1-1-2015

Statistical processing : mean size perception of partially occluded sets

Jae-Eun Jenny Lee
Baruch College

Statistical processing: Mean size perception of partially occluded sets

Jae-Eun Jenny Lee

Submitted to the Committee of Undergraduate Honors at Baruch College of the City University of New York on May 4th, 2015 in partial fulfillment of the requirements for the degree of Bachelor of Arts in Cognitive Science with Honors in Psychology

Professor Tatiana Aloi Emmanouil
Faculty Sponsors

Professor Eric Mandelbaum
Departmental Honors Committee

Professor Steven Young
Departmental Honors Committee

Table of Contents

Abstract

Research suggests that observers perceive ensemble characteristics, which allows them to accurately evaluate statistical properties of a set. In set perception, observers are remarkably accurate and fast in computing the mean size of a set of objects. In this experiment, we introduce partial occlusion, which is extremely common in natural visual environments, to the framework of set perception. Though portions of the objects are covered, we hypothesized that partial occlusion would not affect the accuracy of statistical processing, based on evidence for object completion in early stages of vision and use of the perceived size to calculate mean size. If the visual system did not account for the occluded portions of the objects, accuracy should drop when the sets are partially occluded. Our data showed that there was no significant difference in accuracy across unoccluded and occluded conditions. Whether this was due to the visual system reconstructing the partially occluded circles or the visual system calculating the whole area with the information provided by the fragments is still an open question.

Introduction

Flowers in a garden, cars on a road, people in crowds — in a glance, human observers easily classify these objects into categories and recognize overall features of a group. Yet the details of how this information about the objects and their commonalities is extracted cannot be easily explained. Though commonly known in perceptual psychology, the way we perceive the visual world feels so continuous and complete that we can forget that our perception is not equal to our actual surroundings. The visual system does not have the resources to fully process each detail in our surroundings, as shown by studies on attentional and working memory capacity (Cowan, 2001; Zhang & Luck, 2008). These studies suggest that both attention and working memory are limited to handling only a few objects at a time. If visual attention and working memory cannot sufficiently encode the properties of all the visual stimuli in a scene, then the visual system must rely on tools that overcome these limitations. Given the sheer amount of information provided by our environments, it would make sense for the visual system to have efficient mechanisms that extract certain information. Ensemble representation and perceptual completion of occluded objects are believed to be products of such mechanisms. Before describing the experimental design, this section will review the development of and debates in ensemble representation studies, contextualizing the experiment.

Ensemble Characteristics

When observers perceive some abstract feature of a set that summarizes the statistical properties of the individual members, they are perceiving 'ensemble characteristics' (also said to be forming summary statistics). For example, from a group of moving dots, the observers can extract the overall direction of movement (Watamniuk & Duchon, 1992). The visual system can "summarize" the motion of all the individual dots by providing a single average speed of the set, which represents the ensemble as a whole. Because of the redundancy and regularities in a visual scene, encoding multiple objects as a single group rather than as individuals is a cost-effective method of information extraction. Direction of motion is not the only ensemble characteristic studied; human observers can also extract the speed of motion (Watamaniuk, S. N., & Duchon, A, 1992), average orientation (Dakin & Watt, 1997, Parkes, Lund, Angelucci, Solomon & Morgan, 2001), and mean size (Alvarez, 2011; Chong et al., 2008; Chong and Treisman, 2005b, etc.).

The capacity human observers possess in mean size perception is more surprising than their abilities in other forms of statistical processing. This result was first published by Ariely (2001). In his set of three experiments, Ariely presented a set of sixteen circles for 500 ms (milliseconds), and then presented a test stimulus (one or two circles) for 500 ms. The first two experiments, the participants were asked to decide which circles presented as test stimuli were members of the set. The results showed that participants were unable to differentiate between members and non-members. However, this was not because participants were not retaining any information about the set; circles were likelier to be judged as members the closer their size was to the mean. In the third experiment, Ariely tested their sensitivity in mean size discrimination, shifting the focus from individual characteristics of the circle to ensemble characteristics. For sets with dissimilar sizes, the discrimination threshold was 6 to 12%, and an even lower 4 to 6% for sets of similar sizes. Observers' processing of mean sizes were very precise and quickly formed, and that precision either was independent of the number of items in the set or improved with increasing set size. This remarkable ability has led to numerous studies, making mean size the most intensively studied ensemble characteristic. The accuracy of observers' perceptions of the mean size has been confirmed in other studies (Chong & Treisman, 2003, 2005a, 2005b; de Fockert & Marchant, 2008). Given that absolute motion and absolute orientation are coded by specific neurons, the extraction of these particular ensemble characteristics could be simply explained by these receptors in the early visual system (the sites in the brain that first processes visual information coming from the retina). However, neuronal receptors for absolute size do not exist, so some argue that explaining how we perceive average size demands a different theory (Marchant, Simons, & de Fockert, 2013).

