Research Article

Psychological Science
21(4) 560–567

DOI: 10.1177/0956797610363543
http://pss.sagepub.com

Perceptually Averaging in a Continuous Visual World: Extracting Statistical Summary Representations Over Time

Alice R. Albrecht and Brian J. Scholl
Yale University

Abstract

Beyond processing individual features and objects, the visual system can also efficiently summarize scenes—for example, allowing observers to perceive the average size of a group of objects. Extraction of such *statistical summary representations* (SSRs) is fast and accurate, but researchers do not yet have a clear picture of the circumstances in which they operate. Previous studies have always used discrete input—either spatial arrays of shapes or temporal sequences of shapes presented one at a time. Real-world environments, in contrast, are intrinsically continuous and dynamic. We investigated the ability to compute average size in displays of objects (or sometimes a single object) that changed continuously, expanding and contracting over time. The results indicated that perceptual averaging can operate continuously in dynamic displays—sampling multiple times during a single continuous transformation with no discrete boundaries. Moreover, some dynamic changes (expansion) influence the resulting perceptual averages more than others (contraction), perhaps because of attentional capture. These results collectively illustrate how SSRs may be well adapted to dynamically changing real-world environments.

Keywords

ensemble features, size perception, attentional capture, looming

Received 5/4/09; Revision accepted 7/23/09

It is natural to think about visual perception in terms of the processing of individual elements, such as features, objects, and scenes. In real-world visual experience, however, it is also useful to summarize the distribution of information averaged across many such elements. For example, suppose you are picking apples at an orchard. In choosing a tree, it is not sufficient to evaluate the size and ripeness of a single apple, and it may not be feasible to evaluate each apple on every tree. Instead, it is most useful to summarize each candidate tree—for example, in terms of the average size and ripeness of its apples. Such averaging can be done numerically (for a review, see Pollard, 1984), but here we are interested in the possibility that these types of ensemble features, which we call *statistical summary representations* (SSRs), can be extracted efficiently in the course of visual processing. In fact, the visual system seems able to extract SSRs—with surprising speed and accuracy—over many different dimensions of visual scenes, including size (e.g., Ariely, 2001), length (e.g., Weiss & Anderson, 1969), inclination (e.g., Miller & Sheldon, 1969), motion direction (e.g., Dakin & Watt, 1997), speed (e.g., Watamaniuk & Duchon, 1992), orientation (e.g., Parkes, Lund, Angelucci, Solomon, & Morgan, 2001), and spatial position (e.g., Alvarez & Oliva, 2008). The visual system can even extract higher-level properties, such as the average emotion of a group of faces (e.g., Haberman & Whitney, 2007, 2009). For the purposes of this article, we focus on perception of average size.[1]

An Example: Perception of Average Size

The phenomenon of size averaging has spurred a considerable amount of recent research. For example, when faced with a display of discs of varying sizes, observers are able to accurately report their average size, even while unable to judge whether particular individual sizes were present (Ariely, 2001). This phenomenon also works as a striking demonstration: This task can sound daunting, but it is surprisingly easy to simply *see* the average. This ability is even more fascinating given that it persists when the display is presented very briefly (down to 50 ms), and with considerable variation in the

Corresponding Authors:
Alice R. Albrecht or Brian J. Scholl, Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205
E-mail: alice.albrecht@yale.edu or brian.scholl@yale.edu

number and density of discs, and the types of distributions from which their sizes are drawn (Chong & Treisman, 2003, 2005b). This ability is present even in patients who cannot consciously perceive more than one or two objects at a time (Demeyere, Rzeskiewicz, Humphreys, & Humphreys, 2008).

Additional research has shown that this ability can operate selectively, averaging over only subsets of objects (e.g., those of a particular color; Chong & Treisman, 2005b), but cannot operate as efficiently over multiple dimensions simultaneously (Emmanouil & Treisman, 2008). Size averaging is also influenced both by size illusions (which bias the input to the process; Im & Chong, 2009) and by attention (being facilitated by global but not focused attention, and being biased toward individual attended items; Chong & Treisman, 2005a; de Fockert & Marchant, 2008). The algorithm by which averaging occurs remains uncertain, however, as does the number of individual objects that are actually sampled (Ariely, 2008; Chong, Joo, Emmanouil, & Treisman, 2008; Myczek & Simons, 2008).

