Cognition 234 (2023) 105384

Subcortical encoding of summary statistics in humans

Yuqing Zhao[a,b,1], Ting Zeng[b,c,1], Tongyu Wang[a,b], Fang Fang[d,e,f], Yi Pan[a,*], Jianrong Jia[a,b,*]

[a] *Department of Psychology, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China*
[b] *Center for Cognition and Brain Disorders, The Affiliated Hospital of Hangzhou Normal University, Hangzhou 310015, Zhejiang, China*
[c] *School of Psychology, Jiangxi Normal University, Nanchang 330022, Jiangxi, China*
[d] *School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China*
[e] *IDG/McGovern Institute for Brain Research, Peking University, Beijing 100871, China*
[f] *Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China*

ARTICLE INFO

Keywords:
Summary statistics
Subcortical structures
Monocular channels
Interocular suppression

ABSTRACT

Statistical encoding compresses redundant information from multiple items into a single summary metric (e.g., mean). Such statistical representation has been suggested to be automatic, but at which stage it is extracted is unknown. Here, we examined the involvement of the subcortex in the processing of summary statistics. We presented an array of circles dichoptically or monocularly while matching the number of perceived circles after binocular fusion. Experiments 1 and 2 showed that interocularly suppressed, invisible circles were automatically involved in the summary statistical representation, but only when they were presented to the same eye as the visible circles. This same-eye effect was further observed for consciously processed circles in Experiment 3, in which the estimated mean size of the circles was biased toward the information transmitted by monocular channels. Together, we provide converging evidence that the processing of summary statistics, an assumed high-level

2009). Likewise, presenting items in a rapid serial visual presentation sequence (Joo et al., 2009; McNair, Goodbourn, Shone, & Harris, 2017) or in the periphery (Morgan, Chubb, & Solomon, 2008) such that conscious perception of the items is significantly weakened has little impact on the statistical averaging performance. As such, the specialized mechanism for summary statistical processing may be distinct from the well-established processing hierarchy for individual object identification and recognition (DiCarlo, Zoccolan, & Rust, 2012).

In summary statistical representation, the visual system derives the summary statistical representation of the stimulus set without retaining representations of the individual items within the set (Ariely, 2001). This may provide a foundation for gist perception (Alvarez, 2011; Cha & Chong, 2018; Maule, Witzel, & Franklin, 2014), in which the global configurations of hierarchically structured visual patterns are perceived before the local elements. Intriguingly, many visual features showing such global processing characteristics, such as global shape (Soloveichick, Kimchi, & Gabay, 2021) and numerosity (Collins, Park, & Behrmann, 2017), have been suggested to be encoded in the subcortex. For instance, when the visual cortex is lesioned, subcortical relays can extract attributes from the visual stimuli for crude pattern discrimination (Perenin & Jeannerod, 1979; Weiskrantz, Warrington, Sanders, & Marshall, 1974). By analogy, the summary statistical representation may recruit subcortical structures to convey the summary and crude information about the stimulus set as well (Williams, Morris, McGlone, Abbott, & Mattingley, 2004). It is thus reasonable to hypothesize that subcortical structures are involved in summary statistical processing, supporting automatic and even unconscious extraction of global information.

To investigate the processing stage of summary statistical representation and the possible contribution of subcortical structures to the statistical representation, we utilized two psychophysical methods to differentiate between cortical and subcortical processing. Taking advantage of interocular suppression, which suppresses information processing in various stages of visual hierarchy (Fang & He, 2005; Jiang & He, 2006; Wunderlich, Schneider, & Kastner, 2005; Yuval-Greenberg & Heeger, 2013) while retaining information transmitted by subcortical pathways (Fang & He, 2005; Pasley, Mayes, & Schultz, 2004; Williams et al., 2004), Experiments 1 and 2 first investigated the summary statistical representation of interocularly suppressed, and thus invisible, size information. If the subcortex is involved in the summary statistical representation, the interocularly suppressed circles should be represented in the brain and affect the explicit mean size perception.

In a similar paradigm, Joo et al. (2009) did not find any effect of interocularly suppressed stimuli on the mean size judgment. This may be because the stimuli were presented dichotically to different eyes. Notably, binocular visual signals are separated in pre-striate regions and merged mostly in the layer IV of the primary visual cortex (Baker, Grigg, & von Noorden, 1974; Burkhalter & Van Essen, 1986; Horton, 1990; Menon, Ogawa, Strupp, & Uğurbil, 1997). To confirm the involvement of the subcortex, Experiment 2 used an adapted eye-of-origin paradigm (Gabay, Nestor, Dundas, & Behrmann, 2014), which manipulated the eye of origin of the suppressed and unsuppressed stimuli by presenting circles of different sizes either to the same eye or to different eyes. Experiment 3 then extended the eye-of-origin manipulation to stimuli at the conscious level. These experiments revealed summary statistical representations of interocularly suppressed information and a same-eye effect in summary statistical representation at the unconscious and conscious levels. These results suggest that the subcortex is automatically recruited in the summary statistical representation.

2. Experiment 1

When two incompatible images are presented simultaneously to the corresponding locations in two eyes, they will compete for perceptual dominance. When one image is perceptually dominant, such as with high contrast, transient motion, and high complexity (Blake, 2001), the other image becomes unperceived. This phenomenon, known as interocular suppression, allows us to measure the automatic processing of invisible stimuli in summary statistical representation. Here we took advantage of interocular suppression to render low-contrast circles in the nondominant eye invisible by presenting high-contrast, dynamic masks at the corresponding locations in the dominant eye (Fang & He, 2005).

Experiment 1 used a 2 (presentation types: dichoptic or monocular) × 2 (mask types: mask-big or mask-small) within-subject design (Fig. 1). In the dichoptic condition, an array of eight circles was presented to the nondominant eye while three high-contrast, dynamic masks (Yang & Blake, 2012) were presented to the dominant eye. Masks were applied to either the three largest circles (i.e., mask-big) or the three smallest circles (i.e., mask-small) in the stimulus array. As a control, in the monocular condition, the three largest or smallest circles were directly replaced by three masks such that five circles and three masks were presented monocularly to the nondominant eye. Despite the differences in the stimulus presentation, all conditions commonly contained five visible circles. To measure the mean size perception, the participants were required to estimate the mean size of the visible circles by adjusting the radius of a probe circle.