One explanation about how statistical processing may work is sampling strategies, where the visual system focuses attention on a select number of items. Although the information extracted is more than what would be possible to find in by focusing on each and every item serially, the argument for sampling strategies is that a select few items can be processed within the capacity of working memory, and the visual system may rely on them to extrapolate the statistics of the entire set (Myczek & Simons, 2008). If this theory was true, then statistical processing exists within the limits working memory and attentional bottleneck rather than employing a different technique to overcome the limitations imposed by the bottleneck. Though studies show that sampling strategies achieve performance close to that of human observers (Myczek & Simons, 2008), similar results do not necessarily mean that sampling strategies model the actual mechanisms in cognitive processing (Chong, Joo, Emmanouil, & Treisman, 2008).

A few key findings about human observers' capability to extract mean size has led to belief in automaticity in processing. The idea that statistical processing is automatic means processing does not require conscious effort for the visual system to compute summary statistics. Even while observers maintain high accuracy in judging mean size, they do not have discrete representations of each individual member of the set (Ariely, 2001). They are more complex than circles, but this also holds true for human faces and their high-level properties such as emotion, gender (Haberman & Whitney, 2007), and identity (de Fockert & Wolfenstein, 2009). For example, observers accurately pinpoint the mean emotion of a set of faces when they viewed sets of 16 for 500 ms or less, while still being unable to differentiate non-members from members (Haberman & Whitney, 2009). So these results indicate that observers have access to the higher-level representation of the set, and there need not be conscious access to basic information about the individual items for mean extraction. Additionally, increasing the number or density of visual objects in a set has little effect on performance (Ariely, 2001, Chong & Treisman, 2003, Chong & Treisman, 2005a, Chong & Treisman, 2005b, De Fockert & Marchant, 2008). Both the increase in number and crowding would overload attention, since human observers are limited to focusing attention on a few objects. Instead of heavy attentional load, researchers argue that statistical processing places little strain on resources (Joo et al., 2009). Based on these facts and also on evidence that cueing does not seem to affect accuracy, researchers argue that mean extraction is automatic and parallel (Chong and Treisman 2005a).

The evidence is not conclusive for either side. Regardless of whether statistical processing is completed through sampling or formed outside focus of attention, there is general agreement that judgments of mean size are based on some ensemble representation (Alvarez, 2011). Besides exploring concentration of attention in statistical processing, previous studies have focused on how properties, such rapid temporal presentation (Joo et al., 2009), exposure time (Whiting & Oriet, 2011), resistance to object substitution masking (Jacoby, Kamke, & Mattingley, 2013), item heterogeneity (Marchant, Simons, & de Fockert, 2013), and range of size variation (Allik, et al., 2014) affect mean size perception. This current work is related to the debate around automaticity and the properties which affect statistical processing, but does not directly respond or follow to past studies. Instead, it explores a different and untouched point on the frontier of statistical processing. We relate statistical processing to the visual scenes in natural environments. In our daily lives, we extract mean sizes of objects that are overlapping or hidden from view so this work questions how partial occlusion affects the extraction of that ensemble characteristic from a set.

Partial Occlusion

If we see cars in a parking lot, we rarely see the complete and clear view of any one of them. In the natural world, partially occluded objects are the norm rather than the exception. However complex the scene may be, it seems to us that we see cars rather than a collection of tires, oddly cut windows, and curved metal frames. This phenomenon of inferential information filling in partially occluded objects has been studied since decades ago (Koffka, 1935). Because an incomplete object would look disconcerting in our seamless experience of the world, the perceptual system often “fills in” what is missing using previous knowledge. Everyday vision is more creative and constructive than our perception would let on, a fact exploited by Gestalt illusions such as the one pictured in the figure below.

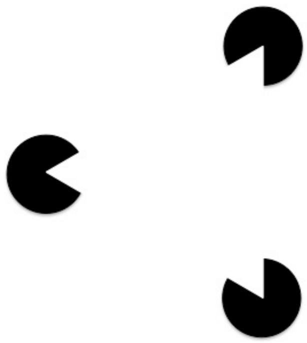

Figure 1. Kanizsa's Triangle. Instead of simply perceiving three pac-man shapes on a flat surface, the visual system interprets this image as a white triangle lying atop a background of three full circles. The circles are completed amodally, and "invisible" contours for the triangle are drawn in by illusory boundary completion. Within the edges of the illusory figure, surface brightness filling-in creates a three-dimensional perceptual experience.

