measure. Experimental group observers' intentions to hold the surround as figure were not significantly more effective in the second half of the experiment than they had been in the first half. Their intentions to hold the center as figure were less effective, as were the intentions of observers in the control group. The fact that this decrease in the effectiveness of intentions to hold the center as figure was found in both groups indicates that it was not due to knowledge regarding the shape depicted by the upright surround.

These effects were reflected in the ANOVA by a marginal interaction between half and region, $F(1, 14) = 3.89$, $p = .069$. No effects of group or interactions involving group were obtained. In particular, the three-way interaction among half, region, and group was not significant ($p > .16$). The ANOVA did reveal a main effect of figure type, $F(1, 14) = 6.96$, $p < .02$; intentions were more effective when observers viewed Figure 2B than Figure 2A.

Comparison of Experiments 1 and 3. ANOVAs comparing the results of Experiments 1 and 3 supported the hypothesis that different patterns of results were obtained for the center and surround regions across group and half in the two experiments. The ANOVAs showed significant interactions among half, region, group, and experiment, $F(1, 28) = 9.79$, $p < .005$, and $F(1, 28) = 11.60$, $p < .002$, for the *I* and ($I - U$) responses, respectively, and a marginal four-way interaction in the *U* responses, $F(1, 28) = 3.59$, $p < .07$. The failure to find a significant four-way interaction in the *U* responses may again be due to floor effects.

Discussion

The observers in the experimental group of Experiment 3 received the same treatment as the observers in the experimental group of Experiment 1, with the exception that observers in Experiment 3 continued to view inverted versions of the stimuli (rather than upright versions) in the second half of the experiment. Therefore, the failure of Experiment 3 to replicate the critical obtain and maintain effects of Experiment 1 demonstrates that those effects were not mediated by contributions from an orientation-independent shape recognition route.

The finding that there was no increase in experimental group observers' ability to maintain the surrounds as figure, as indicated by a larger increase in their surround-as-figure percepts than in those of control group observers, indicates that neither semantic nor nonstructural representations are sufficient to lengthen the duration that the surround is perceived as figure. This finding is particularly compelling because supplementing structural representations with semantic representations has been shown to improve other aspects of encoding (e.g., Schacter, Cooper, & Delaney, 1990; Schacter, Cooper, Delaney, Peterson, & Tharan, 1991). In addition, and more important, it indicates either that structural representations cannot be accessed once a reversal has occurred or that access to structural representations following a figure–ground reversal are irrelevant to those processes that mediate figure–ground reversals. Therefore, the finding that neither the obtain nor the maintain effects of Experiment 1 were found in Experiment 3 may indicate that there is a critical time period, terminated by a reversal, during which shape recognition processes must be completed if they are to influence figure–ground computations. The normalization process that intervenes before inverted shapes are matched to the correct orientation-specific representations may require more than the critical amount of time.

In addition, the results of Experiment 3 join with those of Experiment 2 in ruling out demand character interpretations for Experiment 1, such as the recovery attempts explanation; the spatial attention and/or eye movement explanation; and the motivation explanation. All of these strategies should operate just as well with inverted versions of the stimuli as with upright versions of the stimuli.

Experiment 4

In Experiment 4, we sought to replicate the obtain and maintain effects obtained in Experiment 1 and to examine observers' performance with stimuli at five different orientations between upright and inverted. In Experiment 4, all observers were aware of the shape denoted by the upright surround.

If a time-consuming normalization process precedes the match between disoriented shapes and the proper orientation-specific representation, and if the critical maintain and obtain results found in Experiment 1 depend on a match to the orientation-specific representation, then we should expect to find those effects only when observers view stimuli that are upright or are disoriented from upright by amounts small enough so that the normalization process can be completed within the critical period before a reversal occurs. For stimuli disoriented by greater amounts, the evidence from naming experiments suggests that the normalization process should take increasing amounts of time, and hence should be increasingly unlikely to be completed within the critical period for figure–ground reversal computations.

Method

Subjects. The subjects were 10 students at the University of Arizona who took part in the experiment to fulfill a research participation requirement for an introductory psychology course. All subjects had normal or corrected-to-normal vision.

Stimuli and apparatus. The stimuli were the black and white drawings shown in Figures 2A and 2B, shown at five different clockwise orientations with respect to the upright: 0°, 45°, 90°, 135°, and 180°. The stimuli were centered on a white background and were presented at the same distance as the stimuli in Experiments 1–3. The apparatus used was the same as that used in Experiments 2 and 3.