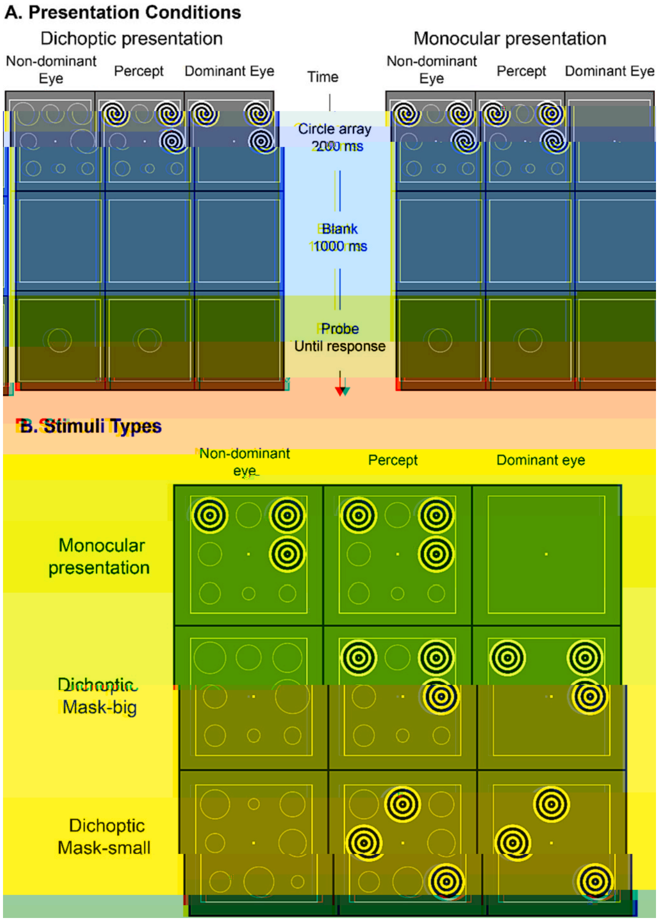

Fig. 1. Experimental conditions and stimuli in Experiment 1. A. The presentation conditions and the sequence of events for each trial. In the dichoptic presentation (left), the circle array and the masks were separately presented to the nondominant and dominant eyes. In the monocular presentation (right), the circle array and the masks were both presented to the nondominant eye. Conscious perception of the circles was obtained from the input to the nondominant eye in both conditions. Participants adjusted the probe size to match the mean size of the perceived circle array. B. Stimuli used in the experiment. In the dichoptic condition, the three masks were applied to either the three biggest or the three smallest circles in the circle array.

2.1. Materials and methods

2.1.1. Participants

Considering that no previous studies compared different manipulations over interocularly suppressed stimuli in summary statistical processing, we used a relatively large sample size of thirty-four participants (mean age 20.97 years, range 18–25 years; 20 females) in Experiment 1. All participants had normal or corrected-to-normal vision. They were naive to the purpose of the experiments and received monetary compensation for participation. This study was approved by the Departmental Ethical Committee of Hangzhou Normal University.

2.1.2. Apparatus

We used MATLAB (The MathWorks) and Psychotoolbox-3 toolbox (Brainard, 1997; Pelli, 1997) to generate and display visual stimuli and to record behavioral responses. The stimuli were presented on a 21-in. CRT with a refresh rate of 75 Hz and a resolution of 1024 × 768 pixels. The experiment was conducted in a dimly lit room, where participants viewed the screen with both eyes through a stereoscope so that the left and right halves of the screen were projected to the left and right eyes, respectively. Participants adjusted the stereoscope to fuse the two visual fields before starting the experiment. A chin rest was used to constrain the viewing distance to 70 cm.

2.1.3. Stimuli

A pilot study was conducted to compare the effectiveness of interocular suppression using four types of masks: the high-contrast and dynamic checkerboard, the textured noise (Fang & He, 2005), the Mondrian patterns (Tsuchiya & Koch, 2005), and the bull's eye. Possibly because the stimuli and the mask shared similar spatial patterns (Hong & Blake, 2009; Yang & Blake, 2012), the bull's eye was found to be the most effective. We thus used the bull's eye (radius: 2.73° in visual angle) that was composed of black and white rings (black: 0.88 cd/m^2; white: 71.45 cd/m^2; width: 0.13°) alternating in luminance at a frequency of 10 Hz as the mask in the present experiments.

Each eye was presented with a white rectangular frame (8.56° × 8.56°) with a white fixation dot (71.45 cd/m^2; 0.17°) in the center of the visual field (13.79 cd/m^2) throughout all trials. The rectangular frame for each eye was subdivided into an imaginary 3 × 3 matrix (Fig. 1A). In the dichoptic presentation condition, eight circles (32.18 cd/m^2) with varied sizes were presented to the nondominant eye (see Fig. 1B). Three bull's eye masks were presented to the dominant eye, at the corresponding locations of the three largest (mask-big condition) or smallest (mask-small condition) circles. In the monocular presentation condition, three masks and five visible circles were presented to the nondominant eye, while no circles were presented in the masked locations. A hole-in-the-card test was used to determine the dominant eye of each participant.

In both the dichoptic and monocular presentations, the mean radius of the visible circles was randomly chosen from four constant sizes of 1.05°, 1.2°, 1.35°, and 1.5°. The radii of the five visible circles were chosen between 0.6 and 1.4 times their average radius randomly. The dichoptic presentation further contained three invisible circles: in the mask-small condition, the radii of the three invisible circles were chosen between 0.8 and 1 times the smallest value of the visible circles randomly; in the mask-big condition, the radii of the three invisible circles were chosen between 1 and 1.2 times the biggest value of the visible circles randomly. This process rendered the average radius of the eight circles 20% smaller and larger than the average radius of the five visible circles in the mask-small and mask-big conditions, respectively. The initial radius of the probe circle was set to be 1.26° in visual angle.

