

# Perceiving parts and shapes from concave surfaces

ANTHONY D. CATE

Veterans Affairs Northern California Health Care System, Martinez, California

AND

MARLENE BEHRMANN

Carnegie Mellon University, Pittsburgh, Pennsylvania

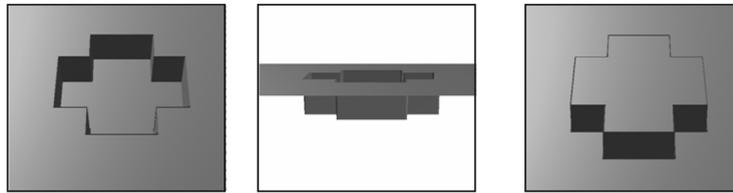
“A hole is nothing at all, but it can break your neck.” In a similar fashion to the danger illustrated by this folk paradox, concave regions pose difficulties to theories of visual shape perception. We can readily identify their shapes, but according to principles of how observers determine part boundaries, concavities in a planar surface should have very different figural shapes from the ones that we perceive. In three experiments, we tested the hypothesis that observers perceive local image features differently in simulated 3-D concave and convex regions but use them to arrive at similar shape percepts. Stimuli were shape-from-shading images containing regions that appeared either concave or convex in depth, depending on their orientation in the picture plane. The results show that concavities did not benefit from the same global object-based attention or holistic shape encoding as convexities and that the participants relied on separable spatial dimensions to judge figural shape in concavities. Concavities may exploit a secondary process for shape perception that allows regions composed of perceptually independent features to ultimately be perceived as gestalts.

The label *hole* can be applied to numerous kinds of structures (for a philosophical review, see Casati & Varzi, 1994). In this study, we focus on the specific case of concave regions that are bounded by a planar surface, as in Figure 1. This kind of stimulus is useful both because it resembles the kind of common real-world hole that can break your neck (e.g., an unexpected pothole) and, more important, because the depth dimension is orthogonal to the 2-D planar outline shape of the bounding cusp. We exploited the latter property to compare the perception of local surfaces that form either concave or convex versions of 3-D shapes that project the same bounding contour (Experiment 1) and to examine how the perception of a 2-D outline changes when it is defined by a 3-D concavity or a 3-D convexity (Experiment 2); in addition, we compared observers' judgments of 3-D concavities and convexities as whole objects (Experiment 3). In this study, depth information is conveyed using purely pictorial (monocular) cues; the transformation used to change apparent depth (rotation in the picture plane) does not alter the low-level properties of the images themselves (i.e., image reflectance, contrast, and spatial frequencies are unchanged) but only leads the observer to interpret them differently. Thus, any effects of depth are the consequences of the depth interpretation and not of the distal cues themselves, thereby differing, for example, from a reversal of the sign of disparity in stereograms. Even though the stimuli were

created purely in 2-D, we use the term *3-D concavity* to distinguish regions of a planar surface that appear concave with respect to depth from *2-D concavities*, which are the concave portions of a 2-D shape boundary that lie in the plane itself. Finally, it should be noted that although these investigations were focused on perceptual differences, there are other very salient functional differences between 3-D concavities and convexities. For example, hand-sized solid masses and holes elicit very different kinds of actions (cf. Palmer, Davis, Nelson, & Rock, 2008; Subirana-Vilanova & Richards, 1996); even the concave and convex aspects of the same object, such as a bowl, require different gripping actions.

## Perceived Shapes of Surface Concavities and Convexities

One of the gestalt principles for predicting which region of an image will be perceived as a figure is the *convexity principle* (Kanizsa & Gerbino, 1976): All things being equal, a 2-D region bounded by convex curvature in its contour will be perceived as the figure. This principle was extended by Stevens and Brookes (1988), who found that the presence of locally concave features (concave cusps) acted over and above global contour concavity to cause regions to be identified as grounds. Since such grounds tend not to be recognized when presented later as isolated shapes (i.e., without the figural part of the original image; Palmer



**Figure 1.** The type of hole that is the focus of this study. **Left:** A concavity bounded by a flat plane, with a closed bounding contour. **Middle:** The concavity shown as an indentation in a planar sheet. **Right:** The concavity shown rotated 180° in depth so that it appears as a convex bulge with the same 2-D bounding contour.

et al., 2008), in this case contours that are primarily concave would likely be associated with poor shape encoding.

However, the shape of the region corresponding to a concavity in depth bounded by a planar surface (as in Figure 1) appears unambiguously figural. In this 3-D case, concavities can produce the same (or at least very similar) shape percepts as the solid regions that would fill these holes. Such an interpretation is similar to findings from recent research on the perception of 2-D holes (apertures in planes). Indeed, when observers perceive shapes from 2-D holes (apertures in planes), they remember these shapes just as well as they do solid planar shapes (Palmer et al., 2008). When the shape described by a hole's outline is perceived, the observers may perceive the hole to have the same shape as its corresponding solid object, either directly based on the shape of the empty region itself (Nelson & Palmer, 2001; Palmer et al., 2008) or indirectly based on the shape of the surrounding solid material (Bertamini & Croucher, 2003). In any case, 2-D holes appear not to be perceived via the same processes as figural shapes associated with solid material. Investigations have shown that switching from a convex to a concave interpretation of a planar contour reverses the figure-ground status of the two sides of the contour (Baylis & Driver, 1993; Bertamini & Croucher, 2003) and changes the perceived part structure of a region (Barenholtz, Cohen, Feldman, & Singh, 2003; Bertamini & Farrant, 2005; Cohen, Barenholtz, Singh, & Feldman, 2005). This, in turn, leads to performance differences for solid shapes and aperture holes in different kinds of psychophysical tasks (Bertamini & Lawson, 2006). This is all in spite of the fact that, regardless of whether a region is perceived to be a hole or a surface, the bounding contour is the same and thus has the potential to convey the same amount of image information (Bertamini & Croucher, 2003).

A 3-D concavity is unambiguously associated with solid material (the sides and bottom of the hole) in addition to the empty space it encompasses. What may differ, then, between the perception of shape from 3-D concavities and convexities is the type and amount of information that observers glean from the same image when a surface appears to encompass solid material versus empty space. In the present study, we focused on three key questions related to the idea that similar figural shapes are ultimately perceived from 3-D concavities and convexities in spite of differences in how observers perceive their local features.

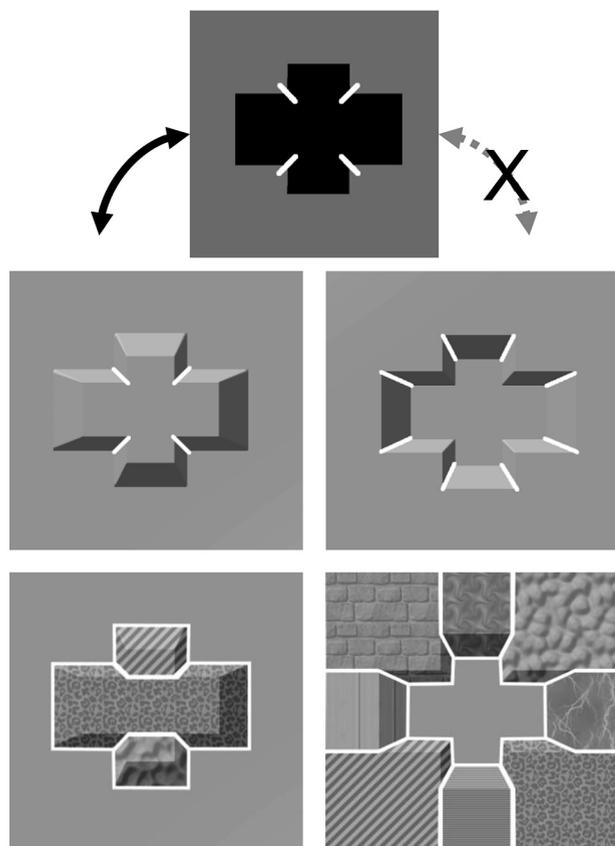
Proceeding in a local-to-global order of analysis, we first asked whether the perceived relationships between local image elements in 3-D concavities are different from those in 3-D convexities. Second, since the 2-D shape projected by the bounding cusp of 3-D convexities and concavities can be the same, are there differences in the perceived part structure of identical 2-D shapes when they are associated with 3-D concave and convex regions? Third, is the same shape gestalt in fact perceived equally well from 3-D convexities and concavities?

## EXPERIMENT 1

### Effects of 3-D Depth on Perceived Part Structure

An a priori reason for believing that local shape features are encoded differently from 3-D concavities and convexities comes from the minima rule of Hoffman and Richards (1984). The minima rule posits that minima of curvature are crucial to parsing objects: "Divide a surface into parts at loci of negative minima of each principal curvature along its associated family of lines of curvature" (p. 74). Essentially, an object can be divided into parts with boundaries formed on the maximally concave sections of each curved contour present on the object. A key implication of the minima rule is that similar 3-D concave and convex regions ought to have different part boundaries.<sup>1</sup> The middle row of Figure 2 shows the same cross-shaped 3-D region portrayed as both a concavity and a convexity. Superimposed lines indicate local minima of curvature.<sup>2</sup> Note that there are twice as many minima in the concave version. This is because the 2-D cross shape itself has unequal numbers of concave and convex extrema of 2-D curvature. The minima rule can be applied to 2-D shapes as well as to 3-D shapes, and the positions of the minima of curvature for a 3-D convexity and its 2-D projection align (Figure 2), but this is not the case for 3-D concavities. Hence we can hypothesize that the shape of a 3-D concavity will not be perceived as holistically as its 3-D convex counterpart (because it is subdivided by more part boundaries), nor will it share the same general pattern of within- and between-parts effects (Barenholtz & Feldman, 2003) with its 2-D outline shape (in contrast to a 3-D convexity).

In Experiment 1, we used a flanker task, inspired by studies of object- and part-based attention (Barenholtz & Feldman, 2003; Kramer & Jacobson, 1991; Vecera,



**Figure 2.** Top: Minima of curvature marked on a 2-D cross shape. These minima of curvature correspond to minima that cut through the sides of a 3-D convexity but not to those of a 3-D concavity, which are complementary. Middle: These minima of curvature marked on equivalent convex and concave 3-D versions of the cross shape. There are twice as many minima cutting through the sides in the concave version. Bottom: Part structures (indicated by different image textures) based on the convex and concave minima and on the shortcut rule of Singh, Seyranian, and Hoffman (1999). It is unclear precisely how the concave image’s part boundaries radiate out past the cusp of the concavity. However, since the cusp itself is a maximum of curvature, the facets oriented in depth, which form the sides of the concavity, belong with regions of the surrounding plane, rather than with the surface forming the bottom of the hole.