Some authors have suggested that the averaging process in some experiments may involve sampling only a few items in the display (Myczek & Simons, 2008). Others have countered that this possibility would require different post hoc heuristics for different experiments, that it fails to account for the inaccurate individual membership judgments, and that participants are actually quite poor at implementing such heuristics when asked directly to do so (Chong et al., 2008). We also note that such heuristics would have to operate unconsciously under some circumstances, as it seems possible to average without consciously employing them. In any case, it seems critical to carefully consider the possible operation of such heuristics in averaging experiments—but, as noted later, such heuristics do not apply in any direct fashion to the present study, in which there was often only a single size present in the display at any one time. Of course, sampling is still an important (and even necessary) part of averaging over time, but such stimuli preclude the operation of many "shortcut" strategies, such as simply sampling the largest and smallest items and averaging only those.

The Current Study: Averaging in Continuous Environments

To date, research on size averaging has always employed discrete objects, presented either in static spatial arrays or in temporal sequences, one at a time. The real world, in contrast, is intrinsically continuous and dynamic. Imagine, for example, that you see a herd of animals nearby on the savannah. Are they heading toward you? To answer this question, it would be more helpful to perceive the average direction of motion of the herd than that of a single animal (which could be misleading). In fact, SSRs are computed over dynamic information, such as motion direction, and speed (Dakin & Watt, 1997; Watamaniuk & Duchon, 1992; Williams & Sekuler, 1984). Beyond this, though, it is also crucial in many natural situations for such processes to operate continuously over time, for at least two reasons. First, the averages being computed may often be moving targets (literally, in the case of the herd of animals, which may suddenly turn your way!), and an average that is only seconds old may already be obsolete. Second, even with static features, such as size, observers often will not encounter the relevant population in a single glance. As a result, averages may have to pool across time as well as space, and over individual elements that are in flux.

Whereas some previous studies have explored perceptual averaging over time (Chong & Treisman, 2005a; Im & Chong, 2009; Morgan, Watamaniuk, & McKee, 2000; Oriet & Corbett, 2008; Weiss & Anderson, 1969), they have always used unchanging discrete elements presented one at a time. In contrast, the present study explored the ability to compute average sizes of continuously changing populations, such as a single object that is continuously expanding and contracting. Such computation could be especially challenging because averaging requires some discrete samples, yet there is no reason to expect that such discrete representations would be extracted from a continuously changing stimulus of this sort during the course of visual processing. In particular, note that within a given expansion or contraction of a continuously changing object, there is no discrete feature or transient available to trigger the creation of a discrete visual representation. The only discrete moments in such stimuli (beyond the initial and final frames) are those sizes that serve as *inflections* in the size profile of the object over time (i.e., the final size of an expansion that is followed by a contraction, or the final size of a contraction that is followed by an expansion). We call these local minima and maxima in our study the "anchor points" of the display, and one possibility is that perceptual averaging of continuously changing stimuli is mediated by sampling only the discrete anchor points. True continuous averaging, in contrast, would require sampling the size of the object even during individual transformations.

In Experiment 1, we tested whether perceptual averages could be constructed from "moving targets": Observers viewed a single disc that changed size continuously over time. In Experiment 2, we explored whether observers were sampling continuously from a single disc as it changed size. Finally, in Experiment 3, we investigated whether the averaging process is biased to sample selectively from certain types of dynamic transformations more than others. Inspired by studies of attentional capture due to looming stimuli, we tested whether averages were biased toward the sizes of expanding objects, compared with those of contracting objects. (Note that this is an example of a question that could not be investigated with the previous experiments using discrete objects.) Collectively, these studies tested how well adapted this type of SSR is to dynamically changing real-world environments.

Experiment 1: Averaging Values That Are Moving Targets

In an initial test of whether continuous transformations impair the computation of average size, observers were presented with a single disc that continuously oscillated among different anchor sizes, and were asked to judge its average size over time.[2]

Method

Participants. Ten naive observers with normal or corrected-to-normal visual acuity participated in a 1-hr session in exchange for course credit. Observers were offered an extra $5 "prize" if their average accuracy exceeded 90% (a challenging criterion achieved by a single observer).

Apparatus. The stimuli were presented on a uniform gray background using custom software written using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997). Observers sat without head restraint approximately 50 cm from a 17-in. monitor in a dark room.