2.1.4. Procedure

At the beginning of the experiment, a white fixation dot centered at a white rectangular frame was presented to each eye for 500 ms. Each trial began with presenting a stimulus array containing eight or five circles and three masks dichotically or monocularly for 200 ms. Participants were asked to estimate the mean size of the perceived circles. After a 1000-ms blank display, a probe circle was presented to the nondominant eye, the same eye as the visible circles (Fig. 1A). Participants adjusted the size of the probe circle to match their perceived mean size of the circles in the stimulus array. The "up" and "down" arrow keys in the keyboard were used for adjustments in small steps (0.03°), and the "left" and "right" arrow keys were used for adjustments in large steps (0.1°). Participants pressed the "enter" key to confirm their response and to move on to the subsequent trial. No time limit was set for the adjustment. After completing 20 practice trials, each participant performed 640 experimental trials, which took about 1 h. A forced break was inserted after every 64 trials.

2.1.5. Data analysis

First, we estimated the distribution of the size differences between the adjusted size and the actual mean size of the visible circles across trials. Trials whose size differences exceeded 3 standard deviations were treated as outliers and gliers could occur at the subcortex. One

3.2. Results and discussion

3.2.1. Summary statistical representation in the monocular channels

We separately calculated the RMSD for the mask-big and mask-small conditions in the same-eye and different-eye conditions. As shown in Fig. 4, in the same-eye condition, the RMSD was not significantly larger than zero in the mask-big condition (t(29) = 0.09, p = 0.929, Cohen's d = 0.02) but was significantly smaller than zero in the mask-small condition (t(29) = 2.03, $p < 0.001$, Cohen's d = −0.88). In the different-eye condition, the RMSD was not significantly different from zero in either the mask-big (t(29) = 1.23, p = 0.229, Cohen's d = 0.22) or the mask-small (t(29) = −0.37, p = 0.715, Cohen's d = −0.07) condition. Therefore, the perceived mean size was smaller for the dichoptic presentation, which contained small invisible circles, than for the monocular presentation, only in the single-eye presentation.

A 2 (eye sources: same-eye or different-eye) × 2 (mask types: mask-big or mask-small) repeated-measures ANOVA showed a significant main effect of the eye source (F(1,29) = 4.29, p = 0.047, partial η^2 = 0.13), with the RMSD being generally larger in the different-eye condition than in the same-eye condition. The main effect of the mask type was also significant (F(1,29) = 8.05, p = 0.008, partial η^2 = 0.22), with the involvement of the big invisible circles significantly increasing the RMSD. Critically, the interaction between the two factors was significant (F(1,29) = 5.29, p = 0.029, partial η^2 = 0.15). Follow-up analyses showed that the RMSD in the mask-big condition was larger than that in the mask-small condition (t(29) = 3.65, p = 0.001, Cohen's d = 0.67) when the visible and invisible circles were presented to the same eye. In contrast, when the visible and invisible circles were presented to different eyes, the RMSD was not reliably influenced by the invisible circles, as indicated by the lack of a significant difference between the mask-big and mask-small conditions (t(29) = 1.35, p = 0.188, Cohen's d = 0.25). Interestingly, the effect size of the mask-small condition was larger than that of the mask-big condition in the same-eye presentation. This result replicates that in Experiment 1, indicating that the masked small circles have a larger effect on the mean size estimation than the masked big circles.

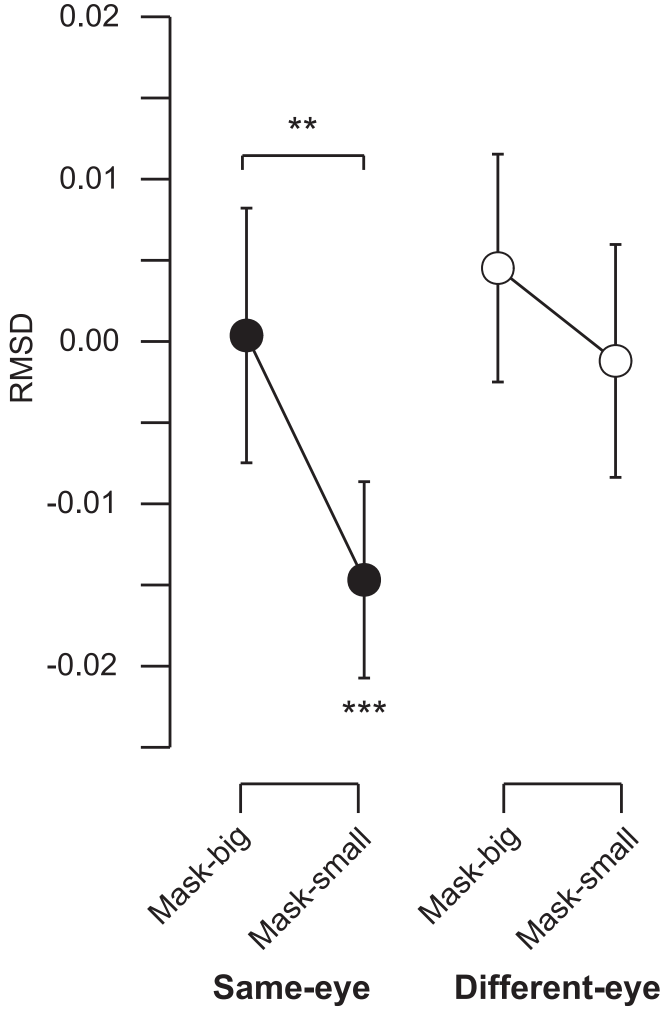

Fig. 4. Results of Experiment 2. The RMSD was larger in the mask-big condition than in the mask-small condition, but only when the visible and invisible circles were presented to the same eye. Error bars represent the 95% confidence interval. ** $p < 0.01$, *** $p < 0.001$.