Procedure. The procedure was the same as that used in the previous experiments, except that before the experimental trials, all observers were shown the upright versions of the two stimulus figures, and the shapes denoted by the upright surrounds were pointed out to them. Each observer participated in 20 experimental trials, one trial with each stimulus at each orientation with each hold instruction. Half the observers viewed Figure 2A in the first block; the other half viewed Figure 2B in the first block. Observers participated in two consecutive trials with each stimulus at each orientation, one trial with each hold instruction. Hold instructions were counterbalanced both within and across subjects. Orientation order was counterbalanced between subjects in a Latin square design. The left–right mapping of the response buttons to "center" and "surround" responses was counterbalanced across subjects.[6]

Results

As shown in Figures 3A–C, the I, U, and $(I - U)$ effects found with upright figures in Experiment 1 were replicated with upright stimuli in Experiment 4, and were increasingly attenuated as the stimuli were increasingly disoriented from upright.

Intended responses. As can be seen in Figure 3A, the results obtained with 0° and 180° stimuli replicated those obtained in the upright and inverted conditions of Experiment 1: The durations of intended surround-as-figure percepts were longer for upright (0°) stimuli than for inverted (180°) stimuli ($p < .025$); and the durations of intended center-as-figure percepts were longer for inverted stimuli than for upright stimuli ($p < .05$). For the most part, intermediate durations of both intended surround-as-figure percepts and intended center-as-figure percepts were found at intermediate orientations. In both the center and the surround reports, the 180° data depart from the linearity expected by extrapolating from the results from the other orientations. Consequently, the comparisons between the 0° and 135° orientations are even stronger than those between the 0° and the 180° orientations. Similar departures from linearity have been found in naming experiments (cf. Jolicoeur, 1985).

The crossover between longer mean durations of intended surround-as-figure percepts to longer mean durations of intended center-as-figure percepts occurred between 45° and 90°. Thus, the high denotative regions of stimuli disoriented

[6] No subjects were excluded from this experiment. In Experiments 1–3, we had excluded subjects who had recognized the inverted denotative shape. Five subjects in the control groups and 1 in the experimental group had been excluded on the basis of this criterion. This was not an issue in Experiment 4 because all subjects were informed of the meaningful shape denoted by the surround. Moreover, the results of Experiment 3 suggest that exclusion was not necessary in the previous experiments.