Stimuli. Each trial contained three types of discs: (a) a white *test disc* that changed size over time and served as the target of observers' averaging, (b) a blue *judgment disc* that observers adjusted to make their response (always presented at the center of the display, with an initial diameter that varied randomly from 0.88° to 7°), and (c) a red *feedback disc* that subsequently appeared to indicate the correct average size. The test disc expanded and contracted continuously over time, moving through nine anchor points, each of which specified a different absolute diameter. On every trial, the smallest possible anchor point was chosen randomly from between 0.7° and 4.56°, and the largest from between 4.56° and 8.8°. The nine actual anchor points were then randomly chosen from 24 linearly spaced disc sizes within this overall range. On half of the trials, the test disc was always presented in a single central location. On the other half, the spatial position of this disc varied randomly over time, such that the disc moved linearly and continuously to a new random position inside an imaginary bounding box (4.23° × 1.88°) as its size passed through each anchor point.

Procedure. The test disc was presented at the beginning of each trial for 250 ms, after which it expanded and contracted a total of eight times, moving through nine anchor points. Each transformation took 750 ms, regardless of the magnitude of the difference between the starting and ending anchor points, and the size always changed at a linear rate. When the test disc reached its final anchor point, it disappeared and was replaced by a uniform gray screen for 1 s, and then the judgment disc appeared in the center of the display.

Observers were instructed to adjust the judgment disc until it matched the average size of the test disc over time. They adjusted the size of the judgment disc via key presses corresponding to either large (0.7°) or small (0.14°) adjustments, and then pressed another key to record their response. As a means of ensuring that observers made careful adjustments, a response on a given trial was not accepted until the observer had made at least one large and one small adjustment. After a response was recorded, two types of feedback were immediately provided: The feedback disc appeared in the display (either atop or behind the judgment disc, so that both were visible), indicating the correct continuous average, and two percentages appeared near the top of the display, indicating the observer's accuracy for that trial and overall accuracy for the experiment up to that point. The feedback remained visible until a key was pressed, and then a uniform gray screen was presented for 1 s before the next trial began.[3]

Observers completed 200 trials (preceded by 4 practice trials, the results of which were not recorded) and were encouraged to take a break every 50 trials, as indicated by periodic alert screens.

Results and discussion

Averaged across trials, accuracy was roughly equivalent to that reported in previous studies: The absolute value of the deviation between the reported and correct average sizes was 12% (SD = 2%). These deviations did not differ depending on whether the test disc remained stationary or moved randomly, $t(9) = 0.005$, $p > .99$. They were composed of both underestimates and overestimates, such that the average deviation taking the direction of the errors into account was an overestimate of only 0.8% (SD = 4%).

The observers' task was to average continuously, and thus the correct response on each trial was determined by averaging the size of the test disc across each frame of the dynamic display. However, given the nature of these displays, this continuous average (5.05°) did not differ from the average of the nine anchor points alone (5.05°). As a result, this initial experiment was not meant to demonstrate any sort of continuous averaging per se: Observers' responses were accurate, but could have been based on only the nine anchor points. What this initial experiment does show, though, is that the averaging process (whatever its input in this experiment) continues to operate in the presence of continuous dynamic transformations, and does not require discrete stimuli as input.[4] This demonstration allowed us to explore, in Experiment 2, whether observers are able to average continuously, taking into account more information than the nine anchor points.

Experiment 2: Continuous Averaging

In Experiment 1, observers had to report the average size of a single changing disc. Over what population did this averaging process operate? It is impossible to tell whether observers averaged continuously, or only over the nine anchor points, because the resulting correct averages were necessarily so similar. In Experiment 2, we unconfounded these possibilities by manipulating the speed at which the test disc transformed.

We divided each transformation—either an expansion or a contraction from one anchor point to the next—into two halves, and varied how much time each half took. In particular, one half of each transformation always took five times as long as the other half. In the *larger-continuous-average* condition, the longer half of each transformation was always the half that had the larger continuous mean (i.e., the first half for contractions and the second half for expansions). In the *smaller-continuous-average* condition, these values were reversed: The longer half of each transformation was the second half for contractions and the first half for expansions. These two conditions, depicted in Figure 1, were realized via matched pairs of trials (encountered at different random times during the experiment), with both trials in a pair using the same nine anchor points.

As in Experiment 1, observers were asked to report the continuous average—that is, the size of the test disc averaged across each frame of the display. If, despite these instructions, observers' averages were computed only over the anchor points, then the speed manipulations would not alter their responses (because the anchor points were identical in the two conditions). However, if observers' averages were computed by sampling more continuously—over more than the anchor points—then their responses would be larger for the larger-continuous-average condition and smaller for the smaller-continuous-average condition.