The results above not only validated the automatic processing of interocularly suppressed information in summary statistical processing but also suggested a subcortical involvement of statistical representation indexed by the same-eye effect. Notably, the same- vs. different-eye comparison also avoided the influence of incomplete interocular suppression on the masked circles. The findings so far have revealed unconscious statistical processing in the monocular channel. Next, we asked whether the monocular processing of summary statistics can occur during conscious visual processing.

4. Experiment 3

Experiment 3 explored whether there remained the same-eye effect in the summary statistical processing when all circles were visible. No masks were applied to make any circles invisible any more. An adapted eye-of-origin paradigm was used (Gabay, Burlingham, & Behrmann, 2014; Karni & Sagi, 1991; Soloveichick et al., 2021). Specifically, among the eight circles in the circle array, the four largest circles were presented to one eye, and the four smallest circles were presented to the other eye. To isolate the same-eye effect, the probe circle was presented either to the same eye as the large circles or to the same eye as the small circles (Fig. 5). We hypothesized that if the summary statistical representation is processed in the monocular channels, the adjusted size of the probe circle should be biased toward the mean size of the circles presented to the same eye as the probe circle. Because the stimuli in the ensemble array were the same in the two conditions and the observers were not explicitly aware of the eye of origin of the stimuli (Blake & Cormack, 1979), this paradigm was free from contaminations by factors related to physical differences or task difficulties in the two conditions.

4.1. Materials and methods

4.1.1. Participants

Experiment 3 recruited 30 participants (mean age 20.9 years, range 18–24 years, 26 females). The sample size was determined based on the same criterion as in Experiment 2.

4.1.2. Apparatus and stimuli

The same apparatus and stimuli were used as in Experiments 1 and 2 except as noted below. The screen was a 19-in. CRT running at a refresh rate of 100 Hz. The viewing distance was 67 cm to match the visual angle of the visual field in Experiments 1 and 2. The stimuli were presented on a gray background (43.86 cd/m^2) throughout the experiment (Fig. 5). Eight circles of varied sizes, all visible, constituted the circle array. The average radius of the eight circles had four sizes: 1.09°, 1.25°, 1.41° or 1.56°. The radius of each circle was randomly chosen to be between 0.4 and 1.6 times the mean radius. The four largest and four smallest circles were presented to different eyes. The probe circle with an initial radius of 1.26° was presented either to the eye of the large circles (big-eye condition) or to the eye of the small circles (small-eye condition). The edge of each circle was composed of an outer white line (luminance: 82.90 cd/m^2) and an inner black line (luminance: 0 cd/m^2), making the average luminance of the edge lines roughly equal to the background luminance.

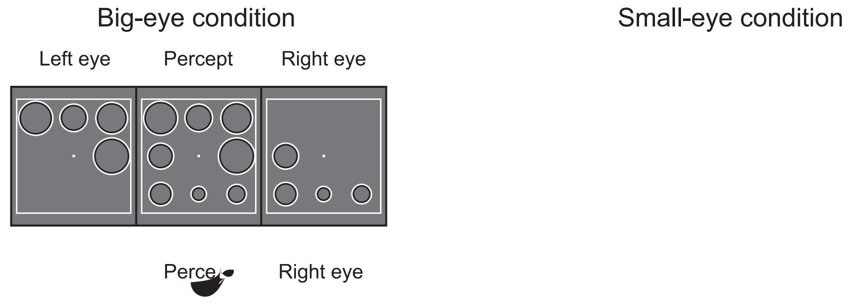

Fig. 5. Presentation conditions in Experiment 3. Four big circles were presented to one eye, and four small circles were presented to the other eye (the right eye in this illustration). In the big-eye condition, the probe was presented to the eye where the big circles were presented (left). In the small-eye condition, the probe was presented to the eye where the small circles were presented (right).

4.1.3. Procedure

The experiment was a one-factor (probe location: big-eye or small-eye) within-subject design. Each trial began with the presentation of a fixation dot in a rectangular frame to both eyes for 200 ms, followed by a 200-ms presentation of the circle array. After a 1000-ms blank display, a probe circle was presented until response. Two types of trials were randomly intermixed during the experiment. In the big-eye trials, the probe circle was presented to the eye to which the larger circles were presented. In the small-eye trials, the probe circle was presented to the eye to which the small circles were presented. Participants adjusted the size of the probe circle to match the perceived mean size of the circle array and pressed the "enter" key to confirm their response. The eye sources for the big and small circles were matched within participants. After completing 20 practice trials, each participant performed 640 experimental trials. The whole experiment took about 1 h. A forced break was inserted after every 64 trials.

4.1.4. Data analysis

The outlier-excluding procedure was the same as above. RMSD was defined as the difference between the estimated mean size and the actual mean size divided by the actual mean size: $RMSD = (\text{Estimated mean size} - \text{Actual mean size}) / \text{Actual mean size}$. The RMSD was calculated in the big-eye and small-eye conditions separately.

4.2. Results and discussion

4.2.1. Same-eye effect of summary statistical processing in conscious processing

We defined the RMSD as the difference between the estimated mean size and the actual mean size divided by the actual mean size. We calculated the RMSD scores in the big-eye and small-eye conditions. The RMSDs in the two conditions were both significantly higher than zero (big-eye: t(29) = 4.77, $p < 0.001$, Cohen's d = 0.87; small-eye: t(29) = 4.54, p < 0.001, Cohen's d = 0.83). These results are consistent with previous findings that the mean size perception is positively biased as the number of items in the stimulus set increases (Kanaya, Hayashi, & Whitney, 2018; Marchant, Simons, & de Fockert, 2013) and that increased variability between items leads to an overestimation of the mean size (Semizer & Boduroglu, 2021). Critically, we found that the RMSD in the big-eye condition was significantly larger than that in the small-eye condition (Fig. 6; t(29) = 3,49, $p = 0.002$, Cohen's d = 0.64), revealing a same-eye effect in mean size perception in the conscious processing. The same-eye effect implies that the summary statistical representation occurs early in visual processing, which may rely on subcortical structures.