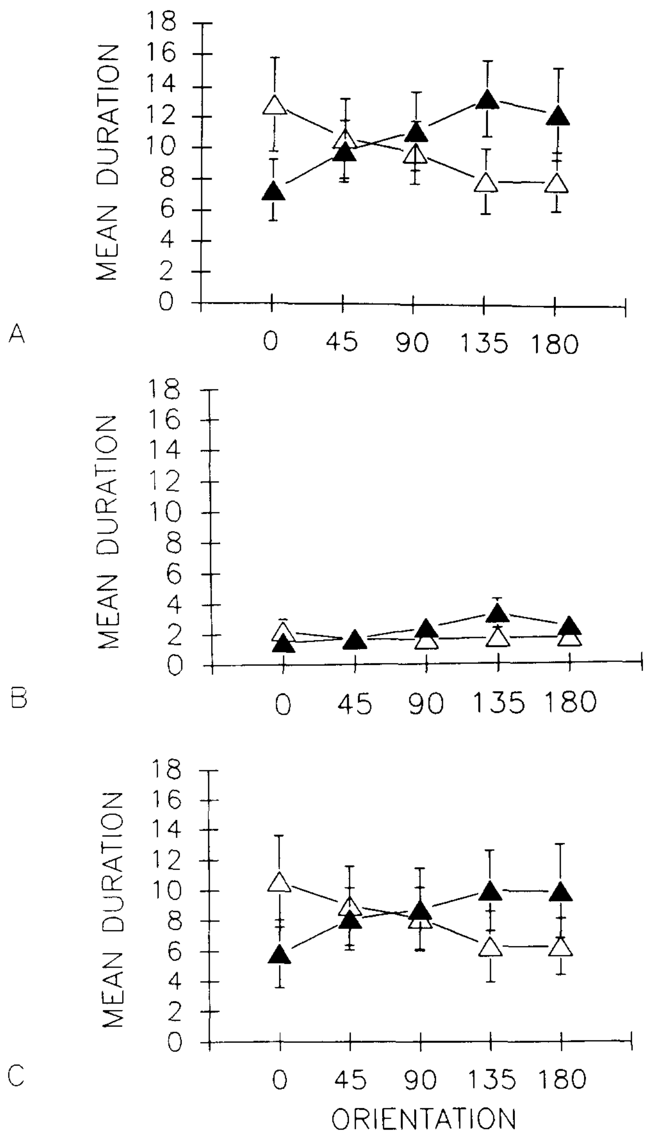

Figure 3. Panel A is the mean duration of an intended percept (*I*) plotted as a function of stimulus orientation for both the center-as-figure and the surround-as-figure percepts. (Closed triangles represent the center reports; open triangles represent the surround reports.) Panel B shows the same for unintended percepts (*U*). Panel C depicts intention effects (*I* – *U*) for "hold surround" and "hold center" trials plotted as a function of stimulus orientation.

by approximately 90° from upright were neither easier to maintain as figure nor easier to obtain as figure than the low denotative regions.

An ANOVA reflected the general orientation trend in an interaction between orientation and region, $F(1, 14) = 5.81$, $p = .001$. The ANOVA also showed an interaction between figure type and region, $F(1, 14) = 12.61$, $p < .007$, reflecting the fact that Figure 2A is more strongly biased toward the center-as-figure interpretation than Figure 2B. In addition, a three-way interaction among figure type, orientation, and region, $F(1, 14) = 3.22$, $p < .03$, demonstrated that for Figure 2B the crossover from longer surround-as-figure percepts to longer center-as-figure percepts occurred between 90° and 135°, whereas for Figure 2A the crossover occurred between 0° and 45°.

Unintended responses. Similar patterns were found in the unintended responses, as can be seen in Figure 3B. Again, the center was seen as figure for longer durations when the stimuli were inverted rather than upright ($p < .025$). In addition, the surround was seen as figure for longer durations when the stimuli were upright rather than inverted, although the statistical comparison was not significant ($p < .10$). The surround was seen as figure significantly longer in the upright orientation than in the 135° orientation, however ($p < .05$).

The ANOVA showed these effects in a two-way interaction between orientation and region, $F(1, 14) = 6.35$, $p < .001$. In addition, a three-way interaction among figure type, orientation, and region obtained in the *U* responses, $F(1, 14) = 3.65$, $p < .02$, demonstrated that the crossover points for the two figures differed (they were the same as in the *I* responses).

Intention effects. Observers' intentions influenced perceived organization in all conditions (all *p*s < .05). However, intentions to hold the surround as figure decreased in effectiveness and intentions to hold the center as figure increased in effectiveness as the stimuli were increasingly disoriented from upright. Intentions to hold the surround as figure were more effective at 0° than at 180° ($p < .025$), whereas intentions to hold the center as figure were more effective at 180° than at 0° ($p < .05$). These conditions serve as replications of the effects of Experiment 1. In addition, intentions to hold the surround as figure were significantly more effective than intentions to hold the center as figure only for upright stimuli.

A two-way interaction between orientation and region revealed these effects to be significant, $F(1, 14) = 4.74$, $p < .004$. In addition, an ANOVA showed that intentions to hold the surround as figure were more effective on Figure 2B, and intentions to hold the center as figure were more effective on Figure 2A, as reflected in an interaction between figure type and region, $F(1, 14) = 13.18$, $p < .006$.

Discussion

The comparison between the 0° and 180° conditions in Experiment 4 revealed maintain and obtain effects similar to those found in Experiment 1. At stimulus orientations between 0° and 135°, surround-as-figure percepts gradually transformed from relatively long durations to relatively short durations, as compared to center-as-figure percepts. These are exactly the effects one would expect were the obtain and maintain effects attributable to outputs from an orientation-specific representation with coordinates that match those of the upright stimuli: As the stimuli were increasingly disoriented from their canonical upright, it may have become increasingly unlikely that any process compensating for the mismatch between viewer-centered and canonical coordinates could be completed within the critical period for figure–ground reversals. Thus, Experiment 4 provides strong support for the hypothesis that input from orientation-specific structural representations can influence figure–ground reversal computations.

In addition, Experiment 4 joins with Experiments 2 and 3 in ruling out explanations of our effects in terms of eye movements, attention movements, recovery attempts, and inputs from nonstructural representations. None of these explanations would predict a gradual tapering off of the effects as the stimuli were increasingly disoriented from upright.