Behrmann, & Filapek, 2001; Vecera, Behrmann, & McGoldrick, 2000), to test whether 3-D concavities and convexities influence attention to local image elements differently, and, by extension, whether their surfaces are perceived more like independent parts or not. The dependent measure was a difference score comparing performance on two types of trials, congruent (both surfaces same color) and incongruent (surfaces different colors), as was used in Kramer and Jacobson. The degree to which congruence between the surfaces’ colors aided performance and the degree to which incongruence hurt performance indicated how well the surfaces were perceived as an integrated unit.

The perceptual integration of various pairs of surfaces was tested by asking participants to judge the color of a target surface while ignoring the color of a distractor sur-

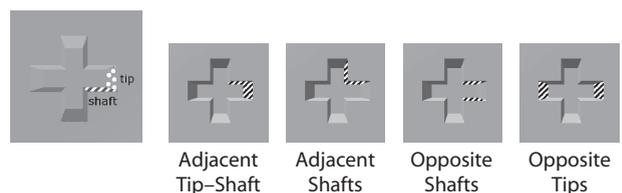
face. Each individual surface on the bounding wall of a cross-shaped region was used equally often in each role, and both adjacent and nonadjacent pairs of surfaces appeared in the experiment. The bounding wall of the 3-D cross-shaped region was composed of two qualitatively distinct types of surface: *tips* (the surfaces forming the ends of the crossbars) and *shafts* (the surfaces aligned with the long dimensions of the crossbars). Note that for a cross shape, there are necessarily twice as many shafts as tips (eight vs. four). Accordingly, there are eight possible adjacent tip–shaft combinations, four adjacent-shaft combinations, four opposite-shaft combinations, and two opposite-tip combinations. Examples of these four kinds of surface pairings are shown in Figure 3.

In Experiment 1, we measured the extent to which different task-irrelevant background images imposed perceptual unity on two spatially distinct parts of the display. We hypothesized that a 3-D convex background would produce a main effect of greater perceptual unity across all pairs of surfaces, in accord with the idea that attending to one region of an object facilitates attention to other spatially distant regions that belong to the object (e.g., Egly, Driver, & Rafal, 1994; Lavie & Driver, 1996). We hypothesized that such object-based attention effects would be absent with 3-D concave backgrounds and that, in this case, intervening 3-D curvature minima might actually render different surfaces more perceptually independent than would be predicted on the basis of their spatial separation alone.

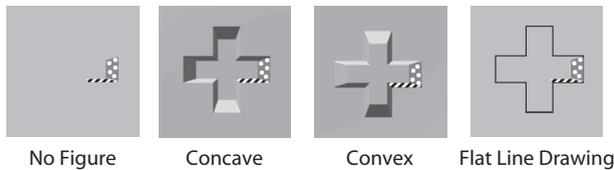
**Method**

**Participants.** Twenty undergraduates participated for course credit. All of the participants had normal or corrected-to-normal vision and were between 18 and 22 years of age.

**Stimuli.** The stimuli for a given trial included three images: a background image, a cue image, and a probe image. The cue and probe images were made by adding color to the background images in systematic ways. Four background images were used and are illustrated in Figure 4: no figure, concave, convex, and flat line drawing. The no-figure image was a blank green metallic surface rendered at a resolution of 400 × 400 pixels, using MATLAB (MathWorks Inc., Natick, MA). Its width and height both spanned 8 cm (approximately 9.2° at 50 cm viewing distance) on the screen. The concave background image described a cross-shaped region with equal horizontal and vertical extents of approximately 6 cm (6.9° at 50 cm) on the screen, with arms approximately 2 cm (2.3°) in width. The convex background image was produced by rotating the concave background image 180° in the picture plane. Many authors (e.g., Kleffner & Ramachandran, 1992) have shown that the polarity of a shape-from-shading image’s 3-D depth reverses when the apparent direction of



**Figure 3.** Schematic illustration of surface pair combinations used. Only surface pigmentation was altered to highlight surfaces in the actual stimuli; hatching is drawn for clarity.



**Figure 4.** Schematic illustration of stimulus backgrounds used in Experiment 1. Hatched and dotted regions represent differently colored surface patches for clarity. All examples show the adjacent tip–shaft surface pair for clarity of comparison.

a light source is inverted. The flat line drawing image consisted of a closed black line drawn on the no-figure background that described the planar boundary of the cross-shaped region. In all background types, the surface was illuminated by a simulated light source above and to the left of the figure; this position has been found to be optimal for inducing perception of depth-from-shading (Sun & Perona, 1998). The light source in the testing room (a halogen bulb desk lamp) was positioned above and to the left of the monitor.

The cue images were created by coloring two of the surfaces on the bounding wall of the concave region—one gold and one white. The same cue images were used with all background types, but they were superimposed differently on the different backgrounds. The convex cue images were made by rotating the concave cue images 180° in the picture plane. The no-figure and flat line drawing cue images were made by copying the colored surface patches from a concave cue image and superimposing them on a blank or line drawing background (see Figure 4). Note that this method preserved all of the luminance and shading properties of the colored patches that were rendered as part of a concavity. To reduce the depth-from-shading cues associated with the colored patches, these entire images were rotated 90° clockwise after the patches were added.

The gold-colored surface (henceforth, the *target cue*) was used to indicate where a target would appear in the probe image, and the white-colored surface (the *distractor cue*) indicated where a distractor would appear. Figure 3 shows examples of the four types of surface location relationships. Two cue images were created for each unique pair of surfaces by counterbalancing the relative location of the target and distractor patches. Thus, for each background type, 16 total cue images were created for the adjacent tip–shaft condition, 8 for the adjacent-shafts condition, 8 for the opposite-shafts condition, and 4 for the opposite-tips condition, for a total of 36 images. Note that in the adjacent tip–shaft condition the two surfaces were different sizes, since a tip had a larger surface area than a shaft. Since a target cue was equally likely to appear on a tip as on a shaft in the adjacent tip–shaft images, all trials involving the adjacent tip–shaft condition were binned together for analysis.

The probe images were created in the same way as were the cue images, except that in lieu of gold and white, the two surfaces were colored with red and blue. A probe image could include any of the four possible permutations of the two color values. Thus, for each of the 36 cue images of a given background type, 4 probe images were created. This yielded a total of 144 probe images per background type.

**Procedure.** The testing session consisted of two stages. In the first stage, all stimuli were of the no-figure type, so that performance in this stage would not be influenced by the expectancy of a global cross-shaped contour. In the second stage, concave, convex, and flat line drawing trials were randomly interleaved.

During the first stage, the participants' only task was to identify the color of the target patch in the probe image. The participants were instructed to direct their attention to the gold-colored surface and to report which color this surface took on in the probe image. The participants were told that the white-colored surface indicated where an additional red or blue patch would appear in the probe image and that this additional patch was a distractor that should be ignored.

The participants pressed a key on the keyboard to initiate a trial. The cue image appeared in the center of the screen, surrounded by a gray background, and remained on the screen for 1 sec. The text "Get ready . . ." appeared at the bottom of the screen for the same interval. A blank screen ensued for 300 msec, followed by the centrally presented probe image, which remained until the participants pressed one of two keys, which were labeled with red and blue stickers. The response and reaction times (RTs; relative to probe image onset) were recorded. A feedback message ("Correct!" or "Incorrect!") then appeared. If the response was correct, the participants could proceed to the next trial, but if not, the computer beeped, the keyboard was locked, and the participants were required to wait for 5 sec before the next trial could be initiated.

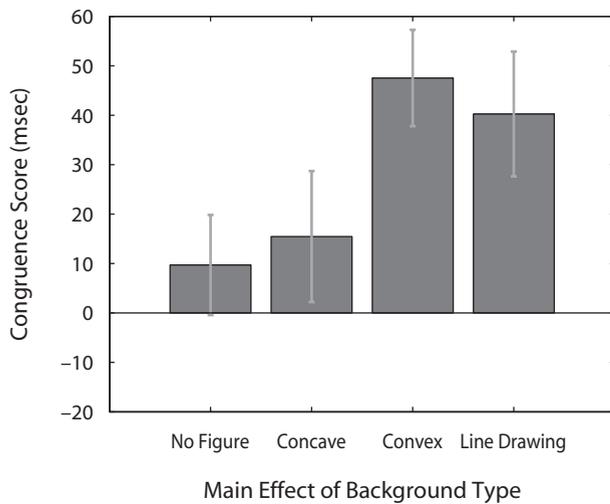
During the second stage, the participants performed an additional task at the beginning of each trial: They judged the apparent depth of a background image containing no colored patches. At the start of a trial, a background image appeared and remained visible until the participants pressed one of three keys to indicate whether the background appeared to be concave, convex, or a flat line drawing. This response and its RT were recorded. The participants used the left hand to press these three keys, in order to leave the right hand free to respond for the main task. The participants were instructed to indicate what depth the image appeared to have upon first inspection. After this first response, the background image disappeared and was replaced with a cue image. The remainder of the trial was identical to the procedure used in the first stage of the experiment.

All experiments were conducted using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) for MATLAB. The participants performed 144 trials in the first stage, with 1 trial for each of the possible probe images, and 432 trials in the second stage (144 trials for each of the four background types). Prior to the experiment, the participants read printed instructions with diagrams, and performed 10 practice trials, randomly selected from the first stage. During the second stage, three breaks were given, and, upon completion, the participants were debriefed.

## Results

For the first stage, trials with incorrect responses were excluded from the RT analysis. Second-stage trials were excluded when either the background image response (concave/convex/flat) or the probe image response was incorrect. The participants with probe image error rates greater than 10% for all trials, or with an error rate greater than 20% for any one cell (see below), were excluded from the analyses. Four participants were excluded on this basis, leaving 16 (9 male, 7 female). The mean error rate for the remaining participants was very low (2.6%) and was not analyzed further.

