AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2023, Vol. 152, No. 5, 1502–1526
https://doi.org/10.1037/xge0001349

Independent Mechanisms for Processing Local Contour Features and Global Shape

Nicholas Baker and Philip J. Kellman
Department of Psychology, University of California, Los Angeles

The visual system can extract the global shape of an object from highly variable local contour features. We propose that there are separate systems for processing local and global shape. These systems are independent and process information differently. Global shape encoding accurately represents the form of low-frequency contour variations, whereas the local system encodes only summary statistics that describe typical features of high-frequency elements. In Experiments 1–4, we tested this hypothesis by obtaining same/different judgments for shapes that differed in local features, global features, or both. We found low sensitivity to changed local features that shared the same summary statistics, and no advantage in sensitivity for shapes that differed in both local and global features compared to shapes that differed only in global features. This sensitivity difference persisted when physical contour differences were equated and when shape feature sizes and exposure durations were increased. In Experiment 5, we compared sensitivity to sets of local contour features with matched or unmatched statistical properties. Sensitivity was higher for unmatched statistical properties than for properties sampled from the same statistical distribution. Experiment 6 directly tested our hypothesis of independent local and global systems using visual search. Search based on either local or global shape differences produced pop-out effects, but search for a target based on a conjunction of local and global differences required focal attention. These findings support the notion that separate mechanisms process local and global contour information and that the kinds of information these mechanisms encode are fundamentally different.

Public Significance Statement
In this paper, we explored the idea that shape information is encoded by two distinct mechanisms in the visual system. The human visual system appears to form precise spatial descriptions of objects' global shape information, which has to do with larger regions of an object's boundary. However, local information, involving smaller, higher frequency fluctuations of position along a contour, appears to be encoded, not with precise location information, but by a more compact statistical summary.

Keywords: Gestalt psychology, shape perception, form perception, local vs. global processing, global shape, object recognition

What is the identity of the object shown in Figure 1A? The question elicits two distinct but equally automatic responses. On the one hand, the object is clearly a cloud, as evidenced by its puffy white texture, its bright reflection of the sun, and its position in the sky. On the other hand, something about the shape of the cloud signals a horse in our perceptual systems. It would be simple to say that surface qualities and context support a cloud percept, while shape cues support a horse percept, but the contour features physically present in the image bear very little resemblance to the set of physical contour features typical in an actual horse. In Figure 1B and C, we find a horse in a similar pose to the cloud horse and compare the edge map (Canny, 1986) for the same local area in both images. The wisps and curls present in the cloud are visibly absent in a real horse.

Any correspondence between the cloud in Figure 1 and a real horse must be at a higher level of abstraction than the extraction of local contour features. As the Gestalt psychologists observed long ago, the representations we ultimately form of a shape are not a simple conjunction of the local elements present during sensation (Koffka, 1999). Wertheimer (1923/1938) outlined several principles by which distinct local elements could be organized to form a relational whole. Although similarity was among these cues, the physical characteristics of these individual elements mattered much less to perception of overall form than how they were arranged with respect to each other. Under these principles, two contours with very different local

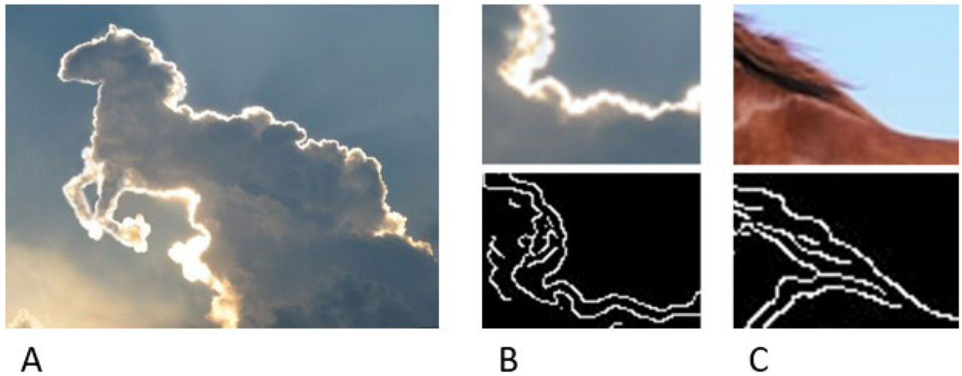
This article was published Online First March 13, 2023.
Nicholas Baker https://orcid.org/0000-0002-0673-2486
Study materials available by request to Nicholas Baker at Email: nbaker1@luc.edu. Data available at https://doi.org/10.17605/OSF.IO/V4E2H. Portions of this work were presented at the 2020 annual meeting of the Configural Processing Consortium and at the 2021 annual meeting of the Vision Sciences Society. Portions of this work also appeared in Nicholas Baker's doctoral thesis (2020). This study was not preregistered.

Nicholas Baker served as lead for formal analysis, investigation, software, writing–original draft. Philip Kellman served as lead for supervision and contributed equally to conceptualization, and served in a supporting role for writing–original draft. Nicholas Baker and Philip Kellman contributed to methodology, writing–review and editing equally.

Correspondence concerning this article should be addressed to Nicholas Baker, who is now at Department of Psychology, Loyola University, 1000 W Sheridan Road, Chicago, IL 60626, United States. Email: nbaker1@luc.edu

Figure 1

Example of Objects with Different Local Contour Features but Similar Global Shape



Note. (A) Cloud shaped like a horse. (B, C) Neck and shoulder blades for the cloud horse and a real horse, with corresponding edge maps. The edge maps were generated by MATLAB's implementation of the Canny edge detection algorithm. Cloud image reprinted with permission from https://www.flickr.com/photos/christigain/5636888777. See theonline article for the color version of this figure.

elements can be perceived as the same as long as the relations between their constituent elements are sufficiently similar.

There is a great deal of evidence from work on human perception that the visual system extracts global properties of shape before accessing features of individual elements. For example, when viewers are shown an S made up of small H's or vice versa, they first perceive the letter formed by the composition of elements and perceive the identity of the composing elements only later (Navon, 1977). Shape representations also appear to be insensitive to changes in positions of elements provided that the curvature of contour they define is preserved (Baker & Kellman, 2018). Research into the perception and recognition of line drawings has also found that simplified pictorial representations of shape are encoded more rapidly and accurately as the drawing's fidelity to a photographic image deteriorates (Biederman & Ju, 1988; Hochberg & Brooks, 1962). Such findings suggest that the visual system extracts abstract form based on global relations in the physical stimulus, abstracting away from local details.

These findings from human perception stand in stark contrast with some recent studies on object recognition in deep convolutional neural networks (DCNNs). Comparisons between DCNNs' and humans' use of shape in object recognition have shown a large divergence in each systems' sensitivity to local and global information. Baker et al. (2020) found that when silhouettes of objects were part-scrambled silhouettes such that many local contour features were preserved but global shape was destroyed, deep networks continued to classify the altered shapes with equal accuracy and confidence as the unscrambled original images, but when we disrupted local contour features by adding a serrated edge to the boundary of objects while preserving global form, networks' performance deteriorated to chance levels. By contrast, humans had difficulty classifying the part-scrambled images, but had no trouble with the images with changed local contour features and preserved global shape (Baker et al., 2018). When we introduced specialized training to bias networks towards global shape classification, networks learned to filter over larger variations in the contour but developed no sensitivity to an object's global shape (Baker et al., 2020). The large influence of local contour features on deep network classification suggests by contrast that humans have special capabilities for separating local and global shape information.

Many plausible models of shape description, however, are not well suited to handle variations in local contour features in a single system. For example, consider theories of part decomposition that separate objects into parts between local concavities (Barenholtz et al., 2003; Hoffman & Richards, 1984). Typical objects like a real horse will be organized into a set of parts that more or less aligns with a semantic part decomposition for legs, neck, body, and tail. For the cloud horse, small wisps and bumps around the outline create breaks in good continuation that would give rise to a very different (and more numerous) set of parts than would be seen in any real horse outline. In other work, we have proposed a theory of contour shape perception in which areas of similar curvature are represented by a single segment of constant curvature (Baker et al., 2021; Baker & Kellman, 2021). One issue in this approach concerns information at different scales. If curvatures are analyzed only with fine-scale detectors on the physical contour boundary, a constant curvature segmentation of shapes with high-frequency contour information would be made of many tiny primitives and would certainly not match up with a topographically similar shape in which the high-frequency contour features were omitted.

One way the visual system could separate local contour features from global shape is by including a constraint for simplicity when encoding a shape representation. Feldman and Singh (2006) implemented such a method in their Bayesian estimation for shape skeletons. In the classical medial axis transform (Blum, 1973), the skeletal representation of a shape is purely data-driven, and a small bump or protrusion along the contour will always be captured by an axial branch. The maximum *a posteriori* skeletal transform proposed by Feldman and Singh forces the algorithm to tradeoff between simplicity in curvature and number of branches (a prior) and fidelity to the original contour (a likelihood), resulting in a representation that aims to capture the essential topography of a shape without including all local variations to the bounding contour.

Research on shapes formed from radial frequency (RF) patterns has also lent important insight into how the visual system might encode global shape irrespective of variation in local contour features. RF patterns are sinusoidal modulations of differing amplitude and frequency that can be independently added to a circle to change its shape (see Figure 2). Wilkinson et al. (1998) added RF patterns to circles to evaluate perceivers' sensitivity to global (low RF) and local (high RF) contour features. They found that participants had accurate recognition for shapes with any RF of 6 or less. Recognition for shapes with radial frequencies greater than 6 was considerably more error-prone and deteriorated monotonically with larger RFs. Converging evidence for special sensitivity in low RF shapes was later found in a 2IFC task in which subjects discriminated an RF shape from a circle. Sensitivity to a difference from a circle was found to be better than predicted by local probability summation (i.e., the probability that participants were attending to a region modulated by the RF pattern during presentation) for RFs between 3 and 5, but not for high-frequency patterns (Loffler et al., 2003). The authors theorized that the visual system encodes shape as a combination of RF patterns and that only low-frequency RFs are ultimately used in our abstract representation of the shape.

Further evidence for separate global processing of shape was obtained by research on the discriminability between RF3 and RF6 patterns. Participants could distinguish the two patterns from each other better than would be predicted by probability summation when they could see more than one cycle of the RF pattern, but not when they were shown only a single RF pattern (Dickinson et al., 2013; Green et al., 2018). These findings suggest that a global feature, such as the angular separation between contour features was more useful for shape discrimination than individual, local contour features.

In a separate paradigm, Prins et al. (2007) introduced jagged edges to open contours to test the influence of local elements on the encoding of contour shape. They found that the visual system disregarded much of the local edge orientation information in their stimuli to encode a more global representation of the unclosed contour in question.

Clearly, the visual system has a robust capacity to extract a representation of overall shape that can support matching across variations in high-frequency local contour features that are inconsequential to the object's global topography. It would be a mistake, however, to think that local contour features are thrown out altogether. When we see a cloud that looks like a horse, we do not actually believe we are seeing a horse. This is true even if we remove all surface and context information and attend only to the object's bounding contour, as evidenced by our visual memory for puffs and wisps around the cloud horse's boundary in Figure 1. Any explanation for how we extract global shape from high-frequency local contours should also explain what information about local contour variation does get preserved in visual representations.

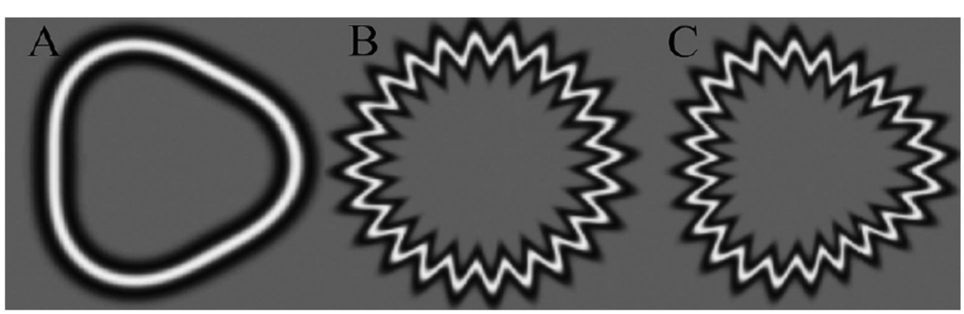
Distinguishing Global and Local Contour Shape Representations

In the current work, we propose a two-system theory of shape representation that includes separate mechanisms for local and global processing. The local processing system is responsible for encoding high-frequency variations along an object's contour, while the global processing system encodes low-frequency topographical information about the shape. These systems are distinguished in two ways. First, in our theory, these systems operate largely independently from each other such that changes to the high-frequency contour features of an object do not affect our representation of its overall shape, nor does our representation of the object's shape interact with our description of local contour features. Second, we posit that the two systems not only have very different levels of specificity but differ fundamentally in *what* they represent. While the global processing system appears to encode a description of the object's overall shape based on positions of contours in space, the local system encodes summary statistics about the local contour features, possibly estimating a distribution from which the contour variations were sampled. This second characteristic implies that the local system does little encoding of the specific locations of specific perturbations along the contour.

The idea that features are not represented individually but as a group has been proposed in other areas of object representation, such as what people choose to present in line drawings. When drawing a skyscraper, for example, one child chose not to individually draw every window, instead drawing a few accurately and writing "etcetera" for the rest (Arnheim, 1971). Kennedy (1974) found that both children and adults omit repetitive details in line drawings once a few detailed exemplars have been drawn that can be extrapolated to the others. One could well imagine that the information

Figure 2

Shapes Generated by the Addition of Radial Frequency Patterns (from Bell et al., 2007)



Note. (A) Circle deformed by the addition of a pattern with three cycles (RF3). (B) Circle deformed by the addition of an RF24 pattern. (C) Circled deformed by the addition of RF3 and RF24 patterns.

that is represented in visual memory uses similar simplifications to free up perceptual resources for other visual tasks.

Some evidence for independent systems has already been found in research on RF shapes showing that global shape is analyzed across features in low-frequency channels while local shape is analyzed by adding together information from individual features (Loffler et al., 2003). Research has also shown that the addition of low-frequency features does not affect participants' sensitivity to differences between a high-frequency contour and a circle or vice versa (Bell et al., 2007). Research on processing of different RF patterns has given many valuable insights about differences between local and global shape processing, but the statistics of contours defined in this way tend to be much more constrained than in typical object contours. For example, the bounding contour of a poodle or a pine tree has much more variety in its local contour variations than would be captured by a small number of RF patterns. For this reason, we developed a different system for generating high-frequency contour noise along an object's boundary that does not have the regularity of a circle modified by a relatively small number of sinusoidal modulations.

Overview of the Experiments

In Experiment 1, we showed participants two shapes sequentially and had them perform a forced-choice task to decide if the shapes were the same or different. We tested for independence between local and global shape processing mechanisms by generating shape pairs that differed in local contour features, in global contour features, or in both local and global contour features. We then assessed whether the inclusion of local and global contour differences conferred an advantage in detecting a difference in shape above what was conferred by one kind of difference. In Experiment 2, we once again used a forced-choice same/different task, this time controlling for the overall physical similarity between pairs of shapes that differed in local and global contour features. We tested whether observers have differential sensitivity to local and global contour differences when the physical similarity between the two conditions was matched. The results of Experiments 1 and 2 indicated that local contour features were represented much less precisely than global features in the visual brain.

In Experiments 3 and 4, we tested two alternative explanations for participants' poor sensitivity to local shape differences in Experiments 1 and 2. In Experiment 3, we tested whether the individual local features in Experiments 1 and 2 were too small for accurate detection by increasing the overall size of both the local and global features. In Experiment 4, we tested whether local features are robustly represented but require longer encoding time.