General Discussion

The four experiments reported in this article demonstrated that figure–ground reversal computations weigh inputs reflecting the goodness of the fit between stimulus regions and orientation-specific structural memory representations. Orientation-independent shape representations were unable to influence figure–ground reversal, even when those representations may have been supplemented by semantic representations. Thus, our experiments are the first to establish that figure–ground reversal is open to input from one specific shape recognition route—the orientation-specific shape recognition route—and to rule out inputs both from another shape recognition route, the orientation-independent distinctive features route, and from other less well-specified means through which the meaningfulness (or denotivity) of a region might influence figure–ground organization (e.g., the recovery attempts explanation).

Previous investigators have been concerned with whether a region's *connotivity*, rather than its *denotivity*, can influence it's likelihood of being seen as figure. For example, Murphy (1947) proposed that the connotations of shapes, as manipulated by their associations with reward or punishment, can serve as input to figure–ground organization, and Schafer and Murphy (1943) provided an empirical demonstration. But the results of Schafer and Murphy's experiment and other experiments examining whether "needs" or "values" could influence perception (e.g., Bruner & Goodman, 1947) proved to be equivocal and unreliable (e.g., Hochberg, 1970; Osgood, 1953; Pastore, 1949; Rock & Fleck, 1950; Smith & Hochberg, 1954; Wallach, 1949). Consequently, these experiments did not dislodge the prevailing view that shape recognition processes are initiated only for those regions already determined to be figure rather than ground on the basis of stimulus analysis (Kanizsa & Luccio, 1987; Kohler, 1929; Wallach, 1949).

Other investigators examined Rubin's claim that prior experience per se influences figure–ground organization (e.g., Dutton & Traill, 1933; Rock & Kremen, 1957). The figure–ground displays typically used in these experiments consisted of two equal-area regions, both of which were nonsense shapes (i.e., were low in denotivity). Observers first participated in a cover task in which they viewed one of the two regions of the figure–ground displays for a short period of time (e.g., 2 s in Rock and Kremen's experiment), for eight repetitions, then participated in an irrelevant intervening task, and finally reported about the organization perceived in brief presentations (e.g., 1 s) of the full figure–ground displays. The perceived organization of the figure–ground displays was not affected by this amount of prior experience. Subsequent experiments showed that although 4 and 8 prior exposures to one alternative of these figure–ground displays were not sufficient to influence figure–ground organization (Cornwell, 1964; Rock, 1975, p. 384), 16 and 24 prior exposures may have been sufficient (Cornwell, 1964). When considered in the light of our findings, these findings raise questions about what is required to establish a structural memory representation.

In other experiments, using a different procedure and different stimuli (e.g., the Schafer-Murphy stimulus in which the regions on either side of the figure–ground contour denoted faces rather than nonsense shapes), Epstein and Rock (1960) demonstrated that recent prior experience can influence figure–ground organization. This finding caused them to confront the following paradox, first articulated by Hoffding (1891): Recency (i.e., prior experience) effects are assumed to be mediated by similarity matches between a perceptual description of a stimulus shape and a memory representation. However, because the region to only one side of a figure–ground contour can have shape (i.e., the figural region; the other region is the shapeless ground), then the critical perceptual description cannot exist until after figure–ground relationships are resolved. The crux of the paradox is the question of how prior experience can possibly influence figure–ground organization, given that no shape description is available until after figure–ground organization is determined. (Neisser, 1967, Rock, 1975, and Wallach, 1949, all contain cogent discussions of this issue.)

Epstein and De Shazo (1960) reconciled recency effects with the Hoffding paradox by proposing that before one figure–ground organization emerges consciously, the regions to both sides of the contour are alternately assigned the contour (and therefore are described perceptually) in rapid succession below the threshold of consciousness. If one organization is preferred for any reason, it is stabilized and perceived consciously. Reasons for preference include a match to a memory representation and/or a computation that one region is smaller in area than the other, and so on. This hypothesis, called the "perceptual oscillation hypothesis" (see also Rock, 1975, p. 289) retained the prevailing view that shape recognition processes can be conducted only following figure–ground organization by positing both conscious and unconscious figure–ground organizations. Although the perceptual oscillation hypothesis has not survived in contemporary shape recognition theories, the notion that the determination of figure–ground relationships occurs before the initiation of shape recognition processes (i.e., before the operation of the contour partitioning processes that delimit the parts of recognition from inside the figural region) has endured (e.g., Biederman, 1987; Hoffman & Richards, 1985; Marr, 1982; but see Lowe, 1985; Witkin & Tenenbaum, 1983, for a solution that avoids contour partitioning).

