Perceived group size is determined by the centroids of the component elements

Alexandria M. Boswell Department of Psychology, University of Nevada, Reno, Nevada, USA

Peter J. Kohler Department of Psychology, York University, Toronto, Ontario, Canada
Centre for Vision Research, York University, Toronto, Ontario, Canada

J. Daniel McCarthy Department of Psychology, Western Oregon University, Oregon, USA

Gideon P. Caplovitz Department of Psychology, University of Nevada, Reno, Nevada, USA

To accomplish the deceptively simple task of perceiving the size of objects in the visual scene, the visual system combines information about the retinal size of the object with several other cues, including perceived distance, relative size, and prior knowledge. When local component elements are perceptually grouped to form objects, the task is further complicated because a grouped object does not have a continuous contour from which retinal size can be estimated. Here, we investigate how the visual system solves this problem and makes it possible for observers to judge the size of perceptually grouped objects. We systematically vary the shape and orientation of the component elements in a two-alternative forced-choice task and find that the perceived size of the array of component objects can be almost perfectly predicted from the distance between the centroids of the component elements and the center of the array. This is true whether the global contour forms a circle or a square. When elements were positioned such that the centroids along the global contour were at different distances from the center, perceived size was based on the average distance. These results indicate that perceived size does not depend on the size of individual elements, and that smooth contours formed by the outer edges of the component elements are not used to estimate size. The current study adds to a growing literature highlighting the importance of centroids in visual perception and may have implications for how size is estimated for ensembles of different objects.

Introduction

Although seemingly trivial, the process by which we perceive the size of an object is, in fact, quite complex, and understanding the principles that underlie size perception remains a central focus of vision science. To perceive the size of an object, the visual system has to determine which elements of the scene should be grouped together to form the object and how big those elements are when combined (Gori & Spillmann, 2010). The processes that resolve each of these questions in isolation from each other are fairly well understood, but their interaction is less studied. The goal of the study presented here is to investigate the principles that underlie the computation of perceived size for an object formed by the constructive processes of perceptual grouping. Specifically, we ask what characteristics of the local elements influence the perceived size of the perceptually grouped object to which they belong.

Because our perception of a three-dimensional world is derived from two-dimensional retinal images, it is fundamentally impossible to directly perceive the physical size of an object from the projected retinal image. The size of the retinal image depends both on the physical size of the object and its distance from the viewer (Emmert, 1881). To overcome this fundamental challenge, the visual system combines information about the retinal size of the object with other visual cues. For example, consider a familiar person in the distance – our percept of how tall they are will be based on a combination of retinal image size and

Citation: Boswell, A. M., Kohler, P. J., McCarthy, J. D., & Caplovitz, G. P. (2021). Perceived group size is determined by the centroids of the component elements. *Journal of Vision*, *21*(13):1, 1–12, https://doi.org/10.1167/jov.21.13.1.

https://doi.org/10.1167/jov.21.13.1 Received April 19, 2021; published December 1, 2021 ISSN 1534-7362 Copyright 2021, The Authors

other variables, including their perceived distance, their relative size compared to other people and objects nearby, and prior knowledge about their true height (Roberts, Harris, & Yates, 2005; Ross & Plug, 1998).

Intuitively, retinal size appears to be the most straightforward piece of information that the visual system needs to extract from the scene to compute perceived size. This is certainly true for objects where boundaries are explicitly designated, such as the continuous contour that defines the shape of a circle. Given the retinotopic organization of the early stages of visual processing it seems plausible that retinal image size for a simple stimulus, such as a circle, is relatively straightforward to represent. By operating, for example, on the spatial-extent, radius, area, or circumference, any number of analyses of the retinal image could provide direct access to its retinal size. It is well known, however, that not all of the objects we perceive are explicitly defined. A fundamental characteristic of our visual experience is that it is based in large part on constructive processes that can generate object representations from spatially and temporally disparate and localized visual elements (McCarthy, Erlikhman & Caplovitz, 2017; McCarthy, Kohler, Tse, & Caplovitz, 2015a; McCarthy, Strother, & Caplovitz, 2015b). This makes it possible for our visual system to deal with ambiguities brought on by occlusion and the limitations of our visual system. The visual image of a flower projected through a picket fence contains very little direct information about the shape or even existence of the flower. On the retina, the flower does not appear as a whole object. Instead, it is a series of segments that are separated by the fence. By integrating the disparate sources of information corresponding to the visible portions of the flower, the constructive processes of perceptual grouping serve an important role in enabling us to experience a stable and sufficiently accurate representation of the world around us – namely the presence of a flower with a given size and shape that is present behind the fence (Kellman & Shipley, 1991; McCarthy et al., 2015a; McCarthy et al., 2015b; McCarthy et al., 2017; Palmer, Kellman, & Shipley, 2006).

The way individual elements in the visual scene become grouped can depend on their shape or size. For example, when looking at a table which is set for dinner, the brain may group the objects by shape. Hence, glasses will be grouped together when a person is pouring the drinks to easily find the glass in the visual field. Similarly, plates may be grouped together by their circular shape while food is being served. Alternatively, salad plates can be grouped together in a separate group than dinner plates because of the difference in size. These grouping techniques are flexible so that we can use them to our advantage when interacting with objects in our visual field.

Objects are grouped by good continuation when occluded edges seem to meet behind an occluding object

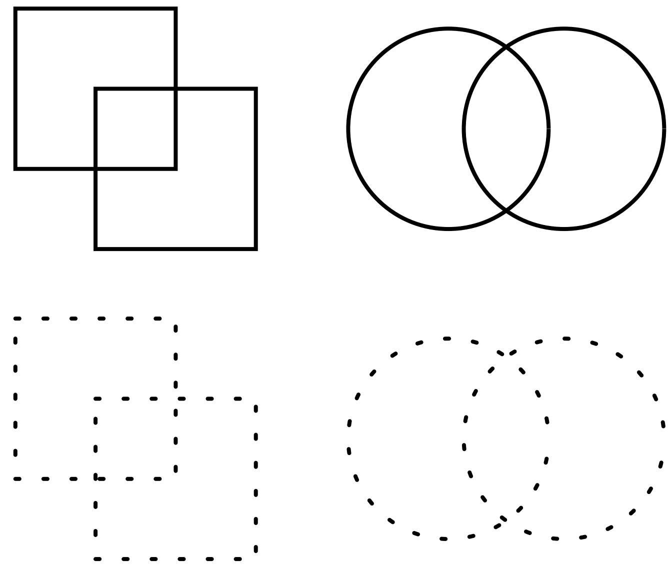

Figure 1. In the top row, objects are grouped by good continuation such that two squares are perceived instead of two abutting L-shapes, and two circles are perceived rather that abutting half-moons. This is true even when objects are not defined by continuous contours.