In Experiment 5, we tested the hypothesis that our descriptions of high-frequency shape are statistical rather than fully descriptive. We compared sensitivity to contour changes that were sampled from the same distribution from which the first set of contour features were sampled with sensitivity to changes when the new features were sampled from a different distribution. In Experiment 6, we sought direct evidence for independence between local and global shape processing. We used a visual search paradigm that both provided converging evidence for the separation of local and global systems and, following predictions from Feature Integration Theory, tested a different behavioral prediction made by our independent systems hypothesis.

Experiment 1

Experiment 1 tested sensitivity to changes in a contour that resulted in different global shape, in different local contour features, or different local and global contour features. We used a forced choice same/different paradigm using briefly displayed novel shape contours. Of particular interest to us were differences in sensitivity to global shape changes and to changes to both the local and global features of the shape. We expected that if both kinds of features were processed in a single perceptual system, there should be some additive effect on sensitivity when both features were changed. On the other hand, if they were processed by separate systems and one system dominated the shape recognition task, there might be no added benefit to sensitivity when the other feature was also changed.

Method

Participants

Twenty-four undergraduates (17 female, seven male, $M_{\text{age}} = 21.21$) from the University of California, Los Angeles participated in Experiment 1 for course credit. All participants had normal or corrected-to-normal vision. The first seven participants completed the study in the laboratory under controlled conditions, while the other 17 completed the study online through Pavlovia due to social distancing orders related to COVID-19. When analyzed separately, similar patterns of results were observed in both the online and in-person groups. All procedures completed by participants in this and all subsequent experiments were approved by the UCLA Office of the Human Research Protection Program.

Display and Apparatus

The participants tested in the laboratory were seated 70 cm from a 20-in. View Sonic Graphic Series G225f monitor. The monitor was set to 1,024 × 768 resolution, with a refresh rate of 100 Hz. Each square pixel of the display subtended 1.89′ of arc. Participants were seated in a dark room with no windows and their heads were kept in position by a chinrest.

For the online experiment, we instructed participants to sit a comfortable distance from the screen. Stimulus sizes were dynamically adjusted to cover the same proportion of the screen regardless of participants' display resolution.

Stimuli

All stimuli were shown as black outlines on a gray background. The stimulus was shown in the center of the screen, extending over an average of 37.5% of the horizontal space and 60% of the vertical space on the screen.

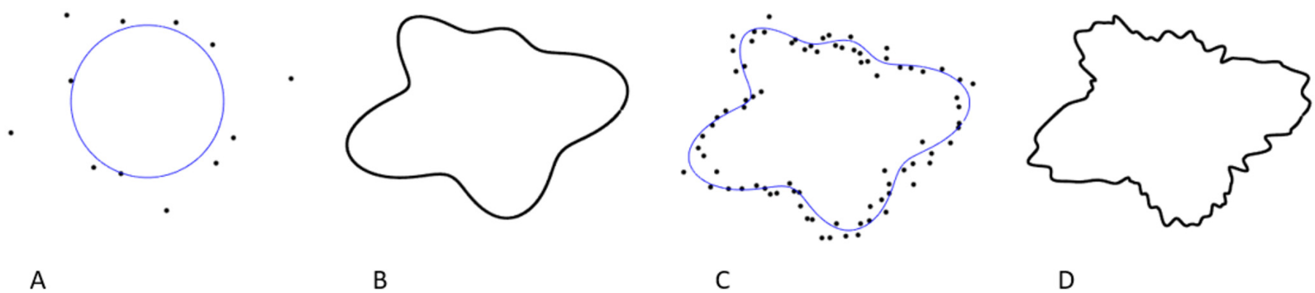
Experiment 1 included three conditions, with separate stimuli generated for each condition. In all conditions, a novel shape was generated by moving 12 control points toward or away from the center of a circle a random distance, then fitting cubic splines through the 12 control points in polar space.

In the Local change condition, we added contour features to the boundary by moving 80 control points on the shape boundary toward or away from its center. The average distance these control points were moved was 1/10th the distance they were moved when generating the global shape. We initially evenly spaced the control points along the boundary, then jittered some of them a small distance so that the spacing was not truly uniform. We then fit cubic splines through the 80 control points to create local contour features as we did for the 12 control points to create global contour features. A schematic for the formation of the stimuli is shown in Figure 3. To create different pairs in the local change condition, we simply inverted the direction of each control point so that bumps that extended away from the center reversed to extend toward the center and vice versa. This technique preserves as many statistical properties of the features as possible while still introducing a large amount of physical difference to the bounding contour.

In the Global change condition, we did not add any local contour features to the novel shape. We generated a different pair by moving one of the 12 control points a random distance between 7% and 17% percent of the total length of the contour toward or away from the shape's center. We then randomly selected a point adjacent to the one we moved and shifted it toward or away from the center whatever distance was needed so that the total length of the different display was the same as the original shape (see Baker & Kellman, 2018 for more detail).

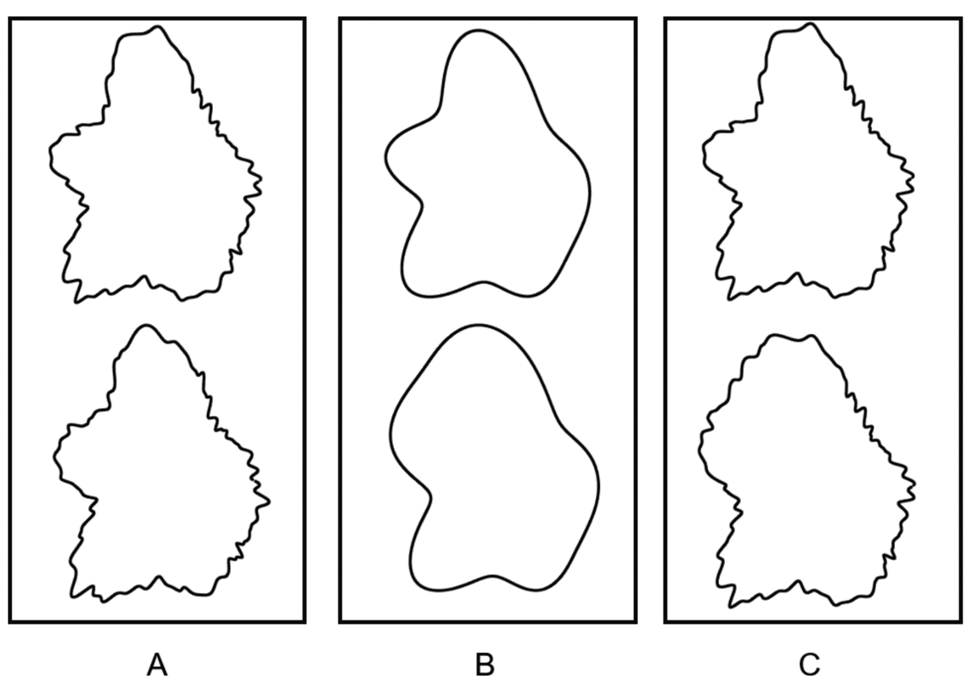
In the Combined local and global change condition, we generated pairs of shapes as in the global change condition, but we added local contour features to both members of the pair as in the local change condition. Figure 4 shows pairs of shapes for all three conditions.

Figure 3
Schematic for Global and Local Shape Generation



Note. (A) First, a circle is deformed by moving 12 control points away from the center. (B) Then, cubic splines are fit through the 12 points. This creates a shape with global features but no local features. (C) 80 control points along the shape's contour are moved toward or away from the center of the shape. (D) Cubic splines are fit through the 80 new control points. See the online article for the color version of this figure.

Figure 4
Stimulus Pairs for Experiment 1



Note. Left column: Shape pairs that differed in local features. Middle column: Shape pairs that differed in global features. Right column: Shape pairs that differed in both local and global features.

Design

The experiment consisted of three conditions, Local, Global, and Combined, with 80 trials per condition. In half of the trials for each condition, the same shape was shown in the first and second presentation. In the other half, different shapes were shown as described above. All trials were randomly interleaved. Participants received five practice trials with feedback before beginning the main experiment.

Procedure

In each trial, participants were first shown a fixation cross for 330 ms, after which the first shape was shown for 150 ms. Following the presentation of the first shape, a mask was displayed for 500 ms to block any apparent motion cues (Braddick, 1973) or access to a visual icon (Smithson & Mollon, 2006). The second shape was then shown for 1,000 ms. The second shape's orientation was always slightly different from the first, regardless of whether it was a same or different trial. We rotated the shape 10–30 degrees in a random direction. After the second shape had been shown, it was masked and we displayed a response screen in which participants were asked to decide if its outline was exactly the same as the first shape, irrespective of any orientation differences. Participants received no feedback during the main experiment. Sample trials are shown in Figure 5.

Dependent Measures and Data Analysis

We analyzed the results in terms of a signal detection theory (SDT) measure of sensitivity (d'), where a correct detection that the second shape was different from the first was counted as a hit and an incorrect response that the second outline was different was counted as a false alarm. The same trials were identical for the Local and the Combined conditions, so we combined them when computing false alarm rates.

Transparency and Openness

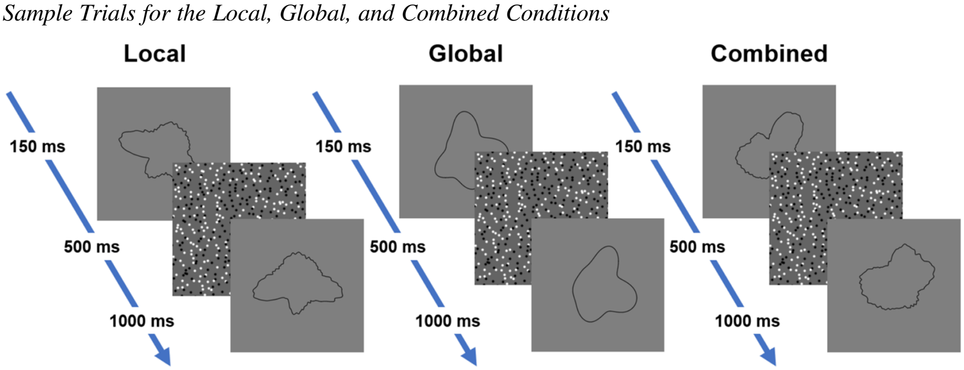
All data from this and other reported experiments is publicly available on OSF (DOI https://doi.org/10.17605/OSF.IO/V4E2H, https://osf.io/v4e2h/). All stimuli and code for generation is available by request to Nicholas Baker.

Results

Average sensitivity for each of the three conditions is shown in Figure 6. The results show a large difference in sensitivity between changes to the local contour features of the shape (Local condition) and changes to the global topography of the shape (Global condition). They also show no difference between the Global and Combined conditions. A one-way repeated measures ANOVA confirmed a main effect for condition type, $F(2, 46) = 26.32$, $p < .001$, $\eta^2_{partial} = 0.534$. Paired samples t-tests confirmed a significant difference in sensitivity between the Local condition and the Global condition, $t(23) = 9.67$, $p < .001$, Cohen's $d = 1.97$, 95% CI for difference = [0.761, 1.17], and between the Local condition and the Combined condition, $t(23) = 5.13$, $p < .001$, Cohen's $d = 1.05$, 95% CI for difference = [0.570, 1.33]. These effects remained significant after correcting for multiple comparisons. A one-sample t-test found sensitivity to local shape changes to be greater than zero, $t(23) = 6.83$, $p < .001$, Cohen's $d = 1.42$, 95% CI = [0.221, 0.411]. A paired samples t-test for the Global and the Combined conditions revealed no significant difference, $t(23) = 0.09$, $p = .92$, 95% CI for difference = [−0.313, 0.343]. Raw accuracies showed the same patterns, with similar performance in the Global and Combined conditions (71.4% and 71.4% correct, respectively) exceeding performance in the Local condition (55.0%).

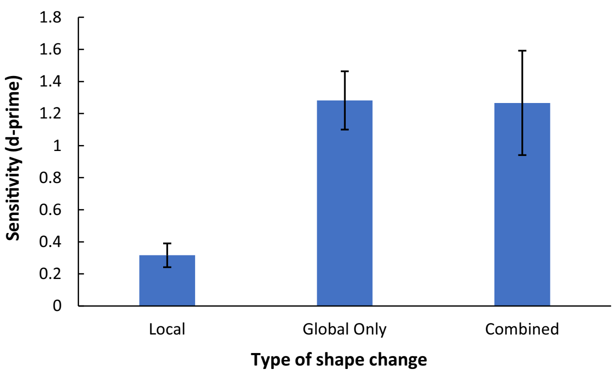
We also found that participants used a different criterion when judging shapes' local contour features. We calculated the criterion $\widehat{\lambda}_{center}$ as a measure of the bias to respond "same" or "different" relative to a neutral criterion midway between the means of the signal and noise distributions (for discussion, see Wickens, 2001). We reasoned that if local contour features were coded in a statistical fashion, then there should be a higher criterion (greater tendency to say "same") in the Local condition, where different pairs shared the same underlying contour statistics. The criterion was indeed reliably higher in the Local condition ($\widehat{\lambda}_{center} = 0.47$) than in the Global condition ($\widehat{\lambda}_{center} = 0.30$), $t(23) = 2.39$, $p < .02$, and also higher than in the Combined condition ($\widehat{\lambda}_{center} = -0.001$), $t(23) = 5.13$, $p < .001$. A more direct measure of the tendency to say "Same" to a different pair may be the *miss rate* (p["same" | different]) in each condition. The miss rate in the Local condition (0.61) reliably exceeded the miss rate of 0.37 in the Global condition, $t(23) = 6.79$, $p < .001$; and the Local condition also reliably exceeded the miss rate of 0.28 in the Combined condition, $t(23) = 5.32$, $p < .001$. The

Figure 5
Sample Trials for the Local, Global, and Combined Conditions



Note. See the online article for the color version of this figure.

Figure 6
Results from Experiment 1



Note. Sensitivity (d') to types of shape changes in the three experimental conditions is shown. Error bars show the 90% confidence interval of the means. See the online article for the color version of this figure.

Global condition miss rate also exceeded that of the Combined condition, $t(23) = 2.07$, $p < .05$.

Discussion

The results of Experiment 1 showed a clear difference in performance for trials in which the shape's global topography changed compared to trials in which the global form remained the same and only local contour features changed. Observers' sensitivity to change was markedly better in both conditions that included global shape change than when only local changes occurred.

The logic behind Experiment 1 is similar to previous work in the RF literature which gave evidence for the independence of local and global shape mechanisms (Bell et al., 2007). Many of these studies focus on discrimination of RF shapes near threshold from the circular. The noncircular shape deviates from the circular through regular sinusoidal modulations, although some previous work used other kinds of contour modulation (Dickinson et al., 2015). The results of Experiment 1 are consistent with these earlier findings and extend the pattern of separate local and global contour processing to irregular shapes whose features are generated by cubic spline fitting.

Our findings are also consistent with classification imaging work which shows that participants' perceptual decisions are influenced more by low-frequency signals than high-frequency signals (Kurki et al., 2014; Wilder et al., 2018). In addition to lending converging evidence for greater global shape sensitivity, our results extend findings in both the RF and classification imaging research to shape discrimination tasks. In our study, rather than tasking participants with distinguishing between a stimulus with shape and a "blank" stimulus (i.e., a circle or ellipse with no additional shape features), participants were tasked with detecting a difference between two related stimuli. Participants' much greater sensitivity to differences in the global shape than local shape suggests that they are not only better at detecting these features but are also better at using these features to compare and distinguish between similar shapes. These results converge with other work comparing shape discrimination in low and high RF patterns (Dickinson et al., 2013; Green et al., 2018).