Can the Hoffding dilemma be resolved within a contour partitioning approach without the prior determination of figure–ground relationships? We argue that it can, and moreover, that current theoretical approaches to shape recognition can be altered to accommodate our findings, by eliminating the requirement that the contour partitioning process operates from inside the figural region only. We propose that figure–ground contours may be partitioned from both sides simultaneously before figure–ground relationships reverse. This proposal is at least logically possible, given that luminance edges can be detected preattentively (e.g., see Marr & Hildreth, 1980). If figure-ground contours can be partitioned from both sides prior to reversal, then the outcomes of the resultant shape recognition processes could serve as inputs to figure-ground reversal computations.

In our view, figure-ground reversal computations are interactive, weighing inputs from the structural representation system as well as from routines that analyze the stimulus variables. Inputs from other systems may be weighed as well (cf. Harvey, Peterson, & Gibson, 1990; Hochberg, 1971; Sejnowski & Hinton, 1987). Our findings suggest that these different inputs are not all given equal weight in these computations; in particular, the weight given to shape recognition inputs may be disproportionately large.

Are Inputs From Structural Representations Sufficient?

Inputs to figure-ground reversal computations from structural representations may have been supplemented by input from semantic representations accessed by the structural representations. It has been demonstrated that pictures rapidly access semantic representations specifying their category membership and/or their functions, at least when the experimental task requires classification (Carr, McCauley, Sperber, & Parmlee, 1982; Durso & Johnson, 1979; Riddoch & Humphreys, 1987). In our task, then, rapidly available semantic information may have bolstered the structural representation input. Thus, although our experiments show that structural representation input is necessary for both the obtain and the maintain effects, they do not show that it is sufficient. With respect to that point, however, we note that other experiments that use tasks in which responses can be based on structure alone have shown that access to semantic representations may not be automatic (Kelter et al., 1984). Figure-ground reversal may be more like the tasks used by Kelter et al. than those used by Carr et al. in that input regarding the shapes denoted by a region may be relevant to figure-ground reversal computations, but input regarding the connotations of those shapes may be irrelevant.

Reversal And/Or Initial Organization?

The effects reported here were obtained in a situation in which observers reported about alternations in figure-ground organization during 30-s trials. Consequently, these experiments cannot address the question of whether the initial organization of figure-ground stimuli reflects contributions from structural representations. We believe, however, that reversal phenomena provide a good model for initial access (Peterson & Gibson, 1991a; Rock, 1975, p. 289). In particular, the mechanism we propose as a precursor for the shape recognition contributions—partitioning a luminance contour simultaneously from both sides—is one that can operate on initial exposure to a stimulus as well as with continued exposure. In other research using brief masked exposures of a small set of figure-ground stimuli, Peterson and Gibson (1991b) found evidence supporting the idea that shape recognition inputs contribute to initial figure-ground formation as well. However, further experiments using the brief exposure paradigm with a larger set of stimuli are necessary before we can be certain whether these effects are spontaneous or require knowledge of the denotative alternative (cf. Girgus, Rock, & Egatz, 1977; Rock & Mitchener, in press).

Mechanisms of Intention

In the experiments reported here, we repeatedly found that observers' intentions to hold the surround as figure were more effective in those situations in which orientation-specific structural memory representations could be accessed (or could reach some threshold of activation) within some critical time period. Conversely, observers' intentions to hold the center as figure were more effective in those situations in which input from structural representations favoring the surround was not, we believe, available within that critical time period. These findings are consistent with two mechanisms of intentional mediation that cannot be distinguished at this point.

First, intention might be operating by providing top-down activation (or "priming") to a structural representation (cf. Carpenter & Grossberg, 1987). The effects of intentional priming should not be evident until bottom-up activation is present (Carpenter & Grossberg, 1987). Accordingly, intentions to hold the surround as figure would be expected to be greater for upright stimuli than for inverted stimuli, because the match between the inverted surrounds and the orientation-specific structural memory representation may not be completed within the critical period for figure-ground reversal.