Method

This experiment was identical to Experiment 1 except as noted here. Twelve new observers participated in exchange for either course credit or monetary compensation. Observers completed 160 experimental trials (preceded by 4 practice trials, the results of which were not recorded)—80 larger-continuous-average trials and 80 smaller-continuous-average trials (using identical anchor points), all randomly ordered. Each transformation lasted 1,200 ms, with one half lasting 1 s, and the other half lasting 200 ms. In the larger-continuous-average condition, the longer half was always the second half for expansions and the first half for contractions. In the smaller-continuous-average condition, the longer half was always the first half for expansions and the second half for contractions.

Results and discussion

One observer was excluded from the analyses for having an overall error magnitude more than 2.5 standard deviations from the mean. Because each set of anchor points was used twice—once for each condition—the overall correct average size of the nine anchor points was identical for the two conditions (5.08°). As a result of the speed manipulation, however,

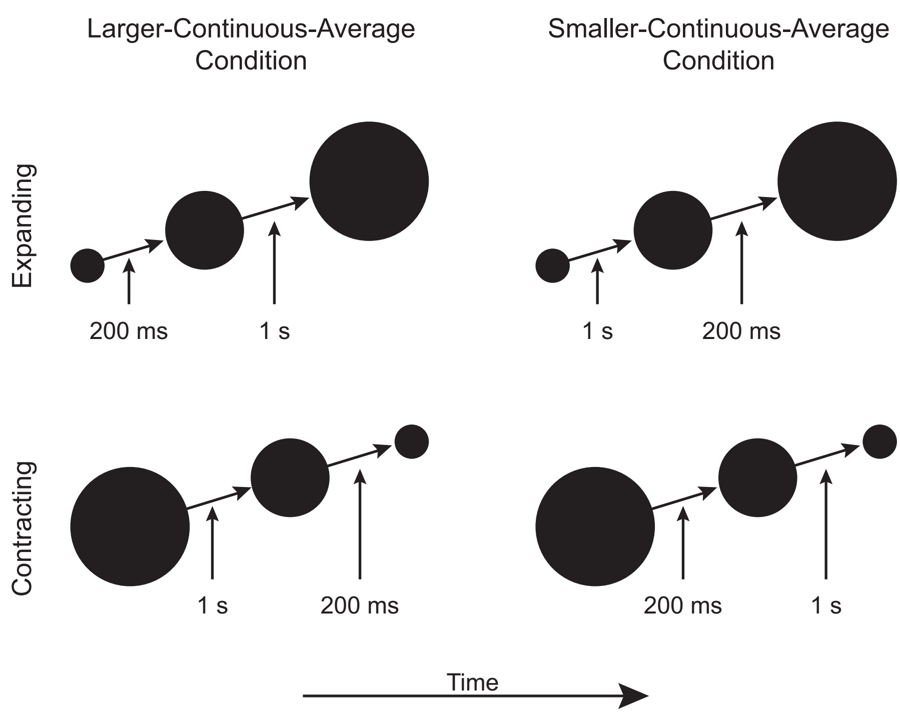

Fig. 1. Depictions of the dynamic manipulations used to test continuous averaging in Experiment 2. Each trial contained a single disc that changed continuously over time. Each transformation (either an expansion or a contraction) was divided into two halves—one of which took much longer than the other, so that the disc spent relatively more time being larger (in the larger-continuous-average condition) or smaller (in the smaller-continuous-average condition). See the text for details. Each trial involved eight such transformations, and the stimulus (shown here in black) was a white changing disc displayed on a uniform gray background.

the overall correct *continuous* averages differed between the larger-continuous-average condition (5.35°) and the smaller-continuous-average condition (4.79°). The key question in this experiment was whether observers' reported continuous averages would differ between the two conditions, and in fact this difference was robust, $t(10) = 16.06$, $p < .01$. As Figure 2 shows, this difference was also in the predicted direction: Reported averages were larger in the larger-continuous-average condition and smaller in the smaller-continuous-average condition. Moreover, the reported averages were extreme enough that the average in each condition differed significantly from the correct average of the nine anchor points alone. The reported continuous mean was 4.53% larger than the correct average in the larger-continuous-average condition, $t(10) = 4.27$, $p < .01$, and 2.22% smaller than the correct average in the smaller-continuous-average condition, $t(10) = 2.63$, $p = .03$.

These results indicate, for the first time, that the computation of SSRs can operate effectively over continuous visual input. In particular, these results show that the averaging process samples multiple times during a single continuous transformation with no discrete boundaries. Future quantitative research may be able to investigate just how often observers sample in such situations.