(Figure 1A). Importantly, this also occurs when objects are not defined by a continuous contour, but a grouping of individual elements (see Figure 1B). Grouping by similarity focuses on a specific feature of objects in order to group them together, such as object color or size (Wertheimer, 1938). Changing characteristics of these elements may alter the grouping paradigm.

Because the perceptual grouping depends on the feature characteristics of the individual elements, we are interested in investigating whether properties, such as size and shape, of a grouped object interact with characteristics of the local elements that define it. For example, will an object formed by grouping large elements appear larger than one formed out of small elements?

The following experiments were conducted to investigate if characteristics of individual elements can influence the perceived size of the objects they form through perceptual grouping. We hold many characteristics of the perceptually grouped object constant while systematically altering characteristics of the local elements. We hypothesize that the perceived size of an array may be determined by either (1) the distance between the center of the array and the outer edges, (2) the inner edges, or (3) the centers of the individual elements that make up the array. The results of these experiments provide insight into how the visual system constructs the perceived size of a perceptually grouped object.

Methods

Participants

Participants in each experiment (experiment 1: $n = 6$, experiment 2: $n = 6$, experiment 3: $n = 5$, experiment 4: $n = 5$, and experiment 5: $n = 4$) were students at the University of Nevada, Reno, Nevada. Each participant was naïve to the goals and specific designs of the experiments, reported normal or corrected-to-normal vision, and provided informed consent according to the guidelines of the Department of Psychology and the Institutional Review Board of the University of Nevada, Reno, Nevada. Some participants were given course credit for their participation. In total, eight individuals participated in the study with all but two participating in all five experiments.

Stimuli and general procedure

Stimuli were created using MATLAB (MathWorks Inc., Natick, MA, USA) and the Psychophysics Toolbox (Brainard, 1997), were presented on a 19 inch(refresh rate = 85 hz) CRT monitor and viewed at a distance of 57 cm. Stimuli consisted of circular (experiments 1 and 2) or square (experiments 3, 4, and 5) arrays of eight equally spaced white (109.4 cd/m^2) elements presented on a black background (0.055 cd/m^2). The elements themselves were either filled circles (experiment 1) or triangles (experiments 2–5). On every trial, a test array and reference array were simultaneously presented for 500 ms such that their centers were positioned 9 degrees (13.8 degrees for experiment 1) of visual angle to either the left or right of a centrally located fixation spot (0.2 degrees visual angle). On each trial, a small amount of random jitter (up to 1.4 degrees) was applied to the center of each array to ensure relative size judgments were formed by comparing the perceived size of the two arrays without any contribution from potentially confounding cues such as their positions relative to the edges of the monitor. Observers were instructed to maintain central fixation and, in a two alternative forced choice manner, indicate via press of a button which of the two arrays appeared larger: the one on the left or the one on the right. No further instructions were given to the subjects and no feedback was provided regarding their responses. The sides on which the test and reference arrays were presented were randomly determined for each trial. Each experiment applied the method of constant stimuli: on each trial, the size of the reference array was held constant, whereas the size of the test array was pseudo-randomly selected from a predetermined set of array sizes. The details of array and element sizes are provided below, as they are specific to each experiment.

Experiment 1: The influence of element size on the perceived size of a circular array

Experiment 1 was designed to dissociate three plausible hypotheses for how observers judge the size of an array. Specifically, the perceived size of an array may be determined by the distance between the center of the array and the outer edges, the inner edges, or the centers of the individual elements that make up the array. In order to dissociate between these hypotheses, this experiment examined the perceived size of arrays constructed from elements of different sizes: small, medium, and large. Each hypothesis makes a distinct prediction about how element size should influence the perceived size of the array. If observers use the inner edges of the elements, then when the distance from the center of the arrays to the centers of the elements are matched, the arrays with the small elements should be seen as the largest. If observers use the outer edges of the elements, then the test arrays with the large elements should be seen as the largest. Finally, if observers use the centers of the elements, then all three types of arrays should have the same apparent size.

In this experiment, the reference array consisted of eight, equally spaced, filled circles each with a radius of 0.95 degrees. Each element was positioned so that the distance between its center and the center of the circular array was exactly 6.0 degrees. For simplicity, we subsequently refer to the center-to-center distance as the radius of the array. As illustrated in Figure 2, three different sets of test arrays were examined: those consisting of small circles (0.16 degrees), medium circles (0.95 degrees), or large circles (1.74 degrees). The size of the medium circles was the same as the size of the circles that make up the reference array. This provides a natural control condition upon which to evaluate how accurately observers can discriminate array sizes in general. On any given trial, the radius of the test array was pseudo-randomly selected from one of nine

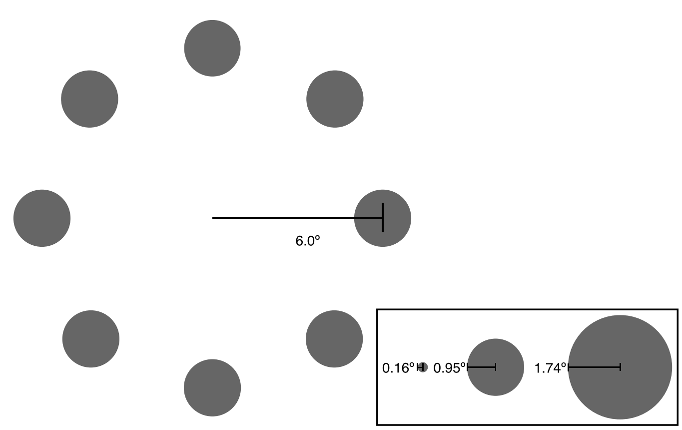

Figure 2. Arrays of circles used in experiment 1. The three circle sizes used are shown in the box.

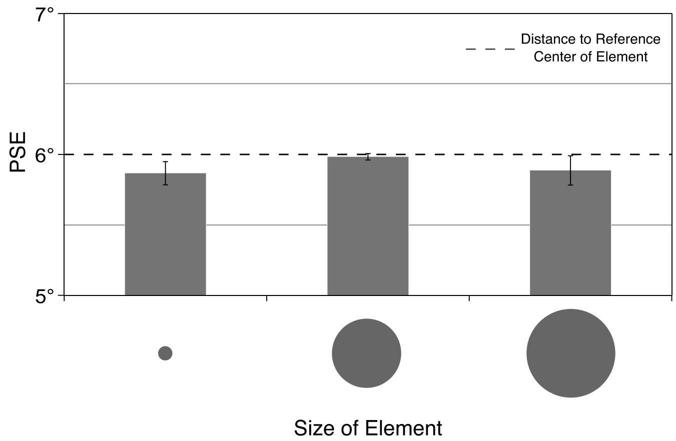

Figure 3. Results of experiment 1. Points of subjective equality averaged across participants. The circular elements used for the three conditions are shown below each bar plot. There was no difference between the different conditions, suggesting that participants were using the center of each element to judge the array size. Error bars are ±1 standard error of the mean.

possible sizes (4.70 degrees, 5.50 degrees, 5.80 degrees, 5.90 degrees, 6.00 degrees, 6.10 degrees, 6.20 degrees, 6.50 degrees, and 7.30 degrees) such that there were 20 trials of each size for each of the three circle-size conditions for a total of 540 trials.