The mean RTs for the preliminary task of identifying the concave, convex, and flat line drawing backgrounds were 881 ( $SEM = 91$ ), 920 ( $SEM = 80$ ), and 727 ( $SEM = 73$ ) msec, respectively. A repeated measures ANOVA revealed a main effect of background type [ $F(2,30) = 17.55, p < .001$ ], and a Tukey's HSD test yielded a critical  $T$  value of 85 msec, meaning that the flat line drawing responses were significantly faster than those for the concave and convex backgrounds, which did not differ significantly from each other. The mean accuracies for the concave, convex, and flat line drawing backgrounds were very similar at 97.3%, 96.4%, and 96.7%, respectively [ $F(2,30) = 0.55, p = .58$ ]. Any differences between the concave and convex conditions in the main task, therefore, were not attributable to longer viewing times or unequal difficulty between conditions.



**Figure 5.** Main effect of stimulus background on congruence scores.

A factorial repeated measures ANOVA with background type (four levels: no figure, concave, convex, flat line drawing), surface relationship (four levels: adjacent tip–shaft, adjacent shafts, opposite shafts, opposite tips), and congruence (two levels: congruent, incongruent) as factors was conducted with the congruent–incongruent difference score as the dependent measure. Before computing the congruence score, all trials were binned into 32 cells for the analysis (4 × 4 × 2). These cells were the basis for excluding participants with greater than 20% errors in any single cell. The median RT for all trials within a cell was calculated, and the median congruent RT was subtracted from the median incongruent RT to obtain the congruence score. The congruence scores for the no-figure trials were used as a baseline measure of the perceptual unity of two surfaces inherent from their sizes and proximity and were subtracted from the congruence scores for the other background image conditions. These baseline congruence scores were 20 msec (*SEM* = 20 msec) for the adjacent-shafts condition, 29 msec (*SEM* = 20 msec) for the adjacent tip–shaft condition, 7 msec (15 msec) for the opposite-shafts condition, and –23 msec (*SEM* = 22 msec) for the opposite-tips condition. The negative score in the opposite-tips conditions may indicate that when a spatially distant pair of surfaces share the same color, the distractor may impede performance by drawing attention away from the target rather than helping by providing a redundant feature (cf. Lamy, Leber, & Egeth, 2004; Theeuwes, De Vries, & Godijn, 2003).

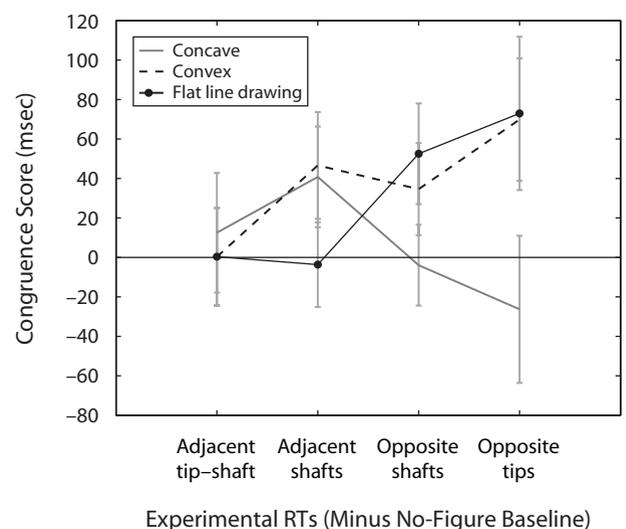
The ANOVA revealed a main effect of background type [ $F(2,30) = 3.42, p = .046$ ]. Planned pairwise comparisons revealed that the concave condition had a significantly lower congruence score than the convex condition [ $t(15) = 2.36, p = .033$ ], but neither the concave versus flat line drawing [ $t(15) = 1.89, p = .078$ ] nor the convex versus flat line drawing [ $t(15) = 0.62, p = .55$ ] comparison was

significant. An ANOVA examining the raw RTs for these three background types showed no main effect [ $F(2,30) = 1.32, p = .28$ ] and no interaction [ $F(6,90) = 0.35, p = .91$ ], so the congruence score effects were not likely to be by-products of baseline performance differences for the background types. Figure 5 shows the non-baseline-adjusted congruence scores for all four conditions.

There was also a significant two-way interaction between background type and surface relationship [ $F(6,90) = 3.39, p = .0046$ ]. A post hoc Tukey’s HSD critical *T* value of 89 msec showed that the only significant differences were in the opposite-tips relationship condition, in which the convex and flat line drawing backgrounds yielded higher congruence scores (i.e., greater interference from the distractor surface) than the concave background. Figure 6 shows the baseline-adjusted scores grouped by background type and surface relationship.

**Discussion**

3-D concavities and convexities engage attention to local image elements differently, even when they define the same 2-D outline shape. Two aspects of the results stand out. First, unlike the convex and 2-D line drawing backgrounds, the 3-D concave backgrounds added very little perceptual unity to the display on the whole, barely more than the no-figure condition (Figure 5). Second, the concave and no-figure backgrounds both yielded negative scores in the opposite-tips condition, which may indicate that these backgrounds allowed a spatially distant distractor to impede performance by drawing attention away from the target when it was congruent (cf. Lamy et al., 2004; Theeuwes et al., 2003), in the manner of an object distinct from the one containing the target feature. Although similar 2-D cross shapes can be perceived in the three figural background types, the figural



**Figure 6.** Congruence scores based on median reaction times (RTs). Scores from the no-figure baseline condition have been subtracted out. Error bars represent the standard errors of the means.

goodness produced by 3-D concavities does not affect visual attention in the manner of most objects.

## EXPERIMENT 2

### Effect of 3-D Depth on Integration of 2-D

#### Part Structure

Experiment 1 showed that attention does not spread globally across 3-D concavities the way it does in 3-D convexities, which suggests that when figural shape is perceived in concavities, it might be perceived in a less holistic way than in convexities. To assess whether figural shape judgments differ for 3-D concavities and convexities, in the next experiment participants made judgments about the shape of contours bounding 3-D concave and convex regions when the 2-D outlines of these shape-from-shading images were identical.

Examining judgments of 2-D shape is a reliable means for assessing the perceptual ramifications of different 3-D structures. In an influential series of experiments, Enns and Rensink (1990, 1991) found that visual search is efficient when 2-D line drawings convey vivid 3-D structure, even when similar but flat-looking drawings yield inefficient search slopes. An even stronger demonstration of the interaction between the orthogonal dimensions of 2-D and 3-D structure comes from Z. Liu, Jacobs, and Basri (1999), who found that concavity in a shape's 2-D outline affects how well observers integrate different parts of a shape across gaps in 3-D depth. Observers were better at noticing a depth disparity separating two halves of a planar shape when the two halves were linked by a concave 2-D contour (like an hourglass) rather than by a convex contour (like a barrel), suggesting that surfaces connected by a convex bulge were perceptually grouped by a better gestalt. Thus, a concave contour appears to allow the properties of different parts to be perceived independently, even when the properties in question are not directly related to the shape of the 2-D contour itself. Conversely, we consider here whether concavity in 3-D (i.e., depth) leads a 2-D outline to be perceived less holistically.

In Experiment 2, we used the Garnerian orthogonal insertion task of Lederman, Klatzky, and Reed (1993) to determine the integrality or separability of parts of a 2-D planar shape. In this task, participants view a set of stimuli and classify them on the basis of the value of a single

stimulus dimension. After this, variability is introduced along a second dimension that is completely independent from the first dimension (orthogonal insertion). The participants continue to classify the stimuli solely on the basis of the value of the first dimension, while ignoring variation along the second dimension. Poor performance after the orthogonal insertion (an orthogonal insertion effect) is evidence for dimensional integrality, because the participants were unable to attend to the first dimension without being distracted by the irrelevant dimension. If performance remains unchanged, then the dimensions are deemed to be separable. This difference is probably quantitative rather than qualitative, since the separable-integral difference itself is probably a continuum rather than a dichotomy (Garner, 1974).

The two dimensions manipulated in this experiment were the appearance of the arms and the ovoid body of a shape (Figure 7). Participants classified the stimuli on the basis of the arms, initially without and then with the presence of trial-to-trial variation in the body shape. Orthogonal insertion effects were predicted to be stronger for 3-D convexities than for 3-D concavities, reflecting greater integration of all of the 3-D convexities' features into a whole.

#### Method

**Participants.** A total of 36 undergraduates participated for course credit. The participants had normal or corrected-to-normal vision and were between 18 and 22 years of age.

Since we measured how responses were affected by an unexpected change to the stimulus in this experiment, each participant performed one session, viewing only a single stimulus type. Thus, the experiment was conducted with stimulus type as a between-subjects factor. Three groups of participants performed the experiment: 12 participants (7 male, 5 female) saw concave stimuli, 12 (6 male, 6 female) saw convex stimuli, and 12 (6 male, 6 female) saw ambiguous (90° rotated) stimuli.

**Stimuli.** The stimuli were 3-D-rendered images of novel shapes composed of two tapered arms protruding from an ovoid body (see Figure 7). One set of images was used to portray all three 3-D stimulus types: concave, convex, and ambiguous stimuli. All shapes were originally rendered as concavities recessed from a planar background surface, with light sources emanating from the top left of the shape, using a ray tracing program (POVray, Williamstown, Victoria, Australia). Convex stimuli were produced by rotating the concave stimuli 180° in the picture plane. The concave stimuli were rotated 90° clockwise to produce ambiguously shaded control stimuli.