Note that if the intention effects are mediated by a priming mechanism such as this, then intention would operate before the interactions among the variables relevant to figure-ground computations. A priming mechanism can explain observers' increased success at holding the upright rather than the inverted surround as figure, but it does less well at explaining observers' considerable success at holding the inverted rather than the upright center as figure.

A second potential mechanism places the effects of intention after the various inputs to figure-ground computations have interacted to produce some value indicating each region's eligibility to be figure, but before the figure-ground contour is assigned to one region. Intention might then operate by selecting one of these two output values. By assuming that stronger signals are easier to select, this mechanism can account for observers' success at holding both the inverted center and the upright surround as figure, because the composite signal favoring the center as figure would be greater for inverted stimuli and the composite signal favoring the surround as figure would be greater for upright stimuli. The problem for this account is specifying how intention selects one output signal or another. One possibility is that intention operates by selectively attending to the spatial location corresponding to the intended alternative in some visual buffer, as Kosslyn, Flynn, Amsterdam, and Wang (1990) have suggested.

Of course, intention might operate independently through either or both of these mechanisms. Harvey et al. (1990) found that both denotivity and fixation location (the latter was not separated from spatial attention location) served as independent inputs to figure-ground reversal. Hence, intention might itself be a complex mechanism, enveloping multiple strategies. Future research will be directed to separating the priming and selection mechanisms of intention.

References

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

Bruner, J. S., & Goodman, C. C. (1947). Value and need as organizing factors in perception. *Journal of Abnormal and Social Psychology, 42,* 33–44.

Carpenter, G. A., & Grossberg, S. (1987). A massively parallel architecture for a self-organizing neural pattern recognition machine. *Computer Vision, Graphics, and Image Processing, 37,* 54–115.

Carr, T. H., McCauley, C., Sperber, R. D., & Parmlee, C. M. (1982). Words, pictures, and priming: On semantic activation, conscious identification, and the automaticity of information processing. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 757–777.

Clowes, M. B. (1971). On seeing things. *Artificial Intelligence, 2,* 79–116.

Cooper, L. A., & Shepard, R. N. (1973). Chronometric studies of the rotation of mental images. In W. G. Chase (Ed.), *Visual information processing* (pp. 75–176). New York: Academic Press.

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

Corballis, M. C., Zbrodoff, N. J., & Roldan, C. E. (1976). What's up in mental rotation? *Perception & Psychophysics, 19,* 525–530.

Cornwell, H. G. (1964). Effect of training on figure-ground organization. *Journal of Experimental Psychology, 68,* 108–109.

Durso, F. T., & Johnson, M. K. (1979). Facilitation in naming and categorizing repeated pictures and words. *Journal of Experimental Psychology: Human Learning and Memory, 5,* 449–459.

Dutton, M. B., & Traill, P. M. (1933). A repetition of Rubin's figure-ground experiment. *British Journal of Psychology, 23,* 389–400.

Eley, M. G. (1982). Identifying rotated letter-like symbols. *Memory & Cognition, 10,* 25–32.

Epstein, W. (1982). Percept-percept couplings. *Perception, 11,* 75–88.

Epstein, W., & De Shazo, D. (1960). Recency as a function of perceptual oscillation. *American Journal of Psychology, 74,* 215–223.

Epstein, W., & Rock, I. (1960). Perceptual set as an artifact of recency. *American Journal of Psychology, 73,* 214–228.

Girgus, J. J., Rock, I., & Egatz, R. (1977). The effect of knowledge of reversibility on the reversibility of ambiguous figures. *Perception & Psychophysics, 22,* 550–556.

Gogel, W. C., & Tietz, J. (1974). The effect of perceived distance on perceived motion. *Perception & Psychophysics, 16,* 70–78.

Graham, C. (1929). Area, color, and brightness difference in a reversible configuration. *Journal of General Psychology, 2,* 470–483.

Harrower, M. (1936). Some factors determining figure-ground articulation. *British Journal of Psychology, 26,* 407–424.

Harvey, E. M., Peterson, M. A., & Gibson, B. S. (1990, April). *The mechanisms of intention in figure-ground perception.* Poster presented at the meeting of the Rocky Mountain Psychological Association, Tucson, AZ.

Hebb, D. O. (1949). *The organization of behavior.* New York: Wiley.

Hochberg, J. (1970). Attention, organization, and consciousness. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis* (pp. 99–124). New York: Appleton-Century-Crofts.

