

The role of part structure in the perceptual localization of a shape

Rutgers University has made this article freely available. Please share how this access benefits you.
Your story matters. [\[https://rucore.libraries.rutgers.edu/rutgers-lib/24053/story/\]](https://rucore.libraries.rutgers.edu/rutgers-lib/24053/story/)

This work is the **VERSION OF RECORD (VoR)**

This is the fixed version of an article made available by an organization that acts as a publisher by formally and exclusively declaring the article "published". If it is an "early release" article (formally identified as being published even before the compilation of a volume issue and assignment of associated metadata), it is citable via some permanent identifier(s), and final copy-editing, proof corrections, layout, and typesetting have been applied.

Citation to Publisher Denisova, Kristina, Singh, Manish & Kowler, Eileen. (2006). The role of part structure in the
Version: perceptual localization of a shape. *Perception* 35(8), 1073-1087.

Citation to *this* Version: Denisova, Kristina, Singh, Manish & Kowler, Eileen. (2006). The role of part structure in the
perceptual localization of a shape. *Perception* 35(8), 1073-1087. Retrieved from
[doi:10.7282/T3639N42](https://doi.org/10.7282/T3639N42).

Terms of Use: Copyright for scholarly resources published in RUcore is retained by the copyright holder. By virtue of its appearance in this open access medium, you are free to use this resource, with proper attribution, in educational and other non-commercial settings. Other uses, such as reproduction or republication, may require the permission of the copyright holder.

Article begins on next page

The role of part structure in the perceptual localization of a shape

Kristina Denisova, Manish Singh, Eileen Kowler[¶]

Department of Psychology, Rutgers University, 152 Frelinghuysen Road, Piscataway, NJ 08854, USA; e-mail: denisova@rci.rutgers.edu, manish@ruccs.rutgers.edu, kowler@rci.rutgers.edu

Received 18 April 2005, in revised form 22 August 2005; published online 9 June 2006

Abstract. The process of object localization may be accomplished with respect to a particular reference location, such as the center of gravity, COG (eg Vishwanath and Kowler, 2003 *Vision Research* **43** 1637–1653). Here, we investigated how part structure affects an object's reference location. The reference location was evaluated with a measure of the illusory displacement of an internal target element embedded within a larger object (Morgan et al, 1990 *Vision Research* **30** 1793–1810). To examine whether the reference location is different for shapes with part structure, two shapes were tested: circle (small and large; no part structure) and bell (shape with two parts, one larger than the other). Results were examined with respect to two predictions: either the location of an object is based on its shape as a whole, disregarding part structure (ie a single, overall COG), or the parts are processed separately (different COGs).

With the circles, the results showed a systematic illusory displacement of the internal target toward the COG. With the bell, the illusion was significantly weaker than with both circles—even though the main part of the bell had the same size as the small circle, and its horizontal axis had the same extent as the large circle. Moreover, the distance judgments for the bell were consistent with a (weaker) reference point being located at the COG of the larger part, rather than at the COG of the entire bell. These results show that the part structure of a shape plays a role in the representation of its location, and that for complex shapes the perceived location of an embedded element depends more on the parts within which it is embedded, rather than on the whole shape.

1 Introduction

The environment contains a multitude of objects of different shapes, and the visual system enables us to extract the necessary information from the surrounding environment for purposes such as navigation, avoiding predators, or finding lost items. In order to accomplish many of its tasks successfully, the visual system must first localize the relevant objects in space. Here, we investigate the effect of the shape of an object on the localization process.

Different attributes or aspects of an object can contribute to its perceived location. A reference point for assigning a location to an object is often the centroid of the shape (also known as center of gravity, COG) (Vishwanath and Kowler 2003, 2004; Watt and Morgan 1983a, 1983b; Westheimer and McKee 1977; Whitaker et al 1996, 2004). The visual system can estimate the COG by averaging the luminance distribution of a target shape, or by averaging its local contrast energy (Guez et al 1994; McGowan et al 1998; Melcher and Kowler 1999; Vishwanath and Kowler 2003).

Other factors can also play a role. In a task where observers had to aim a saccade to an eccentric cluster of dots, McGowan et al (1998) found that the saccades landed near the COG (average dot location). The small departures from the COG were accounted for by a model in which the COG computation is not always uniform throughout the shape, but may depend on the differentially weighted output of contrast detectors that were affected by the activity of their neighbors. Melcher and Kowler (1999), using different shapes as saccadic targets, found that saccadic landing position

[¶] Author to whom all correspondence should be addressed.

was based on a representation of the shape, regardless of the distribution of luminance within the shape.

Perceptual studies have shown that the reference location of an object may also be affected by visual attributes such as luminance peaks, salient features, or texture (eg Burbeck 1991; McGraw et al 2003; Vishwanath and Kowler 2003). There may also be interactions among these cues. For example, the perceived locations of the carrier and envelope of a Gabor patch may interact, with their relative influences depending on spatial separation (Whitaker et al 2004).

How is a reference location computed for shapes with multiple parts? Research on shape representation indicates that the visual system represents a complex shape not as an unstructured template, but as a structured hierarchy of parts. It segments complex shapes into component parts, and organizes shape representation in terms of these parts and their spatial relationships (Biederman 1987; Hoffman and Richards 1984; Hoffman and Singh 1997; Marr and Nishihara 1978; Palmer 1977; Singh and Hoffman 2001).

Part-based representations have the advantage of being robust to shape changes involving, for instance, the articulation of parts, because the shape of the object can be dissociated from the specific configuration its parts currently happen to have (consider the various configurations that a hand may take). Hoffman and Richards's (1984) *minima rule* postulates that the visual system uses tangent discontinuities and negative minima on a shape (points of locally highest curvature in concave regions) to segment it into parts. When two independent parts join to form a single object, their intersection generically produces a concave discontinuity. In turn, the visual system is able to detect and use such points of concave discontinuity and negative minima (smoothed versions of concave discontinuities) to parse an object into distinct parts. For example, while the letter E is a single object, it is also naturally seen as containing distinct parts in specific spatial relationships: a long vertical line, with three horizontal lines protruding from it on the right side. Part-based representations have been shown to explain a number of phenomena in visual shape perception, including the switch in perceived shape that accompanies figure-ground reversals (Baylis and Driver 2001; Driver and Baylis 1996; Hoffman and Singh 1997), object priming (Biederman and Cooper 1991; De Winter and Wagemans 2005), the perception of transparency (Singh and Hoffman 1998), the perception of shape orientation (Cohen and Singh 2006), and performance in comparing two probes along the bounding contour of a shape (Barenholtz and Feldman 2003).