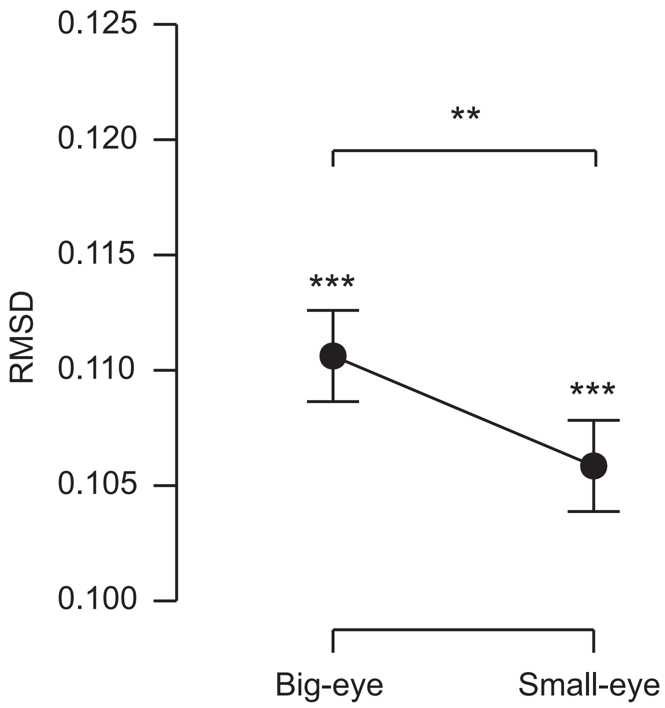

Fig. 6. Results of Experiment 3. The perceived size was larger when the probe was presented to the eye where big circles were presented than to the eye where small circles were presented. Error bars represent the 95% confidence interval. ** $p < 0.01$, *** $p < 0.001$.

5. General discussion

The present study investigated the processing stage of summary statistics in the human brain. Taking advantage of interocular suppression, Experiment 1 revealed that the mean size estimation automatically counted in the invisible, thus unconsciously processed, size information. Using an adapted eye-of-origin paradigm, Experiment 2 further constrained this unconscious processing to information accessed through the monocular channels, providing evidence that the subcortical brain structures may be involved in the summary statistical representation. Finally, Experiment 3 demonstrated that the size information presented to the monocular channels also affected the conscious estimation of the mean size. Together, these findings convergently suggest that summary statistical representation occurs at an early stage of visual processing.

The same-eye effect in statistical processing further hints that the subcortical structures may be recruited for summary statistical representation.

The unconscious statistical processing revealed by Experiments 1 and 2 suggests that summary statistical representation occurs at an early stage of visual processing. This result explains the findings that the summary statistical representation is completed as fast as and even faster than the processing of single items (Epstein & Emmanouil, 2021) and advocates a specialized mechanism in the brain for statistical representation. Unconscious processing has often been used to implicate subcortical processing. This is based on the finding that interocular suppression reduces the neural response in cortical regions, including V1 (Yuval-Greenberg & Heeger, 2013) and higher-level visual areas (Fang & He, 2005; Jiang & He, 2006), but preserves the neural response in subcortical structures (Pasley et al., 2004). In such cases, interocular suppression appears well suited to evaluate the subcortical processing of summary statistical information independent of cortical processing.

Beyond unconscious summary statistical processing, the present study showed a same-eye effect of the summary statistical representation. Specifically, Experiment 2 found that the invisible stimuli affected the mean size judgment only when presented to the same eye as the visible stimuli. The finding that the invisible stimuli presenting to a different eye were not counted in the statistical representation suggested that the process was completed before the binocular fusion. However, it is possible that subcortical processing is recruited adaptively during the processing of unconscious information, which relies much less on the cortical brain areas than conscious information processing (Fang & He, 2005; Jiang & He, 2006). On the other hand, the cortical areas may be recruited to process the statistical information during conscious perception (Cant & Xu, 2012; Tark et al., 2021). To exclude this possibility, Experiment 3 applied a modified eye-of-origin paradigm on conscious statistical processing (Collins et al., 2017; Gabay, Nestor, et al., 2014). Our finding of a same-eye effect in the conscious processing of statistical information suggests that monocular channels are involved in the statistical processing regardless of whether the information is accessible to conscious processing or not. Previous studies have also revealed less interocular suppression for the Magnocellular pathway than Parvocellular pathway (He, Carlson, & Chen, 2005; Zadbood, Lee, & Blake, 2011), but this incomplete suppression was excluded by the different effects of invisible circles in the same- than different-eye conditions in Experiment 2.

In a similar study, Joo et al. (2009) also utilized interocular suppression to render a subset of circles invisible. However, they did not observe the influence of invisible sizes on size averaging. Because their items were presented across two eyes, their comparisons corresponded to the different-eye condition in the present study. As such, these results and our findings convergently suggest that the mean size processing is not preferentially processed in binocular visual channels. Notably, while Joo et al. (2009) used the two-alternative forced choice (2AFC) method, the present study used the adjustment method to measure the differences between the reference and probe sizes. The adjustment method might be more sensitive than the 2AFC method because it not only measures which size is larger but also provides a magnitude of the size difference. Taking advantage of the adjustment method, we further showed that the influence of the invisible circles on the mean size perception was partial. Specifically, while the mean difference between the invisible and visible sizes was 20%, the perceived differences were below 4%. This finding provides important information that future studies exploring mean size perception using subliminal stimuli should enlarge the mean differences between the stimuli sets.

It might be suspected that the suppressors (i.e., the bull's eyes) used in the experiments contributed to the size averaging. This explanation can hardly be true because the suppressors should have had a general influence on size averaging while the effects we observed were condition sensitive. For example, Experiment 2 showed that the interocularly suppressed circles were involved in the mean size computation only when presented to the same eye as the visual circles. In a similar vein, the mask-big and mask-small conditions produced different effects on the mean size judgments, although they were masked by the same suppressors. These results could not be attributed to the general effect of the bull's eyes.