For each participant, psychometric curves were derived by computing the percentage of trials, and the test array was perceived as larger than the reference array for each test array size. This was done independently for each of the three test-array conditions. Sigmoidal shaped curves $f(x) = 100x[e^{b1+xb2}/1+e^{b1+xb2}]$ were then fit to the average data for each subject. These curves were then used to interpolate, for each subject, the point of subjective equality (PSE; defined as the point at which the observer reported that the test array was larger 50% of the time) for each of the three test-array conditions. As can be observed in Figure 3, there are no differences among the PSEs derived from each of the three conditions (repeated measures ANOVA: $F(2,10) = 0.613$, $p = 0.561$, $\eta p2 = 0.109$). This result suggests that observers are likely using the distance between the center of the array and the centers of the elements, and not their inner or outer edges to judge the size of the array.

Experiment 2: Arrays made with triangular elements

The results of experiment 1 suggest observers use the center of the array elements when judging the overall size of an array. It is not clear, however, what constitutes the "center" of an array element. It could be the geometric center (i.e. centroid), but an alternative hypothesis is that observers use the average distance between the inner and outer edges of the array elements (i.e. half the distance between the inner and outer edges). Because the elements in experiment 1 were circles, for which both measures are identical, the results of the experiment cannot dissociate these two possibilities.

In this experiment, the circular elements were replaced with isosceles triangles. As illustrated in Figure 4A, when oriented radially, the centroid of an isosceles triangle is distinct from the average distance between its inner and outer edges (i.e. the half-height). Experiment 2 was designed to differentiate these two alternatives by investigating the perceived sizes of arrays constructed with inward or outward pointing triangles. If observers use the centroid to judge the size of the arrays, then the inward pointing triangle arrays should appear larger than outward pointing ones. If on the other hand, observers are using the half-height of the triangles, then the two sets of arrays should appear to have the same size.

The overall procedure was the same as experiment 1, with a few changes. Instead of the arrays consisting of filled circles, the elements that made up each array were filled isosceles triangles (1 degree base × 1 degree height). With these dimensions, the distance between the centroid and the half height of the triangle was 0.17 degrees. The triangles that formed the reference array always pointed inward toward the center of the array and the distance from the center of the reference array to the half-height of each triangle was held constant at 4 degrees on each trial. As illustrated in Figure 4B, there were two sets of test arrays: one in which the triangles pointed inward and one in which the triangles pointed outward. The choice to have the reference array consist of inward pointing triangles in this and later experiments was arbitrary. On each trial, the size of the test array as measured from the center of the array to the half-height point on any given triangle was pseudo-randomly chosen from this list: 2.00 degrees, 3.00 degrees, 3.50 degrees, 3.75 degrees, 4.00 degrees, 4.25 degrees, 4.50 degrees, 5.50 degrees, and 7.00 degrees. As in experiment 1, the sides on which the test and reference arrays were presented were randomly determined on each trial. There were 20 trials for each size for both sets of test array conditions, for a total of 360 trials.

As in experiment 1, the percentage of trials in which the test array was perceived as larger than the reference arrays was computed for each trial type. For each participant, this was done independently for each of the two test-array conditions. PSEs were derived using the same procedure as in experiment 1. As can be seen in Figure 5, the array of inward pointing triangles was perceived to be larger than the array of outward pointing triangles. A paired samples *t*-test of the PSEs of each participant indicated that there was a significant

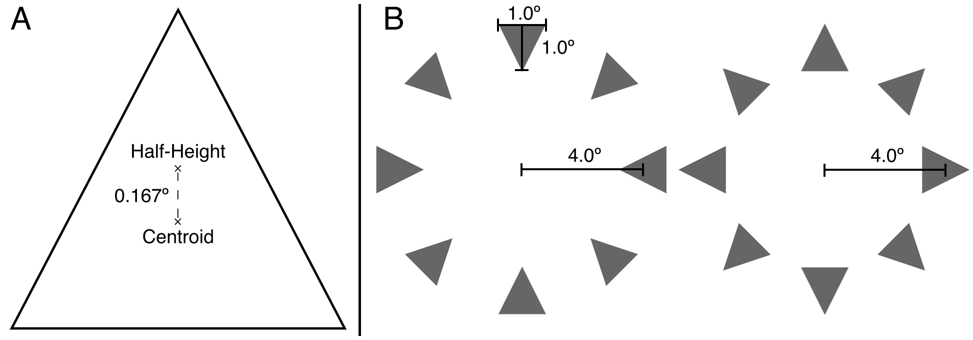

Figure 4. (**A**) The difference between the triangle's bisector midpoint (half-height) and it's centroid. (**B**) Example array from experiment 2, featuring circular arrays with triangular elements. There were two element conditions: pointing towards the center of the array or pointing away from the center.

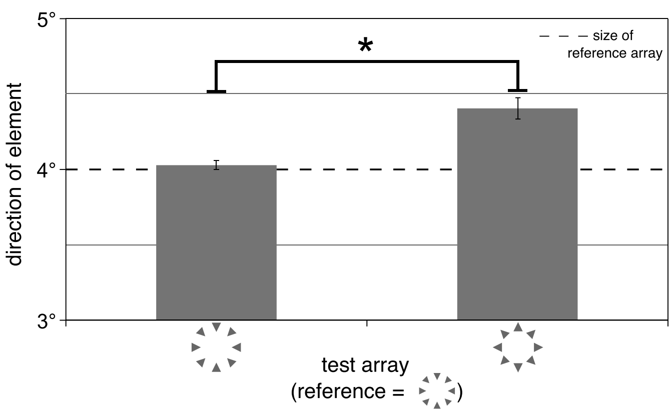

Figure 5. Results of experiment 2. Points of subjective equality averaged across participants. There was a significant difference between the two conditions, suggesting that arrays with outward pointing triangles were seen as being smaller than arrays with inward pointing triangles. Error bars are ±1 standard error of the mean.

difference between the arrays with inward pointing triangles (M = 4.032, SD = 0.072) and the arrays with outward pointing triangles (M = 4.408, SD = 0.174; t(5) = −3.989, $p = 0.010$, d = 1.628). The mean difference between the two PSEs was 0.376 degrees, approximately 10% of the reference array radius. Hence, an array of outward pointing triangles needs to be physically 10% larger than an array of inward pointing triangles in order to be perceived as having the same size, indicating that inward pointing triangles are perceived as larger. This suggests that observers are basing their judgments on the centroid rather than the half-height location of each triangle.