Geometric variables that determine visual part structure may also influence the underlying averaging process for COG computation. For an object that has two parts, one large and one small, an averaging process may either compute the COG of the whole shape, treating the two parts homogeneously, or it may explicitly take into account the part structure of the shape, by computing the COGs of the two parts separately.

Vishwanath and Kowler (2003) studied perceptual and saccadic localization of shapes with different part structure: letters O and L. They found that observers located the perceptual centers of the stimuli very close to the COG of the whole shape. The salient feature of the letter L (the junction where two lines intersect) did not affect the localization of the perceptual center.

The process of averaging, or pooling, object information to encode its location may affect the accuracy of basic tasks, such as estimating the distance between elements located within or near other objects. For example, Burbeck and Hadden (1993) found that the presence of an additional flanking line positioned near two target lines increased the perceived separation between the two lines. Morgan et al (1990) reported an illusion of distance perception for a pair of target elements embedded within objects. Observers were asked to indicate whether the distance between two elements embedded within

circular clusters of dots was ‘smaller’ or ‘larger’ than a standard reference distance. Morgan et al found that observers’ estimates of inter-element distances were systematically biased by the distances between the cluster centers. The illusion was attributed to the effect of the centroid (average position) of each cluster within which the target elements were embedded (Morgan et al 1990), and it arises because observers are systematically biased to locate each element toward the centroid of the surrounding cluster.

1.1 *The present study*

The main goal of the present experiment was to investigate the nature of the illusion of distance perception reported by Morgan et al (1990) using shapes with and without part structure (figure 1). The presence of part structure may complicate distance perception because there are different potential COGs that could serve as reference locations: the COG of the whole shape, or the COGs of the individual parts.

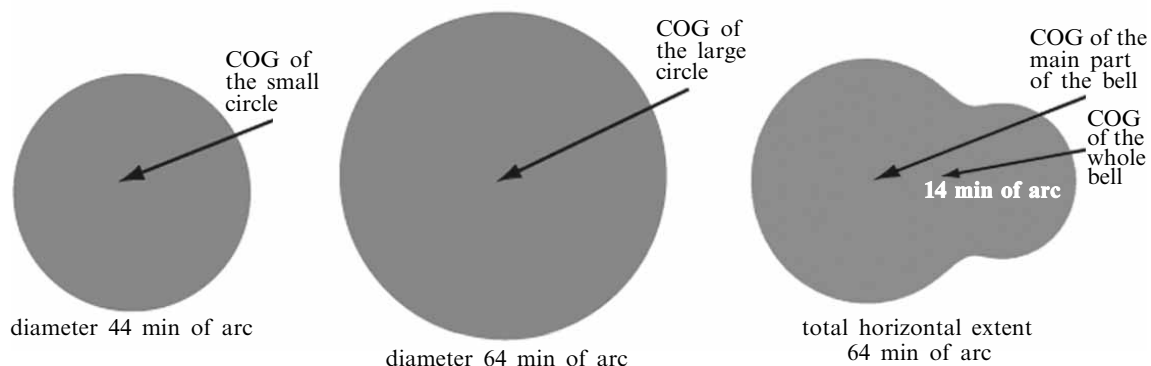


Figure 1. Example of experimental shapes: small circle, large circle, and the bell shape. The diameter of the small circle is 44 min of arc, the diameter of the large circle is 64 min of arc, and the total horizontal extent of the bell shape is 64 min of arc. Note that the experimental shapes were white ($3.5 \log$ units above foveal threshold), presented on a dim gray (0.14 cd m^{-2}) background.

The reference location of a shape depends on the pooling process involved. If the pooling is uniform across the shape, then different areas within the shape would contribute equally toward the computation of the reference location. Such a process is independent of structural aspects of the shape (eg part structure). On the other hand, the reference location may be determined by a non-uniform pooling process that assigns different weights to different areas within the shape. In this case, the reference location may be subject to local influences: in a shape with part structure it may be determined largely by the COG of one of the parts. Thus, in this study we investigated whether the illusion of Morgan et al (1990) is influenced by the COGs of the parts, rather than the COG of the whole object. Such an outcome would show that the perceived location of an object is not necessarily represented by a single reference location (the COG of the whole shape), but may be a product of a more complex pooling process that incorporates part structure.

2 Methods

2.1 *Observers*

Four Rutgers University students (LP, MR, JF, SA), paid volunteers, participated in the experiment. All were naive about the purposes of the study. All had normal or corrected-to-normal vision.

2.2 *Apparatus*

Stimuli were generated with MATLAB (Mathworks) and the Psychophysics Toolbox extensions (Brainard 1997; Pelli 1997). Stimuli were presented on a high-resolution 15 inch monitor of a Dell PC with a 75 MHz refresh rate. Viewing distance was 116 cm.

2.3 Stimuli

A sample sequence of display frames is shown in figure 2. Three consecutive frames are shown. In each, the stimuli were presented against a dark-gray background (0.14 cd m^{-2}). Frame 1 contained the fixation cross, frame 2 contained the critical display, frame 3 was a blank response frame.