There are two forms of visual completion: modal and amodal. Modal completion depends on observers perceiving a contour where there is no contrast. The central illusory figure in the Kanizsa's Triangle above is due to modal completion, but the black "circles" are an example of amodal completion. The perception of complete objects behind an occluder is referred to as amodal completion (Michotte, Thinès, Costall, & Butterworth, 1991), and this experiment will focus on this version of perceptual completion. Although there is no indication of continuation where the object has to be perceptually completed, observers perceive the portion(s) of the object to be one whole. There is still debate around when exactly this occurs (Guttman, Sekuler, & Kellman, 2003; Rauschenberger et al., 2004), but amodal completion likely arises after early stages of visual encoding and is involved with feedback with higher-level mechanisms (Wokke et al, 2012).

Together?

Chong and Treisman (2003) have argued that visual scenes may have "illusions of completeness" due to statistical processing filling in missing details of a scene, but there have not yet been studies that focus on how partial occlusion may affect statistical processing of sets. Previous studies have looked at how well human observers can ignore an irrelevant subset out of the whole presented set. Chong and Treisman (2005b) showed colored shapes, and asked participants for the mean size of the shapes of one particular color. The display could be of three kinds: a single color display, a double color display with the relevant color cued, or a double color display with no cue. Participants showed little difference across conditions, further evidence for automaticity and a parallel process. Another study by Oriet and Brand (2013), which challenged Chong and Treisman's conclusion that averaging processes can be applied in parallel to two subsets, showed bars of differing orientations (horizontal or vertical). Participants were asked the judge the length of either the horizontal bars only or vertical bars only. The researchers concluded that subjects were unable to completely exclude irrelevant items when calculating the mean size. Participants accurately responded when the mean lengths of the horizontal subset and vertical subset were the same, but when the subset means were incongruent, participants' answers were skewed toward the overall mean. This pattern held regardless of whether they viewed for 200 ms or an unlimited length of time. Overall, while properties such as color may allow pre-attentive segregation of members, this is not the case for orientation (and potentially other properties). While these are significant findings relevant to the current study, this work is different in an important way. Interpreting a partially occluded display does not simply involve ignoring an irrelevant subset; the occluders should be understood as blocking part of the set from view. If the occluders are labeled an "irrelevant set", then the set is not recognized as being partially occluded. Partially occluded sets are perceptually completed very quickly under most circumstances, but statistical processing may automatically parse the occluders out of the display by color, leaving the only the visible fragments to calculate the mean.

There is reason to think this is not the case; previous research suggests that perceived rather than physical size is used in computing the mean size. Im and Chong (2009) demonstrated how utilizing Ebbinghaus illusion influenced participants' perception of the mean sizes of sets.

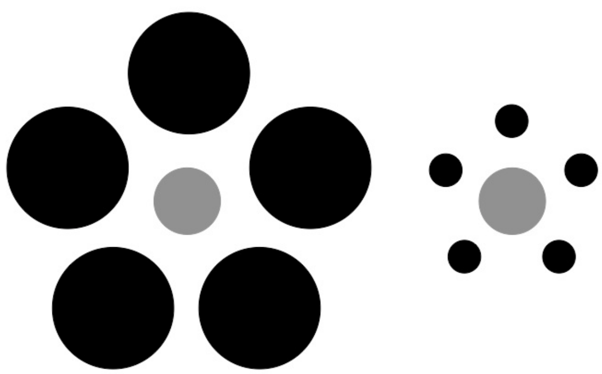

Figure 2. Ebbinghaus illusion. The contrasting sizes of the circles that surround the central circles create the illusion that the central circles are different sizes. However, the central circles are identical.

All target circles in the sets were presented as the central circle in an Ebbinghaus configuration for 250 ms. The experiment found that the strength of illusion affected participants' accuracy. Rather than coding the actual physical sizes of the circles and calculating a mean, observers judged the individual target circles to be smaller or larger than the physical size based on the illusion. This effect was stronger in the second version of the experiment, which was designed to create a more effective perception of the illusion.