The observed difference of 0.376 degrees of visual angle is very close to the predicted difference if observers are basing their judgments on the centroids of the array elements (0.334 degrees for the triangles used here). Remarkably, the difference between the observed and predicted data corresponds to less than two screen pixels. The data thus provide very strong support for the conclusion that observers use the centroids of the elements to judge perceived size. The following experiments were designed to investigate whether the findings of experiments 1 and 2, derived for circular arrays, would generalize to noncircular arrays.

Experiment 3: Square arrays of triangles

In this experiment, the procedure was essentially the same as experiment 2 with a few changes. The shape of the arrays was changed to form a square instead of a circle (see Figure 4B). The array was still made of triangle shaped elements as in experiment 2, and the elements were again manipulated to either point toward or away from the center of the array. The size of each array was defined as the distance from the center of the array to the center (the midpoint of the bisector) of one corner object. The reference array always had inward pointing elements and its size was held constant across trials at 4.44 degrees of visual angle. The test array could be made up of either inward or outward pointing triangles and had sizes that were pseudo-randomly chosen from a list of nine (2.66 degrees, 3.55 degrees, 3.99 degrees, 4.22 degrees, 4.44 degrees, 4.66 degrees, 4.88 degrees, 5.33 degrees, and 6.21 degrees of visual angle). As illustrated in Figure 6A, in order to achieve the square shape, the non-corner elements were positioned such that the entire array could be inscribed by a square. As in experiments 1 and 2, there were 20 trials for each size for both sets of test array conditions for a total of 360 trials. The sides on which the test and reference arrays were presented were randomly determined on each trial.

Again, the point of subjective equality was calculated for each participant per type of stimulus. A paired samples *t*-test of the PSEs of each participant indicated that there was a significant difference between the

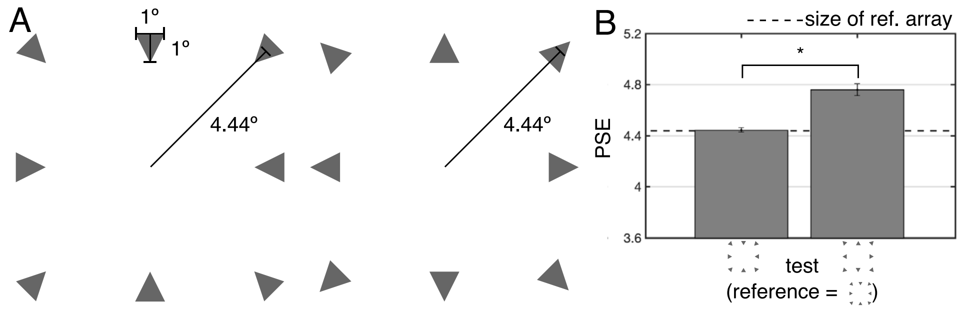

Figure 6. (**A**) Example array from experiment 3, featuring square arrays with triangular elements. There were two element conditions: either pointing toward or away from the center of the array. (**B**) Point of subjective equality for participants in experiment 3. There was a significant difference between the two conditions, suggesting that arrays with outward pointing triangles were seen as being smaller than arrays with inward pointing triangles. Error bars are ±1 standard error of the mean.

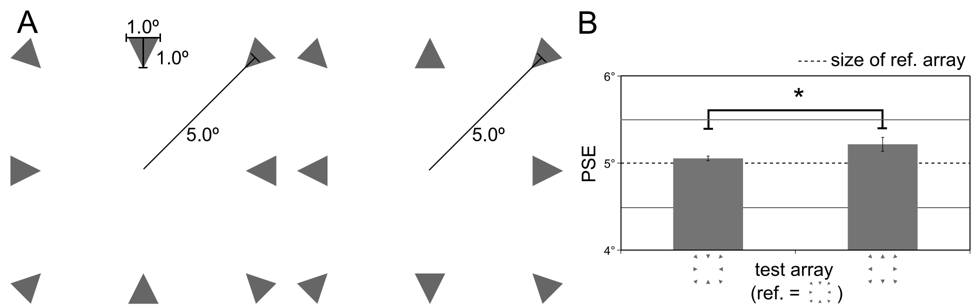

Figure 7. (**A**) Example array from experiment 4, featuring square arrays with triangular elements. There were two element conditions: either all elements pointing toward the center or the non-corner elements pointing away from the center of the array. (**B**) Point of subjective equality averaged across participants. There was a significant difference between the two conditions, suggesting that arrays with outward pointing non-corner triangle elements were seen as being smaller than arrays where all elements were pointing inward. The mean difference here is much smaller but consistent with what would be expected if a participant used the average distance of the element centroids. Error bars are ±1 standard error of the mean.

arrays with inward pointing triangles (M = 4.443, SD = 0.035) and the arrays with outward pointing triangles (M = 4.760, SD = 0.093; $t(4) = -8.369$, $p < 0.01$, d = 3.743), shown in Figure 6B. These results indicate that the array made of outward pointing triangles again appeared smaller than the array with inward pointing triangles. As in experiment 2, the mean difference (0.317 degrees) is very similar to what would be expected (0.334 degrees) if observers are basing their judgments on the centroid. This suggests that, at least for the case of a square array, the result derived in experiment 2 using circular arrays can be generalized to non-circular shapes.

Experiments 4 and 5: Arrays of elements with nonuniformly distributed centroids

These experiments are designed to determine how the perceived size of an array is computed when the centroids of the individual elements are not uniformly distributed (i.e. when the individual triangle elements are not all oriented in the same direction relative to the center and exterior of the array). At least three plausible hypotheses exist as to how this question may be resolved. It may be the case that the perceived size of the array is predicated on the distance between the center of the array and the location of the most distal, proximal, or the average across element of the centroids. Experiments 4 and 5 were designed to dissociate these hypotheses. In the following two experiments, we examine the case in which each of the test arrays are made up of both inward and outward pointing triangles.