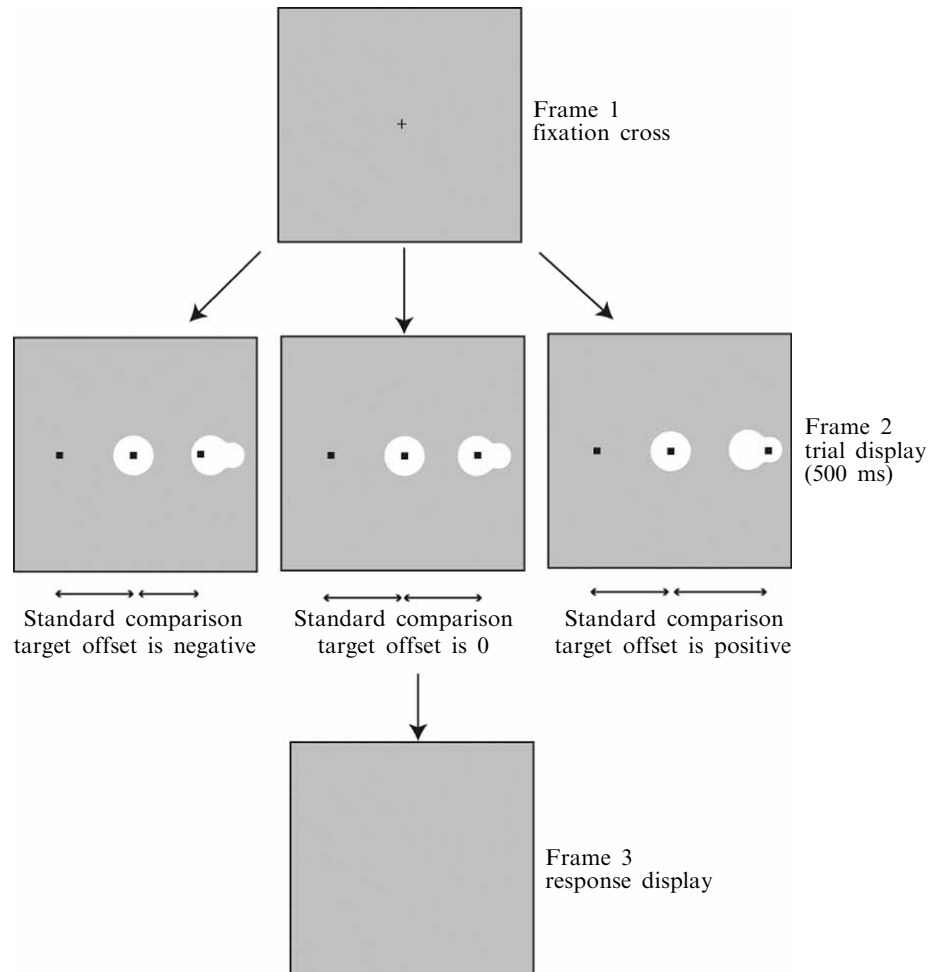


Figure 2. Sequence of frames. Frame 1 contains the fixation cross. Frame 2 is the critical frame containing target squares embedded in a bell shape. The offset of the target squares from the center of the shape could be negative, zero, or positive, where negative values represent offsets toward the central circle, and positive away from the central circle. Frame 3 is the response display. All displayed items were jittered (5% to the left or right) to prevent use of the monitor frame as reference. In a baseline condition no shapes were presented.

Figure 2 contains three examples of the critical display. Consider first the example shown in the center. The frame contains three small black test squares, 4 min of arc on a side. The central test square is located at the center of a white circle (44 min of arc diameter, 3.5 log units above foveal threshold). The right-hand test square is located within the critical shape. The critical shape was either the bell, or a circle (diameter 44 or 64 min of arc). The bell shape was created with two disks (diameter 44 and 28 min of arc) that overlapped by 8 min of arc. The junctions formed by the overlap were then smoothed by convolving the bounding contour of the resulting shape with a 1-D Gaussian (see, eg, Mokhtarian and Mackworth 1992) in the vicinity of the junctions. With the junctions smoothed, the bell was perceived as a single object with two parts (rather than an overlap of two separate objects). The total horizontal extent of the bell was 64 min of arc. The central test square was always located at the center

of the central circle. The distance between the left-hand test square and the central test square was denoted as the standard distance (100, 120, or 140 min of arc, randomly assigned). The distance between the right-hand square (located within the test shape) and the central test square was the comparison distance. The comparison distance resulted from the combination of two factors:

(i) Target offset position: the offset of the right-hand test square from the COG of the surrounding test shape (either the COG of the main part of the surrounding bell, or the COG of the surrounding circle). The three examples of critical frames in figure 2 show three different target offset positions: no offset (center), negative target offset (left), or positive target offset (right).

(ii) Shape separation: the distance between the COG of the critical shape (either the COG of the main part of the bell, or the COG of the circle) and the center of the central test square. The shape separation was defined as an increment on the standard distance, and could take values of ± 20 , ± 15 , ± 10 , ± 5 , or 0 min of arc (randomly assigned).

Other critical displays were generated that were the same as the ones described above, except that the positions of the flanking shapes were reversed so that the test shape (and hence the comparison distance) was on the left.

Psychometric functions were derived, representing the proportion of trials in which the comparison distance was judged larger than the standard distance, as a function of shape separation. Different psychometric functions were derived for the different target offset positions. This is analogous to the presentation of Morgan et al (1990).

A small jitter (5% to the left or right) was applied simultaneously to all displayed elements to discourage use of the monitor frame as a useful reference. A baseline condition was also tested, in which only the target squares were presented, without any surrounding shape.

2.4 Procedure

Observers viewed the displays binocularly with the head position constrained by a chin-rest. The sequence of events is shown in figure 2. At the beginning of each trial, a fixation cross was presented at the center of the monitor, and the observer was asked to look at the fixation cross and press any key when ready to start the trial. After the key-press, the critical display appeared (500 ms) followed by the blank frame. Observers had to respond whether the left-hand or right-hand distance between the small black targets was smaller.

The observer had as much time as needed to make the decision. The response was indicated by pressing one of two keys on the keyboard: Z, designated as the “smaller” response key if the distance between the two leftmost squares was smaller, or M, designated as the “smaller” response key if the distance between the two rightmost squares was smaller. No feedback was given. Once the observer provided a response, the fixation cross reappeared in preparation for the next trial. At any point during the experimental sessions, observers had the option to press Q on the keyboard to end the session.

The experiment was run in sessions of 100 trials. The different critical shapes and the baseline condition were tested in separate sessions. In half the sessions the standard was on the left and in half on the right. The order of testing of the critical shapes was randomized. Subjects LP, MR, and JF were tested in 150 sessions each; SA was tested in 300 sessions.

3 Results

3.1 Data analysis

For experimental conditions involving the test shapes (circle or bell) psychometric functions were generated for each target offset position (ie the offset of the target squares relative to the COG of the circles, or the COG of the main part of the bell).

For the baseline condition, psychometric functions were generated for target position values that were equivalent to those in the experimental conditions (for convenience, these values will also be referred to as the ‘target offset positions’ even though no surrounding shape was present).

At each target offset position, the corresponding PSE indicates the amount of displacement of the critical shape (circle or bell) needed to achieve the 50% point (threshold). Note that the displacement is measured relative to the standard distance, so that a displacement of zero means that the shape separation was equal to the standard distance.

3.2 Predictions

Figure 3 illustrates two possible experimental outcomes graphically for two different target offset positions (-18 and $+18$ min of arc). In the ‘veridical’ outcome (top panels) the comparison distance is physically equal to the standard distance. The target offsets (-18 or $+18$ min of arc) are compensated for by an equal displacement of the circle in the opposite direction. Thus, in this case, the PSE would be equal to either $+18$ min of arc (left panel) or -18 min of arc (right panel). A plot showing PSE as a function of the target offset position would thus have a slope of -1 .