Considering that the monocular segregation of visual input is retained up to the striate cortex, the same-eye effect in the present results might lie in the cortex (i.e., V1) (Horton, 1990; Menon et al., 1997) instead of the subcortex. Specifically, while there are relatively few monocular neurons beyond area V1 (Bi et al., 2011), layer IV of V1 still contains eye-specific, monocular neurons. An alternative explanation for the same-eye effect is that it emerged from V1 rather than pre-striate subcortical regions. However, this is unlikely the case for the summary statistical processing in our study. First, the neurons in V1 have small receptive fields that are impossible to integrate visual inputs from a large visual field, let alone integrate circles from two visual hemifields as in our experimental display. Second, several recent fMRI studies have demonstrated that while the summary statistical presentation can be decoded from neural representations in higher visual areas (V3 and higher) for orientation (Tark et al., 2021), emotion (Im et al., 2017), and visual object (Cant & Xu, 2012), they cannot be decoded in V1. For these reasons, our results advocate the subcortical processing of summary statistical information.

The subcortical statistical processing does not contradict the contribution of the cortical regions (Cant & Xu, 2012; Tark et al., 2021) to statistical processing. It is possible that the subcortical and cortical regions play distinct roles at different stages in the statistical summarization of the visual information. In fact, the small effect size in Experiment 3 may be a consequence of the involvement of both subcortical and cortical processing. Since all the stimuli used in Experiment 3 are accessible to conscious processing, the effect of the subcortical processing may be overshadowed by the cortical processing. Nonetheless, the significant difference between the big-eye and small-eye conditions still reflects an effect of subcortical processing. Notably, the results in Experiment 3 are also consistent with the monocular cueing effect (Kim & Chong, 2022; Solomon & Morgan, 2018) that two eyes representing different mean sizes biased the adjusted size of the probe circle presented to the same eye.

Our finding that the subcortex is capable of statistical processing raises a further question of how the subcortical and cortical structures interact to determine the statistical representation. Several possibilities exist. The same-eye effect and its subcortical neural correlates might act independently of the cortical modulations. For instance, the subcortical structures could form a fast, coarse, and global statistical representation of the visual scene. At the same time, the cortical areas may play roles in scrutinizing the visual display across the canonical visual hierarchy, which is slower in speed but with more detailed representations. On the other hand, at least in human adults, summary statistical information processed subcortically is transmitted to the cortex to interact with attention.

The subcortical structures are known for their roles in conveying crude but rapid signals, such as threat detection (Jiang & He, 2006; Pasley et al., 2004), reaching and grasping (Fang & He, 2005), before reaching conscious awareness to facilitate fast decision and action. So far, whether the monocular bias is specific to summary statistical processing or is general to size perception has not been answered directly. Based on the recent finding that the monocular bias is specific to the perception of the global configuration but not the local shape (Soloveichick et al., 2021), it is likely that the monocular bias may be specific to the integration process. This question can be answered in the future by investigating the interocular transfer of the size adaptation effect. By revealing the involvement of subcortical structures in summary statistical representation, the present study advances our understanding of the functions of subcortical processing. However, while the present study indicates a subcortical origin for statistical processing, we cannot determine which specific subcortical structure(s) is involved in the summary statistical processing. The superior colliculus (Basso & May, 2017) and pulvinar (Rikhye, Wimmer, & Halassa, 2018) are key areas of possible subcortical processing. Future studies could further explore the role of these structures, as well as their interactions with the cortex, in summary statistical processing.

In conclusion, the three independent experiments in the present study convergently suggest that the adult human subcortex is involved in summary statistical processing. Further research based on these findings would establish a broader understanding of the interplay between the cortical and subcortical structures in summary statistical processing.

CRediT authorship contribution statement

Yuqing Zhao: Investigation, Formal analysis, Writing – original draft, Writing – review & editing. **Ting Zeng:** Investigation, Writing – original draft. **Tongyu Wang:** Investigation, Software. **Fang Fang:** Conceptualization, Supervision. **Yi Pan:** Supervision. **Jianrong Jia:** Conceptualization, Funding acquisition, Supervision, Formal analysis, Writing – original draft, Writing – review & editing.

Data availability

The data of this article are publicly available at: https://osf.io/fx6tm/.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant 32000735 to J.J.) and the Natural Science Foundation of Zhejiang Province, China (Grant LY23C090001 to J.J.).

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences, 15*(3), 122–131. https://doi.org/10.1016/j.tics.2011.01.003

Alvarez, G. A., & Oliva, A. (2009). Spatial ensemble statistics are efficient codes that can be represented with reduced attention. *Proceedings of the National Academy of Sciences, 106*(18), 7345–7350.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science, 12*(2), 157–162. https://doi.org/10.1111/1467-9280.00327

Baek, J., & Chong, S. C. (2020). Ensemble perception and focused attention: Two different modes of visual processing to cope with limited capacity. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-020-01718-7

Baker, F. H., Grigg, P., & von Noorden, G. K. (1974). Effects of visual deprivation and strabismus on the response of neurons in the visual cortex of the monkey, including studies on the striate and prestriate cortex in the normal animal. *Brain Research, 66*(2), 185–208. https://doi.org/10.1016/0006-8993(74)90140-1

Basso, M. A., & May, P. J. (2017). Circuits for action and cognition: A view from the superior colliculus. *Annual Review of Vision Science, 3*(1), 197–226. https://doi.org/10.1146/annurev-vision-102016-061234

Belia, S., Fidler, F., Williams, J., & Cumming, G. (2005). Researchers misunderstand confidence intervals and standard error bars. *Psychological Methods, 10*(4), 389–396. https://doi.org/10.1037/1082-989X.10.4.389

Bi, H., Zhang, B., Tao, X., Harwerth, R. S., Smith, E. L., & Chino, Y. M. (2011). Neuronal responses in visual area V2 (V2) of macaque monkeys with strabismic amblyopia. *Cerebral Cortex, 21*(9), 2033–2045. https://doi.org/10.1093/cercor/bhq272

Blake, R. (2001). A primer on binocular rivalry, including current controversies. *Brain and Mind, 2*, 5–38.