Experiment 4 was essentially the same as experiment 3, except, as illustrated in Figure 7A, the test arrays either had all of their elements pointing inward or had their non-corner elements pointing outward. The reference array elements all pointed inward and had a size of 5.00 degrees visual angle, as measured from the center of the array to the bisector midpoint of the corner element. The size of the test array was chosen from a list of nine (3.00 degrees, 4.00 degrees, 4.25 degrees, 4.75 degrees, 5.00 degrees, 5.25 degrees, 5.75 degrees, 6.00 degrees, and 7.00 degrees). There were 20 trials for each size for both sets of test array conditions

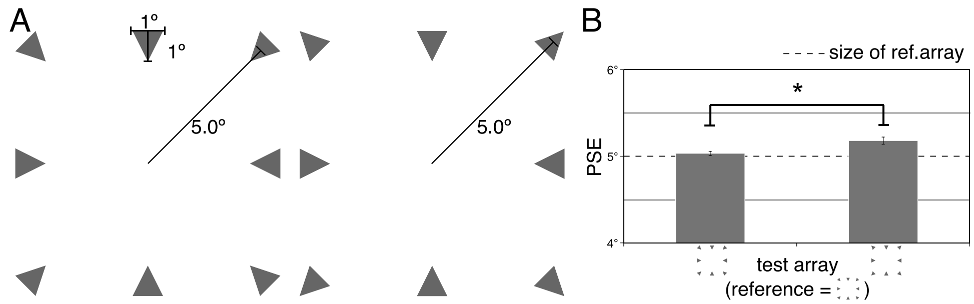

Figure 8. (**A**) Example array from experiment 5, featuring square arrays with triangular elements. There were two element conditions: either all elements pointing toward the center or the corner elements pointing away from the center of the array. (**B**) Point of subjective equality averaged across participants. There was a significant difference between the two conditions, suggesting that arrays with outward pointing triangles were seen as being smaller than arrays with inward pointing triangles. Like in experiment 4, the results here suggest that participants are using the average location of the centroids of the triangles to judge the size of the array. Error bars are ±1 standard error of the mean.

for a total of 360 trials. The sides on which the test and reference arrays were presented were randomly determined on each trial.

The point of subjective equality was calculated for each stimulus type. As seen in Figure 7B, a paired samples *t*-test of the PSEs from each participant indicated that there was a significant difference between the arrays with all inward pointing triangles (M = 5.053, SD = 0.041) and the arrays in which the non-corner elements pointed outward (M = 5.219, SD = 0.073; $t(4) = -6.059$, $p < 0.01$, $d = 2.710$). This observation allows us to rule out the possibility that the most distal centroids determine the perceived size of the array, because the most distal centroids are the same for both test arrays. In order to dissociate between the remaining hypotheses, we can compare the observed difference in perceived size to that which each hypothesis predicts. If the perceived size of an array is predicated on the locations of the most proximal centroids, then the difference in perceived size should be the same as that predicted for experiments 2 and 3, 0.334 degrees of visual angle. The observed mean difference here (0.166 degrees) is far from that, but very close to 0.167 degrees, the value expected if the perceived size of an array is based on the average distance between each element's centroid and the center of the array.

This final experiment was designed to test how well the results of experiment 4 generalize when the orientations of the corner elements are manipulated rather than the non-corner elements. As illustrated in Figure 8A, the test arrays here were composed of all inward pointed elements or had their corner elements pointed outward. In all other ways, the experiment was identical to experiment 4.

The point of subjective equality was calculated for each stimulus type. A paired samples *t*-test of the PSEs from each participant indicated that there was a significant difference between the arrays with all inward pointing triangles (M = 5.032, SD = 0.050) and the arrays with outward pointing center triangles (M = 5.183, SD = 0.082; $t(3) = -3.331$, $p < 0.05$, $d = 1.666$; see Figure 8B). As with experiment 4, the observed difference between the conditions (0.151 degrees) was closer to the average-distance hypothesis prediction (0.167 degrees) than to the prediction from the hypothesis that observers base their judgments using the most proximal centroids (0.334 degrees). This provides further support for the idea that when the centroids of the elements that make up an array are not uniformly distributed, the perceived size of the array is based on their average distance from the center of the array.

Bayesian model comparisons

In order to strengthen our interpretation of the five experiments, it is useful to do a more targeted model comparison. In experiment 1, the goal was to test two alternatives to the centroid model, namely that either (1) the inner edge or (2) the outer edge of the array elements is used to estimate array size. Both models predict differences in PSEs across the three different element sizes, which we do not observe ($p = 0.561$). However, our rejection of these alternative models is relying on a null result, with relatively few participants. To address this limitation, we directly tested the inner edge and outer edge predictions, by computing the predicted PSEs for the three element sizes based on the inner and outer edge models. We then subtracted these two sets of predicted PSEs from the observed PSEs, to generated two sets of adjusted PSEs that capture the deviation of the raw data from each model. If a model is correct, the adjusted PSEs should be the same (ideally 0) across conditions. For both models, a Bayesian repeated measures ANOVA finds decisive evidence ($BF_{10} > 100$) for the hypothesis that the adjusted PSEs are different across the three element sizes. The data from experiment 1 thus allows us to definitively rule out the inner edge and outer edge models, even if we cannot definitively say that the estimates are the same across element sizes, as predicted by the centroid model.

In experiments 2 to 5, we present further support for the centroid model. In experiments 2 and 3, the alternative model is that observers are using the half-height of the triangles, rather the centroid. The half-height model would result in the reference and test arrays appearing to have the same size, so the significant difference we observe between the two arrays disproves the half-height model. We can further quantify the extent to which the two models predict the observed results, by computing the PSE difference between reference and test, and subtracting the difference predicted by each hypothesis. These estimates now capture the deviation of the observed data from each model, and the better model is the one that produces estimates closer to zero. The centroid estimates (M = 0.124, SD = 0.118) are closer to zero than the half-height estimates (M = 0.349, SD = 0.175), and a Bayesian paired *t*-test on the absolute value of the estimates across both experiments find very strong evidence of a difference between the estimates (BF_{10} = 49.777), indicating that the centroid model is a better predictor of the data.

For experiments 4 and 5, the main model is that observers are using the average centroid distance, with alternative models being that observers base their judgments on (1) the most proximal centroids or (2) the most distal centroids. For both experiments, the distal model is disproved by the significant difference between the two arrays. We can use the same approach that we used for experiments 2 and 3 and test the average-centroid model against the distal and proximal models. Across both experiments, the average centroid estimates (M = 0.053, SD = 0.044) are closer to zero than the proximal (M = 0.175, SD = 0.071) and distal (M = 0.159, SD = 0.071) estimates. We find decisive (BF_{10} = 297.132) evidence of a difference between the average centroid and proximal hypotheses, and substantial evidence of a difference between the average centroid and distal hypotheses (BF_{10} = 4.174). Our model comparison thus provides strong evidence, across all five experiments, that the average centroid model offers a better fit to our data than any of the alternative models we consider.