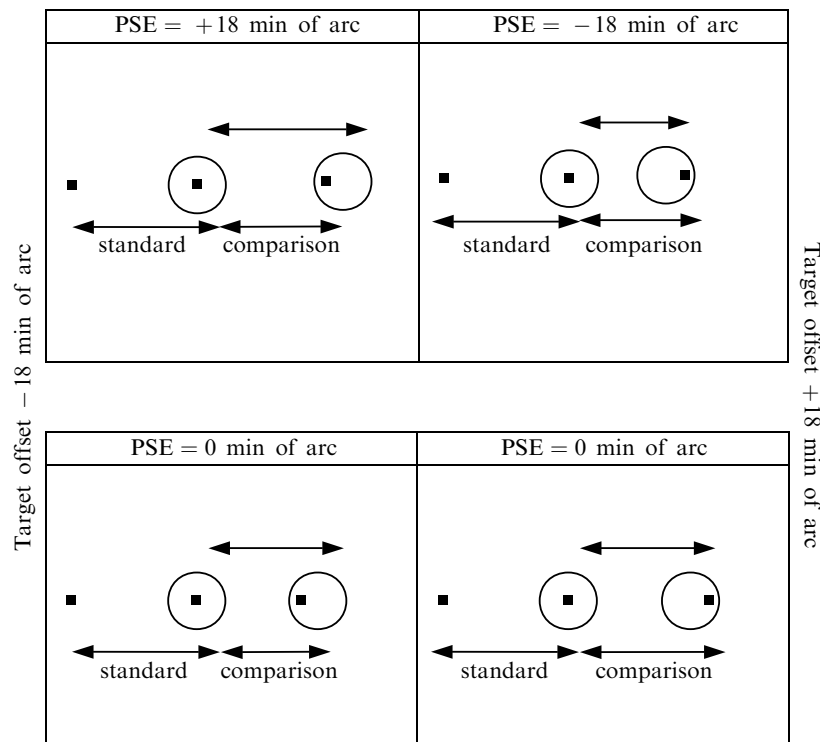


Figure 3. Illustration of possible outcomes (PSEs) for target offset positions -18 min of arc (left) and $+18$ min of arc (right) for the circle condition. Top panel shows veridical judgments and bottom panel shows illusory judgments. In each panel, left arrow below the shapes shows standard distance; right arrow shows comparison distance. The arrow above the shapes indicates distance between COGs of the shapes.

The bottom panels in figure 3 show cases where the shape separation (ie the distance between the centers of the circles) is equal to the standard distance, regardless of target offset position. This is the predicted outcome if the perceived separation of the target squares depends entirely on the COG of the surrounding shape. A plot of the PSE as a function of target offset position would thus be a horizontal line. Readers may be able to see the illusion in figure 3. In the bottom panel, distances between the squares appear about equal; however, one of the distances is considerably smaller than the other one.

3.3 Baseline

Figure 4 shows the PSEs obtained in the baseline condition, when the target squares were presented alone (ie not embedded within any shape), as a function of the target offset position. Linear regression (least-squares fit) was performed on the PSEs. Table 1 shows the corresponding slopes, intercepts, 95% confidence intervals, R^2 , F , and p values for each observer. The slopes of the linear fits for all observers were close to -1 . As expected, these data conform to the ‘veridical’ prediction: target position was the sole predictor of the distance judgments. These baseline data are useful for showing that observers were able to accurately compare the distances between the targets in the absence of any surrounding shape at all eccentricities tested.

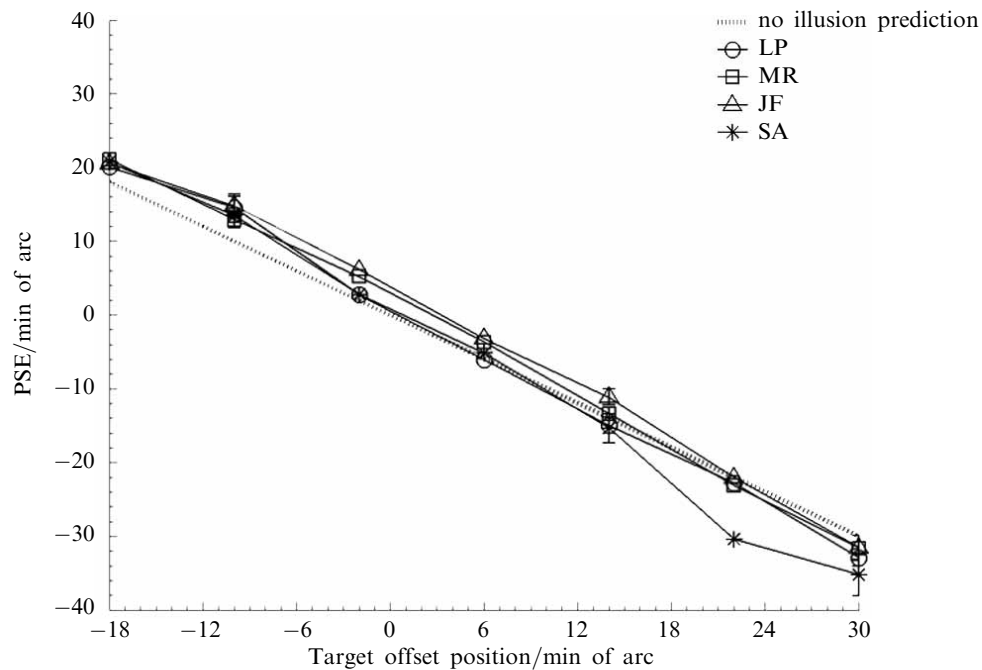


Figure 4. PSE as a function of the target offset position for baseline condition. Prediction of ‘no illusion’ illustrates a veridical judgment: at a given target offset position, the standard distance is seen to be equal to the comparison distance. The slope of this function is -1 . The error bars show ± 1 standard error.

3.4 Small (44 min of arc) and large (64 min of arc) circles

Figure 5 shows how the presence of a surrounding shape influenced the judgments. The PSEs for the small and large circles were relatively constant as a function of the target offset position. This result is consistent with the expected illusion, indicating that a surrounding circle had a large influence on the perceived distance between the targets. Specifically, the distance estimates were biased toward the separation between the circle centers. The slopes of the linear fits to the PSE data for the circles (shown in table 1) were small, but significantly different from 0, showing that the illusion was not perfect.

Figure 5 also shows that the PSEs were larger for the 64 min of arc diameter circle than for the 44 min of arc diameter circle. Table 1 shows that the intercepts of the linear fits to the large-circle data were about 9 min of arc larger than those for the small-circle data. The differences between the intercepts were significant for all observers. The elevated intercepts indicate that there was a tendency to underestimate the comparison distance when the target was in the large circle, regardless of target offset position.