Under our two-system theory of shape representation (Loffler et al., 2003), local contour features are not represented individually, but as a statistical distribution. The results of Experiment 1 furnish evidence for such a hypothesis in more than one way. First, sensitivity was poor in the Local condition. Note that for "different" pairs in the Local condition, local contour features, including their positions and curvatures, differed at essentially all locations. If local contour positions and orientations were encoded in detail, we would have expected relatively easy detection of differences for "different" pairs and high levels of discrimination performance. On the other hand, in this experiment, the statistics of the first and second shape were matched in terms of the mean and standard deviation of the amplitude of features as well as the frequency of features along the contour. Poor sensitivity to actual contour sameness or difference in the local condition suggests little encoding of precise contour features but is consistent with encoding of an ensemble of statistical properties of the distribution from which the features were sampled. Inspecting two contours that differed by local features in Figure 4 (leftmost column), it is easy to see, with simultaneous presentation, that the two are not the same. However, when representing the shape in visual memory, these differences appear to be abstracted over. The most common error in the Local condition was responding "same" to pairs that were actually different. This tendency was considerably greater than in the other conditions, as shown by reliable differences in criterion and by the 0.61 miss rate. The data are consistent with the idea that the sameness of statistical representation of local contour perturbations is often sufficient to produce perceived (or represented) sameness.

Whereas sensitivity was low for the local change condition, statistical analyses confirmed that participants had significantly higher than zero sensitivity to local feature changes. This result forces us to loosen somewhat our hypothesis about local contour features being described only as a few statistical parameters. A more plausible explanation is that people primarily encode statistical information about the local contour features, but they are able to encode a few individual contour features with greater specificity. Non-zero sensitivity in pairs with statistically matched local contour features would then be explained by the probability that participants detected a change in an attended local feature. This notion is somewhat similar to the idea of local probability summation from work with RF patterns (Bell et al., 2009; Loffler et al., 2003; Wilkinson et al., 1998). Our conjecture is that the encoding of individual features is effortful and done only when it would be directly helpful for a perceptual task at hand.

The other major result from Experiment 1 that supports our two-system theory is that participants showed the same sensitivity to changes in only the global features of a shape as they did to changes to both global and local features of the displays. This result also points to the idea that two displays with different local features but the same local feature statistics are difficult to perceive as different. The extremity of this result is somewhat surprising given participants' non-zero sensitivity to local contour features in the Local condition. In a simple summation model applied to the Combined condition, even if the probability of detecting a local change is low, it should still increase participants' overall sensitivity to a change in the Combined condition. For example, if the probability of detecting a local change is 30% and the probability of detecting a global change is 70%, then the probability of detecting either a local or a global change should be 79%. No such additivity was observed in Experiment 1.

One interpretation of this finding is that global information dominates local information when both cues are present. When participants need to detect a change based on only local features, they have some (albeit low) chance of success; however, they appear to fail to use this information when global differences are present. It is important to note that because the conditions were interleaved, participants did not know whether the second shape would have only local contour differences or local and global differences while forming a representation of the first shape. They must, then, have encoded what local information they could about the first shape in both conditions and discarded it only after seeing the second shape. If so, this provides some evidence that the visual system processes local contour features and global shape features relatively independently, only choosing which description to use in its perceptual decision once the second shape has been seen. The present results are consistent with a scheme in which the visual system first checks for different outputs in the global shape system and only compares outputs from the local contour system if no global difference is detected.

The evidence from Experiment 1 for independent local and global systems converges with previous evidence that local and global information are processed independently. Kimchi and Palmer (1985) used a Garner interference task (1974) to show that local contour elements composing a shape did not interfere with global form perception provided that the local elements were sufficiently small. In the literature on RF patterns, the results from Loffler et al. (2003) showing that only low-frequency signals benefit from global pooling also support the claim that local and global processing are fundamentally different mechanisms. The results from Bell et al. (2007) showing that adding a low-frequency (global) signal had no effect on a detection task for a high-frequency shape (and vice versa) directly support the notion that local and global processing are independent of each other.

The independence interpretation does not fully explain why there is not some performance gain from checking for local differences when the visual system does *not* detect a global difference in the Combined stimulus condition. The primary answer lies in the suggestion from the data that the encoding of local contour features does not generally occur; to a first approximation, only some statistical summary is represented. In Experiment 1, where local changes disrupted the local contour position at essentially every location, the use of local information for discrimination was still severely limited by the fact that the local variations were sampled from the same statistical distributions. Although performance in the local change condition is consistent with the occasional encoding of some specific local contour features, the low d' value associated with local features suggests that, if people responded in an unbiased way, they would report a local contour difference in 56% of trials. Against chance level performance of 50%, this suggests that accurate detection of a local contour difference occurred in only 6% of trials. There are a number of possible explanations for why observers might occasionally encode some local contour information apart from the statistical summary. We return to this issue in the discussion following Experiment 2 and the General Discussion. For the moment, the overall conclusion is that the performance in Experiment 1 was generally consistent with the global–local divergence of encoding and the idea that the local system apprehends a statistical summary. The results were generally inconsistent with any idea of local encoding that preserves much information about specific contour fluctuations and their locations apart from global shape.

Based on our theoretical framework, we would predict different results if two shapes' local contour features were sampled from *different* distributions. In that case, there should be both improved discrimination for pairs with instances differing in local contour characteristics as well as some evidence of additives effect in a combined local and global change condition. This prediction was tested in Experiment 3 below.

Experiment 2

In Experiment 1, we found that participants were less sensitive to changes in local contour features than they were to global shape. We interpreted this as a difference in the descriptive specificity of independent local and global shape processing systems. An alternative explanation is that the differences between pairs in the global condition were simply larger than differences in the local condition and were therefore easier to detect. In Experiment 2, we tested this alternative explanation by equating the overall physical dissimilarity between local and global shape changes and comparing subjects' sensitivity to each. We predicted that if the two systems were distinct and represented information differently, then subjects' sensitivity to global shape changes should still be higher than their sensitivity to local contour feature changes, even if the overall physical dissimilarity was the same.

Method

Participants

Eighteen undergraduates from the University of California, Los Angeles (12 female, five male, $M_{\text{age}} = 21.5$) participated in Experiment 2 for course credit. Sixteen of the participants in Experiment 2 also participated in Experiment 1. All participants had normal or corrected-to-normal vision and completed the study online through Pavlovia.

Display and Apparatus

Since the experiment was conducted online, display conditions varied slightly from participant to participant. Subjects were instructed to sit at a comfortable distance from the screen, and stimulus sizes were adjusted to cover the same proportion of the screen regardless of participants' display resolution. We allowed these variations for obvious practical reasons during the Covid-19 pandemic, but also because the perceptual abilities under study here should be robust across a range of ordinary screen sizes and viewing distances.

Stimuli

All stimuli were shown as black outlines on a gray background. Each stimulus was shown in the center of the screen, extending over about 37.5% of the horizontal space and 60% of the vertical space on the screen.

Pairs of locally and globally different shapes were generated as in Experiment 1. The only difference in how shapes were generated was that we reduced the amount the control point was shifted in the Global condition from 7% to 17% in Experiment 1 to 2.7% to 6.7% in Experiment 2. This was done to better equate the physical contour difference between local and global shape pairs. To that end, we also used a measure of physical contour similarity to ensure that the shape pairs with local contour differences were as dissimilar from each other as the shape pairs with global contour differences. Similarity was measured as the ratio of the overlapping areas to the non-overlapping areas for both contours. The equation we used was as follows (Baker & Kellman, 2021):

$$\frac{\frac{\text{Shape 1 and 2 overlap}}{\text{Total area of Shape 1}} + \frac{\text{Shape 1 and 2 overlap}}{\text{Total area of Shape 2}}}{2}$$

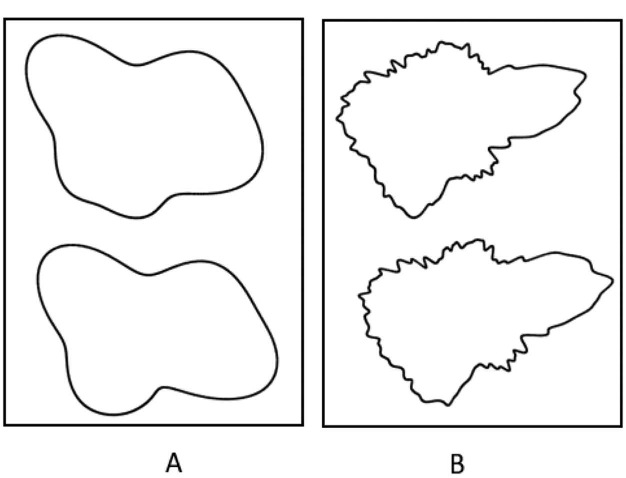
We generated 120 shape pairs for each of the two conditions. The average total shape difference was 4.52% for pairs of shapes in the Global change condition and 4.62% for pairs in the Local change condition. Sample shape pairs for Experiment 2 are shown in Figure 7.

In addition to controlling for the physical similarity between pairs of locally and globally different shapes, we used ShapeComp, a measure of perceptual similarity developed by Morgenstern and colleagues, which uses a variety of shape representation models to estimate the perceived dissimilarity between shapes across multiple dimensions (Morgenstern et al., 2021). We found that in terms of Morgenstern et al.'s estimate of perceived similarity, the average Euclidean distance between local pairs was 1.42 times greater than the distance between global pairs (22.56 vs. 15.79). If participants still have more trouble distinguishing between members of local pairs, this would support the hypothesis of statistical encoding, rather than precise spatial localization, of local contour features.

Design

Experiment 2 had two conditions: a Local condition in which shapes differed in local features, and a Global condition, in which shapes differed in global features. There were 120 trials in each condition. Local and global trials were randomly interleaved. There were five practice trials with feedback before the main experiment began.

Figure 7
Shape Pairs for Experiment 2



Note. Left column: Pairs of shapes that differed in global features (A). Right column: Pairs of shapes that differed in local features (B).

Procedure

The procedure for Experiment 2 was the same as in Experiment 1 except that there were no conditions in which both local and global contour features changed. For trials in which local contour features varied, both instances in a stimulus pair always had the same overall shape and could differ only in local contour features. Likewise, if the first stimulus had no local contour features, the second stimulus could differ only in global shape properties.

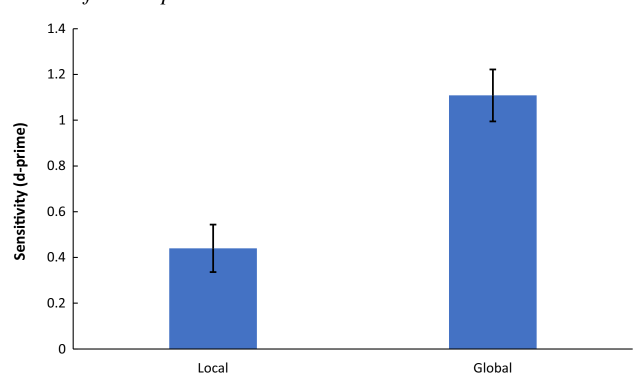
Results

As in Experiment 1, the results were analyzed in terms of sensitivity, where a hit was a correct detection of a shape change, and a false alarm was an incorrect report of a change in shape. Sensitivity to both conditions is shown in Figure 8. A paired samples *t*-test confirmed a significant difference in sensitivity to global shape differences vs. local shape differences, $t(17) = 7.15$, $p < .001$, Cohen's $d = 1.69$, 90% CI for difference = [0.51, 0.82].

Discussion

In Experiment 2, we once again compared sensitivity to local and global shape changes. We created pairs of shapes that differed in either local contour features or global form and equated the physical similarity between pairs in both conditions. These were tested in a sequential forced-choice same/different paradigm. Despite having equally different physical contours overall, pairs in the local change condition were significantly less discriminable from each other than pairs in the global change condition. These results clarify our interpretation of the results in Experiment 1. Participants' lower sensitivity to local shape changes cannot be explained by a smaller amount of overall physical contour difference, but by a perceptual difference in how local features and global features are encoded. We would expect that, if one system processed both local and global contour information in the same way, then sensitivity to a difference in shape should depend only on the physical similarity of the two shapes. On the other hand, if local and global contour features are processed by distinct systems, the kind of information encoded in

Figure 8
Results from Experiment 2



Note. Error bars show the 90% confidence interval for each condition. See the online article for the color version of this figure.

one system might be more discriminative between shape pairs than the other.

The outputs of ShapeComp (Morgenstern et al., 2021) predicted that the locally different shape pairs should be more discriminable from each other than the globally different shape pairs. The data from Experiment 2 suggest the opposite pattern: sensitivity is greater for globally different pairs than locally different pairs. This suggests that some aspect of shape representation is driving performance in our stimuli that is not fully explained by the representational dimensions included in ShapeComp. We believe that the statistical nature of the visual system's representation of local contour features is responsible for this difference. While both the physical dissimilarity measures we employed and ShapeComp are influenced by variation in individual local contour features, we hypothesize that only variations that change the *statistical* description of local features will influence the detection of local contour changes.

Still, the comparison based on overall physical difference is not a perfect one. The kinds of differences that local and global shape changes produce on the shape contour are necessarily different from each other. While a change to global shape typically involves a relatively large perturbation of the contour in a single region, changes to local shape are individually smaller and distributed over a larger percentage of the object's boundary. One reasonable question is whether local changes of the magnitude tested here pose difficulties for detection. Local changes as tested here were designed to be readily apparent, but there are various hypotheses relating to the probability of detection that could be raised here. The following considerations are relevant to some of them. First, as Figure 7 shows, when viewed simultaneously, the local contour features for pairs of shapes were visibly different from each other. This suggests that participants' worse performance for local change is not because the differences are below some threshold of detectability. Second, it might be argued that each local difference is harder to detect than any global shape difference in these experiments. Whereas that is likely the case, it is important to note that in our experiments, every pair of displays that differed by local features did not differ by one such difference, but differed *everywhere* in local contour positions and variations. The relevant question may not be whether each local difference is as salient as a local one, but how pervasive local differences may be missed, leading to the erroneous classification of two displays as "same." Third, a different idea is that numerous local changes may strain the capacity of a shape representation system. Although that is true, accurate discrimination between locally different pairs did not require the representation of all local shape positions and features; in every case, representing even *one* featural difference would have been sufficient. Notwithstanding these considerations, it remains possible that some combination of the smaller size of local feature changes, observers' strategies, and the time limitations of our task produced issues of detectability, rather than representation, in Experiments 1 and 2. This possibility is explored in Experiment 3.

These observations and the results of Experiment 2 are directly compatible with an account in which local features are perceptible but are not encoded individually when we represent an object's contour shape. As in Experiment 1, local contour features in Experiment 2 were changed by inverting the curvature polarity of the bumps, the statistical properties of locally different shape pairs should have been very similar. Participants' low sensitivity to local contour feature changes is consistent with the visual system encoding a statistical description of local contour features, but not encoding individual features specifically. It is true that the sensitivity in local trials was significantly different from chance, but d' values of 0.4 suggest that participants only detected the change in 8% of locally different trials. Possibly, some task-specific strategy or serendipitous alignment of local and global features allowed participants to answer accurately in the small proportion of trials in which a local change was detected. In both Experiments 1 and 2, a local contour change involved a change in all the contour's all local features. If participants were encoding specific local information in any but a very small percentage of trials, their sensitivity to local changes should have been much higher, as specific comparison of any set of local contour features would lead to a detection of a contour difference.