Comparing precision of judgment across conditions

We deliberately did not provide explicit instructions to our participants about how to do the task. It is therefore possible that observers were using different strategies in different conditions, which might explain our pattern of results. To test for this possibility, we computed a measure of precision as the standard deviation of the fitted psychometric function for each participant in each experiment. If participants used the same strategy in all conditions, precision should be the same across conditions. We ran the same statistical tests on precision as those that were run on the PSEs. For experiment 1, we ran a repeated measures ANOVA on precision, which revealed significant differences across conditions (F(2,10) = 5.917, $p = 0.020$). Post hoc tests revealed that precision was higher for test arrays with medium circles than for test arrays of small circles (t(5) = 2.561, adjusted $p = 0.057$) and higher for test arrays with medium circles than test arrays with large element circles (t(5) = 3.512, adjusted $p = 0.025$). There was no significant difference between large and small element circles (t(5) = 0.708, adjusted $p =$ 0.495). There are several reasons why reduced precision may have been observed when the elements in the reference and test arrays were different sizes (reference medium and test small or large). One possibility is that participants were using a mixed strategy for making their judgments. For example, on some trials using the centroids, and on other trials using the inner or outer edges. Across multiple trials, such a mixed strategy could on average yield equivalent results as simply using the centroid on every trial. Alternatively, it could be that the process of extracting and comparing the centroids for larger and/or smaller elements, or making comparison between different elements more generally, is simply more difficult than doing so when all elements are the same. These alternative hypotheses can be more specifically examined in experiments 2 to 5 in which the array elements all had the same size. For experiments 2 to 5, paired *t*-tests comparing precision between the two conditions in each experiment, revealed no significant differences across four experiments (smallest $p = 0.077$). This shows that, for the array size judgments our participants are doing, comparing array sizes of triangles of the same or different orientation does not influence precision, but does influence PSE, as demonstrated above.

In summary, although we have some evidence that strategy used by participants may vary among the conditions in experiment 1, we have no such evidence for any of the other experiments, and it is difficult to come up with a reasonable model for how differences in strategy would produce the pattern of results observed for the PSE. We conclude that our results are likely not driven by a difference in strategy, and that in the absence of explicit instruction, observers appear to adopt the centroid strategy.

Discussion

This study investigated how the visual system determines the perceived size of perceptually grouped objects. The experiments were designed to determine whether the characteristics of local elements can affect the perceived size of grouped objects, using two different shaped global objects. The results unequivocally demonstrate that the visual system uses the centroid of local elements to compute the perceived size of the array of objects to which they belong. This finding indicates that local element shape does indeed contribute to global size perception, a novel contribution to the existing body of literature addressing how perceptually grouped objects are perceived.

The stimuli used in experiment 1 were circles for which the centroid is equidistant from all points along the contour. This means that, although the large and small elements were aligned on their centroids, they were also aligned on other geometric dimensions, including, as can be seen in Figure 2, the half-way point between the innermost and outermost portions of their contour (relative to the center of the array). Therefore, whereas the results of experiment 1 indicate that perceived size is not based on spatial extent, they leave open the possibility that the visual system is using other sources of information besides the centroid. We addressed this in subsequent experiments that used triangular elements which disambiguated the centroid from other geometric dimensions and found that perceived size could be almost perfectly predicted from the distance between the centroid of each array elements and the center of the array. In contrast, half-height, the half-way point between the innermost and outermost portions of element contours, was not predictive of perceived size. This result generalized across two different array shapes. The results of experiments 4 and 5 further indicate that when the elements were positioned in such a way that the centroids along the global contour were not all the same distance from the center of the array, perceived size was based on the average distance. The current results indicate that the centroid of the elements is fundamental to the visual system's analysis of size for perceptually grouped objects.

It is worth noting that these results are not what one might expect a priori. The results of experiment 1 show that element size has little or no effect upon the judgment of array size. This result is surprising because the spatial extent subtended by a group of large elements is greater than that subtended by smaller elements when they are aligned on their centers. When viewing these stimuli, most notably the circular array with triangular elements, the outer edges of the triangles form illusory contours in the shape of a circle. It would be reasonable to expect that judgments of the size of the array would be based on this contour. However, the Binding Ring illusion demonstrates that the outermost portion of a circular array of individual elements is perceived as smaller when superimposed with a continuous contour that intersects the centroids of individual elements (McCarthy, Kupitz, & Caplovitz, 2013). Providing a continuously defined contour thus appears to increase the influence of the centroid on size judgments of individually grouped elements, highlighting its importance in size perception. For all combinations of component elements and global shapes used in the experiments presented here, the centroids of the elements are more important to the perceived size of a grouped object than any illusory contour.

Perceptual grouping of local image elements into a global percept is fundamental for vision. The available evidence indicates that grouping processes are mediated by higher-level visual areas, such as the lateral occipital and parietal cortices (Grassi, Zaretskaya, & Bartels, 2018; McCarthy et al., 2015a; Murray, Kersten, Olshausen, Schrater, & Woods, 2002; Zaretskaya, Anstis, & Bartels, 2013), and often associated with downregulation of activity in early visual areas (Fang, Kersten, & Murray, 2008; Grassi et al., 2018; Murray et al., 2002), likely mediated by feedback connections (Grossberg, Mingolla, & Ross, 1997; Plewan, Weidner, Eickhoff, & Finsk, 2012). This is consistent with predictive coding models of visual perception in which representations of the component elements lower in the visual processing hierarchy are discarded once grouping has taken place (Rao & Ballard, 1999).

It is well-known that perceptual grouping can influence the mechanisms that operate on component elements and have drastic effects on judgments of perceived speed (Caplovitz & Tse, 2007; Kohler, Captovitz, & Tse, 2014; Kohler, Caplovitz, & Tse, 2009; Verghese & McKee, 2006; Verghese & Stone, 1996), motion fading (Kohler, Caplovitz, Hsieh, Sun, & Tse, 2010), and stereoscopic depth (Hou, Lu, Zhou, & Liu, 2006; Liu, Jacobs, & Basri, 1999; Mamassian & Zannoli, 2020).

The inverse question of how features of the component elements influence the global percept has been addressed more rarely, and often in the context of displays of moving objects. The perceived shape of an object formed by the grouping of an array of drifting Gabor patches can be influenced by the speed and direction of the component drift (Whitney, 2006). Similarly, the orientation of local Gaussian blobs translating across the visual field can influence the perceived shape of an object they are grouped into (McCarthy, Cordeiro, & Caplovitz, 2012). Moreover, adding translational motion to classic illusions, such as the Ebbinghaus Illusion, results in a more robust illusory size distortion (Mruczek, Blair, & Caplovitz, 2014). Experiments 2 to 5 in the current study shows that for a static display, a feature of the component elements, triangle orientation, can influence the perceived size of the global percept. Our analyses show that this is because the distance of the centroids from the center of the array is the relevant feature when computing perceived size of the array.