Table 1. Linear regression analysis for target offsets -18 to $+18$ min of arc. CI refers to 95% confidence interval of slope and intercept; p -values < 0.05 show that the linear model is significantly better than the null model.

Observers	Slope, CI	Intercept, CI	R^2	F	p
Subject LP					
baseline	-1.12 (-1.20 to -1.04)	1.09 (-0.22 to 2.40)	0.99	1.41	< 0.01
small circle	-0.18 (-0.31 to -0.04)	3.79 (2.16 to 5.42)	0.69	11.33	< 0.02
large circle	-0.20 (-0.30 to -0.09)	9.98 (8.76 to 11.20)	0.83	24.60	< 0.01
bell	-0.52 (-0.62 to -0.42)	-4.05 (-5.23 to -2.86)	0.95	156.61	< 0.01
Subject MR					
baseline	-1.11 (-1.16 to -1.06)	2.00 (1.15 to 2.85)	0.90	3.30	< 0.01
small circle	-0.35 (-0.43 to -0.27)	1.21 (0.24 to 2.18)	0.96	127.75	< 0.01
large circle	-0.23 (-0.31 to -0.15)	8.86 (7.91 to 9.80)	0.92	56.30	< 0.01
bell	-0.52 (-0.62 to -0.44)	0.89 (-0.16 to 1.93)	0.97	209.13	< 0.01
Subject JF					
baseline	-1.10 (-1.19 to -1.01)	2.85 (1.31 to 4.39)	0.99	990.93	< 0.01
small circle	-0.31 (-0.41 to -0.22)	-7.05 (-8.18 to -5.92)	0.94	74.06	< 0.01
large circle	-0.38 (-0.51 to -0.25)	0.36 (-1.20 to 1.92)	0.92	56.32	< 0.01
bell	-0.54 (-0.77 to -0.33)	-10.26 (-12.87 to -7.66)	0.84	36.37	< 0.01
Subject SA					
baseline	-1.12 (-1.36 to -1.09)	0.31 (-2.02 to 2.64)	0.99	530.83	< 0.01
small circle	-0.31 (-0.53 to -0.09)	-5.04 (-7.70 to -2.39)	0.72	13.19	< 0.02
large circle	-0.26 (-0.40 to -0.13)	9.15 (7.52 to 10.79)	0.83	24.88	< 0.01
bell	-0.55 (-0.84 to -0.26)	-9.81 (-13.30 to -6.33)	0.74	20.38	< 0.01

3.5 Bell versus circles

The PSEs obtained for the bell shape are also shown in figure 5. Table 1 shows the slopes and intercepts of the best fitting lines to the data in figure 5, for target offset positions ranging between -18 and $+18$ min of arc. These positions are located within the main part of the bell. The slopes of the linear fits for the bell fell between those for the large and small circles and those for the baseline condition. Table 1 shows that the slopes for the bell were significantly different from those for both the small and large circles for all observers (see confidence intervals). These results show that the illusion (ie influence of the surrounding shape on the target) was significantly weaker with the bell shape than with either of the circles. It should be noted that the sizes of the two circles were selected such that the diameter of the small circle was the same as the diameter of the main part of the bell, and the diameter of the large circle matched the length of the horizontal axis of the bell (along which measurements were obtained). It is, therefore, particularly striking that the illusion for the bell did not resemble that for either of the two circles. This result thus shows that the structure of the surrounding shape had a strong influence on the perceived distance between the enclosed target elements.

The intercepts for the bell were similar to those for the small circle. Thus, the consistent underestimation of the comparison distance seen in the case of the large circle was not observed for the bell, even though it had the same horizontal extent as the large circle. This suggests that the bell was perceptually localized with respect to the COG of the main part rather than the COG of the whole bell.

3.6 Other locations in the bell

To examine the influence of the surrounding shape at other target offset positions, separate regression analyses were also performed for different ranges within the bell: around the COG of the whole bell (target offset range 9 to 18 min of arc), and within

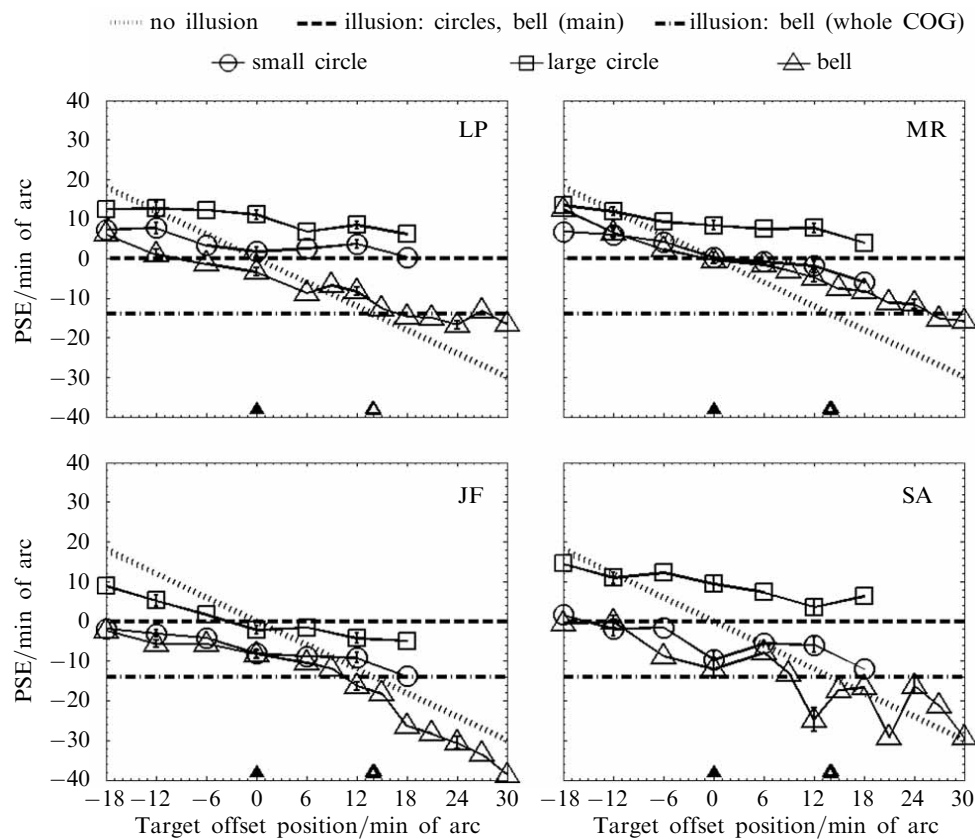


Figure 5. PSEs as a function of the target offset position for the three critical shapes: small circle, large circle, and bell. Filled triangle on the abscissa shows the locations of the COG of the circles, and the COG of the main part of the bell; open triangle at 14 min of arc shows the COG of the whole bell. Dashed and dotted lines show predicted outcomes. If there is no illusion the slope of the function will be -1 . The illusion predicts a slope of 0 , with intercepts determined by the COG of the shape (predicted intercept = 0 for the circles or the main part of the bell, and -14 min of arc for the COG of the whole bell). Standard errors are smaller than the plotting symbols.

the small part of the bell (target offset range 18 to 30 min of arc). Table 2 summarizes the results of these linear fits.