Experiment 3

In Experiments 1 and 2, we found evidence that visual representations of shape are more sensitive to global shape features than local shape features. Equating the total amount of shape change between locally different shapes and globally different shapes in Experiment 2 provided some evidence that people's greater sensitivity to global shape could not be accounted for by the total amount of shape difference in the pairs that differed in their global features. Another possibility remains, however. Individual local features in Experiments 1 and 2 may have been so small that they were at or below the participants' threshold for detection in the conditions of our task. If this were the case, then the lower sensitivity we observed to local shape changes might not be explained by a difference in how the two kinds of features are processed but by an inability to register the local features of the shape while they were being displayed. Some evidence from past research on local and global processing supports this view. Kimchi and Palmer (1985) found that local elements did not interfere with the perception of global form provided that they were sufficiently small, but if their size was increased, the local elements did interfere with the perception of global form.

In Experiment 3, we tested this alternative interpretation of the data from Experiments 1 and 2. We reasoned that if the main cause of the difference in participants' performance on the local vs. global task was the size of the individual local features, then increasing the overall size of the shape displays should increase participants' performance on the local task because the features would be larger. We repeated Experiment 2 with shapes that were 50% larger to test whether increasing the size of local features meaningfully increased participants' sensitivity to local shape differences.

Method

Participants

Ten (nine male, one female, $M_{age} = 26.9$) participants completed this study. Nine of the participants were recruited online through Prolific and were compensated monetarily for their participation. The tenth participant was one of the authors (Nicholas Baker). All participants had normal or corrected-to-normal vision and all except Nicholas Baker were naïve to the purpose of the study.

Display and Apparatus

As in Experiment 2, the experiment was completed online, so screen size and distance were not fixed. The experiment adjusted the absolute size of stimuli so that it covered the same proportion of the screen for different monitors.

Stimuli

The stimuli used in Experiment 3 were identical to those used in Experiment 2 except that their dimensions on the screen were multiplied by 1.5 so that the shapes occupied a maximum of 56.25% of the horizontal width of the screen and 90% of the vertical width of the screen (vs. 37.5% of the horizontal extent and 60% of the vertical extent in Experiment 2).

Design

The design of Experiment 3 was identical to the design of Experiment 2, consisting of 120 Local condition trials and 120 Global condition trials. In each condition, the shape presented second was different in half of the trials.

Procedure

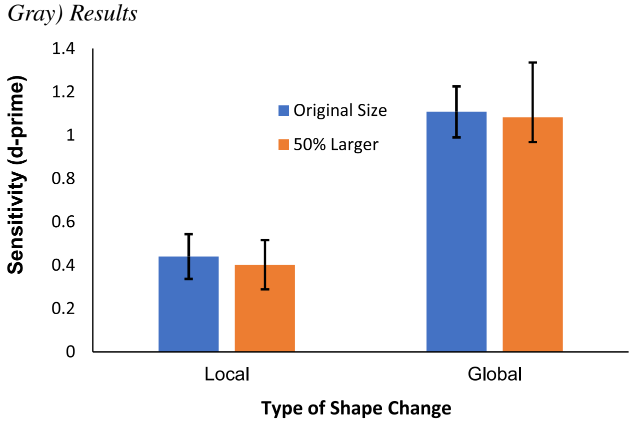
The procedure of Experiment 3 was identical to the one used in Experiment 2. Participants were shown a fixation cross for 330 ms, after which the first shape was shown for 150 ms. The first shape was covered by a pattern mask for 500 ms, followed by the second shape which was shown for 1,000 ms and then masked again for 500 ms. Following the second mask, participants were asked to decide whether the second shape was identical to the first irrespective of a change in orientation.

Results

We once again analyzed participants' sensitivity to a change in local contour features vs. global shape. The results are shown in Figure 9. Participants' sensitivity to local changes with larger shapes was very similar to the sensitivity observed in Experiment 2 with smaller shapes. The same is true of participants' sensitivity to global changes. As in Experiment 2, we found a significant difference between global and local shape sensitivity, $t(9) = 5.01$, $p < .001$, Cohen's $d = 1.9$. We tested for an interaction between the size of the shape and participants' sensitivity to local and global changes with a 2 × 2 mixed ANOVA with shape change (Local vs. Global) as a within-subjects factor and stimulus size (original size vs. 50% larger) as a between-subjects factor using the data from Experiments 2 and 3. The ANOVA revealed no significant interaction between the performance on the shape change conditions and stimulus size, $F(1, 26) = 0.01$, $p = .94$, $\eta^2_{\text{partial}} < 0.001$. We used between-subjects t-tests to analyze differences in performance for local changes and global changes at the two different sizes. We found that stimulus size differences produced no significant difference in either condition; local: $t(21) = 0.40$, $p = .69$, Cohen's $d = 0.16$; global: $t(21) = 0.15$, $p = .89$, Cohen's $d = .07$.

Figure 9

Experiment 2 (Blue/Dark Gray) and Experiment 3 (Orange/Light Gray) Results



Note. The error bars show the 90% confidence interval of the mean. See the online article for the color version of this figure.

Discussion

Although the total amount of physical change was equated between local and global features in Experiment 2, by necessity the global changes were larger than the local changes. This raised the possibility that the local changes introduced in our first two experiments were below some threshold of detection and that the difference we found in participants' sensitivity to each kind of change depended more on a detection problem than differences in the way local and global contour features are represented.

Experiment 3 addressed this possibility by increasing the overall size of presented stimuli by 50%. As a result, more of the local features should be above the threshold for detection. If it were the case that sensitivity differences were driven merely by the small size of local features, sensitivity to local feature changes was predicted to be higher for the larger shapes. On the other hand, if participants' insensitivity to local contour features was really driven by a difference in how local features are perceptually represented, we expected that changing the size of the stimuli would have little effect on performance.

The results of Experiment 3 appear to support the notion that the lower sensitivity to local shape features observed in Experiments 1 and 2 was caused primarily by a difference in what the visual system represents, not in what is extractable from the physical stimulus. Participants' pattern of performance with 50% larger shapes in Experiment 3 was almost identical to the participants' performance in Experiment 2. Detection of global shape changes was significantly easier than detection of changes in individual local contour features despite the increased size of local features. While we might not expect differences between local and global shape sensitivity to fully disappear under this manipulation, if participants' insensitivity to local feature changes in Experiments 1 and 2 had been mainly due to the small size of the local features, we would have expected increased sensitivity from the 50% size increase tested in Experiment 3. The data from Experiment 3 did not appear to support this prediction. Both the overall levels of performance and the differences in sensitivity between global and local changes were remarkably similar across the two experiments.

To ensure that the presented shape fully fits on the viewer's display, we could only increase the total shape size by 50% in Experiment 3. Although it would be possible for local information in both the displays in Experiment 2 and their 50% enlargement in

Experiment 3 to be below some detection threshold or require special scrutiny to encode, this is unlikely given the results we will discuss later (Experiments 5 and 6) showing that local information is registered even in the smaller displays, albeit in a statistical summary rather than in a representation of precise spatial position.

Experiment 3's findings raise interesting questions about how the visual system separates local contour features and global form from each other. One simple way that local contour features could be separated from the global form is by making use of the multiscale edge detectors in early visual areas and sorting features as local or global based on absolute spatial frequency. This explanation seems implausible in its basic form because it follows that retinal image size would be a primary determinant of how contour shape is represented and therefore makes the unlikely prediction that our contour representations would change greatly with viewing distance. The data from Experiment 3 also cast doubt on this simple explanation, as some local features that were below threshold in Experiment 2 would be above threshold in Experiment 3. We return to this issue in the General Discussion.

Experiment 4

The results of Experiment 3 suggest that participants' insensitivity to local contour features is not restricted to cases in which the size of the individual features is below some threshold. An important question remains about what happens to the local shape information later in perceptual processing. We found little evidence that humans' shape representations include information about individual local elements from Experiments 1–3. An interesting question is whether this result is somewhat specific to the presentation time used. Past research using hierarchical displays (e.g., a large letter comprised smaller letter elements) suggested an interplay of observers' local and global shape representations with time of exposure (Navon, 1977).

Kimchi (1998) tested whether configural or local features had a more significant priming effect for various exposure durations of a priming display. She found that short exposures (up to 150 ms) resulted in greater configural priming than local priming. When the exposure time of the priming stimulus was increased to 100–390 ms, individual local elements had a greater priming effect than the overall configuration of the display. The priming effect reversed back to favor global configurations with longer exposures (690 ms). These findings suggest that global information dominated at brief and long exposure times, but there may be intermediate exposure times at which individual local features are represented.

In Experiment 4, we tested whether participants' insensitivity to local contour features in our previous experiments is really due to a different way of processing local vs. global shape features or if this result is specific to the exposure duration we chose in Experiments 1–3. Our 150 ms presentations are right in the range where other research involving local and global information might lead one to expect that global information would eclipse local information. What would happen if the exposure time were increased to 350 ms, when some local information may figure more prominently in observers' representations? We tested participants' sensitivity to local and global changes at both 350 ms, where past research has shown that local features are spontaneously individuated, and at 690 ms where global features appeared to reassert dominance in the hierarchical display work. Our goal was to determine whether, as in Kimchi (1998), increasing participants' exposure time to stimuli with both local and global contour features would result in greater sensitivity to changes in individual local contour features relative to sensitivity to global shape changes.

Method

Participants

We conducted two studies, one in which participants saw the first stimulus for 350 ms and one in which they saw the first stimulus for 690 ms. For the 350 ms study, we collected data from ten participants (five male, five female, $M_{age} = 29.6$). Nine of the participants were recruited online through Prolific and were compensated monetarily for their participation. The tenth participant was one of the authors (NB). For the 690 ms study, ten (six male, four female, $M_{age} = 22.8$) participants completed this study. For the 350 ms study, nine participants were recruited through Prolific and one was an author (NB). All participants had normal or corrected-to-normal vision and all except NB were naïve to the purpose of the study.

Display and Apparatus

Since the experiment was completed online, screen size and distance were not fixed. The experiment adjusted the absolute size of stimuli so that it covered the same proportion of the screen for different monitors.

Stimuli

The stimuli used in Experiment 4 were identical to those used in Experiment 2.

Design

The design of Experiment 4 was identical to the design of Experiment 2, consisting of 120 Local condition trials and 120 Global condition trials. In each condition, the shape presented second was different in half of the trials. The experiment differed from Experiment 2 only in the exposure time of the first display (350 ms for 10 participants and 690 ms for 10 participants).

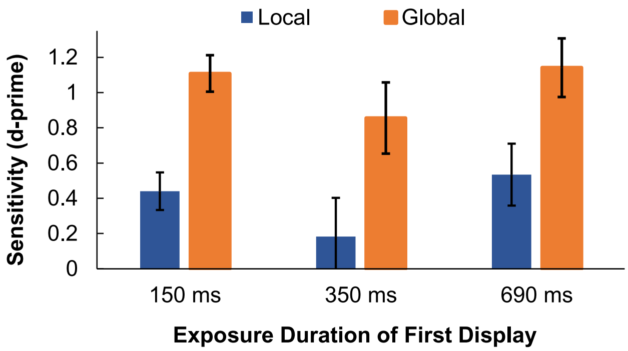
Procedure

The procedure of Experiment 4 was identical to the one used in Experiment 2 except for exposure times to the first shape display. Participants were shown a fixation cross for 330 ms, after which the first shape was shown for 350 or 690 ms in the two conditions, respectively (instead of 150 ms in Experiment 2). The first shape was covered by a pattern mask for 500 ms, followed by the second shape which was shown for 1,000 ms and then masked for 500 ms. Following the second mask, participants were asked to decide whether the second shape was identical to the first irrespective of a change in orientation.

Results

The results of Experiment 4 are shown in Figure 10, along with the earlier results from Experiment 2 (150 ms presentation time) for comparison. One participant's data from the 690 ms condition was not included in the analysis due to at-chance performance in all conditions. Including this participant's data did not meaningfully alter the results.

Figure 10
Experiment 4 Results



Note. The data for 150 ms are from Experiment 2. The error bars show the 90% confidence interval of the mean. See the online article for the color version of this figure.

As in Experiment 2, we found a significant difference between performance in the Local and Global conditions with both 350 ms, $t(9) = 3.77$, $p = .004$, Cohen's $d = 1.19$, and 690 ms of exposure time, $t(8) = 5.10$, $p < .001$, Cohen's $d = 1.54$. To analyze whether varied exposure times had an effect on the relationship between participants' performance on the global task vs. the local task, we conducted a 2 × 3 mixed ANOVA with shape change (local vs. global) as a within-subjects factor and exposure duration (150, 350, or 690 ms) as a between-subjects factor, including the data for 150 ms from Experiment 2. The ANOVA revealed no significant interaction between the performance on the shape change conditions and exposure time, $F(2, 34) = 0.07$, $p = .93$, $\eta^2_{partial} < 0.01$. There was a significant main effect for local vs. global shape, $F(1, 34) = 73.84$, $p < .001$, $\eta^2_{partial} = 0.69$, but not for exposure duration, $F(2, 34) = 2.81$, $p = .074$, $\eta^2_{partial} = 0.14$.

Individual comparisons revealed a marginally significant difference in local sensitivity for 150 ms vs. 350 ms, $t(26) = 2.09$, $p = .05$, Cohen's $d = 0.82$, and for 350 ms vs. 690 ms, $t(17) = 2.13$, $p = .05$, Cohen's $d = 0.98$, but no significant differences between local performance for 150 ms vs. 690 ms, $t(25) = 0.80$, $p = .43$, Cohen's $d = 0.33$. There was no significant difference in global sensitivity between the 150 ms condition and either the 350 ms condition, $t(26) = 1.43$, $p = .16$, Cohen's $d = 0.57$, or the 690 ms condition, $t(25) = 0.29$, $p = .82$, Cohen's $d = 0.09$, nor was there a significant difference between the 350 ms condition and the 690 ms condition, $t(17) = 1.10$, $p = .29$, Cohen's $d = 0.51$.

Discussion

In Experiment 4, we tested an alternative interpretation of our data that global information is not necessarily represented *better* than local information but rather, the time course of its representation is different. This hypothesis was inspired by work by Kimchi (1998, 2015), using a different paradigm, that showed that the local constituent elements of a display can have equal priming potency as more global shape properties, but these local elements require more encoding time before their effects manifest.

We tested for the possibility that local contour features in our displays would be represented in a way more similar to what was observed for global contour features by extending the presentation duration of the first display from 150 to 350 ms, the encoding time at which local priming matched or slightly exceeded global priming in Kimchi's (1998) study, and at 690 ms, the encoding time at which configural information reasserted dominance. In our experiment, the results showed little effect for exposure time on participants' relative sensitivity to local and global shape properties. Performance was slightly (but not significantly) lower in both the Local and Global conditions at 350 ms (Experiment 4) than in the corresponding conditions at 150 ms (Experiment 2). Statistical analysis showed no interaction between exposure duration and type of shape change when we analyzed the data from Experiments 2 and 4 together. These results suggested that overall performance may vary modestly with exposure duration (or due to the different subject groups in the three exposure duration conditions) but there was no indication of differences in the relative strength of global and local information across the different exposure durations in our paradigm.