Hochberg, J. (1971). Perception I. Color and shape. In J. W. Kling and L. Riggs (Eds.), *Woodworth and Schlosberg's experimental psychology* (3rd ed., pp. 395–474). New York: Holt, Rinehart & Winston.

Hochberg, J. (1972). The representation of things and people. In E. H. Gombrich, J. Hochberg, & M. Black (Eds.), *Art, perception, and reality* (pp. 47–94). Baltimore: Johns Hopkins University Press.

Hochberg, J. (1974). Higher order stimuli and interresponse coupling in the perception of the visual world. In R. B. McLeod & H. L. Pick (Eds.), *Perception: Essays in honor of James J. Gibson* (pp. 17–39). Ithaca, NY: Cornell University Press.

Hochberg, J. (1978). *Perception* (2nd ed). Englewood Cliffs, NJ: Prentice-Hall.

Hochberg, J., & Peterson, M. A. (1987). Piecemeal organization and cognitive components in object perception: Perceptually coupled responses to moving objects. *Journal of Experimental Psychology: General, 116,* 370–380.

Hoffding, H. (1891). *Outlines of psychology.* New York: Macmillan.

Hoffman, D. D., & Richards, W. A. (1985). Parts of recognition. In S. Pinker (Ed.), *Visual cognition* (pp. 65–96). Cambridge, MA: MIT Press.

Humphreys, G. W., & Riddoch, M. J. (1984). Routes to object constancy: Impliciations from neurological impairments of object constancy. *Quarterly Journal of Experimental Psychology, 36A,* 385–415.

Humphreys, G. W., Riddoch, M. J., & Quinlan, P. T. (1988). Cascade processes in picture identification. *Cognitive Neuropsychology, 5,* 67–103.

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

Jolicoeur, P. (1988). Mental rotation and the identification of disoriented objects. *Canadian Journal of Psychology, 42,* 461–478.

Jolicoeur, P., & Milliken, B. (1989). Identification of disoriented objects: Effects of context on prior presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 200–210.

Kanizsa, G., & Luccio, R. (1987). Formation and categorization of visual objects: Hoffding's never refuted but always forgotten argument. *Gestalt Theory, 9,* 11–127.

Kelter, S., Grotzbach, H., Freiheit, R., Hohle, B., Wutzig, S., & Diesch, E. (1984). Object identification: The mental representation of physical and conceptual attributes. *Memory & Cognition, 12,* 123–133.

Koffka, K. (1935). *Principles of gestalt psychology.* New York: Harcourt, Brace.

Kohler, W. (1929). *Gestalt psychology.* New York: Liveright.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres. *Psychological Review, 94,* 148–175.

Kosslyn, S. M., Flynn, R. A., Amsterdam, J. B., & Wang, G. (1990). Components of high level vision: A cognitive neuroscience analysis and accounts of neuropsychological syndromes. *Cognition, 34,* 203–277.

Kunnapas, T. (1957). Experiments in figural dominance. *Journal of Experimental Psychology, 53,* 31–39.

Lowe, D. G. (1985). *Perceptual organization and visual recognition.* Boston: Kluwer Academic.

Maki, R. H. (1986). Naming and locating the tops of rotated figures. *Canadian Journal of Psychology, 40,* 368–387.

Marr, D. (1982). *Vision.* San Francisco: W. H. Freeman.

Marr, D., & Hildreth, E. (1980). Theory of edge detection. *Proceedings of the Royal Society of London, Series B, 207,* 187–217.

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

McCracken, P. J., Gogel, W. C., & Blum, G. S. (1980). Effects of post hypnotic suggestion on perceived egocentric distance. *Perception, 9,* 561–568.

Murphy, G. (1947). *Personality.* New York: Harper.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton-Century-Crofts.

Osgood, C. E. (1953). *Method and theory in experimental psychology.* New York: Oxford University Press.

Palmer, S. E. (1989). Reference frames in the perception of shape and orientation. In B. E. Shepp and S. Ballesteros (Eds.), *Object perception: Structure and process* (pp. 121–161). Hillsdale, NJ: Erlbaum.

Palmer, S. E., Rosch, E., & Chase, P. (1981). Canonical perspectives and the perception of objects. In J. Long & A. Baddely (Eds.), *Attention and performance IX* (pp. 135–151). Hillsdale, NJ: Erlbaum.