Experiment 3: Selective Averaging of Behaviorally Urgent Objects

A central fact about visual perception is that the retinal input contains far too much information to process efficiently, and so the visual system must be intrinsically selective—prioritizing some parts of the input over others. Observers sometimes deliberately attend to certain objects, and previous work has shown that such endogenous selection can be used to constrain SSRs. For example, observers can choose to compute the average size of only discs of a certain color in a multicolored display (Chong & Treisman, 2005b), and explicitly directing observers' attention to a particular disc will bias the resulting average in the direction of that disc's size (de Fockert & Marchant, 2008). Selection can also be exogenous, however, as when attention is automatically captured (see Egeth & Yantis, 1997; Most, Scholl, Clifford, & Simmons, 2005). Beyond simple physical transients, such as the sudden onset of an object (Yantis & Jonides, 1984), higher-level properties that reflect behavioral urgency can also be the basis for attentional capture. In particular, looming stimuli (but not receding stimuli) have been shown to automatically capture attention, even when looming and receding are implemented via two-dimensional expansion and contraction (Franconeri & Simons,

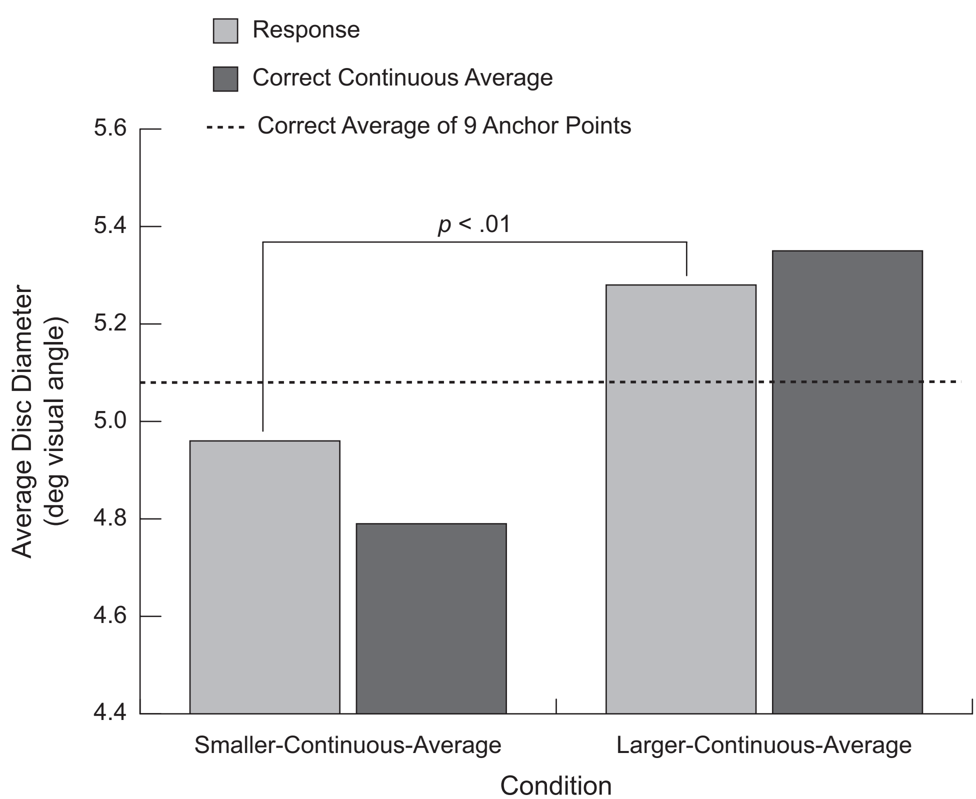

Fig. 2. Results of Experiment 2: average reported average disc diameter and average correct continuous average, for both the larger-continuous-average and the smaller-continuous-average conditions. Because the conditions employed identical anchor points, the correct average of the anchor points alone (depicted by the dashed horizontal line) was the same in the two conditions.

2003; see also Franconeri & Simons, 2005; von Mühlenen & Lleras, 2007). Are SSRs also influenced by attentional capture—perhaps automatically prioritizing behaviorally urgent objects?

In Experiment 3, we tested whether looming (expanding) and receding (contracting) objects differentially influence perceptual averaging when both types of transformations are present in a multiobject display (see Fig. 3). We hypothesized that sampling occurs selectively, and that expanding discs will be sampled with a higher likelihood or more frequently than contracting discs, as a result of looming-induced capture. To test the automaticity of this effect, we employed a situation in which such a bias would impair performance.