Why does exposure time appear to play such a small role in our experiment compared to previous findings in local vs. global encoding? Our own stimuli and paradigm differed from those used by Kimchi in several important ways. First, the individual local features in our experiments were randomly sampled from a distribution, not uniform in shape as in Kimchi's (1998) study. Consequently, encoding the individual local features in our shapes would require representing a large number of individual local features, while encoding the local elements in the displays used by Kimchi could be accomplished by categorizing all the local elements as either checkers or columns. We believe that Kimchi's (1998) findings are in fact consistent with our hypothesis that the local features are not encoded individually but statistically. Priming based on local features in Kimchi's study could be accomplished without representing each square or rectangle individually, but by encoding a statistical description of the local elements in the display.

Another key difference is that in our displays, the local elements were only along the shape's bounding contour, whereas in Kimchi's displays, they formed the internal texture of the shape. It seems possible that improved performance observed by Kimchi was due to some kind of texture processing that is not evoked by our displays.

Experiment 5

Experiments 1–4 supported our global–local theory by furnishing evidence that local contour feature changes were poorly represented. Experiments 2 and 3 tested other possible interpretations, such as the possibility that local changes were not detectable, and found evidence of the same pattern of results even when the size of figures, and of their local contour perturbations, were increased by 50% and when the total amount of difference was equated for global and local stimuli. These results cast doubt on the idea that there is a single general contour shape processing system, but they do not comprise a direct test of our proposal that local contour features are encoded via a statistical summary. In Experiment 5, we addressed this issue directly by investigating participants' sensitivity to changes in local contour features in the presence of changes to their summary statistics. Our hypothesis was that when the new set of contour features was sampled from a different distribution, participants would be more sensitive to the change than when the new set of features was sampled from the same distribution.

Method

Participants

Twenty-one (18 female, two male, one gender not specified, $M_{\text{age}} = 20.9$) participants from the University of California, Los Angeles completed this study online for course credit. All participants had normal or corrected-to-normal vision.

Display and Apparatus

Since the experiment was completed online, screen size and distance were not fixed. The experiment adjusted the absolute size of stimuli so that it covered the same proportion of the screen for different monitors.

Stimuli

All stimuli were shown as black outlines on a gray background. The stimulus was shown in the center of the screen, extending over 37.5% of the horizontal space and 60% of the vertical space on the screen.

We generated four kinds of shape pairs for comparison in Experiment 5: matched statistics, different frequency, different phase, and different amplitude. In the Matched Statistics condition, we sampled a new set of contour features from the same distribution from which the features for the first member of the pair were sampled. Both sets of contour features had 80 control points and the mean of the distribution from which amplitudes were sampled was the same. We also wanted to ensure that shape pairs were not merely sampled from the same distribution but did in fact have matched statistics themselves. To that end, we computed the mean and standard deviation of the amplitude of bumps for both shape pairs and resampled until the differences in their means and standard deviations were both less than 0.01.

In the different *frequency* (Frequency Change) condition, we manipulated the number of bumps along the contour by a Weber fraction of 1.5, such that the original stimulus with 80 bumps was paired with another shape that had either 53 or 120 bumps. The mean and standard deviation of the amplitudes were matched as in the matched statistics condition.

In the different *amplitude* (Amplitude Change) condition, the amplitude of the bumps was increased or decreased by a Weber fraction of 1.5. The frequency of the bumps was kept the same as the first member of the pair. Sample shape pairs for all four conditions are shown in Figure 11.

In the different *phase* (Phase Change) condition, we shifted all the local contour features so that the peaks and troughs were approximately halfway between the peaks and troughs in the first member of the pair (1.67% of the contour's total length). This had the effect of preserving all frequency and amplitude statistics but changing the spatial relationships between the local and global features. Whereas the prediction from a summary statistics view of local contour processing is that changes in the amplitude and frequency would change the contour statistics sufficiently to enhance discrimination, the prediction for phase was the opposite. The essence of the local contour statistics hypothesis is that the visual system does not in general encode the local contour orientations or fluctuations at specific locations. If so, then preserving frequency and amplitude statistics, but moving the positions of particular features to different places along the contour (phase shift) should be difficult to discriminate.

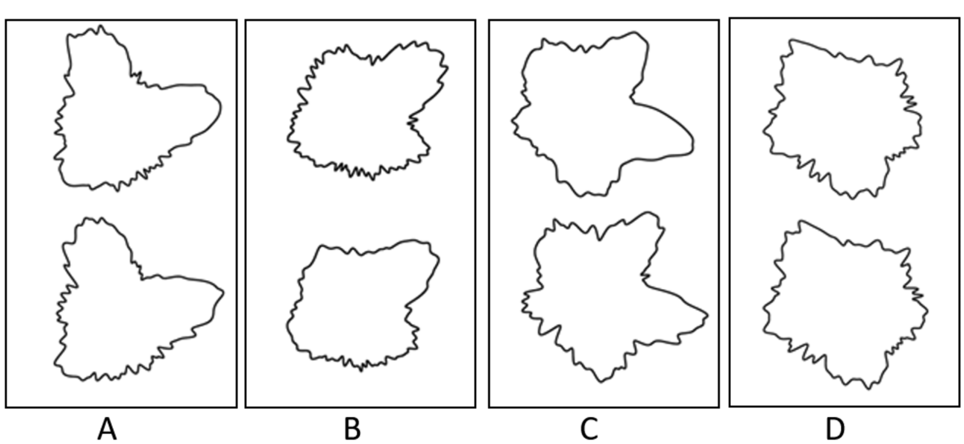
Design

The experiment consisted of four change conditions with 40 trials per condition plus another 160 no-change trials. All conditions were randomly interleaved. Participants completed eight practice trials before beginning the main experiment. At least one trial from each of the four conditions was completed during practice.

Procedure

Apart from the difference in stimuli, the procedure and instructions for Experiment 5 were identical to those used in Experiments 1 and 2. Participants were shown two shapes, one after the other, then asked to determine whether the second shape

Figure 11
Sample Shape Pairs for Experiment 5



Note. Columns from left to right: (A) Matched Statistics condition, (B) Frequency Change condition, (C) Amplitude Change condition, and (D) Phase Change condition.

was exactly the same as the first shape apart from a difference in orientation.

Results

As in Experiments 1–4, we analyzed results in terms of sensitivity, where a hit was a correct detection of a change in shape and a false alarm was an incorrect report of a change in shape. Since there were no differences in the *same* trials between conditions, false alarm rates were calculated as the proportion of "different" responses across all *same* trials in the experiment.

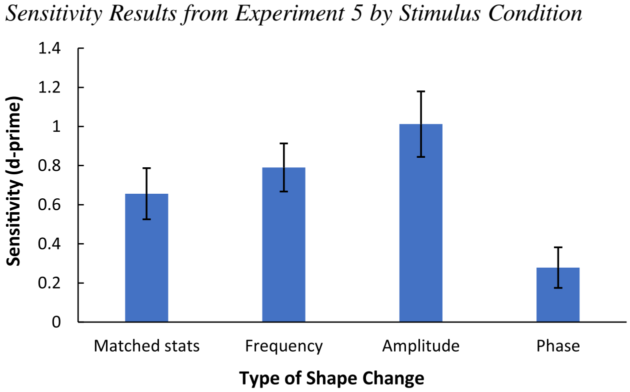
The primary results of Experiment 5 are shown in Figure 12. A one-way repeated measures ANOVA confirmed that there were differences in sensitivity between conditions, $F(3, 60) = 28.84$, $p < .001$, $\eta^2_{\text{partial}} = 0.59$. Paired samples *t*-tests revealed a significant difference between the Amplitude Change and Frequency Change conditions, $t(20) = 2.96$, $p = .008$, Cohen's $d = 0.67$, 95% CI for difference = [0.078, 0.366]; a significant difference between the Frequency Change and Matched Statistics conditions, $t(20) = 2.47$, $p = .022$, Cohen's $d = 0.52$, 95% CI for difference = [0.032, 0.236]; a significant difference between the Amplitude Change and Matched Statistics conditions, $t(20) = 5.35$, $p < .001$, Cohen's $d = 1.16$, 95% CI for difference = [0.226, 0.486]; and a significant difference between Matched Statistics and Phase Change conditions, $t(20) = 4.33$, $p < .001$, Cohen's $d = 0.97$, 95% CI = [0.207, 0.549]. After correcting for multiple comparisons, the difference between the Frequency Change and Matched Statistics conditions became marginally significant.

Discussion

Experiment 5 tested the hypothesis that the system responsible for processing local contour features primarily encodes information about the distribution from which features are sampled, not the properties of individual features. We hypothesized that the statistical information people encode about high-frequency contour features might include information related to the mean and standard deviation of the amplitude of features and the frequency of features along the contour. We therefore predicted that participants would have the greatest sensitivity to local contour feature changes that included a change in the frequency or amplitude of the features, and that participants should be relatively insensitive to changes that did not affect these summary statistics of the features, such as resampling from the same distribution or shifting the same features a small amount along the contour.

Figure 12

Sensitivity Results from Experiment 5 by Stimulus Condition



Note. Error bars show the 90% confidence interval of the mean. See the online article for the color version of this figure.

Our results broadly aligned with these predictions. Sensitivity was highest for amplitude and frequency differences and lower for phase differences and matched statistics. Performance in the Amplitude Change and Frequency Change conditions, but not the other conditions, approached participants' sensitivity to a global shape change in Experiments 1 and 2. This suggests that outputs from the processing system responsible for local features are much more sensitive to differences in the statistical distributions from which local features are sampled than specific descriptions of individual features. These results show that local contour perturbations on the order of those used in Experiments 1 and 2 are not simply poorly encoded; rather, discrimination is poor when local contour features are changed but summary statistics are preserved. In Experiment 5, when summary statistics were varied, performance was reliably better than in the earlier experiments. The latter finding indicates that the issue is not the detectability of local contour perturbations but what gets encoded into shape representations. Local contour perturbations as used in these experiments must be detectable; otherwise, changes in their statistics would be undetectable. The particular positions and topography of local contour features are not encoded as such; rather these detectable properties appear to be utilized as inputs into a summary statistical representation for purposes of shape representation.

Participants performed better for changes in the mean amplitude of contour features than for changes in the frequency of features, and both of these were better detected than local feature differences that had matched statistics. In our design, we equated the Weber fraction for differences in amplitude and frequency, using a ratio of 1.5:1 for both conditions. It appears that the local processing system's sensitivity to the frequency of features is coarser, by this metric, than its sensitivity to the amplitude of local features, resulting in less reliable detection of contour changes brought about by a change in frequency. One reason that the statistical representation for local feature amplitudes might be more precise than local feature frequencies is that frequencies are a second-order relation that depends on the prior detection of amplitude. In order to detect the frequency of local contour features, one must determine the distance between peaks of individual features, which presupposes that the peaks have been detected. By that point, amplitude should already have been obtained.

Another possible explanation for participants' higher sensitivity to amplitude than frequency has to do with how the visual system separates local contour features from global shape. One way the visual system could do this is by comparing the responses of oriented detector outputs at multiple scales, selecting the scale that perceives a relatively small number of contour features, and discarding features below that scale in its global shape representation (see "General Discussion" for a more detailed description of this hypothesis). If the visual system separates global shape from local contour features in this way, the rough amplitude of the local features can be obtained for free simply by looking at the chosen scale of detectors, whereas frequency requires additional processing to be represented statistically.

Sensitivity (d') for contour features that were different but statistically matched to the first shape was 0.66. Performance was better for this Matched Statistics condition than was observed in Experiment 1 ($d' = 0.32$) or Experiment 2 ($d' = 0.44$). We sampled an entirely new set of contour features for each stimulus pair with locally different instances in Experiment 5, whereas in the previous experiments, we inverted the polarity of all contour features. We did this to test a more specific hypothesis about what statistical properties of the local contour the visual system represents. Statistics were matched in terms of mean amplitude, variance of the amplitude, and frequency. These statistics appear to play a large role in subjects' sensitivity to a contour change, but observers' higher sensitivity to contour differences in Experiment 5 may indicate that the system processing local contours encodes types of statistical information about local features beyond those we matched. For example, the visual system might encode more information about the distribution of curvatures within local contour features, although past research indicates low sensitivity to curvature differences of local features in search tasks (Dickinson et al., 2018).

One particularly striking result is how *low* participants' sensitivity was for pairs of shapes in which the local features were phase-shifted a small distance along the shape's contour. Unlike the other three conditions, the features in the phase condition were otherwise identical to those in the first display, so any statistical differences we may not have accounted for in the matched statistics condition would be precisely matched in phase-shifted shape pairs. At the same time, if local contour features were represented precisely in the same system as global features, we would expect high sensitivity to changes in phase as the shift along the contour alters relationships between local and global features. Performance in the Phase Change condition suggests independence between local and global contour features, as there appears to be no precise binding between local features and their position on the shape's global structure.

Experiment 6

In Experiments 1–5, we found evidence that the visual system is more sensitive to the global shapes of objects than to relatively high-frequency, low-amplitude local contour features along an object's boundary. While shape representations appear to include descriptions of global features and relations of parts, local contour features are seemingly described by a small set of statistical properties. Differences in processing for local and global shape information suggest that they are handled by distinct systems. If so, we would predict that these two systems would have distinguishable effects on other perceptual tasks. In Experiment 6, we developed a convergent measure and additional direct test for the independence of local and global shape features in the context of visual search.

In visual search, targets that differ from distractors by a single feature tend to "pop out" from the search array, resulting in a search time that is independent of the number of elements in the array (Treisman, 1986; Wolfe & Horowitz, 2017). However, when the visual system must integrate two independent features, such as shape and color, search time becomes serial, increasing with the size of the array (McElree & Carrasco, 1999; Quinlan & Humphreys, 1987; Treisman & Gelade, 1980; Wolfe & Bennett, 1997). According to Treisman and Gelade's (1980) Feature Integration Theory, basic visual features are extracted automatically and in parallel across the visual field, but independent features of an object can only be integrated together with focused attention.

Experiment 6 tested whether local and global aspects of shape are processed together or independently in human perception. If descriptions of local contour features and global form are represented in independent systems, we reasoned that models of feature integration from visual search might predict that they should require focused attention to be integrated together. If, on the other hand, our theory is incorrect and local and global features are processed in the same system, differences among items in what we have been calling local and global features would not require conjunction search. Readily detectable contour shape differences, either local or global, or their combinations, might all produce more efficient search than is typically seen for conjunctions of properties processed separately.

Method

Participants

Twenty-one (8 female, 11 male, $M_{age} = 31.9$) people participated in this study. About half the participants completed the study for course credit. The other half volunteered to complete the study without compensation. All but two participants were naïve to the purpose of the experiment before completing it. No significant differences were found in the data with these two participants omitted. All participants had normal or corrected-to-normal vision.

Display and Apparatus

All participants completed the experiment online and had variable screen sizes. They were instructed to sit 1.5 times the diagonal length of their screen away from the monitor while completing the experiment but were also told they could adjust if this distance was uncomfortable for them.

Stimuli

All shapes were shown as black outlines on a white background. Each shape extended over up to 16.7% of the horizontal space on the screen and up to 22% of the vertical space. In each trial, there was a target shape and 4, 8, or 12 shapes in a search array. The target was always shown in the top left of the screen outlined by a blue square. In target-present trials, there was also an identical target in the search array. In the target-absent trials, no shape in the search array was identical to the target. The shapes in the search array were displayed in a grid. The four-element search array was in a 2×2 grid, the eight-element search array was in a 4×2 grid, and the twelve-element search array was in a 4×3 grid. Shapes in each search array were positioned so that the mean distance from participants' fixation was equated for all three array sizes. All elements in the search array were rotated by the same magnitude and direction 10–30 degrees off from the target exemplar in the top left.