Pastore, N. (1949). Need as a determinant of perception. *The Journal of Psychology, 28,* 457–475.

Peterson, M. A. (1984). *Measures of selective components in perceptual organization.* Unpublished doctoral dissertation, Columbia University, New York.

Peterson, M. A. (1986). Illusory concomitant motion in ambiguous stereograms: Evidence for nonstimulus contributions to perceptual organization. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 50–60.

Peterson, M. A., & Gibson, B. S. (1991a). Directing spatial attention within an object: Altering the functional equivalence of structural descriptions. *Journal of Experimental Psychology: Human Perception and Performance, 17,* 170–182.

Peterson, M. A., & Gibson, B. S. (1991b). The initial identification of figure–ground relationships: Contributions from shape recognition routines. *Bulletin of the Psychonomic Society, 29,* 199–202.

Peterson, M. A., & Hochberg, J. (1983). Opposed-set measurement procedure: A quantitative analysis of the role of local cues and intention in form perception. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 183–193.

Peterson, M. A., & Hochberg, J. (1989). Necessary considerations for a theory of form perception: A theoretical and empirical reply to Boselie and Leeuwenberg (1986). *Perception, 18,* 105–119.

Peterson, M. A., Rose, P. M., Gibson, B. S., & Vezey, E. (1991). *A new corpus of figure–ground stimuli.* Unpublished manuscript.

Peterson, M. A., & Shyi, G. C. (1988). The detection of real and apparent concomitant rotation in a three-dimensional cube: Implications for perceptual interactions. *Perception & Psychophysics, 44,* 31–42.

Riddoch, M. J., & Humphreys, G. W. (1987). Picture naming. In G. W. Humphreys & M. J. Riddoch (Eds.), *Visual object processing: A cognitive neuropsychological approach.* Hillsdale, NJ: Erlbaum.

Rock, I. (1973). *Orientation and form.* New York: Academic Press.

Rock, I. (1975). *An introduction to perception.* New York: Macmillan.

Rock, I., & Fleck, F. (1950). A reexamination of the effect of monetary reward and punishment on figure–ground perception. *Journal of Experimental Psychology, 40,* 766–776.

Rock, I., & Gutman, D. (1983). The effect of inattention on form perception. *Journal of Experimental Psychology: Human Perception and Performance, 7,* 275–285.

Rock, I., & Kremen, I. (1957). A re-examination of Rubin's figural after-effect. *Journal of Experimental Psychology, 53,* 23–30.

Rock, I., & Mitchener, K. (in press). Further evidence of failure of reversal of ambiguous figures by uninformed subjects. *Perception.*

Rubin, E. (1958). Figure and ground. In D. Beardslee & M. Wertheimer (Ed. & Trans.), *Readings in perception* (pp. 35–101). Princeton, NJ: Van Nostrand. (Original work published 1915)

Schacter, D. S., Cooper, L. A., & Delaney, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119,* 5–24.

Schacter, D. S., Cooper, L. A., Delaney, S. M., Peterson, M. A., & Tharan, M. (1991). Implicit memory for possible and impossible objects: Constraints on the construction of structural descriptions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 17,* 3–19.

Schafer, R., & Murphy, G. (1943). The role of autism in a visual figure–ground relationship. *Journal of Experimental Psychology, 32,* 335–343.

Selfridge, O. G., & Neisser, U. (1960). Pattern recognition by machine. *Scientific American, 203,* 60–68.

Sejnowski, T. J., & Hinton, G. E. (1987). Separating figure from ground with a Boltzman machine. In M. A. Arbib & A. R. Hanson (Eds.), *Vision, brain, & cooperative computation* (pp. 703–724). Cambridge, MA: MIT Press.

Smith, D., & Hochberg, J. (1954). The effect of "punishment" (electric shock) on figure–ground perception. *Journal of Psychology, 37,* 83–87.

Takano, Y. (1989). Perception of rotated forms: A theory of information types. *Cognitive Psychology, 21,* 1–59.

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

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

Wallach, H. (1949). Some considerations concerning the relation between perception and cognition. *Journal of Personality, 18,* 6–13.

Witkin, A. P., & Tenenbaum, J. M. (1983). On the role of structure in vision. In A. Rosenfeld, A. Hope, & J. Beck (Eds.), *Human and machine vision* (pp. 481–543). New York: Academic Press.

Received March 7, 1990
Revision received March 13, 1991
Accepted March 14, 1991 ■