This suggests that statistical processing took place after the higher-level processing it took to interpret the context of the scene (in this case, the ring of circles which induce a misjudgment). Although the target circles were black and the inducers were white, and they were never asked to average the sizes of the inducers, observers could not ignore the inducers when making their judgments. In fact, the study found that participants' answers were influenced in the direction induced by the illusion, so they accounted for the inducers meaningfully. This could mean that statistical processing would not simply ignore the occluders or mandatorily include them as relevant to the mean, but interpret them as hiding parts of the set.

This experiment only starts exploring at what point statistical processing takes place in the stages of visual processing and how much information is needed. Statistical processing could possibly happen at a very elementary stage when the occluded objects have not yet been perceptually completed, but the visual system relies on the available fragments to find mean size. Alternatively, statistical processing could happen at a later stage, when perceptual completion has taken place through more elaborate feedback mechanisms. While these details about statistical processing are open questions, partial occlusion of regular shapes leaves enough available information to find the mean size of the original shapes. Given this and what we know about how perception of size is readily influenced by context and occluded object completion, we hypothesize that accuracy in judging the mean size of a set will not change as a result of partial occlusion.

Method

Participants

Twelve students (five female), age ranging from 19 to 35 years (M = 24.5) from Baruch College took part in this experiment for course credit. All participants had normal or corrected-to-normal vision and gave written informed consent. One participant was excluded because they did not perform the task as instructed.

Stimuli and procedure

The stimuli were created using MATLAB (Psychophysics Toolbox; Brainard, 1997) and presented on a CRT monitor at a screen resolution of 1280 x 1024 and frame rate of 75 Hz. The stimuli were a black fixation cross of 20 x 20 px (pixels), white circles, and black occluders. All were presented against a light grey background. For each trial, a set of sixteen circles appeared on the center of the screen. The circles were of four distinct sizes, each repeated four times. We created a basic set in which each circle size differed from the next largest by a factor of 1.2 with the smallest circle being 30 px. All other sets were created by multiplying the sizes in this basic set by a scale factor (possible factors: 1.0, 1.1, 1.2, 1.3) The circles were presented in random locations of an imaginary 4 x 4 grid, each cell was 100 square px, and the circles' positions within cells were jittered. On different trials, the circles appeared either in front or behind six identical horizontal bars of 480 px width x 18 px height with 45 px vertical gaps in between. The location of the bars was offset with respect to the 4 x 4 grid of the set of circles, with the two centermost bars placed 30 px below and 40 px above the center of the screen respectively. Participants reported the average size of the set by adjusting a single probe that appeared immediately after the set. The probe had eight possible sizes: it was 12% or 24% larger or smaller than the mean size of the set or 12% or 24% larger or smaller of the "visible" mean size (calculated based on the average size of the fragments that would remain visible after occlusion).

Each trial began with a fixation cross for 500 ms, followed by the set of circles for 1000 ms. On half the trials the circles appeared in front the bars (unoccluded condition) while on the other half of the trials the circles appeared partly behind the bars (occluded condition).

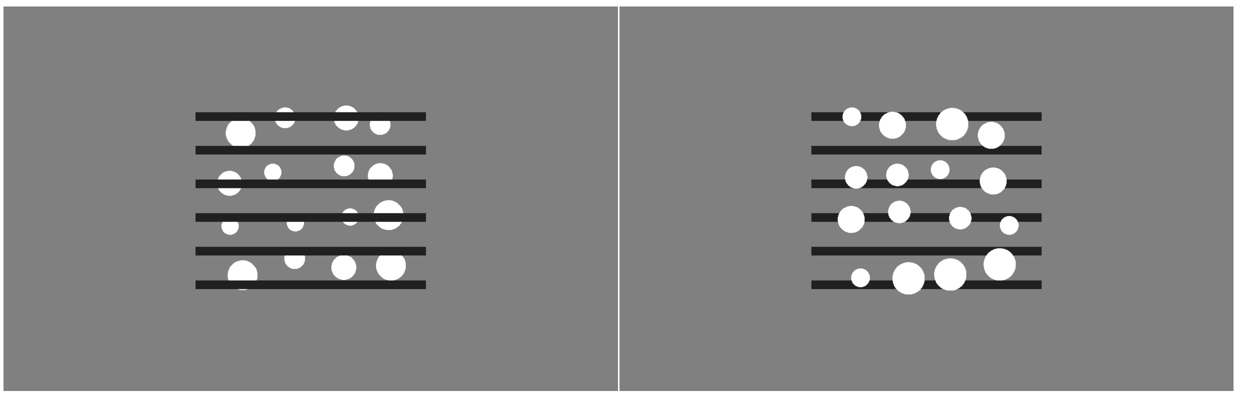

Figure 3. A comparison of the unoccluded (left) and occluded (right) conditions.