We tested three different search array conditions: Local, Global, and Conjunction. In the Local condition, all shapes in the search array had the same global form as the target, but only the target had the same local contour features. The distractors all had different local contour features from each other, but the frequency of their contour features was matched, while the target had a different frequency of features. In half of the target-present trials, the target shape had 25 bumps along its contour while the distractors had 80 bumps, and vice versa in the other half. In terms of actual contour features, all shapes in the array were unique, but in terms of statistics of features, only the target was unique. This arrangement differed from most visual search studies in which distractors are typically uniform or fall into a small number of categories with identical items in each. Our design leveraged the idea that pop-out in this context might be possible based on the common contour statistics of distractors, despite the uniqueness of the local contours of each item in the display (for displays where local information was varied between target and distractors).

In the Global condition, both the target and all shapes in the search array had the same set of local contour features added to them. The local contour features had 25 bumps in half of the trials and 80 in the other half. The target always had a global form different from any of the distractor. The distractors all shared the same global form and were identical to each other.

In the Conjunction condition, the target had a unique combination of local contour features and global form. All shapes in the array had added local contour features, half of which had the same frequency as the target and half of which had a different frequency. Likewise, half of the shapes in the array had the same global form as the target and half had a different global form. No distractor had both the same global form and the same frequency of local contours as the target. A sample search array for each condition is shown in Figure 13.

Design

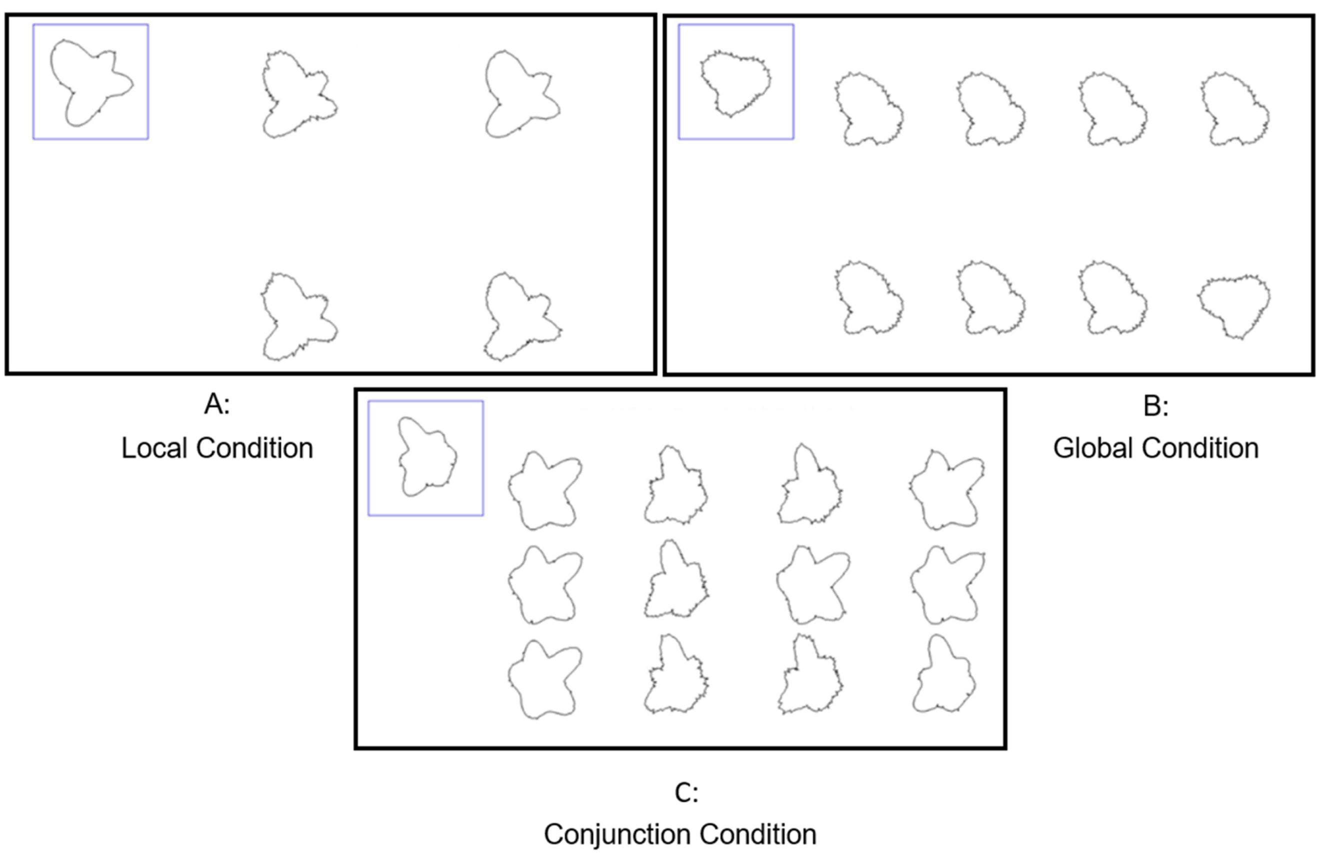
Experiment 6 employed a 3 × 3 (condition type × search array size) within-subjects design. Different condition types were done in blocks with randomly interleaved search array sizes in each block. Each block consisted of 20 target-present trials plus 20 target-absent trials for each of the three array sizes. The order of the three blocks was randomized for each subject to eliminate any systematic fatigue or practice effects. Participants completed six practice trials before beginning the main experiment.

Procedure

In each trial, participants were first shown the search target shape by itself in the top left of the screen enclosed by a blue square. The target remained on the screen for 1,500 ms. It then disappeared and a red fixation cross was presented in the center of the search array location for 500 ms. The fixation cross disappeared, and the search array was shown on the screen. The target in the top left was also shown with the search array for participants to use as a reference if needed. Participants were instructed to decide if the target was present or

Figure 13

Sample Displays for Experiment 6



Note. (A) A 4-element search array in which the target has different local contour features than the distractors. (B) An 8-element search array in which the target has different global form than the distractors. (C) A 12-element search array in which the target has a unique combination of local features and global form. The black rectangles (not present in the experiment) are added in the figure to separate the displays. The blue/dark gray rectangle within each condition was shown during the trial to show subjects the target they were supposed to look for in the search array. See the online article for the color version of this figure.

absent in the search array. They were told to respond affirmatively as soon as they had found the target, but to check all shapes in the array before reporting that the target was absent. Participants were given feedback after each trial telling them if the target was present or absent.

During practice, we showed each of the three different target conditions and each of the three different array sizes. Feedback in practice was more detailed than during the main experiment. After participants responded, the target (if present) was circled in red and an explanation about how the target differed from the distractors was given. This explanation was different for each of the three conditions.

Before each block in the main experiment, participants were given a brief explanation of how the target would differ from distractors in the next block of trials. We did this to reduce any confusion participants might have in what to look for during the first few trials in the new condition.

Dependent Measures and Data Analyses

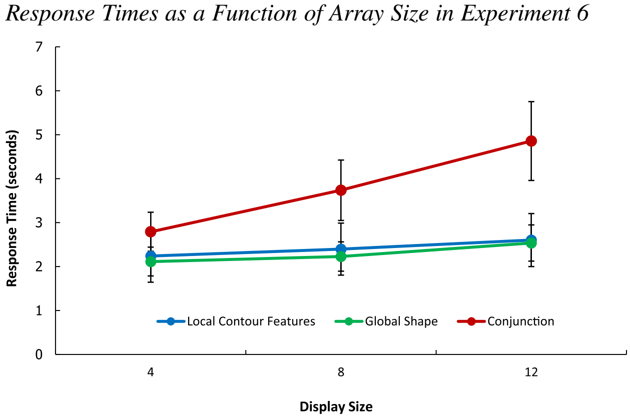
Two subjects were excluded from the final data set because their accuracy was below 65% in all three conditions. For the remaining subjects, we analyzed time to respond as a function of array size for the local, global, and conjunction conditions. We report target-present data only, where search time more systematically varies betweeen feature and conjunction searches (e.g., Treisman & Gelade, 1980; Quinlan & Humphreys, 1987), whereas target-absent data may involve additional issues of search termination criteria that are not of primary interest here (Chun & Wolfe, 1996). All of the main findings of this experiment, however, remain unchanged when target-absent data are also considered.

Results

The results of Experiment 6 are shown in Figure 14. In the Conjunction condition, there appears to be a clear linearly increasing function relating response time to increasing set size of the search array. The Local and Global conditions showed much flatter search functions across different array sizes in the local and global conditions. Following Quinlan and Humphreys (1987), we performed a repeated measures ANOVA and tested for a linear component in within-subjects contrasts for each display condition. For the Conjunction condition, where the target had a unique conjunction of local and global features, an ANOVA confirmed a substantial linear relationship between response time and set size, $F(1, 18) = 40.51$, $p < .0001$. This linear component accounted for 72.2% of the variance; the search slope was 258 ms per item, and the ratio of the slope to the intercept was more than five times larger than in either the local or global condition (0.15 vs. 0.022 and 0.028, respectively).

Figure 14

Response Times as a Function of Array Size in Experiment 6



Note. Error bars represent the 90% confidence interval of the mean. See the online article for the color version of this figure.

In the Local condition, there was a reliable linear component, $F(1, 18) = 16.39$, $p = .001$, which accounted for 40.6% of the variance. In the Global condition, there was also a reliable linear component, $F(1, 18) = 13.92$, $p = .002$, which accounted for 43.6% of the variance. Compared to the slope in the Conjunction condition, both the Local and Global conditions were quite shallow. Search time increased by 46 ms per item in the local condition and 53 ms per item in the global condition. These increases were less than 1/5 of the 258 ms per item slope in the Conjunction condition and minimal relative to the intercept terms for each condition (2,047 and 1,869 ms for the local and global conditions, respectively).

Differences in response times in this experiment cannot be explained by a speed-accuracy tradeoff. In our data, accuracy for the Local, Global, and Conjunction conditions was 93.7%, 95.4%, and 95.1%, respectively. A repeated measures ANOVA confirmed that there were no significant accuracy differences between the conditions, $F(2, 36) = 0.61$, $p = .512$. Some research has found a difference in accuracy for conjunction search trials compared to feature search trials (Treisman, 1993; Wolfe, 1994), but we found no evidence for a loss in accuracy when searching for a conjunction of local and global shape features compared to searching for a single feature.

Discussion

Experiment 6 compared visual search time for shapes that differed in local contour features and/or features of global form. Our hypothesis was that local and global features are distinct and processed independently, and focal attention should be required to integrate both features together (Treisman & Gelade, 1980). An alternative hypothesis is that local and global features are processed by the same system and that global features are simply larger local features. Under this hypothesis, we might expect that search time for a conjunction of local and global features should not be different from search time for local or global features on their own. Participants had flat or nearly flat search times for both local contour features and global form. This suggests that the differences between targets and distractors in the Local and Global conditions comprised features that could be searched for in parallel across the visual field. Although search times did increase slightly with larger array sizes, the increases were small enough relative to the intercepts that they are most likely explainable by the presence of more retinally eccentric shapes in the larger array (Eckstein, 2011; Palmer, 1995; Scialfa & Joffe, 1998) and/or crowding effects from shapes being closer together in larger arrays (Vlaskamp & Hooge, 2006).

The speed at which participants could detect the target in the Local difference condition is especially remarkable because all shapes in the array, including distractors, had unique sets of local contour features. What defined the task and made the target shape pop out among the distractors was not the uniqueness of its high-frequency features but a unique set of *statistical properties* of the high-frequency features. Shapes that had matched statistical properties, but different local elements, could only be differentiated by close scrutiny, but shapes with different statistical properties popped out in the array. In Experiment 5, we found a (statistically adjusted) marginal difference in discriminability for shape pairs with matched statistics and shape pairs with a frequency ratio of 1.5:1. We conjectured that the visual system's sensitivity to differences in frequency might depend on a larger Weber fraction. The results of Experiment 6 support that conjecture: statistical differences in frequency resulted in pop-out with the 3.2:1 ratio used here.

Previous research using a visual search paradigm has found that other kinds of local features cannot be searched for in parallel. Dickinson et al. (2018) compared visual search times for targets that differed from their distractors by the frequency of their RF pattern (a global feature) or by the curvature of frequency-matched contour distortions (a local feature). They found that only global properties, such as the relations between features resulted in parallel search time, whereas the curvatures of individual features needed to be searched through in serial time.

Why, then, do we find that local contour feature differences pop out in our search arrays but not in those used by Dickinson et al. (2018)? The key difference might be the use of a statistical contrast vs. a more specific local feature underlying search in the two tasks. In Dickinson et al., the targets differed in their local contour properties; in a key comparison, rectified RF patterns (i.e., patterns with an orientation discontinuity at their peaks) were searched among regular RF patterns (i.e., patterns that are differentiable at all points). If the local contour system is really concerned with the overall statistics of the contour's local features, it might find that the two displays are generally very similar to each other, made up either of smooth (in the case of regular RF patterns) or mostly smooth (in the case of rectified RF patterns) parts. In its statistical representation of local shape, the visual system may be relatively insensitive to the variations of local contour features used in the Dickinson et al. studies. As a result, it may form a statistical description of the rectified RF shape' local contour features that is very similar to the description of the regular RF shapes' local contour features. If the cornersof the rectified RF shapes factor little into humans' statistical estimations of local shape in these patterns, then Dickinson et al.'s results are consistent with our findings supporting a statistical representation, rather than precise local description, of local contour perturbations.

Support for the hypothesis that local and global shape features are distinct and independent comes from the large increase in search time as a function of array size observed in the conjunction condition. Pop-out ceases when detection of the combined features is required in the search task. Even if one argued that the single-feature visual search tasks in the local and global conditions are not truly parallel, the difference in slope for the conjunction condition indicates an integration of two independent features. Under the alternative hypothesis that local and global features are processed in the same system, we would expect the degree to which visual search varies with array size to depend on the similarity between the target and the distractors (Pashler, 1987), regardless of whether differences are local, global, or both.

In the conjunction search task, the difference between the target and the distractor was *never smaller* than it was in either the global search task or the local search task. It differed from half the distractors by the same amount as in the local search task and half the distractors by the same amount as in the global search task. If global and local features are part of the same description, then, the slope in the conjunction condition might have been expected to be no more than the mean of the slope of the local and global conditions. Instead, it is much greater than either slope, indicating that the two features are distinct, and the visual system must integrate them together in a process that requires focal attention. For example, in a basic ideal observer model of the task using a shape description in which local features and global features were not separated, we would expect the ideal observer to have a faster search rate than in either the local or global task. The total amount of shape difference is strictly larger in the conjunction task than in either of the other two tasks, but search rate is serial. Local and global features appear to be processed separately from each other and integrated together only through focal attention.

General Discussion

The goal of this research was to investigate the relations of shape coding of the global form of an object and descriptions of the local features from which the object's shape is composed. Logically, global shape must be based in some fashion on the local positions of contour elements, but a century of research in perception has shown that what is ultimately represented often depends on relations among elements and often results in little retention of elements present during sensation (Baker & Kellman, 2018; Kanizsa, 1976; Koffka, 1999; Navon, 1977; Pomerantz & Portillo, 2011; Tanaka et al., 1998). At the same time, some information about local contour features is preserved from local sensory information (Erens et al., 1993; Kimchi, 1998; Mamassian et al., 1996). We aimed to clarify the degree to which local and global information are processed together and to better understand what information about a shape's local contour features is represented beyond the visual icon.