All images were rendered at a resolution of 400 × 400 pixels. The stimuli had widths of approximately 9 × 5.7 cm, subtending

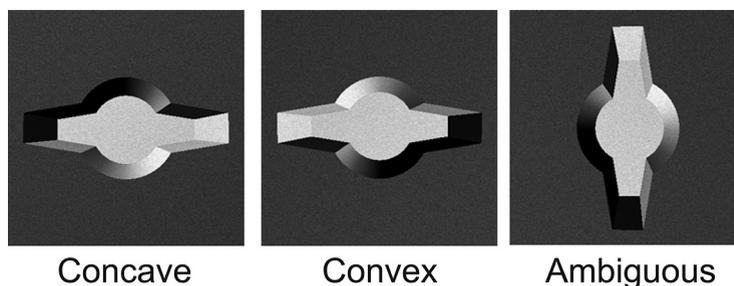
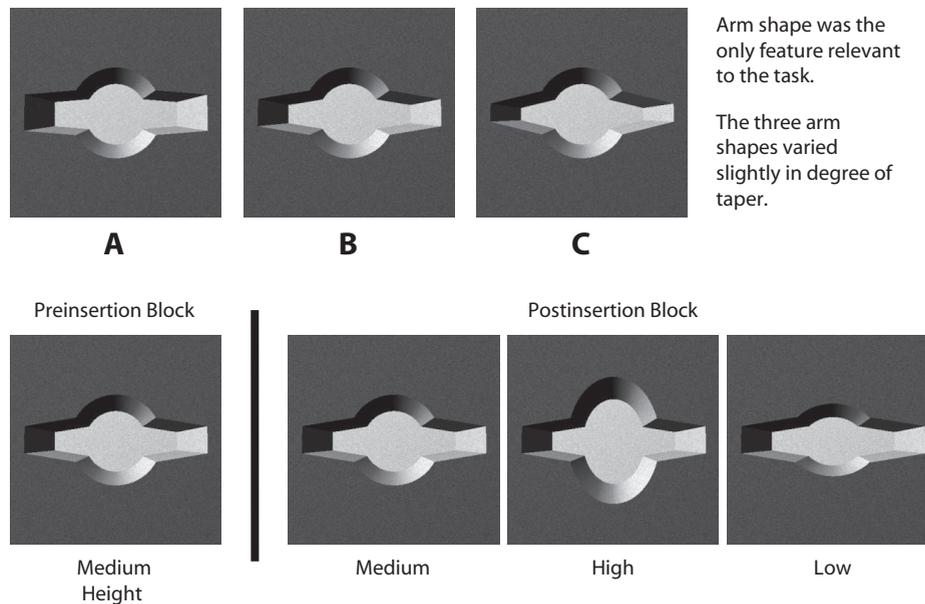


Figure 7. Examples of the three 3-D rendered stimulus types used in Experiment 2. The convex and ambiguous stimuli were created by rotating the concave stimuli in the picture plane.



**Figure 8. Scheme for stimuli used in Experiment 2. Top row: The three different arm shapes that the participants were instructed to identify. All three arm shapes are shown on the same (medium height) body shape for consistency of comparison. Bottom row: Illustration of the orthogonal insertion manipulation of the irrelevant (body shape) dimension. During the preinsertion block, the body shape dimension did not vary; all stimuli had a medium-height body. During the postinsertion block, body shape changed randomly from trial to trial, uncorrelated with arm shape. All postinsertion examples are shown with Arm Shape B for consistency of comparison.**

approximately  $10.2^\circ \times 7.4^\circ$  from a viewing distance of 50 cm. The surfaces within the central region had a lighter tone than the background plane, and both types of surface were finely textured. Pilot testing indicated that the concave texture cues did not prevent the images from being perceived as convex or ambiguous.

The stimuli had one of three different arm shapes (Figure 8, top row). Arm types A, B, and C tapered to end thicknesses subtending  $2.1^\circ$ ,  $1.6^\circ$ , and  $1.1^\circ$ , respectively. In addition, shapes in the postinsertion portion of the experiment could also have one of three different body shapes: high (subtending  $7.3^\circ$ ), medium ( $5.3^\circ$ ), or low ( $4.2^\circ$ ) (Figure 8, bottom row). All stimuli in the preinsertion portion of the experiment had medium body shapes.

**Procedure.** The participants performed a three-alternative forced choice categorization task, in which they categorized each stimulus according to its arm thickness. The participants first studied a display of sample stimuli that illustrated each arm type (all had medium body shapes) and its response label (A, B, and C), and were told that the body shape was irrelevant (Figure 8). Responses were made with the right hand on the keyboard using three adjacent keys. Verbal and written instructions informed the participants about the depth (concave or convex) that the stimuli were supposed to portray. Ambiguous stimuli were simply described as “computer-generated images.”

The participants performed a block of 18 practice trials (6 for each arm type). A stimulus appeared and remained until the participants made a keypress response. If the response was correct, a 1,000-msec intertrial interval (ITI) followed. If the response was incorrect, the word “Incorrect” was displayed in red text for 500 msec during the 1,000-msec ITI. All of the participants correctly identified each stimulus on at least four out of six trials. Prior to the experiment itself, the participants were instructed to respond both quickly and accurately.

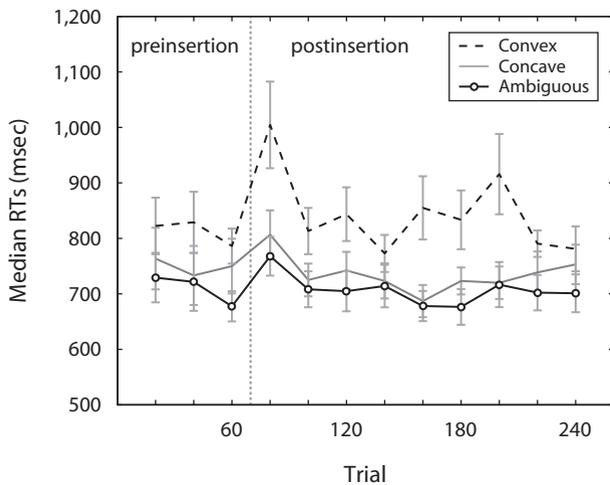
The main task consisted of two blocks of trials. The preinsertion block of 60 trials consisted of 20 presentations of each of the three stimuli, presented in random order. The trial procedure was identical to the practice block, but no feedback was provided. At the end of the first block, the experimenter instructed the participants that

the stimulus body would vary from trial to trial but that their three-way decision was still to be based only on the arm thickness. The postinsertion block began immediately following these instructions. The postinsertion block used nine different stimuli, each repeated 20 times, defined by the crossing of the three body types (low, medium, and high) with the three arm types (A, B, and C). After this, the participants were asked whether the stimuli appeared more concave or more convex, and their responses were recorded.

## Results

Two participants were excluded, one from the concave and one from the convex group, since their RTs or error rates, respectively, exceeded  $2 SD$  of the group mean. All of the participants in the concave and convex stimulus groups reported that the stimuli appeared to have the appropriate depth. Of the participants in the ambiguous group, 5 reported them to appear concave and 7 convex; this judgment difference did not significantly affect performance in the analyses of the data.

RTs from correct response trials were divided into bins of 20 trials, yielding three preinsertion bins and nine postinsertion bins, as is plotted in Figure 9 (see Lederman et al., 1993, for the rationale of this analytic method). A repeated measures ANOVA of the immediately pre- and postinsertion trials (i.e., Bins 3 and 4) included stimulus type as a between-subjects factor and time (preinsertion, postinsertion) as a within-subjects factor. Significant main effects were found for both stimulus type [ $F(2,31) = 4.51, p < .05$ ] and time [ $F(1,31) = 25.26, p < .001$ ], and there was a significant interaction between the two [ $F(2,31) = 4.10, p < .05$ ]. Planned paired-samples  $t$  tests revealed significant orthogonal insertion effects for the convex [ $t(10) = -3.36,$



**Figure 9.** Median reaction times (RTs) for each bin of 20 trials, averaged across participants for Experiment 2. The vertical line after Trial 60 indicates the orthogonal insertion point dividing the pre- and postinsertion blocks. Error bars represent the standard error of the mean.

$p = .0072$ ] and ambiguous groups [ $t(11) = -4.04$ ,  $p = .0019$ ] but not for the concave group [although there was a trend:  $t(10) = -2.11$ ,  $p = .061$ ]. Also, two-sample  $t$  tests showed that the magnitudes of the orthogonal insertion effect were significantly different between the concave and convex groups [ $t(20) = 2.30$ ,  $p = .032$ ] but not between the concave and ambiguous groups [ $t(21) = 0.97$ ,  $p = .34$ ] or between the convex and ambiguous groups [although there was a trend:  $t(21) = 1.93$ ,  $p = .067$ ]. The same ANOVA using accuracy as the dependent measure indicated lower postinsertion accuracy [ $F(1,31) = 4.45$ ,  $p < .05$ ] and no other significant effects.

The bin size of 20 trials was chosen for these analyses because it appeared to capture the peak of the orthogonal insertion effect while containing instances of all the different stimulus types. ANOVAs comparing the median RTs obtained with different bin sizes were performed using the concave and convex participant groups. The critical time  $\times$  stimulus type interaction was significant for the bin sizes of 10 [ $F(1,26) = 4.74$ ,  $p = .039$ ], 15 [ $F(1,26) = 5.43$ ,  $p = .028$ ], and 20 [as was reported above;  $F(1,26) = 5.08$ ,  $p = .033$ ], indicating that the critical finding is not specific to the bin type initially chosen.

## Discussion

The results of Experiment 2 indicate that 2-D figural shape may be perceived by different means when it is associated with 3-D concavities and convexities, even though this depth dimension is orthogonal to the plane of the figure itself. The participants viewing the 3-D convex and the 3-D ambiguous stimuli showed significant orthogonal insertion effects, as was measured by the RT difference between immediately pre- and postinsertion trials. In contrast, there was very little decrement in performance from the orthogonal insertion with the 3-D concave stimuli—significantly less than with 3-D convexities. It appears

that the participants perceived the 3-D convex and ambiguous images in an integrated or holistic manner, since they had difficulty ignoring changes to a task-irrelevant part of the shape.

In summary, the results of Experiment 2 suggest that figural shape perceived from 3-D concavities is derived from a collection of perceptually independent parts,<sup>3</sup> compared with the relatively holistic mode of shape perception seen with 3-D convexities. 3-D depth has high-level effects on perception even when it is an irrelevant and orthogonal dimension in a shape perception task. 3-D depth appears to affect the process of part determination in addition to the grouping across depth disparities reported by Liu et al. (1999), and the grouping of individual surfaces found in Experiment 1.

## EXPERIMENT 3

### The Effects of 3-D Depth on Whole Shape Matching

The results of the first two experiments showed that local surface elements of 3-D concavities, as well as the parts of the 2-D bounding contour shapes of 3-D concavities, are perceived as independent sets of features in contrast to the equivalent features of 3-D convexities. Continuing the local-to-global trajectory of the first two experiments, in Experiment 3 we investigated whether observers perceive the global dimensions of figural shapes differently with 3-D concavities and convexities and asked participants to make judgments about entire shapes, rather than to focus on specific parts of them. If observers really do perceive that 3-D concavities and convexities describe similar 2-D boundary shapes, do they arrive at these similar percepts via the same processes, and are they sensitive to the same global shape information in both cases?