After 200 ms a single probe appeared on the screen. Participants were asked to report the mean size of the full circles by adjusting the size of the probe. The probe could be adjusted by pressing the right and left arrow keys to increase and decrease its size respectively. One keyboard press changed the probe by 1 px in diameter. Participants were instructed to focus on accuracy and to take as much time as they needed. In order to record their final estimate, participants pressed the space button. An example of a single trial is shown in Figure 4.

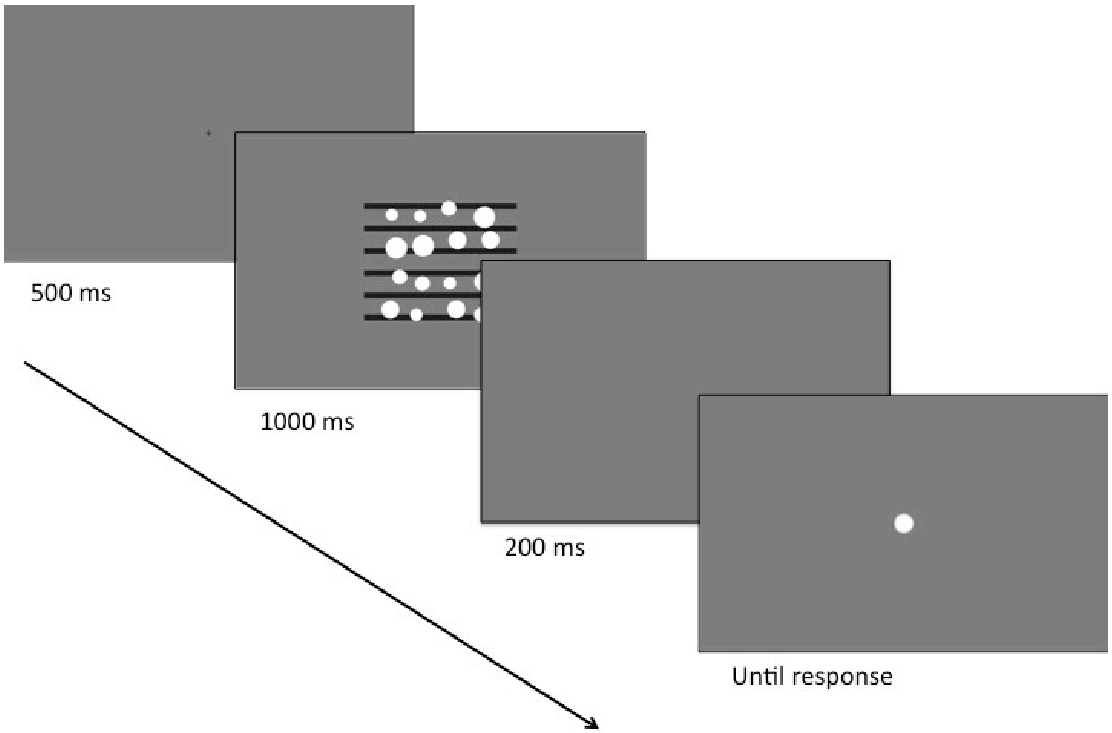

Figure 4. A sample trial, with the timing for each display.

Each participant completed 240 trials, 120 of each condition (occluded unoccluded). Trials were completed in a single session, and participants were informed that they could take a break in between trials, after adjusting the size of the circle and before submitting their response by withholding the space bar response. Participants also completed two practice sessions before the experimental trials, each consisting of 16 trials. The first practice block was intended to familiarize the participants with the statistical processing task. The trials only presented sets of circles without any bars and displayed feedback in terms of the percent error of their estimate after each trial. The trials of the second practice session were identical to the experimental trials. Participants completed practice sessions in the presence of the investigator and could ask questions about their task at any time.

Results

The hypothesis was that partially occluding sets would not affect the extraction of mean size, because observers will somehow account for the occluded portion of the set. If observers do not compensate, the data should show overall smaller estimates in the occluded condition, and therefore more negative overall error. If observers account for partial occlusion, then they should estimate at the same accuracy under both conditions, showing no difference in error. Percentage error was calculated for all trials between the response and actual mean size. For the unoccluded condition, mean overall error was 3.30 (SD = 11.59) and 3.30 (SD = 10.55) for the occluded condition. Mean absolute error was 16.64 (SD = 7.30) for the unoccluded condition and 14.91 (SD = 6.29) for the occluded condition (Figure 5).