Our hypothesis was that local contour features and global shape descriptions are largely processed independently and in separate systems. This theory was partly inspired by past experiments studying the discrimination between simple circles and circles modified by the addition of RF patterns, which showed that detection of the target is different when high (local) and low (global) RF patterns are added to the contour (Bell et al., 2007; Jeffrey et al., 2002). Investigation of the influence of the orientation of individual elements on global contour perception has also suggested that local orientation features are processed separately from global shape (Prins et al., 2007). We also hypothesized that the visual system represents local contour features in a fundamentally different way from global form. While representations of global form are highly descriptive about the curvature, relative size, orientation, and relative positions of various parts, representations of local contour features do not in general specifically describe the individual elements. Instead, we proposed that the visual system captures statistical summaries of the properties of local contour features, it is mainly sensitive to contour differences that change these statistical properties.

Differences in participants' sensitivity to global form and local contour features in Experiment 1 supported the idea that the two kinds of shape features are handled by different systems. Experiment 1 also found that sensitivity was as high for global shape differences as for a combined global and local shape differences. The shapes in the local and global conditions always had strictly more total dissimilarity than the global only condition. If local and global contour features are processed in a single system, we would have predicted larger differences in the contour to correspond to better accuracy in detecting a shape change. A lack of any indication of additivity in local and global features suggested that they are handled in separate systems and do not necessarily interact in recognition tasks.

Experiment 2 followed up on this finding by equating the amount of physical difference between locally and globally distinct shape pairs. If local and global information are processed in the same system, we expected the amount of physical difference between the pairs to be the main predictor for detectability of differences between the two shapes. Participants should therefore have had similar sensitivity in the local and global conditions when similarity was equated. To the contrary, sensitivity was markedly higher in shape pairs that differed in global features than shape pairs that differed in local features, providing evidence that the two kinds of contour features are processed in different systems with different representational priorities. Similar results have been obtained with alternative methods. For example, an analysis of Fourier descriptors using reverse image correlation found that low-frequency shape components played a larger role in shape discrimination than high-frequency components (Wilder et al., 2018).

In addition to physical dissimilarity, we used ShapeComp (Morgenstern et al., 2021) to predict perceived similarity using a combination of many prominent models of shape representation. ShapeComp predicted that locally different shapes should be more discriminable from each other than globally different shapes, the opposite of what we observed in our data. This discrepancy supports the view that local contour features are described statistically, in a fundamentally different way than global features. Models that include local contour properties as individualized features in the visual system's shape representations cannot fully explain performance on visual discrimination tasks.

In Experiments 3 and 4, we tested two alternative explanations for participants' low sensitivity to local contour features. We first tested the hypothesis that the local features in Experiment 2 were below a threshold for detection or were less impactful because they required special scrutiny to encode. We enlarged the stimulus' size on the screen by 50% and compared participants' sensitivity to the enlarged stimuli with the sensitivity observed in Experiment 2. Analyses revealed no difference in participants' sensitivity, and specifically, no gain in representation of local features, with enlarged displays. These results suggest that the visual system does not fail to extract local contour features due to their retinal size but due to a difference in how these kinds of features are represented. The results of Experiment 3 also suggest that the mechanism by which local and global features are processed separately is size invariant within some range, having to do with object-centric concerns like the relative size of local and global features rather than features of the proximal stimulus.

In Experiment 4, we tested the hypothesis that local features are individually represented in our shape descriptions but take more encoding time than global features. We extended the exposure duration of our stimuli from 150 to 350 ms and 690 ms, but once again found no difference in observers' sensitivity to local contour features. The lack of improvement in the local condition with additional processing time supports the notion that perceivers ultimately form a statistical, not an individuated representation of local contour features. When the statistics of the local features were matched as in Experiment 4, observers were unable to detect a difference between shapes even with longer exposure.

Participants' low sensitivity for phase-shifted contour features in Experiment 5 also suggests independent local and global processing systems. In this condition, local contour features were preserved but shifted along the global form of the object. If participants represented local and global features together, then all of the changes in the spatial relationships between the small and large features would result in a very different percept. Instead, the visual system's description of local contour features appears to be independent of their position relative to the global features of the shape. Participants' low sensitivity to phase in the local condition contrasted sharply with the global condition, where the angular positions of curvature features are a highly salient cue (Pasupathy & Connor 1999, 2001, 2002). In Experiments 1–4, we rotated the shapes a small amount between their first and second presentation. This did not appear to affect participants' sensitivity to global shape differences, suggesting that the angular position of global features is an object-centric consideration in global shape encoding, defined by the relative positions of global features, as has been previously proposed (Dickinson et al., 2013, 2018).

Experiment 6 was an additional direct test of independence between local and global systems, using a converging method. Following Treisman and Gelade's (1980) Feature Integration Theory, we predicted that if local and global aspects of shape are distinct feature dimensions, then they should require focal attention to be integrated together. This hypothesis was confirmed in a visual search task in which the target could differ from distractors in local statistical properties, global form, or a conjunction of the two. While targets that differed only in local statistical properties or global form popped out in the search array, targets with a conjunction of local and global features required serial search time. Targets' physical dissimilarity to distractors was just as high in the Conjunction condition as in the Local or Global conditions, so a steeper search slope in the conjunction condition is evidence that the visual system needs to integrate distinct features together to do the task.

Besides implicating separate systems for global and local contour information, the results of our experiments also support the idea that the visual system represents statistical properties of local contour features rather than each feature individually. In Experiments 1–4, local change trials inverted the polarity of all bumps along the object's contour. Even though no local contour feature was the same, participants had difficulty detecting a difference between locally different shape pairs. Participants' poor performance suggests that matching shapes based on local features is not primarily done by probability summation of local detectors (Loffler et al., 2003). Because all bumps were inverted, the representation of even a small number of local contour features would provide high sensitivity to local changes in Experiments 1–4. On the other hand, if the visual system's primary tool for local feature comparison involves looking for a difference in the distribution from which features were sampled, sensitivity should be low for inverted bumps (assuming equal probability of positive and negative curvature bumps).

In Experiment 5, we proposed that the visual system primarily encodes the mean and variance of the amplitude of bumps and the frequency of bumps. We tested this by generating shape pairs that had matched frequency, amplitude, and variance. We compared sensitivity to these statistically matched pairs with shape pairs that had different frequency or different mean and variance of the amplitude. We also compared the statistically matched pairs with pairs whose local contour features were the same but were shifted slightly along the contour. Results confirmed that sensitivity was highest when the summary statistics for shapes' contour features differed in either frequency or amplitude.

Pop-out of the target in the local condition in Experiment 6 lent further support to the hypothesis that the visual system represents statistical properties of local contour elements rather than individually representing each element. Despite all shapes in the array having unique local contour features, we found that participants quickly detected the target in the local condition, irrespective of the size of the array. Though they were not physically identical, all distractors had the same number of bumps along their contour, while the target always had 3.2 times more or fewer bumps. Pop-out for local features appears to depend on different statistical properties of the local contour elements, not differences in the individual elements.

Evidence from previous studies shows that local features do not always pop out in visual search paradigms (Dickinson et al., 2018). This evidence coheres with our account of a statistical representation of local contour variation. Especially when local perturbations are numerous, we might think of the local contour system described here as being concerned with "contour texture." By contour texture, we refer to repeating (although not necessarily uniform) variation along a contour boundary. Contour texture would be an excellent candidate for statistical representation because its features are numerous and not individually diagnostic for the identity of an object. For example, it might be useful to know the basic statistical pattern of leaves on a tree, wool on a sheep, or gears on a wheel, but representing them individually could stress the visual system's representational capacities without contributing much information that is not already captured in a statistical summary. The contour texture hypothesis is particularly attractive because it is consistent with earlier findings that high-frequency RF patterns are represented in a fundamentally different way from low-frequency patterns (Bell et al., 2007; Loffler et al., 2003).

One aspect of our data that our hypothesis about statistical descriptions of local contour features does not explain is why sensitivity is greater than zero in statistically matched shape pairs. One possibility is that the visual system is sensitive to more kinds of statistical features than we equated in Experiments 1–5. Participants' higher sensitivity to the statistically matched shape pairs than the phase-shifted shape pairs in Experiment 5 suggests that this might be the case. Given that even sensitivity to the inverted bumps in Experiments 1–4 is non-zero, though, it seems likely that the visual system can represent some local contour features with specificity, albeit rarely.

Another possibility is that participants adopted a task-specific strategy for scrutinizing and memorizing a few local contour features and checking for their presence or absence in the second shape stimulus. This hypothesis is indirectly supported by the fact that sensitivity to local contour feature changes was lowest in Experiment 1, where shapes with high-frequency contour features could undergo a global change in addition to the local change. If memorizing a single local feature or small number of local contour features is effortful and unnatural, we would expect participants to do it less when other cues for discrimination are potentially available. A different potential explanation is that a local feature sometimes gets encoded as a global feature if, by random chance, it is particularly large or placed in a particularly salient position on the shape's global form, such as a local maximum or minimum. While either strategy could potentially lead to better-than-zero sensitivity to differences between statistically matched local contour features, neither appears to be used often, as the observed sensitivity to local contour feature changes suggested that participants only truly detected a local difference in a small number (less than 10%) of trials.

Why might human shape representations employ a separation of global shape description and local contour characteristics? The reasons involve the functional importance of shape descriptions in perception, cognition, and action. In biological vision, shape is the primary basis for object recognition (Baker et al., 2018; Biederman & Ju, 1988). Shape often provides the most crucial information about an object's functional properties, as well as its conceptual category and name (Imai et al., 1994; Landau et al., 1992). For most of these purposes, it is global shape and relations, not local contour detail, that are most important. The shape information that allows the classification of a tree as a maple tree, rather than a poplar or spruce, is to be found in its overall shape. As one can easily verify by looking at any particular tree, the overall shape impression is based on real information but is also an abstraction: Following the exact contours of leaves and branches and gaps gives both more and less than the overall shape perceived. The zigs and zags of particular edges of leaves or needles provide information that goes beyond the global shape, whereas the overall shape percept extends through small gaps and indentations despite a lack of local information there. With regard to useful categorization and seeing of similarities involving natural kinds in the world, a fully detailed map of exact contour features of each instance of a tree, dog, or cat would be an impediment to useful categorization, as no two instances would match. The same is true when we consider shape descriptions of terrain; while hiking, we want to plan our path based on major topographic features and their relations. Consistent with these ideas, evidence indicates that the primacy of global shape in human processing (Baker & Kellman, 2018; Pomerantz et al., 1977). Conversely, local information furnishes other important information, such as object texture and material properties.

Understanding how the visual system processes global and local information might be crucial to understanding how to build artificial visual systems that more closely match human perceptual capabilities. In current state-of-the-art artificial neural networks, there is strong evidence to suggest that local and global shape properties are not processed separately from each other (Baker et al., 2018, 2020). Without the separation of these properties into independent feature dimensions, these artificial systems tend to make erroneous responses that humans would never make. For example, straightening out local edge features on a bear's back results in it being classified as a warplane, or adding a serrated pattern to a violin results in a bald eagle classification. These errors seem ridiculous to us, but we might be prone to very similar confusions without separate systems to process local and global shape. The results reported here suggest that our global description of the shape is separated from the local elements composing its boundary, which would allow us to still see a bear as a bear if some contour features were straightened or to see similarity between a horse and a cloud despite the numerous wisps and curls along the cloud's bounding contour. Building artificial systems with similar capabilities may also require separate shape-coding systems for local and global information.

The findings of our study also offer valuable insights into the transition from transient, literal descriptions of local stimulus information to more durable, abstract descriptions of shape. Past research has shown that this abstraction takes meaningful processing time to compute and results in a very different representation of the shape than what is present during initial sensory activations (Baker & Kellman, 2018). The results reported here suggest that two different forms of abstraction take place in visual shape encoding. In the local domain, contour features are abstracted by estimating a statistical distribution from which they are likely to be sampled, and only this distribution persists in the abstract shape representation. In the global domain, the visual system extracts contour shape apart from local contour fluctuations, then describes the relative size, orientation, and curvature of each part in an object-centric format, as evidenced by the visual system's invariance to planar transformations in several of our tasks.

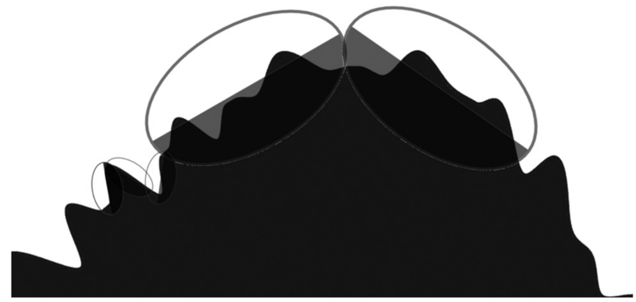
How does the visual system abstract global form from high-frequency contour features? One possibility is that the global processing system uses oriented detectors at different scales to extract the low-frequency features from a shape. Multi-scaled detectors in early visual areas are sensitive to particular orientations in the visual field (Gur et al., 2005; Hubel & Wiesel, 1962). Curvature along a contour could be estimated by the difference in the preferred orientation of nearby detectors. Elsewhere, we have proposed the existence of what we called *arclets*, higher-order neural detectors derived from co-circular oriented units, that span a range of turning angles connecting linked elements and scales (Baker et al., 2021; Garrigan, 2006; Kellman & Garrigan, 2007). The existence of such detectors is consistent with considerable evidence regarding constant-curvature representations in a variety of visual tasks (Baker & Kellman, 2021; Kellman & Garrigan, 2007), and they may play a special role in the attainment of abstract, symbolic descriptions in vision from initially local, transient, and subsymbolic activations (for discussion, see Baker et al., 2021).

For a contour with no high-frequency features, the most precise description of curvature would be obtained from arclets at the smallest scale—as the size of each detector approaches a point, the difference between the curvature estimated by the turning angle and the contour's true curvature approaches zero. Detectors at multiple scales, however, may have important applications in contour representation, including obtaining size invariance for contours scaled up or down (Baker et al., 2021; Kellman & Garrigan, 2007). Another application of larger scale arclets, especially relevant to the present work, might be to encode the curvature of a contour's global shape in the presence of smaller, high-frequency contour features, as illustrated in Figure 15, which shows a zoomed-in portion of a contour with added local contour features. Detectors at the finest scale are sensitive to orientation changes given by local contour features, but larger scale detectors would remain sensitive to more global properties of the shape.

It is unlikely that the visual system processes global information with detectors larger than some fixed scale and local information below that. Small contour features are often processed as part of an object's global shape when they are not accompanied by other similar small features. Likewise, we can still encode the global

Figure 15

Comparison of Two Detectors' Sensitivity to High-Frequency Contour Features



Note. Three small detectors in the bottom left pick out changes along the contour from local features while the larger detectors on the top abstract over them.

form of an object with high-frequency contour features when it is scaled up, which would sometimes result in local elements crossing a threshold to be processed globally. A more likely hypothesis is that the visual system settles on the scale of global detectors depending on the properties of the object's contour. One simple way it could do this is by looking at the profile of curvatures outputted by detectors at multiple scales. Going from the largest scale detectors to the smallest, there should be a point in which the curvature profile dramatically changes in contours with high-frequency features. For the shape in Figure 15, detectors of a scale too large to be much activated by the orientation of local elements will output a curvature profile corresponding to the underlying global shape of the object. However, at a certain point, the scale of the detectors will be small enough that they are influenced by the orientation of local features, resulting in a curvature profile that includes dozens of turns and reverses in curvature polarity. The visual system might use the curvature profile outputted from arclets at the largest scale at which there is still a reasonable fit, or, conversely, the finest scale at which adequately fitting arclets describe the contour's curvature with a limited number of sign changes. We offer this hypothesis at present only as a conjecture to suggest how early neural coding might support separable local and global coding schemes. It does, however, have intriguing connections to other recent ideas about abstract shape representations and how they are obtained (Baker et al., 2021; Baker et al., 2021; Baker & Kellman, 2021). Further research is needed to develop a systematic theory of global shape extraction from multi-scale filters.