Visual search studies indicate that 3-D concave shapes, even those without distinct parts, may be represented differently from 3-D convexities in the human visual system. Under many experimental conditions a 3-D concave target pops out of an array of convex but otherwise similar shapes, but convex targets are hard to detect among concave distractors. This search asymmetry (Treisman & Souther, 1985) suggests that concavity in depth is a salient feature for capturing attention. A concave shape is an unusual feature-present stimulus for a search asymmetry (Kleffner & Ramachandran, 1992), since it can be constructed by physically removing material from the feature-absent stimulus, the convex shape. However, it may be that the salient feature guiding attention is not so much the distal cues indicating concavity, but rather the greater number of psychological units (i.e., dimensions) perceived with a concavity versus a convexity. Popout for concavity has been found with 3-D stimuli that have no discernible divisions into multiple surfaces or parts, as with the shape-from-shading circles used by Kleffner and Ramachandran. It seems unlikely that the feature singleton guiding attention to concavities would be the presence of greater pictorial depth, because B. Liu and Todd (2004) found that pictorial depth tends to be greater in convexi-

ties, not concavities. If anything, this would appear to give convex targets an extra feature. One interpretation of these studies is that concave shapes are perceived to consist of a larger number of basic features than those of convexities.

In Experiment 3, we tested the hypothesis that observers are sensitive to different global dimensions of shapes (i.e., dimensions that span the entirety of the shape) with 3-D concavities and convexities. Specifically, the hypothesis predicts that observers maintain sensitivity to the spatial extents of concavities (e.g., height or width) as separate quantities, whereas these independent dimensions are integrated into a single emergent shape dimension (e.g., aspect ratio) in the case of convexities. This hypothesis holds that the number of global shape features perceived in concavities will always be greater than the number perceived in convexities, because integrating multiple spatial dimensions reduces the number of dimensions ultimately perceived. This basic hypothesis also implies that observers will be relatively sensitive to the absolute size of 3-D concave images but will perceive 3-D convex images via a more scale-invariant process.

In Experiment 3, we used a match-to-sample task to measure observers' sensitivity to the global spatial extents of images and to the features that emerge when these independent dimensions are perceptually integrated—specifically, aspect ratio and surface area. Two properties of the sample images' shapes were manipulated—the spatial extents of height and width—and distorting specific combinations of values for these properties yielded three categories of foils, corresponding to three possible schemes for identifying the stimulus: judging aspect ratio (ratio of height to width), area (product of height and width), and individual spatial extents (height and width considered as individual quantities). Each type of foil kept the stimulus's description constant according to one of the three schemes, despite making changes to the other two schemes. For example, a same-aspect-ratio foil had a different area and also different height and width values than the sample. If the participants relied especially on a given dimension to recognize the shapes, their performance would be poor when foils preserved that dimension.

## Method

**Participants.** Eleven undergraduates, all of whom had normal or corrected-to-normal vision and were between 18 and 22 years of age, participated for course credit.

**Stimuli.** The stimuli were 3-D rendered cross-shaped regions (see Figure 10) that could appear concave or convex depending on their rotation in the picture plane. The two images presented in a trial were either a sample–match pair (images identical) or a sample–foil pair (second image slightly distorted). Different sample images were created by generating cross shapes on the basis of four parameters: the major axis extents of the crossbars and their thicknesses (minor axis extents). Only the major axis extents were used as dependent variables. These parameters were obtained by randomly selecting one of five values for each parameter. The five values spanning the range from 2.5 cm (2.9°) to 6 cm (6.9°) formed the set of possible crossbar major axis extents, and the five values spanning 1.9 cm (2.2°) to 4.8 cm (5.5°) formed the set of possible crossbar minor axis extents. The program for generating the images included constraints that prevented the occurrence of combinations of these parameters that resulted in impossible cross shapes (e.g., the vertical extent of

the horizontal crossbar being greater than the height of the vertical crossbar). The images were rendered in MATLAB at a resolution of  $400 \times 400$  pixels, and the images were simulated to have light sources above and to the left of the figure.

Foil images were created by distorting these cross shapes. The magnitude of the distortions was defined by a parameter  $d$ , which was set to 0.15 after pilot testing. Both positive and negative values of  $d$  were used. The three distinct kinds of distortion described above were used (see Figure 10). For simplicity, the following descriptions mention only positive values of  $d$ .

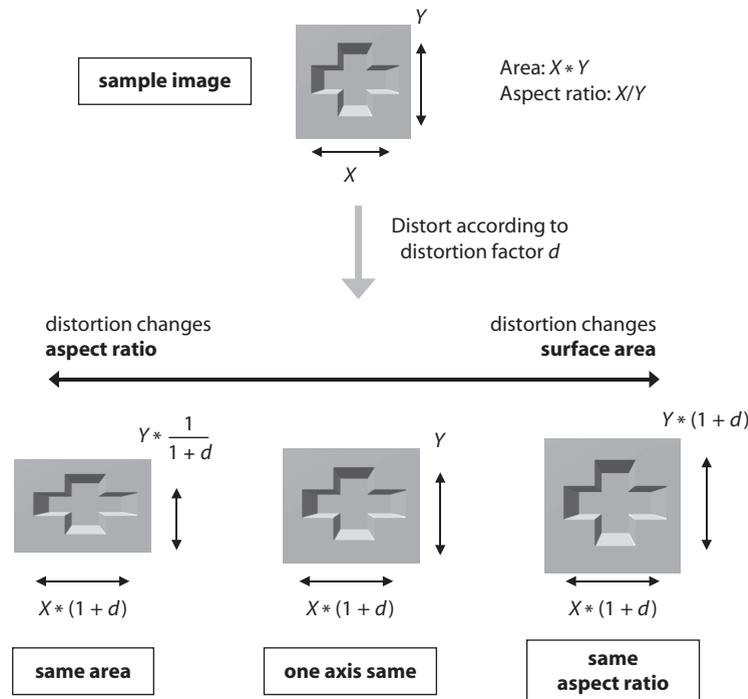
The same-area distortion preserved the 2-D surface area covered by the cross shape by stretching the shape along one axis and compressing it along the other axis. The identities of the stretched and compressed dimensions were counterbalanced, so that the horizontal and vertical extents were stretched and compressed equally often. Although this same-area distortion preserved the original surface area of the shape  $\{(1 + d) * [1/(1 + d)] = 1\}$ , the aspect ratio of the horizontal and vertical extents changed significantly, by a factor of  $(1 + d)/[1/(1 + d)] = (1 + d)^2$ .

The same-aspect-ratio distortion was the converse of same-area distortion: It preserved the aspect ratio of the horizontal and vertical extents of the cross shape by changing both axes by the same amount. Aspect ratio was unchanged,  $(1 + d)/(1 + d) = 1$ ; surface area was changed significantly, by a factor of  $(1 + d)^2$ .

The one-axis distortion was a condition that fell midway between the extremes of same-area and same-aspect-ratio distortion. By distorting only one axis, both the surface area  $[1 * (1 + d) = 1 + d]$  and aspect ratio  $[1/(1 + d)]$  were altered but to a smaller degree than with the other two distortions. The horizontal and vertical axes were manipulated as the distorted dimension equally often. Because the one-axis distortion left one axis completely unchanged, unlike the other two distortions, it provided a chance to test whether the participants were sensitive to the axes as independent dimensions. Our assumption was that if the observers perceived the two dimensions independently, similarity between samples and foils would be reckoned according to a city-block metric encompassing these two dimensions (Garner, 1974). All three kinds of distortions affect the city-block distance between sample and foil ( $|\text{width}_{\text{sample}} - \text{width}_{\text{foil}}| + |\text{height}_{\text{sample}} - \text{height}_{\text{foil}}|$ ), but the one-axis distortion produces the largest such difference. Simply put, the one-axis distortion should be hard to detect if the participants perceive the two axis extents independently, because half of the relevant dimensions are unchanged. The one-axis distortion is the key test for distinguishing whether observers made their judgments on the basis of the individual spatial extents of the shapes themselves instead of on higher order features constructed by integrating them.

Eighty unique trials were generated, comprising 40 *same* and 40 *different* trials. Each trial was presented twice during an experimental block, in random order, for a total of 160 trials. In the concave and convex blocks, the same set of trials was presented in different random orders. The noise mask that intervened between the presentations of the sample and match/foil images was a  $400 \times 400$  pixel image composed of 1,600 randomly colored  $10 \times 10$  pixel squares tiling the image.

**Procedure.** 3-D depth (concave or convex) was treated as a within-subjects factor. All stimuli within a block had the same 3-D depth, and the order in which the conditions were viewed was counterbalanced across participants. The participants performed 10 practice trials, followed by a block of 160 trials for each depth condition. Following a keypress, the sample image appeared on the computer monitor, centered on the screen against a gray background, and remained there for 1 sec. A noise mask of 500-msec duration replaced the sample. The match or foil image replaced the mask and remained on the screen until the participants responded using keyboard keys labeled “same” and “different.” After this keypress, the match/foil image was cleared from the screen and a feedback message (“Correct!” or “Incorrect!”—the latter accompanied by a beep) appeared for 1 sec. RTs and accuracy were recorded by the computer.



**Figure 10.** The three types of distortion used to create foils in Experiment 3. The cross shape that served as the sample image for a trial (labeled “sample image” at top) has horizontal extent  $X$  and vertical extent  $Y$ . The same-area distortion elongated one dimension and shortened the other, which changed aspect ratio  $X/Y$ , but preserved the surface area (proportional to  $X*Y$ ). The same-aspect-ratio distortion elongated the shape along both dimensions equally, which significantly increased surface area but preserved the shape’s aspect ratio. The one-axis distortion produced changes intermediate to these two extremes by stretching only one axis, moderately changing both surface area and aspect ratio. The one-axis distortion is the only case in which one of the shape’s original axis dimensions was kept constant. All of the above examples show concave images and positive values of  $d$ ; negative values of  $d$  were also used. The shapes’ square backgrounds have also been distorted for emphasis, although they were not distorted in the actual stimuli.