Method

This experiment was identical to Experiment 1 except as noted here. Eleven new observers participated in exchange for monetary compensation. No feedback was given. Each display involved the simultaneous presentation of eight discs. In order to determine the discs' locations, we divided the display into an imaginary 4 × 3 grid, creating 12 possible locations. Each disc was randomly assigned a location without replacement, and then its position was randomly jittered (by up to slightly less than the radius of the disc). Over the course of 1.2 s, each disc either expanded or contracted gradually. We divided the eight discs in each display into two groups—a larger group (containing the four largest discs) and a smaller group (containing the four smallest discs). In the *larger-looming* condition, discs in the larger group expanded while discs in the smaller group contracted. In the *larger-receding* condition, discs in the larger group contracted while discs in the smaller group expanded. These trials came in pairs, such that each larger-receding animation was identical to a larger-looming animation played in reverse. Consequently, the continuous average size was identical in the two conditions. In the larger-looming condition, each disc in the larger group had an initial size randomly chosen from between 3.25° and 4.63° (and subsequently expanded to a random size between 4.63° and 6.00°), and each disc in the smaller group had an initial size randomly chosen from between 1.86° and 3.25° (and subsequently contracted to a random size between 0.50° and 1.86°). (In the larger-receding animations, these values were swapped—as these trials were identical to larger-looming animations played in reverse.)

If perceptual averaging is biased toward behaviorally urgent objects, then observers would provide a larger average size in larger-looming trials than in larger-receding trials, even though these conditions had identical correct continuous averages (and also identical minimum and maximum disc sizes, etc.). Observers were instructed to judge the average size of all the discs over time during each trial. Observers completed 130 larger-looming trials and 130 matched larger-receding trials, all randomly ordered (preceded by 10 practice trials, the results of which were not recorded).

Results and discussion

Overall accuracy was lower in both conditions than in the previous experiments (though still roughly equivalent to that in previous studies that used 75% thresholds): Averaged across trials, the absolute value of the deviation between the reported and correct average sizes was 18.55% (SD = 7.31%) in the larger-looming condition and 16.52% (SD = 4.59%) in the larger-receding condition, and these values differed, $t(10) = 3.27$, $p < .01$. Critically, as depicted in Figure 4, observers' reported averages were significantly larger on larger-looming trials than on larger-receding trials, $t(10) = 2.80$, $p < .05$. In this demonstration that attentional capture influences the

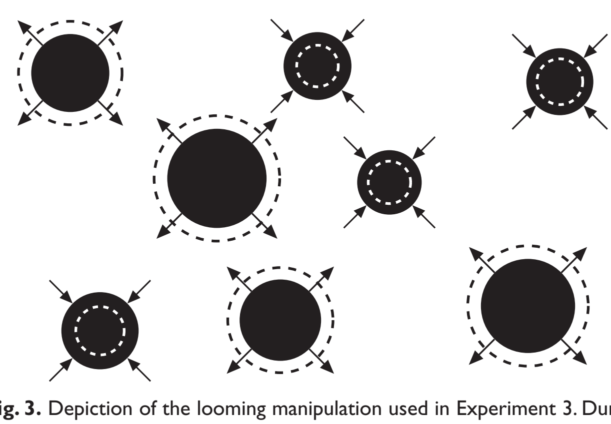

Fig. 3. Depiction of the looming manipulation used in Experiment 3. During each trial, half of the items loomed (expanded), and the other half receded (contracted). The solid black discs represent example starting sizes, and the dashed circles represent example end sizes. The arrows indicate whether each disc is looming or receding. This figure depicts a display from the larger-looming condition, in which the larger items are looming, and the smaller items are receding. (In the larger-receding condition, the transformations were reversed.) In the experiment, the discs were white, presented on a uniform gray background, and no dashed circles or arrows were visible in the displays.