Why is the centroid so important to the visual system? Perhaps its utility is ecologically derived, allowing us to intrinsically compute the centroid location for objects (Bulatov, Bertulis, Bulatova, & Loginovich, 2009; Bulatov, Bulatova, Loginovich, & Surkys, 2015a; Bulatov, Bulatova, Surkys, & Mickiene, 2015b). Indeed, the centroid is important for accurate motion detection of grouped elements (Morgan, Ward, & Cleary, 1994) and visual acuity (Watt, Morgan, & Ward, 1983a; Watt, Morgan, & Ward, 1983b). Rapid access to this information would perhaps be beneficial as input to motor planning of interactions with and manipulations of the object in our gravitational environment (Baud-Bovy & Soechting, 2001). This hypothesis is contradicted, however, by evidence showing that eye fixations were more attracted to the centroid during a viewing task than during a grasping task, and that secondary fixations during grasping tended to go toward grasping locations (Brouwer, Franz, & Gegenfurtner, 2009). Eye movements during natural viewing target the centroid with a high degree of accuracy (Vishwanath & Kowler, 2003) consistent with the idea that the centroid is generally important for vision. A similar and perhaps even more dramatic demonstration of the relative importance of the centroid can be found in observers' inability to identify the half-height of an object in which the centroid and half-height are at two different locations (Anstis, Gregory, & Heard, 2009). There are results suggesting that identifying the centroid location for a cluster of dots adds no additional noise to the discrimination of individual dot locations (Harris & Morgan, 1993; Morgan & Glennerster, 1991; Vos, Bocheva, Yakomoff, & Helsper, 1993). Conversely, however, there is controversy about how accurately individuals can perceive the average size of individual objects in an array (Ariely, 2001; Ariely, 2008; Chong, Joo, Emmmanouil, & Treisman, 2008; Myczek & Simons, 2008). Specifically, visual noise impairs mean judgments of element size and increases logarithmically as the number of elements is increased (Solomon, Morgan, & Chubb, 2011). Hence, whereas estimates of mean orientation and size for individual elements appear to be impacted by visual noise, computation of the centroid is less effortful and less susceptible to sources of visual noise both for individual objects and across ensembles of objects. The current study adds to these findings by showing that the centroid is also critical for computing the perceived size of grouped objects. We did not provide explicit instruction in our experiments, and it is possible that observers could be made to adopt a non-centroid strategy if such instruction were to be given, but we suggest that the easy availability of the centroid may in fact interfere with reporting of non-centroid features, such as the half-height.

It has been proposed that the visual system uses summary statistics to extract the "gist" of a visual scene prior to processing individual elements. This ensemble encoding rapidly groups elements of the visual scene based on shared features to create a global representation (Cant & Xu, 2015) capable of guiding perception before individual scene elements have been analyzed. These grouping processes have been demonstrated to allow for rapid extraction of the centroid in spatially separated elements to provide an estimate of the center of perceptually grouped objects (Morgan & Glennerster, 1991). Moreover, observers can accurately point to briefly exposed dot ensembles after only a brief exposure, again pointing to the importance of the centroid when interacting with environmental objects (Vos et al., 1993). These studies all used circular elements, like the current experiment 1, which raises the question of whether it might be possible to bias extraction of the ensemble centroid by using triangles as the individual elements. The current results suggest that the perceived location of the ensemble centroid would change depending on the orientations of the triangle elements. This is a yet unaddressed question and may provide additional insight into the ecological value of the centroid when interacting with objects via goal-directed action.

Finally, we would like to highlight a possible practical implication of these findings, in the context of designing stimuli consisting of arrays of elements. Software packages, such as the Psychophysics Toolbox, render shapes within predefined bounding boxes. It is common to position the shapes using the center of the bounding box, consistent with a "half-distance" definition of center. Our results suggest that a better practice would be to use the centroid of the shape.

In conclusion, our results confirm that local elements do indeed influence the perceived size of a perceptually grouped object. Perception of grouped objects is a more complicated task than perceiving explicitly defined objects and, as such, the methods by which the visual system analyses these objects are not yet fully understood. The results presented here contribute to a growing body of evidence with the intention of assisting in elucidating the components of this process.

Keywords: perceptual grouping, size perception, local-global interactions

Acknowledgments

The authors thank two anonymous reviewers whose comments helped improve and clarify the manuscript.

Supported by EPSCoR Research Infrastructure awards from the National Science Foundation to GPC under Award Numbers 1632849 and 1632738, by York University's Vision Science to Applications

(VISTA) program funded by the Canada First Research Excellence Fund (CFREF, 2016–2023) and by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada awarded to P.J.K.

Commercial relationships: none.
Corresponding author: Gideon P. Caplovitz.
Email: gcaplovitz@unr.edu.
Address: Department of Psychology, Mack Social Science Bldg., University of Nevada, Reno, 1664 N. Virginia Street, Reno, NV 89557, USA.

References

Anstis, S., Gregory, R., & Heard, P. (2009). The Triangle-Bisection Illusion. *Perception, 38*, 321–332.

Ariely, D. (2008). Better than average? When can we say that subsampling of items is better than statistical summary representations? *Perception Psychophysics, 70*, 1325–1326.

Ariely, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychological Science, 12*, 157–162.

Baud-Bovy, G., & Soechting, J. (2001). Visual localization of the center of mass of compact, asymmetric, two-dimensional shapes. *Journal of Experimental Psychology. Human Perception and Performance, 27*, 692–706.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*, 433–436.

Brouwer, A.-M., Franz, V. H., & Gegenfurtner, K. R. (2009). Differences in fixations between grasping and viewing objects. *Journal of Vision. 9*, 18.

Bulatov, A., Bertulis, A., Bulatova, N., & Loginovich, Y. (2009). Centroid extraction and illusions of extent with different contextual flanks. *Acta Neurobiologiae Experimentalis (Warsaw, Poland), 69*, 504–525.

Bulatov, A., Bulatova, N., Loginovich, Y., & Surkys, T. (2015a). Illusion of extent evoked by closed two-dimensional shapes. *Biological Cybernetics, 109*, 163–178.

Bulatov, A., Bulatova, N., Surkys, T., & Mickienė, L. (2015b). A quantitative analysis of illusion magnitude changes induced by rotation of contextual distractor. *Acta Neurobiologiae Experimentalis (Warsaw, Poland), 75*, 238–251.

Cant, J. S., & Xu, Y. (2015). The Impact of Density and Ratio on Object-Ensemble Representation in Human Anterior-Medial Ventral Visual Cortex. *Cerebral Cortex (New York, N.Y.: 1991), 25*, 4226–4239.

Caplovitz, G. P., & Tse, P. U. (2007). Rotating dotted ellipses: motion perception driven by grouped figural rather than local dot motion signals. *Vision Research, 47*, 1979–1991.

Chong, S. C., Joo, S. J., Emmmanouil, T.-A., & Treisman, A. (2008). Statistical processing: Not so implausible after all. *Perception Psychophysics, 70*, 1327–1334.