Conclusion

We conducted six experiments, the results of which point to a dissociation in human perception between local features along a contour and the global form defined by relations between them. The system that encodes information about local elements does not precisely represent their properties individually, instead estimating a few statistical properties that are shared by local elements. The system that encodes global form represents parts of the object with much greater specificity and spatial precision. Although they are both concerned with representations of shape, the local and global processing systems are distinct from each other and operate independently. Descriptions of the shape that require integration of local and global contour features can only be formed with focal attention as is needed for other distinct visual feature dimensions. Our theoretical separation of global and local shape processing systems and the experimental evidence for this separation may reflect the important differences of global and local shape descriptions for perception, cognition, and action. For similar reasons, this distinction and an understanding of differences in the types of processing involved may have important consequences for the development of artificial vision systems.

Context of Research

This work originates from Nicholas Baker and Philip Kellman's interest in abstraction in visual perception and representation, and the relation between the perception of overall configuration and local structure. We became curious about these issues by observing that humans can effortlessly see overall form as similar in displays whose local information, context, and meaning differ radically. A cloud may look like a fish, or, in a classic example, observers may report seeing the Virgin Mary (or at least a face) in a slice of toast. We seek to understand how the visual system can grasp the overall shape similarities of objects, abstracting this away from local, nonessential features that may differ completely. This issue connects directly to classic Gestalt ideas in perception ("the whole differs from the sum of the parts"), but our work aims to discover the computational processes and mechanisms that lead to abstraction in visual perception. The issue of local vs. global shape is also timely and interesting to us because some recent research into how top-performing artificial intelligence systems, such as deep convolutional neural networks (DCNNs) classify images suggests that the abstraction and priority in biological vision for global form are not inevitable. While humans are much more sensitive to global form, DCNNs are exclusively concerned with local contour features. The contrast suggests that biological vision systems have evolved specialized routines for extracting and representing global shape. One of the aims of the current paper was to begin to understand differences in how local and global features are encoded and to test the possibility that they are handled by independent systems. In future work, we hope to continue efforts to understand differences between the local and global systems, and to develop more detailed models of abstraction and representation in visual processing of separable global and local contour shape properties.

Constraints on Generality

Experiments in this paper were collected from a subject pool of undergraduates at the University of California, Los Angeles and (primarily) from a database of participants from Prolific. The only requirements for participation through Prolific were fluency in English and normal or corrected-to-normal vision. Given the diversity of our participant pool and the nature of our research questions into basic mechanisms of visual perception, we believe that these findings should generalize broadly.

References

Arnheim, R. (1971). *Entropy and art*. University of California Press.

Baker, N., Garrigan, P., & Kellman, P. J. (2021). Constant curvature segments as building blocks of 2D shape representation. *Journal of Experimental Psychology: General*, *150*(8), 1556–1580. https://doi.org/10.1037/xge0001007

Baker, N., & Kellman, P. J. (2018). Abstract shape representation in human visual perception. *Journal of Experimental Psychology: General*, *147*(9), 1295–1308. https://doi.org/10.1037/xge0000409

Baker, N., Kellman, P. J., & Maiello, G. (2021). Constant curvature modeling of abstract shape representation. *PLoS ONE*, *16*(8), Article e0254719. https://doi.org/10.1371/journal.pone.0254719

Baker, N., Lu, H., Erlikhman, G., & Kellman, P. J. (2020). Local features and global shape information in object classification by deep convolutional neural networks. *Vision Research*, *172*, 46–61. https://doi.org/10.1016/j.visres.2020.04.003

Baker, N., Lu, H., Erlikhman, G., Kellman, P. J., & Einhäuser, W. (2018). Deep convolutional networks do not classify based on global object shape. *PLoS Computational Biology*, *14*(12), Article e1006613. https://doi.org/10.1371/journal.pcbi.1006613

Barenholtz, E., Cohen, E. H., Feldman, J., & Singh, M. (2003). Detection of change in shape: An advantage for concavities. *Cognition*, *89*(1), 1–9. https://doi.org/10.1016/S0010-0277(03)00068-4

Bell, J., Badcock, D. R., Wilson, H., & Wilkinson, F. (2007). Detection of shape in radial frequency contours: Independence of local and global form information. *Vision Research*, *47*(11), 1518–1522. https://doi.org/10.1016/j.visres.2007.01.006

Bell, J., Wilkinson, F., Wilson, H. R., Loffler, G., & Badcock, D. R. (2009). Radial frequency adaptation reveals interacting contour shape channels. *Vision Research*, *49*(18), 2306–2317. https://doi.org/10.1016/j.visres.2009.06.022

Biederman, I., & Ju, G. (1988). Surface versus edge-based determinants of visual recognition. *Cognitive Psychology*, *20*(1), 38–64. https://doi.org/10.1016/0010-0285(88)90024-2

Blum, H. (1973). Biological shape and visual science (part I). *Journal of Theoretical Biology*, *38*(2), 205–287. https://doi.org/10.1016/0022-5193(73)90175-6

Braddick, O. (1973). The masking of apparent motion in random-dot patterns. *Vision Research*, *13*(2), 355–369. https://doi.org/10.1016/0042-6989(73)90113-2

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *8*(6), 679–698. https://doi.org/10.1109/TPAMI.1986.4767851

Dickinson, J. E., Bell, J., Badcock, D. R., & de Beeck, H. P. (2013). Near their thresholds for detection, shapes are discriminated by the angular separation of their corners. *PLoS ONE*, *8*(5), Article e66015. https://doi.org/10.1371/journal.pone.0066015

Dickinson, J. E., Cribb, S. J., Riddell, H., & Badcock, D. R. (2015). Tolerance for local and global differences in the integration of shape information. *Journal of Vision*, *15*(3), Article 21. https://doi.org/10.1167/15.3.21

Dickinson, J. E., Haley, K., Bowden, V. K., & Badcock, D. R. (2018). Visual search reveals a critical component to shape. *Journal of Vision*, *18*(2), Article 2. https://doi.org/10.1167/18.2.2

Eckstein, M. P. (2011). Visual search: A retrospective. *Journal of Vision*, *11*(5), Article 14. https://doi.org/10.1167/11.5.14

Erens, R. G., Kappers, A. M., & Koenderink, J. J. (1993). Perception of local shape from shading. *Perception and Psychophysics*, *54*(2), 145–156. https://doi.org/10.3758/BF03211750

Feldman, J., & Singh, M. (2006). Bayesian estimation of the shape skeleton. *Proceedings of the National Academy of Sciences*, *103*(47), 18014–18019. https://doi.org/10.1073/pnas.0608811103

Garner, W. R. (1974). The stimulus in information processing. In H. R. Moskowitz, B. Scharf, & J. C. Stevens (Eds.), *Sensation and measurement* (pp. 77–90). Springer. https://doi.org/10.1007/978-94-010-2245-3_7

Garrigan, P. (2006). *Representation of contour shape* [Doctoral dissertation]. University of California, Los Angeles. https://www.globethesis.com/?t=1458390008950211

Green, R. J., Dickinson, J. E., & Badcock, D. R. (2018). Convergent evidence for global processing of shape. *Journal of Vision*, *18*(7), Article 7. https://doi.org/10.1167/18.7.7

Gur, M., Kagan, I., & Snodderly, D. M. (2005). Orientation and direction selectivity of neurons in V1 of alert monkeys: Functional relationships and laminar distributions. *Cerebral Cortex*, *15*(8), 1207–1221. https://doi.org/10.1093/cercor/bhi003

Hochberg, J., & Brooks, V. (1962). Pictorial recognition as an unlearned ability: A study of one child's performance. *The American Journal of Psychology*, *75*(4), 624–628. https://doi.org/10.2307/1420286

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition*, *18*(1–3), 65–96. https://doi.org/10.1016/0010-0277(84)90022-2

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology*, *160*(1), 106–154. https://doi.org/10.1113/jphysiol.1962.sp006837

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development*, *9*(1), 45–75. https://doi.org/10.1016/0885-2014(94)90019-1

Jeffrey, B. G., Wang, Y. Z., & Birch, E. E. (2002). Circular contour frequency in shape discrimination. *Vision Research*, *42*(25), 2773–2779. https://doi.org/10.1016/S0042-6989(02)00332-2

Kanizsa, G. (1976). Subjective contours. *Scientific American*, *234*(4), 48–52. https://doi.org/10.1038/scientificamerican0476-48

Kellman, P. J., & Garrigan, P. (2007). Segmentation, grouping, and shape: Some Hochbergian questions. In M. A. Peterson, B. Gillam, & H. A. Sedgwick (Eds.), *In the mind's eye: Julian Hochberg on the perception of pictures, films, and the world* (542–554). Oxford University Press.

Kennedy, J. M. (1974). Perception, pictures, and the etcetera principle. In R. B. MacLeod & H. L. Pick (Eds.), *Perception: Essays in honor of James J. Gibson* (pp. 209–226). Cornell University Press.

Kimchi, R. (1998). Uniform connectedness and grouping in the perceptual organization of hierarchical patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *24*(4), 1105–1118. https://doi.org/10.1037/0096-1523.24.4.1105

Kimchi, R. (2015). The perception of hierarchical structure. In J. Wagemans (Ed.), *Oxford handbook of perceptual organization* (pp. 129–149). Oxford Academic.

Kimchi, R., & Palmer, S. E. (1985). Separability and integrality of global and local levels of hierarchical patterns. *Journal of Experimental Psychology: Human Perception and Performance*, *11*(6), 673–688. https://doi.org/10.1037/0096-1523.11.6.673

Koffka, K. (1999). *Principles of Gestalt psychology, 1935*. Lund Humphries.

Kurki, I., Saarinen, J., & Hyvärinen, A. (2014). Investigating shape perception by classification images. *Journal of Vision*, *14*(12), 24–24. https://doi.org/10.1167/14.12.24

Landau, B., Smith, L. B., & Jones, S. (1992, December 1). Syntactic context and the shape bias in children's and adults' lexical learning. *Journal of Memory and Language*, *31*(6), 807–825. https://doi.org/10.1016/0749-596X(92)90040-5

Loffler, G., Wilson, H. R., & Wilkinson, F. (2003). Local and global contributions to shape discrimination. *Vision Research*, *43*(5), 519–530. https://doi.org/10.1016/S0042-6989(02)00686-7

Mamassian, P., Kersten, D., & Knill, D. C. (1996). Categorical local-shape perception. *Perception*, *25*(1), 95–107. https://doi.org/10.1068/p250095

McElree, B., & Carrasco, M. (1999). The temporal dynamics of visual search: Evidence for parallel processing in feature and conjunction searches. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(6), 1517–1539. https://doi.org/10.1037/0096-1523.25.6.1517

Morgenstern, Y., Hartmann, F., Schmidt, F., Tiedemann, H., Prokott, E., Maiello, G., Fleming, R. W., & van den Berg, R. (2021). An image-computable model of human visual shape similarity. *PLoS Computational Biology*, *17*(6), Article e1008981. https://doi.org/10.1371/journal.pcbi.1008981

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, *9*(3), 353–383. https://doi.org/10.1016/0010-0285(77)90012-3

Palmer, J. (1995). Attention in visual search: Distinguishing four causes of a set-size effect. *Current Directions in Psychological Science*, *4*(4), 118–123. https://doi.org/10.1111/1467-8721.ep10772534

Pashler, H. (1987). Target-distractor discriminability in visual search. *Perception and Psychophysics*, *41*(4), 285–292. https://doi.org/10.3758/BF03208228

Pasupathy, A., & Connor, C. E. (1999). Responses to contour features in macaque area V4. *Journal of Neurophysiology*, *82*(5), 2490–2502. https://doi.org/10.1152/jn.1999.82.5.2490

Pasupathy, A., & Connor, C. E. (2001). Shape representation in area V4: Position-specific tuning for boundary conformation. *Journal of Neurophysiology*, *86*(5), 2505–2519. https://doi.org/10.1152/jn.2001.86.5.2505

Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, *5*(12), 1332–1338. https://doi.org/10.1038/972

Pomerantz, J. R., & Portillo, M. C. (2011). Grouping and emergent features in vision: Toward a theory of basic Gestalts. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(5), 1331–1349. https://doi.org/10.1037/a0024330

Pomerantz, J. R., Sager, L. C., & Stoever, R. J. (1977). Perception of wholes and of their component parts: Some configural superiority effects. *Journal of Experimental Psychology: Human Perception and Performance*, *3*(3), 422–435. https://doi.org/10.1037/0096-1523.3.3.422

Prins, N., Kingdom, F. A., & Hayes, A. (2007). Detecting low shape-frequencies in smooth and jagged contours. *Vision Research*, *47*(18), 2390–2402. https://doi.org/10.1016/j.visres.2007.06.006

Quinlan, P. T., & Humphreys, G. W. (1987). Visual search for targets defined by combinations of color, shape, and size: An examination of the task constraints on feature and conjunction searches. *Perception and Psychophysics*, *41*(5), 455–472. https://doi.org/10.3758/BF03203039

Scialfa, C. T., & Joffe, K. M. (1998). Response times and eye movements in feature and conjunction search as a function of target eccentricity. *Perception and Psychophysics*, *60*(6), 1067–1082. https://doi.org/10.3758/BF03211940

Smithson, H., & Mollon, J. (2006). Do masks terminate the icon? *Quarterly Journal of Experimental Psychology*, *59*(1), 150–160. https://doi.org/10.1080/17470210500269345

Tanaka, J. W., Kay, J. B., Grinnell, E., Stansfield, B., & Szechter, L. (1998). Face recognition in young children: When the whole is greater than the sum of its parts. *Visual Cognition*, *5*(4), 479–496. https://doi.org/10.1080/713756795

Treisman, A. (1986). Features and objects in visual processing. *Scientific American*, *255*(5), 114–125. https://doi.org/10.1038/scientificamerican1186-114B

Treisman, A. (1993). The perception of features and objects. In A. D. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, awareness, and control: A tribute to Donald Broadbent* (pp. 5–35). Clarendon Press/Oxford University Press.

Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*(1), 97–136. https://doi.org/10.1016/0010-0285(80)90005-5

Vlaskamp, B. N., & Hooge, I. T. C. (2006). Crowding degrades saccadic search performance. *Vision Research*, *46*(3), 417–425. https://doi.org/10.1016/j.visres.2005.04.006

Wertheimer, M. (1923/1938). Laws of organization in perceptual forms. In W. D. Ellis (Ed.), *A source book of Gestalt psychology* (pp. 71–88). Kegan Paul, Trench, Trubner & Company. https://doi.org/10.1037/11496-005

Wickens, T. D. (2001). *Elementary signal detection theory*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195092509.001.0001

Wilder, J., Fruend, I., & Elder, J. H. (2018). Frequency tuning of shape perception revealed by classification image analysis. *Journal of Vision*, *18*(8), Article 9. https://doi.org/10.1167/18.8.9

Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, *38*(22), 3555–3568. https://doi.org/10.1016/S0042-6989(98)00039-X

Wolfe, J. M. (1994). Guided search 2.0 A revised model of visual search. *Psychonomic Bulletin and Review*, *1*(2), 202–238. https://doi.org/10.3758/BF03200774

Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, *37*(1), 25–43. https://doi.org/10.1016/S0042-6989(96)00111-3

Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*(3), 1–18. https://doi.org/10.1038/s41562-017-0058

Received December 24, 2021
Revision received November 11, 2022
Accepted November 14, 2022 ■