When targets were located near the COG of the whole bell, slopes were steeper than for those around the COG of the main part (table 1) for three observers (LP, MR, and JF). These differences were generally not large enough to reach significance; however, the direction of the results shows that the illusion, if anything, was weaker around the COG of the whole bell than around the COG of the main part.

For target offset positions within the small part of the bell, the observers differed. JF and MR showed a moderate slope (weak illusion), similar to other portions of the bell. LP showed a slope near zero, similar to her performance with the circle. SA's performance was quite variable at these locations. It should be noted that such individual differences were not obtained for targets located near the main part of the bell (table 1). Moreover, these differences are not due to retinal eccentricity, since performance was quite accurate at the same eccentricities in the baseline condition (see figure 4). The individual differences for targets within the small part of the bell suggest that the COG of the small part may not be as effective a reference location as the COG of the main part and, as a result, observers are forced to adopt different strategies to perform the task.

Table 2. Linear regression analysis for restricted ranges of target offsets shown in figure 5. COG of the whole bell: 14 min of arc. COG of the small part of the bell: 28 min of arc. Negative minima located at 18 min of arc. CI refers to 95% confidence intervals of slope and intercept. p -Values < 0.05 show that the linear model is significantly better than a null model.

Target offset range	Slope, CI	Intercept, CI	R^2	F	p
Subject LP					
9 to 18 min of arc	-0.93 (-1.45 to -0.41)	-1.91 (-5.35 to 9.16)	0.97	58.42	<0.02
18 to 30 min of arc	-0.07 (-0.64 to 0.49)	-13.49 (-27.30 to 0.31)	0.05	0.16	<0.71
Subject MR					
9 to 18 min of arc	-0.61 (-1.02 to -0.20)	2.32 (-3.37 to 8.01)	0.95	41.51	<0.02
18 to 30 min of arc	-0.63 (-0.93 to -0.32)	2.68 (-4.76 to 10.12)	0.93	42.77	<0.01
Subject JF					
9 to 18 min of arc	-1.51 (-2.71 to -0.32)	2.17 (-14.45 to 18.80)	0.94	29.59	0.03
18 to 30 min of arc	-1.00 (-1.40 to -0.61)	-7.43 (-17.09 to 2.23)	0.96	64.96	0.01
Subject SA					
9 to 18 min of arc	-0.10 (-3.94 to 3.74)	-16.64 (-70.08 to 36.79)	0.01	0.01	<0.92
18 to 30 min of arc	-0.59 (-2.84 to 1.67)	-8.56 (-63.58 to 46.45)	0.19	0.68	<0.47

4 Discussion

What aspects of an object’s geometry guide judgments of relative distance? The center of gravity (COG) of an object may be used as a natural reference location (Morgan et al 1990; Vishwanath and Kowler 2003). In the current study we used two shapes, a circle and a bell, to determine whether part structure influences the reference location. We evaluated the reference location using a measure of the illusory displacement of a target element within a larger object (Morgan et al 1990). Results indicated that there were illusory displacements of the target elements within all shapes. For the bell shape, however, the illusion was significantly diminished relative to the circles. Thus performance with the bell cannot be characterized simply in terms of a shift in reference point (although the results do suggest a shift of the reference location as well, relative to the bell’s overall COG). The results are considered with respect to three theories of distance judgment: Morgan et al (1990), Whitaker et al (1996, 2004), and Burbeck and Hadden (1993).

4.1 *The mechanism of Morgan et al (1990)*

Morgan et al (1990) proposed a mechanism composed of higher-order ‘eclectic units’ that determines the spatial position of objects and texture patches. Eclectic units are large, closely packed, and overlapping receptive fields (RFs). These RFs collect output from smaller, lower-order RFs that record the position of simple features in an image. The locations of relatively large objects are encoded by converging positional information from multiple lower-order RFs.

When small targets are embedded within larger objects, only the spatial position of the large object may be available, not the spatial positions of the internal targets. In the distance-judgment task, this produces a bias to report distances measured from the COGs of the objects, rather than the actual positions of the small embedded targets. Morgan et al (1990) attribute this bias to a two-stage, hierarchical encoding process (Morgan et al 1990; Watt 1988). In the first stage, positions of small target squares are encoded relative to the positions of the respective larger objects. In the second stage, the distance between the two small target squares is determined by computing the distance between the two larger objects, and the positions of each target square relative to its surrounding object. Indeed, it has been argued that such a hierarchical representational scheme is computationally efficient on information-theoretic grounds (Attneave 1954).

Biases may arise if the second step is not carried out fully. In that case, the relative positions of the target squares would be biased toward the positions of the corresponding larger objects (ie their centroids—Morgan et al 1990). Morgan et al (1990) noted that for larger-scale distance judgments, such as between two target squares embedded within different objects, only the COG of each of the large objects may be available.

4.2 *The mechanism of Whitaker et al (1996, 2004)*

Whitaker et al (1996, 2004) proposed a spatial filter mechanism that pools positional responses at different spatial scales. Whitaker et al (2004) examined the effects of modulating the carrier and envelope of a Gabor patch in an alignment task. Observers were asked to make one of two horizontal-position alignment judgments of a central Gabor patch, relative to two other, vertically flanking, Gabors. The two alignment judgments were: alignment of the envelope or alignment of the carrier. In the envelope-alignment task, the envelope was in physical alignment with the two flanking Gabors, while the carrier was misaligned, and vice versa. The magnitude of envelope and carrier modulations, as well as separation between the stimuli, varied across trials.

Two alignment biases were found: the envelope had the most influence on carrier alignment judgments at large stimulus separations, while the carrier had the most influence on envelope alignment judgments at small stimulus separations. Whitaker et al (2004) noted that observers' responses were influenced by the entire Gabor patch, even when the alignment task required them to selectively attend to either the carrier or the envelope of the Gabor. These biases were consistent with the use of the respective centroids of high-spatial-frequency and low-spatial-frequency defined modulations (Whitaker et al 1996, 2004). Whitaker et al (2004, page 285) pointed out that this may be due to “non-volitional, automatic combinatorial processes within the human visual system”. Whether high-spatial-frequency or low-spatial-frequency attributes affect the reference location depends on the relative separation between the stimuli. Note that this account predicts that Morgan et al's (1990) illusion would diminish at smaller separations.