Emmert, E. (1881). Grossenverhaltnisse der nachbilder. *Klinische Monatsblätter für Augenheilkunde, 19*, 443–450.

Fang, F., Kersten, D., & Murray, S.O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision, 8*, 2.1–2.9.

Gori, S., & Spillmann, L. (2010). Detection vs. grouping thresholds for elements differing in spacing, size and luminance. An alternative approach towards the psychophysics of Gestalten. *Vision Research, 50*, 1194–1202.

Grassi, P. R., Zaretskaya, N., & Bartels, A. (2018). A Generic Mechanism for Perceptual Organization in the Parietal Cortex. *Journal of Neuroscience, 38*, 7158–7169.

Grossberg, S., Mingolla, E., & Ross, W. D. (1997). Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends in Neurosciences, 20*, 106–111.

Harris, J. M., & Morgan, M. J. 1993. Stereo and motion disparities interfere with positional averaging. *Vision Research, 33*, 309–312.

Hou, F., Lu, H., Zhou, Y., & Liu, Z. (2006). Amodal completion impairs stereoacuity discrimination. *Vision Research, 46*, 2061–2068.

Kellman, P. J., & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology, 23*, 141–221.

Kohler, P. J., Caplovitz, G. P., Hsieh, P.-J., Sun, J., & Tse, P. U. (2010). Motion fading is driven by perceived, not actual angular velocity. *Vision Research, 50*, 1086–1094.

Kohler, P. J., Caplovitz, G. P., & Tse, P. U. (2014). The global slowdown effect: Why does perceptual grouping reduce perceived speed? *Attention, Perception Psychophysics, 76*(3), 780-792.

Kohler, P. J., Caplovitz, G. P., & Tse, P. U. (2009). The whole moves less than the spin of its parts. *Attention Perception Psychophysics, 71*, 675–679.

Liu, Z., Jacobs, D. W., & Basri, R. (1999). The role of convexity in perceptual completion: beyond good continuation. *Vision Research, 39*, 4244–4257.

Mamassian, P., & Zannoli, M. (2020). Sensory loss due to object formation. *Vision Research, 174*, 22–40.

McCarthy, J. D., Cordeiro, D., & Caplovitz, G. P. (2012). Local form–motion interactions influence global form perception. *Attention, Perception Psychophysics, 74*, 816–823.

McCarthy, J. D., Erlikhman, G., & Caplovitz, G. P. (2017). Chapter 8 - The maintenance and updating of representations of no longer visible objects and their parts, In: C.J. Howard (Ed.), *Progress in Brain Research, Temporal Sampling and Representation Updating*. New York, NY: Elsevier, pp. 163–192.

McCarthy, J. D., Kohler, P. J., Tse, P. U., & Caplovitz, G. P. (2015a). Extrastriate Visual Areas Integrate Form Features over Space and Time to Construct Representations of Stationary and Rigidly Rotating Objects. *Journal of Cognitive Neuroscience, 27*, 2158–2173.

McCarthy, J. D., Kupitz, C., & Caplovitz, G. P. (2013). The Binding Ring Illusion: assimilation affects the perceived size of a circular array. *F1000Research, 2*, 58.

McCarthy, J. D., Strother, L., & Caplovitz, G. P. (2015b). Spatiotemporal Form Integration: Sequentially presented inducers can lead to representations of stationary and rigidly rotating objects. *Attention Perception Psychophysics, 77*, 2740–2754.

Morgan, M. J., & Glennerster, A. (1991). Efficiency of locating centres of dot-clusters by human observers. *Vision Research, 31*, 2075–2083.

Morgan, M. J., Ward, R. M., & Cleary, R. F. (1994). Motion displacement thresholds for compound stimuli predicted by the displacement of centroids. *Vision Research, 34*, 747–749.

Mruczek, R. E. B., Blair, C. D., & Caplovitz, G. P. (2014). Dynamic illusory size contrast: A relative-size illusion modulated by stimulus motion and eye movements. *Journal of Vision, 14*, 2.

Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 15164–15169.

Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. *Perception Psychophysics, 70*, 772–788.

Palmer, E. M., Kellman, P. J., & Shipley, T. F. (2006). A theory of dynamic occluded and illusory object perception. *Journal of Experimental Psychology: General, 135*, 513–541.

Plewan, T., Weidner, R., Eickhoff, S. B., & Fink, G. R. (2012). Ventral and Dorsal Stream Interactions during the Perception of the Müller-Lyer Illusion: Evidence Derived from fMRI and Dynamic Causal Modeling. *Journal of Cognitive Neuroscience, 24*, 2015–2029.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*, 79–87.

Roberts, B., Harris, M. G., & Yates, T. A. (2005). The Roles of Inducer Size and Distance in the Ebbinghaus Illusion (Titchener Circles). *Perception, 34*, 847–856.

Ross, H. E., & Plug, C. (1998). The history of size constancy and size illusions, in: *Perceptual Constancy: Why Things Look as They Do*. New York, NY, USA: Cambridge University Press, pp. 499–528.

Solomon, J. A., Morgan, M., & Chubb, C. (2011). Efficiencies for the statistics of size discrimination. *Journal of Vision, 11*, 13.

Verghese, P., & McKee, S. P. (2006). Motion grouping impairs speed discrimination. *Vision Research, 46*, 1540–1546.

Verghese, P., & Stone, L. S. (1996). Perceived visual speed constrained by image segmentation. *Nature, 381*, 161–163.

Vishwanath, D., & Kowler, E. (2003). Localization of shapes: eye movements and perception compared. *Vision Research, 43*, 1637–1653.

Vos, P. G., Bocheva, N., Yakimoff, N., & Helsper, E. (1993). Perceived location of two-dimensional patterns. *Vision Research, 33*, 2157–2169.

Watt, R. J., Morgan, M. J., & Ward, R. M. (1983a). The use of different cues in vernier acuity. *Vision Research, Neural Interactions in the Vertebrate Retina, 23*, 991–995.

Watt, R. J., Morgan, M. J., & Ward, R. M. (1983b). Stimulus features that determine the visual location of a bright bar. *Investigative Ophthalmology Visual Science, 24*, 66–71.

Wertheimer, M. (1938). Laws of organization in perceptual forms, in: *A Source Book of Gestalt Psychology*. London, England: Kegan Paul, Trench, Trubner & Company, pp. 71–88.

Whitney, D. (2006). Contribution of bottom-up and top-down motion processes to perceived position. *Journal of Experimental Psychology. Human Perception and Performance, 32*, 1380–1397.

Zaretskaya, N., Anstis, S., & Bartels, A. (2013). Parietal Cortex Mediates Conscious Perception of Illusory Gestalt. *Journal of Neuroscience, 33*, 523–531.