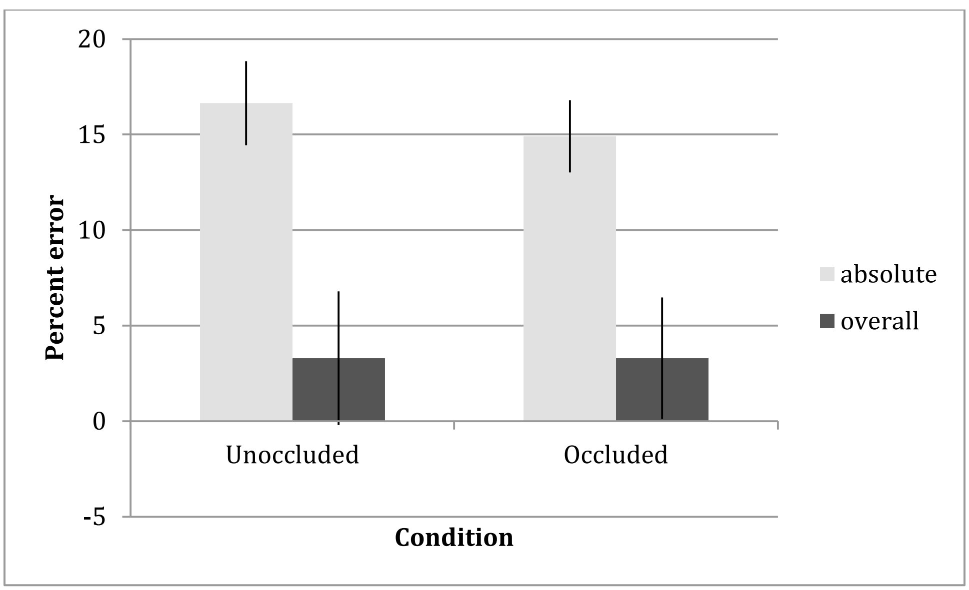

Figure 5. Percent error deviations. Participants tended to overestimate mean size across both occlusion conditions. Although less of the set was visible for the occluded condition, overall error for the occluded condition were not different from the unoccluded condition and mean error tended to be slightly overestimated.

The overall means of the percent error for the trials were analyzed in a t-test, $t(10) = .001$, $p = .999$, yielding no significant statistical difference across occlusion conditions. This suggests that the occlusion did not result in a reduction in overall mean estimates. The comparison of absolute errors yielded a small but significant accuracy advantage in the occluded condition, $t(10) = 2.780$, $p = .019$. This does not challenge the hypothesis observers can statistically process mean size for partially occluded sets, but it simply shows that observers had a much wider range of estimates for the unoccluded condition. This could possibly be explained by the fact that the duration of presentation was relatively long, so participants fixated their attention on individual items. When unoccluded, the individual circles with the smallest and largest sizes, which are not as distinguishable in the occluded condition, may have drawn more focused attention and skewed their judgments. Again, the data supports the argument that instead of calculating a mean size that only takes the visible areas of the set into account, the visual system appears to take the missing portions into account.

In both conditions, there was a slight tendency to overestimate mean size, both in the number of trials of overestimation versus underestimation and in the magnitude of error. The average number of trials per participant with positive error was 63 for the unoccluded and 65 for the occluded. With negative error, the average number of trials was 57 and 55, respectively. As mentioned before, this pattern of overestimation is also reflected in the overall error, which was positive. This tendency could be possibly be due to the presence of the bars. Because the bars are additional shapes on the screen, observers may have been unable completely exclude the bars when processing the circle sizes under both conditions.

Discussion

The results are consistent with Im and Chong's findings (2009), since the bars over the set in the occluded condition were not simply ignored as an irrelevant subset, but accounted for as occluders. Because human observers were not judging mean size using physical shape and size, statistical processing must be taking place later than previously thought. Statistical processing has been argued to happen at the feedforward stage (Hochstein & Ahissar, 2002; Treisman 2006), an early stage of visual feature encoding. However, results from this experiment showed that statistical processing happened after amodal completion of the circles, which is thought to happen only after early stages of encoding. If statistical processing happens after the elaborate processing needed to amodally complete objects, it may mean that statistical processing it is not strictly limited to the feedforward stage.