4.3 *The theory of Burbeck and Hadden (1993)*

Burbeck and Hadden (1993) proposed a position-integration-area theory that accounts for how the presence of additional stimuli (eg an extra line) may affect the spatial scale used by the position-coding mechanism. They asked observers to make separation judgments between two target stimuli in successive intervals. One interval contained only the targets—two horizontal lines, one placed below the other—the other interval contained the targets, as well as an additional, flanking line (the ‘background line’). Observers were asked to report the interval in which the separation between the target lines was larger. Burbeck and Hadden (1993) found that adding a third, flanking line increased the perceived separation between targets. They proposed that adding the third line may have resulted in an increase in the position integration area used by the visual system; this in turn results in an increased perceived distance between the targets (Burbeck and Hadden 1993).

In the present study, the pattern of results was consistent with the previous findings described above in that observers could not veridically extract the position of a small target embedded within a surrounding shape in order to judge the distance between two such targets. Morgan et al (1990), Whitaker et al (1996, 2004), and Burbeck and Hadden (1993) would all predict such a failure of target localization. The comparison of the illusion across the small and large circles and the bell, however, allows us to investigate how the shape of the surrounding object, in particular its part structure, influences the distance judgments (and hence the localization of the object).

4.4 Slopes

If judgments of target separation are determined entirely by a single reference point within the surrounding shape (whether that reference point is the COG of the circle, the COG of the main part of the bell, or the COG of the whole bell), PSEs at all, or most, target offsets should shift toward that reference position. This would yield nearly flat slopes depicting PSEs as functions of target offset position. For the small and large circles, the slopes were in fact relatively flat, thereby indicating that the perceived distance between target points was systematically biased toward a fixed reference point within the circles.

Examination of the slopes for the bell showed that the illusion was different from that for circles. Targets within the main part of the bell (-18 to $+18$ min of arc, table 1) yielded an illusion that was significantly weaker than with the small circle at the same range of target offsets, even though the main part of the bell had the same size as the small circle. Moreover, the illusion for the bell was also weaker than for the large circle, even though the bell had the same extent along the horizontal axis (along which measurements were obtained) as the large circle. For target offset positions near the COG of the whole bell, there was, if anything, a smaller illusion than for targets near the COG of the main part. The significantly diminished illusion for the bell indicates that distance judgments involving the bell cannot be characterized simply in terms of a shift in reference location. Rather, it points to a different kind of localization process being used for multi-part shapes. [It should be noted that a simple shift in reference location would manifest itself as a shift in the intercept of the linear fits (see below), but without an accompanying change in slope.]

4.5 Intercepts

Examination of the intercepts of the linear fits (table 1 and figure 5) is informative about the location of the reference position used within the test shapes. For the small circle, the intercept is near 0, which is predicted if the reference location was near the COG of the circle. For the large circle, however, the intercepts were systematically larger. This indicates a consistent underestimation of the comparison distance, regardless of target offset. This underestimation may reflect the Baldwin illusion (Coren and Girgus 1978). In one of the variants of the illusion, a line is flanked by two shapes: a small square on one side and a large square on the other. When comparing the distance from the midpoint of the line to the two squares, observers report that the distance to the larger square appears smaller. Moreover, the magnitude of the illusion increases with the size difference between the two flanking shapes. Similar to our findings of different intercepts between the small and the large circles, there is a tendency to underestimate the distance to the larger shape.

This illusion may arise owing to the differences in distances between boundaries of the central and flanking circles, or the fact that greater area of the large circle fell nearer to the fovea. This increase in the area of improved resolution may have biased the perceived location of the large circle, if its location depended on a weighted COG.

These explanations are supported by the similarity of the intercepts for the bell and the small circle: for both of these shapes, the distance between the boundaries of the central circle and the flanking shape was the same. The similar intercepts of the bell and small circle also indicate that the reference point defining the location of these shapes was the same. This suggests that the bell was perceptually localized with respect to the COG of its main part, rather than the COG of the whole bell. (As revealed by the slope analysis, however, the COG of the main part was not as effective a reference point for the bell, as was the center of the circle for both the small and the large circle.)

4.6 *Localization of shapes with and without part structure*

The results above indicate that the differences between the bell and the circles arise from a difference in shape—in particular, part structure—rather than in spatial extent or eccentricity. Specifically, the two primary differences observed were: (i) a shift in the location of the reference point of the bell, away from its overall COG, and toward the COG of the main part; and (ii) a weaker illusion of displacement in the case of the bell.

The first result is consistent with a differentially weighted (rather than homogeneous) computation of COG, in which points within the smaller attached part are given a lower weighting, relative to points within the main part. Such an ‘averaging’ with unequal weights would result in the perceived center being shifted closer to the COG of the main part. Indeed, this differentially weighted scheme is consistent with recent findings in the perception of orientation of two-part shapes. Cohen and Singh (2006) found that observers’ orientation estimates of two-part shapes were well explained by a principal-axis computation in which points within the smaller part are assigned a lower weight than those within the larger ‘base’ part.

The second main result was a significantly diminished illusion for the bell, relative to both circles, indicating that the perceptual reference point for the bell was not as effective in localizing the entire bell, as was the center of the circle. There are at least three reasons to expect this. First, the circle has multiple symmetries and it does not contain separate parts over which integration must occur. This makes it likely that the reference point for the bell (based on a differently weighted part-based scheme) cannot be localized as precisely as the center of a circle (of equivalent or lesser extent). A greater uncertainty in the location of the reference point would in turn translate to a weaker illusion of target separation. Second, the presence of part structure introduces the need for a hierarchical representation of the shape of an object. Indeed, most computational models of shape representation incorporate a hierarchical component, with smaller parts nested hierarchically within larger ones (Kimia et al 1995; Marr and Nishihara 1978; Palmer 1977; Rom and Medioni 1993). A hierarchical representation entails that points within a shape are not represented directly with respect to the central reference location, but, instead, are represented indirectly—mediated by hierarchical relationship between the parts. Indeed, the large difference in slope obtained between the small circle and the main part of the bell (which had the same size as the small circle) suggests that the addition of an attached part (which alters the hierarchy of the shape) affects the representation of points within the main part as well. Finally, the mechanisms proposed by Morgan et al (1990) and Whitaker et al (1996, 2004) (see above) would also predict a weaker illusion for the bell. The presence of a smaller part would likely engage smaller spatial filters, closer in size to the target itself, thereby resulting in greater accuracy in localizing the targets—hence a weakened illusion.