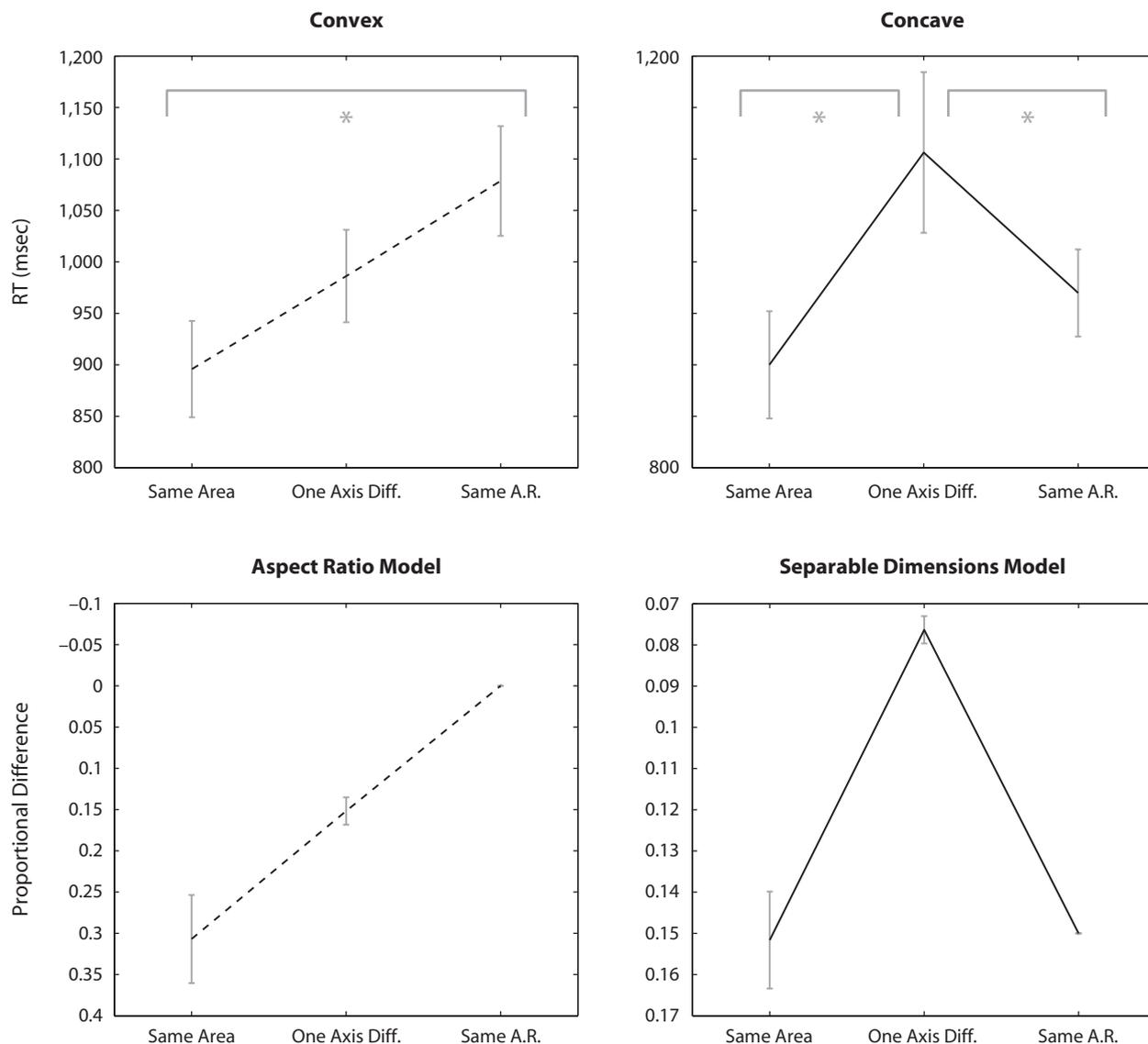
**Modeling.** A simple computational model was used to evaluate the data further in light of the hypothesis that the major axis extents of the cross shapes were either perceived by integrating these dimensions into the single property of aspect ratio (in the case of convexities) or perceived as separable dimensions (concavities). The difficulty of each *different* trial was modeled by calculating the dissimilarity between representations of the stimulus shapes. Dissimilarities were measured in terms of proportional difference (Weber fraction,  $\Delta X/X$ ). Aspect ratio (integral dimension) dissimilarities were thus computed with the formula  $|AR_{\text{sample}} - AR_{\text{foil}}|/AR_{\text{sample}}$ , whereas a city-block dissimilarity measure (separable dimensions model) used the formula  $(|width_{\text{sample}} - width_{\text{foil}}| + |height_{\text{sample}} - height_{\text{foil}}|)/(width_{\text{sample}} + height_{\text{sample}})$ . The changes in aspect ratio and the city-block distance measure that occurred in each experimental trial were measured and averaged with respect to the same factors used in the behavioral RT ANOVA.

## Results

Separate analyses were performed for the *same* and *different* trials, since the effects of the three types of foil distortion (same area, one axis, same aspect ratio) could only be compared in *different* trials. Repeated measures ANOVAs of the *same* trials revealed no significant effect

of 3-D depth with either RT data (concave,  $M = 984$  msec,  $SEM = 67$ ; convex,  $M = 991$  msec,  $SEM = 76$ ) [ $F(1,10) = 0.01$ , n.s.] or error data (concave,  $M = 14\%$ ,  $SEM = 3\%$ ; convex,  $M = 13\%$ ,  $SEM = 2\%$ ) [ $F(1,10) = 0.15$ , n.s.].

Repeated measures ANOVAs were performed on the RT and error data from the *different* trials, using two within-subjects factors: 3-D depth (two levels, concave and convex) and distortion type (three levels, same area, one axis, and same aspect ratio). The RT analysis (Figure 11) showed a significant effect of distortion type [ $F(2,10) = 7.17$ ,  $p < .005$ ] and an interaction between distortion type and 3-D depth [ $F(2,20) = 9.91$ ,  $p < .001$ ], but no main effect of 3-D depth [ $F(1,10) = 0.03$ , n.s.]. A post hoc Tukey’s HSD critical  $T$  value of 151 msec revealed that the mean RTs for the same-area distortion type ( $M = 898$  msec,  $SEM = 48$ ) were significantly faster than means for both the one-axis type ( $M = 1,046$  msec,  $SEM = 56$ ) and the same-aspect-ratio type ( $M = 1,024$  msec,  $SEM = 44$ ). The Tukey’s HSD critical  $T$  value of 114 msec for the distortion type  $\times$  3-D depth interaction indicates that the one-axis distortion type mean was significantly higher than



**Figure 11.** Top row: Mean of individual participants’ median reaction times (RTs) for Experiment 3. Error bars represent the standard error of the mean. Bottom row: Differences in modeled shape representations for the three distortion types based on representing shapes by aspect ratio (left) or by a city-block model in which differences were calculated by considering length and width as separable dimensions (right). Error bars represent the standard deviations for the set of stimulus dimensions used in the trials of the experiment. Ordinates for the model data have been inverted so that small shape differences compare to longer RTs in the behavioral data graphs.

both the same-area and same-aspect-ratio means in the concave data, but only the same-area and same-aspect-ratio means differed in the convex data. The only significant difference between the concave and convex groups was in the one-axis distortion, which had a higher mean with concave stimuli.

The error analysis showed a strong main effect of distortion type [ $F(2,20) = 12.99, p < .001$ ] but no main effect of 3-D depth [ $F(1,10) = 0.27, p = .61$ ] and no significant interactions. A Tukey HSD critical T value of 15.5% for the distortion type main effects revealed that the mean error rate for the same-area distortion type ( $M = 4.7%$ ) was significantly lower than means for both the one-axis type ( $M = 25.2%$ ) and the same-aspect-ratio type ( $M =$

21.8%). It is unclear why a distortion type effect, but not an effect of 3-D depth, was found in the error rate data. In any case, the patterns of the distortion type effects (i.e., collapsed across 3-D depth) closely correspond for both the RT and the error rate data.

The changes in aspect ratio and in the city-block distance measure that occurred during each experimental trial were measured and are plotted below the RT data in Figure 11. These two models reproduced the ordinal patterns of the RT results. To test how well the two models fit the concave and convex RT data, regressors based on the two models were used in a linear mixed effects regression. The model included binary indicator variables for the fixed effect depth (concave or convex), random

effects terms for the aspect ratio and city-block distance model values for each trial, and the interactions of depth with these two model terms. Individual participants were modeled as random factors using the linear mixed effects model function of the R statistics package.

The aspect ratio regressor was highly significant [ $t(1659) = -2.68, p < .01$ ], confirming that a greater difference in aspect ratio led to lower RTs. This effect did not interact with stimulus depth, as can be seen from the nonsignificant beta weight for the interaction between aspect ratio and concave depth [ $t(1659) = 0.57, p > .1$ ]. Concavity itself was not associated with higher RTs in the regression analysis [ $t(1659) = 0.57, p > .1$ ], nor was the city-block distance regressor significant [ $t(1659) = -0.99, p > .1$ ]. However, the concavity  $\times$  city-block interaction term was significant [ $t(1659) = -2.06, p < .05$ ], meaning that the observers were attuned to height and width as independent dimensions only when viewing concave stimuli.

### Discussion

The observers appear to have relied on different dimensions to judge the concave and convex shapes. The observers' performance with convex images closely followed the linear pattern across distortion types predicted by the aspect ratio model, indicating that this one integrated dimension accounts for convex shape perception very well. In contrast, concave performance followed the pattern of the city-block model, indicating that the observers were able to perceive the major axis extents separately. This V-shaped pattern of results indicates that no one integrated dimension, either aspect ratio or surface area, can account for the results.

The principle that concavities are perceived in terms of relatively independent features applies to global dimensions as well as to surface- or part-based ones. The observers appear to have used the major axis extents of 3-D concave regions to judge their shapes, rather than limiting their analysis to features (e.g., aspect ratio) that result from the integration of these one-dimensional spatial quantities. This qualitative difference between the perception of 3-D concavities and convexities is an important complement to the results of Experiments 1 and 2, which comprise mainly quantitative differences on measures of perceptual unity and holistic processing. It seems unlikely that asymmetries in the perception of depth from pictorial cues (e.g., the great depth perceived with pictorial convexities; B. Liu & Todd, 2004) could have produced the results of Experiment 3.