Blake, R., & Cormack, R. H. (1979). Psychophysical evidence for a monocular visual cortex in stereoblind humans. *Science, 203*(4377), 274–275. https://doi.org/10.1126/science.760194

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

Burkhalter, A., & Van Essen, D. (1986). Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *The Journal of Neuroscience, 6*(8), 2327–2351. https://doi.org/10.1523/JNEUROSCI.06-08-02327.1986

Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *Journal of Neuroscience, 32*(22), 7685–7700. https://doi.org/10.1523/JNEUROSCI.3325-11.2012

Cha, O., & Chong, S. C. (2018). Perceived average orientation reflects effective gist of the surface. *Psychological Science, 29*(3), 319–327. https://doi.org/10.1177/0956797617735533

Choi, Y. M., & Chong, S. C. (2020). Effects of selective attention on mean-size computation: Weighted averaging and perceptual enlargement. *Psychological Science*. https://doi.org/10.1177/0956797620943834, 095679762094383.

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. https://doi.org/10.3758/PP.70.7.1327

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

Chong, S. C., & Treisman, A. (2005a). Attentional spread in the statistical processing of visual displays. *Perception & Psychophysics, 67*(1), 1–13. https://doi.org/10.3758/BF03195009

Chong, S. C., & Treisman, A. (2005b). Statistical processing: Computing the average size in perceptual groups. *Vision Research, 45*(7), 891–900. https://doi.org/10.1016/j.visres.2004.10.004

Collins, E., Park, J., & Behrmann, M. (2017). Numerosity representation is encoded in human subcortex. *Proceedings of the National Academy of Sciences, 114*(14), E2806–E2815. https://doi.org/10.1073/pnas.1613982114

DiCarlo, J. J., Zoccolan, D., & Rust, N. C. (2012). How does the brain solve visual object recognition? *Neuron, 73*, 415–434. https://doi.org/10.1016/j.neuron.2012.01.010

Epstein, M. L., & Emmanouil, T. A. (2021). Ensemble statistics can be available before individual item properties: Electroencephalography evidence using the oddball paradigm. *Journal of Cognitive Neuroscience, 33*(6), 1056–1068. https://doi.org/10.1162/jocn_a_01704

Fang, F., & He, S. (2005). Cortical responses to invisible objects in the human dorsal and ventral pathways. *Nature Neuroscience, 8*(10), 1380–1385. https://doi.org/10.1038/nn1537

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Fiser, J., & Aslin, R. N. (2002). Statistical learning of new visual feature combinations by infants. *Proceedings of the National Academy of Sciences, 99*(24), 15822–15826. https://doi.org/10.1073/pnas.232472899

de Fockert, J. W., & Marchant, A. P. (2008). Attention modulates set representation by statistical properties. *Perception & Psychophysics, 70*(5), 789–794. https://doi.org/10.3758/PP.70.5.789

Gabay, S., Burlingham, C., & Behrmann, M. (2014). The nature of face representations in subcortical regions. *Neuropsychologia, 59*, 35–46. https://doi.org/10.1016/j.neuropsychologia.2014.04.010

Gabay, S., Nestor, A., Dundas, E., & Behrmann, M. (2014). Monocular advantage for face perception implicates subcortical mechanisms in adult humans. *Journal of Cognitive Neuroscience, 26*(5), 927–937. https://doi.org/10.1162/jocn_a_00528

de Gardelle, V., & Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proceedings of the National Academy of Sciences, 108*(32), 13341–13346. https://doi.org/10.1073/pnas.1104517108

Gilster, R., & Kuhtz-Buschbeck, J. P. (2011). The Muller-Lyer illusion: Investigation of a center of gravity effect on the amplitudes of saccades. *Journal of Vision, 10*(1). https://doi.org/10.1167/10.1.11, 11–11.

Haberman, J., & Whitney, D. (2007). Rapid extraction of mean emotion and gender from sets of faces. *Current Biology, 17*(17), R751–R753. https://doi.org/10.1016/j.cub.2007.06.039

Han, L. (2021). Holistic ensemble perception. *Attention, Perception, & Psychophysics, 83*(3), 998–1013.

He, S., Carlson, T., & Chen, X. (2005). Parallel pathways and temporal dynamics in binocular rivalry. In *Binocular rivalry* (pp. 81–100). MIT Press.

Hochstein, S., Pavlovskaya, M., Bonneh, Y. S., & Soroker, N. (2015). Global statistics are not neglected. *Journal of Vision, 15*(4), 7. https://doi.org/10.1167/15.4.7

Hong, S. W., & Blake, R. (2009). Interocular suppression differentially affects achromatic and chromatic mechanisms. *Attention, Perception & Psychophysics, 71*(2), 403–411. https://doi.org/10.3758/APP.71.2.403

Horton, J. C. (1990). Arrangement of ocular dominance columns in human visual cortex. *Archives of Ophthalmology, 108*(7), 1025. https://doi.org/10.1001/archopht.1990.01070090127054

Im, H. Y., Albohn, D. N., Steiner, T. G., Cushing, C. A., Adams, R. B., & Kveraga, K. (2017). Differential hemispheric and visual stream contributions to ensemble coding of crowd emotion. *Nature Human Behaviour, 1*(11), 828–842. https://doi.org/10.1038/s41562-017-0225-z

Jia, J., Wang, T., Chen, S., Ding, N., & Fang, F. (2022). Ensemble size perception: Its neural signature and the role of global interaction over individual items. *Neuropsychologia, 173*, Article 108290. https://doi.org/10.1016/j.neuropsychologia.2022.108290

Jiang, Y., & He, S. (2006). Cortical responses to invisible faces: Dissociating subsystems for facial-information processing. *Current Biology, 16*(20), 2023–2029. https://doi.org/10.1016/j.cub.2006.08.084

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). https://doi.org/10.1167/9.9.7, 7–7.

Kanaya, S., Hayashi, M. J., & Whitney, D. (2018).