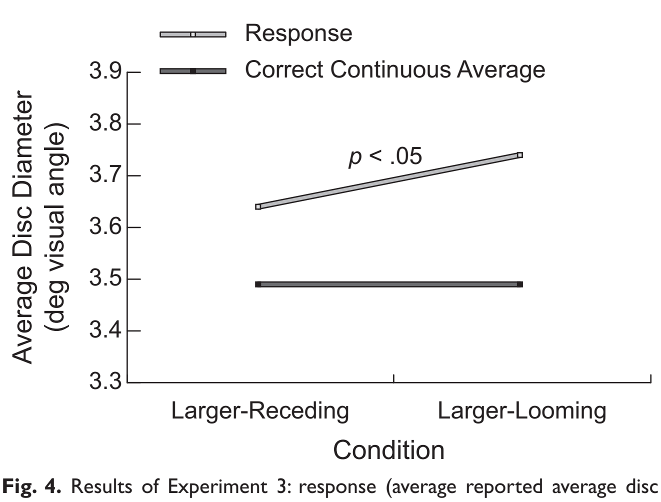

Fig. 4. Results of Experiment 3: response (average reported average disc diameter) and correct continuous average disc diameter as a function of condition.

computation of SSRs, size averaging was biased toward looming (expanding) over receding (contracting) objects. (This attentional capture might also have interfered with the averaging, resulting in lower accuracy than in the previous experiments.) This phenomenon is an example of how a perceptual bias may lead to a type of subsampling in the averaging process (see Myczek & Simons, 2008)—reflecting not an overt shortcut strategy, but rather an automatic visual routine.

General Discussion

The central finding of this study is that perceptual size averaging is not isolated to discrete static displays, but rather can operate in some sense continuously—sampling the sizes of objects multiple times during ongoing transformations with no discrete boundaries. This process operates over arrays of multiple dynamically transforming objects, as well as over single objects changing over time—and the accuracy of the resulting averages appears to be in the same range as that with discrete static displays.

The discovery that SSRs can operate continuously allowed us to ask about other types of dynamic transformations, and to show that perceptual averaging automatically prioritizes some objects—in particular, looming objects, which have been shown to capture attention in other contexts. Though this type of prioritization was erroneous given our task demands (as observers were instructed to average the entire display), such biases may be adaptive in a more general sense. A purpose of SSRs may be to summarize a population when one cannot predict the individual members with which one will be interacting. The behavioral urgency of looming objects in the real world, however, can effectively bias this probability—increasing the likelihood that one will interact with the looming objects, and thus making a summary of those objects especially useful.

Previous research on SSRs has been limited to discrete displays of unchanging properties. The present demonstration of continuous averaging illustrates a way in which such processing may also be well adapted to real-world visual experiences, which are intrinsically continuous and dynamic as a result of ongoing changes in both the observer and the environment.

Acknowledgments

For helpful conversation and comments on earlier drafts, we thank Marvin Chun, Joshua New, two anonymous reviewers, and the members of the Scholl and Chun labs at Yale University.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Notes

1. The existing work on SSRs in perception (including this article) has focused on statistical mean representations—for example, the computation of mean size, as explored here. It is an interesting open question whether other such summary representations, such as measures of the range or variance of set properties, can also be directly perceived. Measures of variance are particularly interesting given that some types of variance do seem to be available to visual processing—for example, the average cluster of a texture distribution, which is essentially a measure of the variance of its density (e.g., Durgin, 1995).

2. Such displays cannot be readily depicted in a static figure, but dynamic animations of the stimuli for the three experiments in this article can be viewed on-line at http://www.yale.edu/perception/SSRs/.

3. We provided feedback to reduce variance. Bauer (2009) made the important point that feedback can influence the nature of the results—for example, leading observers toward one measure of central tendency over another. This is a concern for studies exploring the algorithm by which the averaging process operates (a focus of much of the work in the 1960s and 1970s), but there is no evidence that feedback can have dramatic effects on averaging abilities per se. In the present study, we based feedback on the arithmetic mean of the diameters as a case study of observers' competence at intentional perceptual averaging of this type, without assuming that this is necessarily the type of "average" that observers would choose to use on their own. We chose the arithmetic mean of the diameters in particular because (a) this has been the most commonly used metric in the literature on spatial size averaging, and (b) it does seem to approximate what observers use when averaging without feedback, when averaging both spatial arrays (Bauer, 2009) and temporal sequences (Weiss & Anderson, 1969). Note that the resulting absolute error rates underestimate the accuracy of the averaging process itself, though, because individual disc sizes are also estimated with some degree of error (typically being compressed in subsequent reports; e.g., da Silva, Marques, & Ruiz, 1987; Teghtsoonian, 1965).

4. An additional experiment, not reported here, examined the ability to explicitly average the sizes of nine discs when they were presented in a single static array versus when they were used as the nine anchor points of a single continuously expanding and contracting disc (which froze for 250 ms at each anchor point so that it could be identified). Observers were instructed to report the average of only those nine sizes in both cases, and their accuracy was marginally better in the temporal case, $t(13) = 1.88$, $p = .08$. This result further suggests that sequential averaging of size in the presence of continuous transformations is no less accurate than the more familiar case of spatial size averaging.