Two questions that were not addressed directly in this experiment deserve mention. First, the shapes were very simple and were designed so that their global structure could be captured by a minimal number of spatial parameters. It is unknown what limits there are to the number of independent spatial dimensions that observers might be able to use to judge the shapes of 3-D concavities and how their performance might change when there are many simple dimensions to attend simultaneously. Perhaps one would rely instead on a holistic mode of perception identical to that used to perceive 3-D convexities. Indeed, it

may be difficult for observers even to perceive complex 2-D regions as having concave depth. Nelson and Palmer (2001) reported that when planar holes described familiar and meaningful shape outlines, they were more likely to be perceived as solids rather than holes. The very compelling hollow face illusion (Hill & Bruce, 1993, 1994, 1996) can prevent even the structure of face masks with physical (not just apparent) 3-D concavity from being perceived properly. It seems likely that simply perceiving an image to be convex would be sufficient to produce the holistic perception effects seen here with convex stimuli, regardless of the actual structure of the distal stimulus. However, visual guidance of action is not affected by such illusions of convexity, as was shown recently (Króliczak, Heard, Goodale, & Gregory, 2006).

Performance in the same-area condition was superior to same-aspect-ratio performance with concavities, and the regression term for the aspect ratio model was significant with respect to concave stimuli. This indicates some sensitivity to aspect ratio akin to that found with convexities. There are at least two possible sources for sensitivity to aspect ratio with concavities. First, it could be that observers encode aspect ratio in addition to separable spatial dimensions, although they rely primarily on the latter. Alternatively, they could use the separably encoded dimensions themselves as input to construct a representation of aspect ratio, as a subsequent process in a serial chain. These alternatives may be weighed as possible solutions to the problem of how observers ultimately perceive phenomenologically similar figural shapes with concavities and convexities. The results of Experiment 3 do not answer this question directly but provide new data to constrain possible answers.

### GENERAL DISCUSSION

In this study, we found that observers analyze the local and global properties of 3-D concavities and convexities differently, in spite of the fact that these images portray the same figural shape. The results of three experiments suggest that there may be two pathways for perceiving figural shape in the human visual system, which operate differently on figures whose local elements are or are not grouped by overall 3-D convexity.

In Experiment 1, we examined how well participants could focus attention on one small region of an image without being influenced by features of other local parts. Performance with concave and convex stimuli was significantly different only when these regions were maximally distant from each other, and not in every case predicted by part parsing rules. This result is somewhat in contrast to the findings of Vecera and colleagues (Vecera et al., 2001; Vecera et al., 2000), who found that participants were more accurate at reporting two object attributes when they belonged to the same part than when they belonged to different parts of the same object. It is difficult to compare those studies with the present one, since they used markedly different stimuli (nonpolygonal stick figures) and tasks. Nonetheless, it appears that 3-D concavities and convexities differ in terms of the degree of global, object-

based attentional spreading (facilitation or inhibition due to properties of distant regions of the same figure), but not necessarily in terms of local, part-based attention.

In Experiments 2 and 3, we determined that a specific distinction between the processes for perceiving figural shape in 3-D concavities and convexities is the use of separable versus integral dimensions. The participants could ignore variation in task-irrelevant regions of 3-D concave shapes, whereas they showed significant signs of holistic shape perception when viewing convexities and ambiguously oriented 3-D images. Also, since the observers in Experiment 2 identified parts of the 2-D figural shape's outline equally well with concavities and convexities during the preinsertion block, the differences found between concavities and convexities in Experiment 1 were not likely the result of weaker figural percepts with concavities.

The results of Experiment 3 and the common perception of figural shape from 3-D concavities would seem to indicate that the final outcomes of figural shape perception in 3-D concavities and convexities are identical. However, in Experiment 3, we did not test that hypothesis exhaustively. In Experiment 3, we were able to assess subjective shape perception more directly than we could in Experiments 1 and 2, because the participants were explicitly instructed to judge entire shapes.

Experiment 3 showed that the subjective similarity between the perceived shapes of 3-D concavities and convexities is belied by qualitative differences in how observers perceive them. However, Experiment 3 assessed simple discrimination abilities and not subjective reports of perceived shape. There are many other conceivable aspects of recognition performance that could also distinguish the processes for perceiving shape with concavities and convexities, including, for example, viewpoint dependence of recognition (Biederman & Gerhardstein, 1993; Tarr, Williams, Hayward, & Gauthier, 1998) and the precision of categorization boundaries for different shapes (Siddiqi, Tresness, & Kimia, 1996). Furthermore, the cross shapes used in Experiment 3 were designed to limit the number of dimensions along which the exemplars varied. It will be necessary to investigate performance using complex concave shapes to determine whether such concavities, especially depictions of familiar objects, rely on the same separable dimension processes found here (and also whether the dimensions of their convex versions are integrated to the same extent). Shape cues may recruit different processes when they relate to familiar versus unfamiliar objects (Peterson, de Gelder, Rapcsak, Gerhardstein, & Bachoud-Levi, 2000; Peterson & Gibson, 1994; Peterson, Harvey, & Weidenbacher, 1991). It also appears to be difficult to perceive veridical concavity with complex shapes in the first place (Nelson & Palmer, 2001), especially for stimuli that evoke strongly holistic perception, as in the hollow face illusion (Króliczak et al., 2006). If a hollow face could ever be reliably perceived as concave, would observers perceive it less holistically, or even as a face?

This study allowed us to reject the hypothesis that observers rely on the same process to perceive 2-D shape from 3-D concavities and convexities, but it remains to be determined exactly how a subjective figural gestalt is

assembled from the separable elements that observers encode from concavities. We propose a hypothetical model of shape perception that is based on the existence of parallel processes for images that meet different perceptual grouping criteria. The model proposes that the local organization (e.g., structure of individual vertices and surfaces) and simple global organization (including discrimination of concave from convex regions) of both 3-D convexities and concavities are determined in early occipital areas. This assertion is in accordance with the research on shape-from-shading perception in agnosic patients (e.g., Humphrey, Symons, Herbert, & Goodale, 1996), which shows that even in the absence of object recognition abilities, patients can distinguish shading-defined concavities and convexities. This ability is almost certainly the result of activity in occipital areas (Humphrey et al., 1997; Lee, Yang, Romero, & Mumford, 2002), although dorsal visual regions are sensitive to shape-from-shading cues as well (Taira, Nose, Inoue, & Tsutsui, 2001).

The model proposes that only stimuli that meet certain criteria, notably global convexity, would eventually engage relatively anterior occipitotemporal areas for shape analysis (Farah, 1992; Goodale & Milner, 1992; Grill-Spector, Kourtzi, & Kanwisher, 2001; Kourtzi, Erb, Grodd, & Bühlhoff, 2003; Logothetis & Sheinberg, 1996; Malach et al., 1995; Tanaka, 1996). There, local image elements would undergo additional perceptual organization involving the integration of different intraobject spatial features. This would render the global aspects of the figure more salient and the local elements harder to perceive independently and would also render the specific metric dimensions of the figure less salient. 3-D concave regions, and possibly planar holes, fail to meet these criteria. Instead, their local features would be analyzed in a way that preserves the salience of their spatial properties (e.g., location, extent).

3-D concavities do meet at least one grouping criterion that indicates perceptual coherence above that of a disjoint collection of surfaces or of an image composed of multiple discrete objects: the presence of a closed boundary contour. Stimuli that fail to meet the convexity criteria but do meet this closure criterion could then form the input (in the form of the separable spatial dimensions of their component features) for distinct holistic shape perception processes, which would transform this input into figural shape percepts. In this way, two ultimately similar holistic shape perception processes could operate on either the output of lower level visual areas that meet reliable criteria for objecthood (convexity) on the output of higher level analysis of the spatial properties of features grouped by a closed contour. A good candidate locus for such processing is the dorsal visual pathway, which has been found to play an important role in perceiving spatial dimensions of objects such as position, size, and orientation (Goodale, Milner, Jakobson, & Carey, 1991; Ungerleider, Galkin, & Mishkin, 1983; Valyear, Culham, Sharif, Westwood, & Goodale, 2006). The idea that the perceptual grouping of image features is not a unitary process has been proposed in several domains related to figure-ground parsing and shape perception (Behrmann & Kimchi, 2003; Palmer, Brooks, & Nelson, 2003; Peterson et al., 1991). Our hypo-

thetical second process, which remains to be investigated, would thus constitute a second chance for perceiving figural shape from concave stimuli: an alternative way to recover valuable figural shape information from images that contain more global grouping cues than collections of distinct objects or unorganized features but that differ significantly from the image properties of solid objects.

In summary, this study found that the similar figural percepts observed with 3-D concavities and convexities are belied by very different processes for encoding their shape information. Observers perceive 3-D concavities to be composed of visual elements that are markedly more local than those of equivalent 3-D convexities and perceive the spatial extents of 3-D concavities as separable, not integral, dimensions. These results help to explain the paradox of how figural shape is perceived from holes and suggest that shape perception, in general, may consist of two distinct processes. These results motivate a hypothetical model of dichotomous shape perception processes that can be tested in future studies.

#### AUTHOR NOTE

The authors thank Roberta Klatzky and Carl Olson for their valuable guidance on this study. We also thank Joy Geng, Craig Haimson, John Philbeck, and Rachel Diana for thoughtful discussions and support, and three anonymous reviewers for their insightful and very helpful comments. A.D.C. was supported by a Department of Defense NDSEG fellowship, an APA Dissertation Research Award, and the Center for the Neural Basis of Cognition. M.B. was supported by Grant MH54246 from the National Institutes of Health. Correspondence concerning this article should be addressed to A. D. Cate, Human Cognitive Neurophysiology Laboratory, Veterans Administration, Northern California Health Care System, 150 Muir Rd., 151-I, P.O. Box 2339, Martinez, CA 94553 (e-mail: acate@ebire.org).