Partial occlusion of the set of circles does not appear to reduce mean set estimates, but this experiment does not pinpoint how this occurs. One way in which the missing areas could be accounted for is perceptual completion, which was previously discussed. In this case, the bars are seen as blocking parts of the circles from view, so the visual system could amodally complete all the circles of the set and then average them. This account explains our results, but perceptual completion from partial occlusion does not necessarily have to happen to extract mean size. A second way in which the visual system could find the whole circle size is from the visible circular fragments left by the occluders. The occluders leave enough of the original circle that the information from the fragments alone can be used to calculate the original circle size. If this was the case, the visual system does not need the fragments as a circle to appear partially occluded in order to calculate the size of the whole circle.

Ongoing studies

In order to determine which of these two proposed explanations is correct, we designed and started executing another experiment. This experiment utilizes the same paradigm, but introduces a new occlusion condition. In addition to the unoccluded and occluded condition, a third condition rotates the visible circular fragments 90 degrees in place (See Figure 6). This new "rotated occlusion" condition has the same fragments the normal occlusion condition would have, but without appearing covered by the occluders. The only difference is in orientation, which we control for by introducing vertical occluders for half the trials. After this experiment, we can determine if the visual system depends on the pieces appearing occluded, or if the same accuracy in judging mean size happens when only fragments are shown.

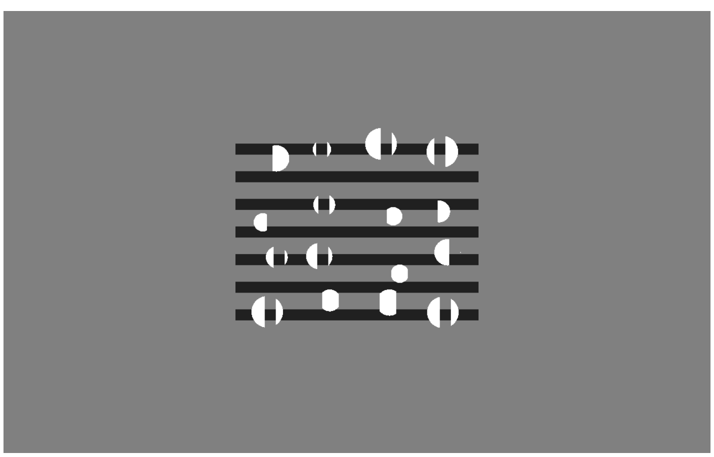

Figure 6. This is the new "rotated occlusion" condition where the same shapes are presented as the occluded condition, but without appearing to be partially occluded since the edges of these circular fragments do not align with the edges of the occluders.

Data for the experiment just described is being collected, and we will be testing the effects of exposure time and complexity of occluders following the end of its collection. The experiment executed in this work used a 1000 ms interval, designed to give participants ample time to process the set. In comparison to other experimental designs, this was a relatively long time, as Ariely (2001) gave a 500 ms exposure time of the set and Im and Chong (2009) 250 ms. To further investigate the relationship between statistical processing of partially occluded sets, we will be testing using shorter intervals of time to verify what happens as exposure time decreases. Statistical processing could potentially stop accounting for the bars as occluders as the visual system loses the time it needs to account for the missing areas.

Even after the completion of the two additional experiments described, the question of whether statistical processing can still occur when easy reconstruction is not possible still remains. The occluders used in those experiments are highly regular (smooth edges, same sizes and locations, evenly spaced), but irregular occluders can create visual scenes that are harder to interpret. The set of circles in these scenes may be completed when the occluders are visible, but removing the occluders (while keeping the circular fragments) leaves shapes that cannot be easily completed, amodally or through calculation. We manipulated the original shape of the occluders into a fence-like shape by adding five large, equally-spaced squares to the bars. These occluders "cut through" the set in a less predictable way; they did not create a row of perfect semi circles.

Even with simple bars and semicircles, error was skewed positively. People tended to overestimate, presumably because of the number of shapes (circles and bars) on the screen that were not supposed to be all averaged together. With more complex occluders, we would expect the much greater error in the positive direction. Initial results indicated that, as we expected, error was positive and higher than the experiment with simple occluders. We require more data to conclude anything much more significant, but perhaps the complexity of the occluders will prevent statistical processing from accounting for partial occlusion, as fragments form less meaningful shapes that are difficult to interpret as fragments of circles Instead, much more of the occluders are counted as part of the “relevant” subset, and added into the calculation of mean circle size as a result. This would be further evidence for automatic statistical processing, since participants would not fully exclude occluders by conscious decision. Regardless of what objects they wanted to choose to focus on, the visual system averaged everything in.