Krzywinski, M., & Altman, N. (2014). Points of significance: Designing comparative experiments. *Nature Methods, 11*(6), 597–598. https://doi.org/10.1038/nmeth.2974

Li, K.-A., & Yeh, S.-L. (2017). Mean size estimation yields left-side bias: Role of attention on perceptual averaging. *Attention, Perception, & Psychophysics, 79*(8), 2538–2551. https://doi.org/10.3758/s13414-017-1409-3

Manning, C., Morgan, M. J., Allen, C. T. W., & Pellicano, E. (2017). Susceptibility to Ebbinghaus and Müller-Lyer illusions in autistic children: A comparison of three different methods. *Molecular Autism, 8*(1), 16. https://doi.org/10.1186/s13229-017-0127-y

Marchant, A. P., Simons, D. J., & de Fockert, J. W. (2013). Ensemble representations: Effects of set size and item heterogeneity on average size perception. *Acta Psychologica, 142*(2), 245–250. https://doi.org/10.1016/j.actpsy.2012.11.002

Maule, J., Witzel, C., & Franklin, A. (2014). Getting the gist of multiple hues: Metric and categorical effects on ensemble perception of hue. *Journal of the Optical Society of America A, 31*(4), A93. https://doi.org/10.1364/JOSAA.31.000A93

McNair, N. A., Goodbourn, P. T., Shone, L. T., & Harris, I. M. (2017). Summary statistics in the attentional blink. *Attention, Perception, & Psychophysics, 79*(1), 100–116. https://doi.org/10.3758/s13414-016-1216-2

Menon, R. S., Ogawa, S., Strupp, J. P., & Uğurbil, K. (1997). Ocular dominance in human V1 demonstrated by functional magnetic resonance imaging. *Journal of Neurophysiology, 77*(5), 2780–2787. https://doi.org/10.1152/jn.1997.77.5.2780

Morgan, M., Chubb, C., & Solomon, J. A. (2008). A "dipper" function for texture discrimination based on orientation variance. *Journal of Vision, 8*(11). https://doi.org/10.1167/8.11.9, 9–9.

Neumann, M. F., Schweinberger, S. R., & Burton, A. M. (2013). Viewers extract mean and individual identity from sets of famous faces. *Cognition, 128*(1), 56–63. https://doi.org/10.1016/j.cognition.2013.03.006

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. https://doi.org/10.1038/89532

Pasley, B. N., Mayes, L. C., & Schultz, R. T. (2004). Subcortical discrimination of unperceived objects during binocular rivalry. *Neuron, 42*(1), 163–172. https://doi.org/10.1016/S0896-6273(04)00155-2

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.

Perenin, M. T., & Jeannerod, M. (1979). Subcortical vision in man. *Trends in Neurosciences, 2*, 204–207. https://doi.org/10.1016/0166-2236(79)90081-X

Rikhye, R. V., Wimmer, R. D., & Halassa, M. M. (2018). Toward an integrative theory of thalamic function. *Annual Review of Neuroscience, 41*(1), 163–183. https://doi.org/10.1146/annurev-neuro-080317-062144

Semizer, Y., & Boduroglu, A. (2021). Variability leads to overestimation of mean summaries. *Attention, Perception, & Psychophysics, 83*(3), 1129–1140. https://doi.org/10.3758/s13414-021-02269-2

Solomon, J. A., & Morgan, M. J. (2018). Precues' elevation of sensitivity is not only preattentive, but largely monocular. *Attention, Perception, & Psychophysics, 80*(7), 1705–1717. https://doi.org/10.3758/s13414-018-1564-1

Soloveichick, M., Kimchi, R., & Gabay, S. (2021). Functional involvement of subcortical structures in global-local processing. *Cognition, 206*, Article 104476. https://doi.org/10.1016/j.cognition.2020.104476

Sweeny, T. D., Haroz, S., & Whitney, D. (2013). Perceiving group behavior: Sensitive ensemble coding mechanisms for biological motion of human crowds. *Journal of Experimental Psychology: Human Perception and Performance, 39*(2), 329–337. https://doi.org/10.1037/a0028712

Sweeny, T. D., Wurnitsch, N., Gopnik, A., & Whitney, D. (2015). Ensemble perception of size in 4-5-year-old children. *Developmental Science, 18*(4), 556–568. https://doi.org/10.1111/desc.12239

Tark, K.-J., Kang, M.-S., Chong, S. C., & Shim, W. M. (2021). Neural representations of ensemble coding in the occipital and parietal cortices. *NeuroImage, 245*, Article 118680. https://doi.org/10.1016/j.neuroimage.2021.118680

Tsuchiya, N., & Koch, C. (2005). Continuous flash suppression reduces negative afterimages. *Nature Neuroscience, 8*(8), 6.

Weiskrantz, L., Warrington, E. K., Sanders, M. D., & Marshall, J. (1974). Visual capacity in the hemianopic field following a restricted occipital ablation. *Brain, 97*(1), 709–728. https://doi.org/10.1093/brain/97.1.709

Williams, M. A., Morris, A. P., McGlone, F., Abbott, D. F., & Mattingley, J. B. (2004). Amygdala responses to fearful and happy facial expressions under conditions of binocular suppression. *The Journal of Neuroscience, 24*(12), 2898. https://doi.org/10.1523/JNEUROSCI.4977-03.2004

Wunderlich, K., Schneider, K. A., & Kastner, S. (2005). Neural correlates of binocular rivalry in the human lateral geniculate nucleus. *Nature Neuroscience, 8*(11), 1595–1602. https://doi.org/10.1038/nn1554

Xu, F., & Garcia, V. (2008). Intuitive statistics by 8-month-old infants. *Proceedings of the National Academy of Sciences, 105*(13), 5012–5015. https://doi.org/10.1073/pnas.0704450105

Yang, E., & Blake, R. (2012). Deconstructing continuous flash suppression. *Journal of Vision, 12*(3). https://doi.org/10.1167/12.3.8, 8–8.

Yuval-Greenberg, S., & Heeger, D. J. (2013). Continuous flash suppression modulates cortical activity in early visual cortex. *Journal of Neuroscience, 33*(23), 9635–9643. https://doi.org/10.1523/JNEUROSCI.4612-12.2013

Zadbood, A., Lee, S.-H., & Blake, R. (2011). Stimulus fractionation by Interocular suppression. *Frontiers in Human Neuroscience, 5*. https://doi.org/10.3389/fnhum.2011.00135