#### REFERENCES

- BARENHOLTZ, E., COHEN, E. H., FELDMAN, J., & SINGH, M. (2003). Detection of change in shape: An advantage for concavities. *Cognition*, **89**, 1-9.
- BARENHOLTZ, E., & FELDMAN, J. (2003). Visual comparisons within and between object parts: Evidence for a single-part superiority effect. *Vision Research*, **43**, 1655-1666.
- BAYLIS, G. C., & DRIVER, J. (1993). Visual attention and objects: Evidence for hierarchical coding of location. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 451-470.
- BEHRMANN, M., & KIMCHI, R. (2003). What does visual agnosia tell us about perceptual organization and its relationship to object perception? *Journal of Experimental Psychology: Human Perception & Performance*, **29**, 19-42.
- BERTAMINI, M., & CROUCHER, C. J. (2003). The shape of holes. *Cognition*, **87**, 33-54.
- BERTAMINI, M., & FARRANT, T. (2005). Detection of change in shape and its relation to part structure. *Acta Psychologica*, **120**, 35-54.
- BERTAMINI, M., & LAWSON, R. (2006). Visual search for a circular region perceived as a figure versus as a hole: Evidence of the importance of part structure. *Perception & Psychophysics*, **68**, 776-791.
- BIEDERMAN, I., & GERHARDSTEIN, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 1162-1182.
- BRAINARD, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, **10**, 433-436.
- CASATI, R., & VARZI, A. C. (1994). *Holes and other superficialities*. Cambridge, MA: MIT Press.
- COHEN, E. H., BARENHOLTZ, E., SINGH, M., & FELDMAN, J. (2005). What change detection tells us about the visual representation of shape. *Journal of Vision*, **5**(4, Art. 3), 313-321. doi:10.1167/5.4.3
- EGLY, R., DRIVER, J., & RAFAL, R. D. (1994). Shifting visual attention between objects and locations: Evidence from normal and parietal lesion subjects. *Journal of Experimental Psychology: General*, **123**, 161-177.
- ENNS, J. T., & RENSINK, R. A. (1990). Influence of scene-based properties on visual search. *Science*, **247**, 721-723.
- ENNS, J. T., & RENSINK, R. A. (1991). Preattentive recovery of three-dimensional orientation from line drawings. *Psychological Review*, **98**, 335-351.
- FARAH, M. J. (1992). Agnosia. *Current Opinion in Neurobiology*, **2**, 162-164.
- GARNER, W. R. (1974). *The processing of information and structure*. Potomac, MD: Erlbaum.
- GOODALE, M. A., & MILNER, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neurosciences*, **15**, 20-25.
- GOODALE, M. A., MILNER, A. D., JAKOBSON, L. S., & CAREY, D. P. (1991). A neurological dissociation between perceiving objects and grasping them. *Nature*, **349**, 154-156.
- GRILL-SPECTOR, K., KOURTZI, Z., & KANWISHER, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, **41**, 1409-1422.
- HILL, H., & BRUCE, V. (1993). Independent effects of lighting, orientation, and stereopsis on the hollow-face illusion. *Perception*, **22**, 887-897.
- HILL, H., & BRUCE, V. (1994). A comparison between the hollow-face and 'hollow-potato' illusions. *Perception*, **23**, 1335-1337.
- HILL, H., & BRUCE, V. (1996). Effects of lighting on the perception of facial surfaces. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 986-1004.
- HOFFMAN, D. D., & RICHARDS, W. A. (1984). Parts of recognition. *Cognition*, **18**, 65-96.
- HUMPHREY, G. K., GOODALE, M. A., BOWEN, C. V., GATI, J. S., VILIS, T., RUTT, B. K., & MENON, R. S. (1997). Differences in perceived shape from shading correlate with activity in early visual areas. *Current Biology*, **7**, 144-147.
- HUMPHREY, G. K., SYMONS, L. A., HERBERT, A. M., & GOODALE, M. A. (1996). A neurological dissociation between shape from shading and shape from edges. *Behavioral Brain Research*, **76**, 117-125.
- KANIZSA, G., & GERBINO, W. (1976). Convexity and symmetry in figure-ground organization. In M. Henle (Ed.), *Vision and artifact* (pp. 25-32). New York: Springer.
- KLEFFNER, D. A., & RAMACHANDRAN, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, **52**, 18-36.
- KOURTZI, Z., ERB, M., GRODD, W., & BÜLTHOFF, H. H. (2003). Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex*, **13**, 911-920.
- KRAMER, A. F., & JACOBSON, A. (1991). Perceptual organization and focused attention: The role of objects and proximity in visual processing. *Perception & Psychophysics*, **50**, 267-284.
- KRÓLICZAK, G., HEARD, P., GOODALE, M. A., & GREGORY, R. L. (2006). Dissociation of perception and action unmasked by the hollow-face illusion. *Brain Research*, **1080**, 9-16.
- LAMY, D., LEBER, A., & EGETH, H. E. (2004). Effects of task relevance and stimulus-driven salience in feature-search mode. *Journal of Experimental Psychology: Human Perception & Performance*, **30**, 1019-1031.
- LAVIE, N., & DRIVER, J. (1996). On the spatial extent of attention in object-based visual selection. *Perception & Psychophysics*, **58**, 1238-1251.
- LEDERMAN, S. J., KLATZKY, R. L., & REED, C. L. (1993). Constraints on haptic integration of spatially shared object dimensions. *Perception*, **22**, 723-743.
- LEE, T. S., YANG, C. F., ROMERO, R. D., & MUMFORD, D. (2002). Neural activity in early visual cortex reflects behavioral experience and higher-order perceptual saliency. *Nature Neuroscience*, **5**, 589-597.
- LIU, B., & TODD, J. T. (2004). Perceptual biases in the interpretation of 3-D shape from shading. *Vision Research*, **44**, 2135-2145.
- LIU, Z., JACOBS, D. W., & BASRI, R. (1999). The role of convexity in perceptual completion: Beyond good continuation. *Vision Research*, **39**, 4244-4257.
- LOGOTHETIS, N. K., & SHEINBERG, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, **19**, 577-621.
- MALACH, R., REPPAS, J. B., BENSON, R. R., KWONG, K. K., JIANG, H., KENNEDY, W. A., ET AL. (1995). Object-related activity revealed by

- functional magnetic resonance imaging in human occipital cortex. *Proceedings of the National Academy of Sciences*, **92**, 8135-8139.
- NELSON, R., & PALMER, S. E. (2001). Of holes and wholes: The perception of surrounded regions. *Perception*, **30**, 1213-1226.
- PALMER, S. E., BROOKS, J. L., & NELSON, R. (2003). When does grouping happen? *Acta Psychologica*, **114**, 311-330.
- PALMER, S. [E.], DAVIS, J., NELSON, R., & ROCK, I. (2008). Figure-ground effects on shape memory for objects versus holes. *Perception*, **37**, 1569-1586.
- PELLI, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, **10**, 437-442.
- PETERSON, M. A., DE GELDER, B., RAPCSAK, S. Z., GERHARDSTEIN, P. C., & BACHOUD-LÉVI, A. (2000). Object memory effects on figure assignment: Conscious object recognition is not necessary or sufficient. *Vision Research*, **40**, 1549-1567.
- PETERSON, M. A., & GIBSON, B. S. (1994). Object recognition contributions to figure-ground organization: Operations on outlines and subjective contours. *Perception & Psychophysics*, **56**, 551-564.
- PETERSON, M. A., HARVEY, E. M., & WEIDENBACHER, H. J. (1991). Shape recognition contributions to figure-ground reversal: Which route counts? *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 1075-1089.
- SIDDIQI, K., TRESNESS, K. J., & KIMIA, B. B. (1996). Parts of visual form: Psychophysical aspects. *Perception*, **25**, 399-424.
- SINGH, M., SEYRANIAN, G. D., & HOFFMAN, D. D. (1999). Parsing silhouettes: The short-cut rule. *Perception & Psychophysics*, **61**, 636-660.
- STEVENS, K. A., & BROOKES, A. (1988). The concave cusp as a determiner of figure-ground. *Perception*, **17**, 35-42.
- SUBIRANA-VILANOVA, J. B., & RICHARDS, W. (1996). Attentional frames, frame curves and figural boundaries: The inside/outside dilemma. *Vision Research*, **36**, 1493-1501.
- SUN, J., & PERONA, P. (1998). Where is the sun? *Nature Neuroscience*, **1**, 183-184.
- TAIRA, M., NOSE, I., INOUE, K., & TSUTSUI, K. (2001). Cortical areas related to attention to 3-D surface structures based on shading: An fMRI study. *NeuroImage*, **14**, 959-966.
- TANAKA, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, **19**, 109-139.
- TARR, M. J., WILLIAMS, P., HAYWARD, W. G., & GAUTHIER, I. (1998). Three-dimensional object recognition is viewpoint dependent. *Nature Neuroscience*, **1**, 275-277.
- THEEUWES, J., DE VRIES, G.-J., & GODIJN, R. (2003). Attentional and oculomotor capture with static singletons. *Perception & Psychophysics*, **65**, 735-746.
- TREISMAN, A., & SOUTHER, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, **114**, 285-310.
- UNGERLEIDER, L. G., GALKIN, T. W., & MISHKIN, M. (1983). Visuotopic organization of projections from striate cortex to inferior and lateral pulvinar in rhesus monkey. *Journal of Comparative Neurology*, **217**, 137-157.
- VALYEAR, K. F., CULHAM, J. C., SHARIF, N., WESTWOOD, D., & GOODALE, M. A. (2006). A double dissociation between sensitivity to changes in object identity and object orientation in the ventral and dorsal visual streams: A human fMRI study. *Neuropsychologia*, **44**, 218-228.
- VECERA, S. P., BEHRMANN, M., & FILAPEK, J. C. (2001). Attending to the parts of a single object: Part-based selection limitations. *Perception & Psychophysics*, **63**, 308-321.
- VECERA, S. P., BEHRMANN, M., & MCGOLDRICK, J. (2000). Selective attention to the parts of an object. *Psychonomic Bulletin & Review*, **7**, 301-308.

#### NOTES

1. Holes themselves were afforded special status as a second kind of part in Hoffman and Richards (1984). Although concave contour discontinuities indicate regions similar to ones in which two solid objects have intersected, convex contour discontinuities (like the cusp of a hole) indicate regions similar to ones in which a solid piece has been removed from a larger surface. Hoffman and Richards call this a *negative part*, and give it the same status as *positive (solid) parts* in their theory. However, they do not specify whether holes themselves can be divided into parts, nor what kinds of contours would form part boundaries in a hole.

2. The abrupt curvature changes at corners are actually discontinuous but correspond to maximal changes in contour orientation. Only the minima that fall on the sidewalls of the figure have been marked, since these are the ones that influence perception of the 2-D planar outline shape, which is the focus of this study.

3. The parts (arms, body) manipulated in Experiment 2 corresponded to the part structure appropriate for a convex or 2-D rendition of this shape; a concave version of this shape would not have part cuts at the arm-body intersection. In spite of this, the concave group participants were still able to restrict their attention to the arm region.

(Manuscript received April 3, 2009;  
revision accepted for publication August 18, 2009.)