Overall, these results underscore the complex nature of the object localization process (ie extraction of a reference location). Schemes involving a homogeneous ‘averaging’ to localize the COG of a shape work well for simple near-convex shapes, but are insufficient for complex multi-part shapes. The specific way in which the visual system computes a reference location for multi-part shapes, the precision with which such a reference point is localized, and the way in which the hierarchy of parts in the representation of a shape influences localization, remain important questions for future research.

Acknowledgments. Supported by the Air Force Office of Scientific Research, Grant AF 49620-02-1-0112, Life Sciences Directorate to Eileen Kowler, and by NSF, Grant BCS-0216944 to Manish Singh. We thank Jacob Feldman, and Brian Schnitzer for their comments and suggestions.

References

- Attneave F, 1954 “Some informational aspects of visual perception” *Psychological Review* **61** 183–193
- 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, 2001 “Shape coding in IT cells generalizes over contrast and mirror reversal, but not figure–ground reversal” *Nature Neuroscience* **4** 937–942
- Biederman I, 1987 “Recognition-by-components: A theory of human image understanding” *Psychological Review* **94** 115–147
- Biederman I, Cooper E E, 1991 “Priming contour-deleted images: evidence for intermediate representations in visual object recognition” *Cognitive Psychology* **23** 393–419
- Brainard D H, 1997 “The Psychophysics Toolbox” *Spatial Vision* **10** 433–436
- Burbeck C A, 1991 “Encoding spatial relation”, in *Vision and Visual Dysfunction* Ed. J Cronly-Dillon, volume 14 *Pattern Recognition by Man and Machine* Ed. R J Watt, chapter 2 (London: Macmillan) pp 8–18
- Burbeck C A, Hadden S, 1993 “Scaled position integration areas: accounting for Weber’s law for separation” *Journal of the Optical Society of America A* **10** 5–15
- Cohen E H, Singh M, 2006 “Perceived orientation of complex shape reveals graded part decomposition” *Journal of Vision* **6** in press
- Coren S, Girgus J S, 1978 *Seeing is Deceiving: The Psychology of Visual Illusions* (New York: John Wiley)
- Driver J, Baylis G C, 1996 “Edge-assignment and figure–ground segmentation in short-term visual matching” *Cognitive Psychology* **31** 248–306
- Guez J, Marchel P, Le Gargasson J, Grall Y, O’Regan J K, 1994 “Eye fixations near corners: evidence for a center of gravity calculation based on contrast, rather than luminance or curvature” *Vision Research* **34** 1625–1635
- Hoffman D D, Richards W A, 1984 “Parts of recognition” *Cognition* **18** 65–96
- Hoffman D D, Singh M, 1997 “Saliency of visual parts” *Cognition* **63** 29–78
- Kimia B B, Tannenbaum A R, Zucker S W, 1995 “Shapes, shocks, and deformations. 1. The components of two-dimensional shape and the reaction-diffusion space” *International Journal of Computer Vision* **15** 189–224
- McGowan J, Kowler E, Sharma A, Chubb C, 1998 “Saccadic localization of random dot targets” *Vision Research* **38** 895–909
- McGraw P V, Whitaker D, Badcock D R, Skillen J, 2003 “Neither here nor there: localizing conflicting visual attributes” *Journal of Vision* **3** 265–273
- Marr D, Nishihara H K, 1978 “Representation and recognition of three-dimensional shapes” *Proceedings of the Royal Society of London, Series B* **200** 269–294
- Melcher D, Kowler E, 1999 “Shape, surfaces and saccades” *Vision Research* **39** 2929–2946
- Mokhtarian F, Mackworth A, 1992 “A theory of multiscale, curvature-based shape representation for planar curves” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **14** 789–805
- Morgan M J, Hole G J, Glennerster A, 1990 “Biases and sensitivities in geometrical illusions” *Vision Research* **30** 1793–1810
- Palmer S E, 1977 “Hierarchical structure in perceptual representation” *Cognitive Psychology* **9** 441–474
- Pelli D G, 1997 “The Video Toolbox software for visual psychophysics: Transforming numbers into movies” *Spatial Vision* **10** 437–442
- Rom H, Medioni G, 1993 “Hierarchical decomposition and axial shape description” *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 823–839
- Singh M, Hoffman D D, 1998 “Part boundaries alter the perception of transparency” *Psychological Science* **9** 370–378
- Singh M, Hoffman D D, 2001 “Part-based representations of visual shape and implications for visual cognition”, in *From Fragments to Objects: Segmentation and Grouping in Vision* Eds T Shipley, P J Kellman *Advances in Psychology* volume 130 (New York: Elsevier Science) pp 401–459
- Vishwanath D, Kowler E, 2003 “Localization of shapes: Eye movements and perception compared” *Vision Research* **43** 1637–1653
- Vishwanath D, Kowler E, 2004 “Saccadic localization in the presence of cues to three-dimensional shape” *Journal of Vision* **4** 445–458
- Wagemans J, De Winter J, 2005 “Part-priming of object naming” *Perception* **34** Supplement, 127 (abstract)
- Watt R J, 1988 *Visual Processing: Computational, Psychophysical and Cognitive Research* (London: Lawrence Erlbaum Associates)

-
- Watt R J, Morgan M J, 1983a “Mechanisms responsible for the assessment of visual location: Theory and evidence” *Vision Research* **23** 97–109
- Watt R J, Morgan M J, 1983b “The recognition and representation of edge blur: Evidence for spatial primitives in vision” *Vision Research* **23** 1465–1477
- Westheimer G, McKee S P, 1977 “What prior uniocular processing is necessary for stereopsis?” *Investigative Ophthalmology & Visual Science* **18** 614–621
- Whitaker D, McGraw P V, Keeble D R T, Skillen J, 2004 “Pulling the other one: 1st and 2nd order visual information interact to determine perceived location” *Vision Research* **44** 279–286
- Whitaker D, McGraw P V, Pacey I, Barret B T, 1996 “Centroid analysis predicts visual localization of first- and second-order stimuli” *Vision Research* **36** 2957–2970

PERCEPTION

www.perceptionweb.com

This article is an advance online publication. It will not change in content under normal circumstances but will be given full volume, issue, and page numbers in the final PDF version, which will be made available shortly before production of the printed version.

Conditions of use. This article may be downloaded from the Perception website for personal research by members of subscribing organisations. Authors are entitled to distribute their own article (in printed form or by e-mail) to up to 50 people. This PDF may not be placed on any website (or other online distribution system) without permission of the publisher.