References

Alvarez, G.A., & Oliva, A. (2008). The representation of simple ensemble visual features outside the focus of attention. *Psychological Science*, *19*, 392–398.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*, 157–162.

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.

Bauer, B. (2009). The danger of trial-by-trial knowledge of results in perceptual averaging studies. *Attention, Perception, & Psychophysics*, *71*, 655–665.

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

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

Chong, S.C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*, 393–404.

Chong, S.C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics*, *67*, 1–13.

Chong, S.C., & Treisman, A. (2005b). Statistical processing: Computing the average size in perceptual groups. *Vision Research*, *45*, 891–900.

Dakin, S.C., & Watt, R.J. (1997). The computation of orientation statistics from visual texture. *Vision Research*, *37*, 3181–3192.

da Silva, J., Marques, S., & Ruiz, E. (1987). Subjective differences in exponents of psychophysical power functions for inferred, remembered, and perceived area. *Bulletin of the Psychonomic Society*, *25*, 191–194.

de Fockert, J.W., & Marchant, A.P. (2008). Attention modulates set representation by statistical properties. *Perception & Psychophysics*, *70*, 789–794.

Demeyere, N., Rzeskiewicz, A., Humphreys, K.A., & Humphreys, G.W. (2008). Automatic statistical processing of visual properties in simultanagnosia. *Neuropsychologia*, *46*, 2861–2864.

Durgin, F.H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149–169.

Egeth, H.E., & Yantis, S. (1997). Visual attention: Control, representation, and time course. *Annual Review of Psychology*, *48*, 269–297.

Emmanouil, T.A., & Treisman, A. (2008). Dividing attention across feature dimensions in statistical processing of perceptual groups. *Perception & Psychophysics*, *70*, 946–954.

Franconeri, S.L., & Simons, D.J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, *65*, 999–1010.

Franconeri, S.L., & Simons, D.J. (2005). The dynamic events that capture visual attention: A reply to Abrams and Christ (2005). *Perception & Psychophysics*, *67*, 962–966.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology*, *17*, R751–R753.

Haberman, J., & Whitney, D. (2009). Seeing the mean: Ensemble coding for sets of faces. *Journal of Experimental Psychology: Human Perception and Performance*, *35*, 718–734.

Im, H.Y., & Chong, S.C. (2009). Computation of mean size is based on perceived size. *Attention, Perception, & Psychophysics*, *71*, 375–384.

Miller, A.L., & Sheldon, R. (1969). Magnitude estimation of average length and average inclination. *Journal of Experimental Psychology*, *81*, 16–21.

Morgan, M., Watamaniuk, S., & McKee, S. (2000). The use of an implicit standard for measuring discrimination thresholds. *Vision Research*, *40*, 2341–2349.

Most, S.B., Scholl, B.J., Clifford, E.R., & Simons, D.J. (2005). What you see is what you set: Sustained inattentional blindness and the capture of awareness. *Psychological Review*, *112*, 217–242.

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.

Oriet, C., & Corbett, J. (2008). Evidence for the rapid extraction of average size in RSVP displays of circles. *Journal of Vision*, *8*(6), Abstract 13. Retrieved December 31, 2009, from http://journalofvision.org/8/6/13/

Parkes, L., Lund, J., Angelucci, A., Solomon, J.A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*, 739–744.

Pelli, D.G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.

Pollard, P. (1984). Intuitive judgments of proportions, means, and variances: A review. *Current Psychological Research & Reviews*, *3*, 5–18.

Teghtsoonian, M. (1965). The judgment of size. *American Journal of Psychology*, *78*, 392–402.

von Mühlenen, A., & Lleras, A. (2007). No-onset looming motion guides spatial attention. *Journal of Experimental Psychology: Human Perception and Performance*, *33*, 1297–1310.

Watamaniuk, S.N.J., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research*, *32*, 931–941.

Weiss, D.J., & Anderson, N.H. (1969). Subjective averaging of length with serial presentation. *Journal of Experimental Psychology*, *82*, 52–63.

Williams, D.W., & Sekuler, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, *24*, 55–62.

Yantis, S., & Jonides, J. (1984). Abrupt visual onsets and selective attention: Evidence from visual search. *Journal of Experimental Psychology: Human Perception and Performance*, *10*, 601–621.