In summary, this paper has explored the effects that partial occlusion of a set has on statistical processing of the mean size. The results indicated that there was no decrease in mean estimates for occluded estimates, but this finding forms the basis for several new questions. There are multiple ways in which the visual system could account for the missing areas (amodal completion versus calculation purely based on visible fragments), the temporal interval necessary to complete statistical processing of mean sizes under these conditions is not yet defined, and it is not completely clear why observers overestimate in this experiment. This study offers up a previously unexplored topic in statistical processing to further investigation and to bring insight into existing debates.

References

Allik, J., Toom, M., Raidvee, A., Averin, K., & Kreegipuu, K. (2014). Obligatory averaging in mean size perception. *Vision Research*, *101*, 34–40

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. Trends in cognitive sciences, 15(3), 122-131.

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

Banno, H., & Saiki, J. (2012). Calculation of the mean circle size does not circumvent the bottleneck of crowding. *Journal of Vision*, *12*(13), 1–15.

Chong, S. C., Joo, S. J., Emmmanouil, T. A., & Treisman, A. (2008). Statistical processing: Not so implausible after all. Perception & Psychophysics, 70(7), 1327-1334.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision research*, *43*(4), 393-404.

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

Chong, S. C., & Treisman, A. (2005). Statistical processing: Computing the average size in perceptual groups. *Vision research*, *45*(7), 891-900.

Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and brain sciences*, *24*(01), 154-176.

Dakin, S. C., & Watt, R. J. (1997). The computation of orientation statistics from visual texture. *Vision research*, *37* (22), 3181-3192.

De Fockert, J. W., & Marchant, A. P. (2008). Attention modulates set representation by statistical properties. *Perception & Psychophysics*, *70*(5), 789-794.

De Fockert, J. W., & Wolfenstein, C. (2009). Rapid extraction of mean identity from sets of faces. The Quarterly Journal of Experimental Psychology, 62(9), 1716-1722.

Guttman, S. E., Sekuler, A. B., & Kellman, P. J. (2003). Temporal variations in visual completion: A reflection of spatial limits?. Journal of Experimental Psychology: Human Perception and Performance, 29(6), 1211.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. Current Biology, 17(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*(3), 718–734.

Hochstein, S., & Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. Neuron, 36(5), 791-804.

Im, H. Y., & Chong, S. C. (2009). Computation of mean size is based on perceived size. *Attention, Perception, & Psychophysics*, *71*(2), 375–384.

Jacoby, O., Kamke, M., & Mattingley, J. (2013). Is the Whole Really More Than the Sum of Its Parts? Estimates of Average Size and Orientation Are Susceptible to Object Substitution Masking. *Journal of Experimental Psychology: Human Perception and Performance*, *39*(1), 233–244.

Joo, S. J., Shin, K., Chong, S. C., & Blake, R. (2009). On the nature of the stimulus information necessary for estimating mean size of visual arrays. Journal of vision, 9(9), 7.

Koffka, K. (1935). Principles of Gestalt Psychology. Harcourt, New York

Marchant, A., Simons, D., & de Fockert, J. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica*, *142*, 245–

250.

Michotte, A., Thinès, G., Costall, A., & Butterworth, G. (1991). Michotte's experimental phenomenology of perception. Hillsdale: L. Erlbaum Associates

Myczek, K., & Simons, D. J. (2008). Better than average: Alternatives to statistical summary representations for rapid judgments of average size. Perception & Psychophysics, 70(5), 772-788.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature neuroscience*, *4*(7), 739-744.

Rauschenberger, R., Peterson, M. A., Mosca, F., & Bruno, N. (2004). Amodal Completion in Visual Search Preemption or Context Effects?. Psychological Science, 15(5), 351-355.

Treisman, A. (2006). How the deployment of attention determines what we see. Visual Cognition, 14(4-8), 411–443.

Utochkin, I., & Tiurina, N. (2014). Parallel averaging of size is possible but range-limited: A reply to Marchant, Simons, and De Fockert. *Acta Psychologica*, *146*, 7–18.

Watamaniuk, S. N., & Duchon, A. (1992). The human visual system averages speed information. *Vision research*, *32*(5), 931-941.

Whiting, B. F., & Oriet, C. (2011). Rapid averaging? Not so fast!. Psychonomic bulletin & review, 18(3), 484-489.

Wokke, Martijn E., et al. "Confuse your illusion feedback to early visual cortex contributes to perceptual completion." *Psychological Science* (2012): 0956797612449175.

Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, *453*(7192), 233-235.