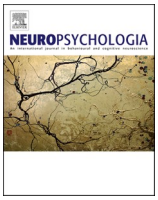
Neuropsychologia 173 (2022) 108290

Contents lists available at ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia



Ensemble size perception: Its neural signature and the role of global interaction over individual items

Jianrong Jia [a,b], Tongyu Wang [a,b], Siqi Chen [a,b], Nai Ding [c,d], Fang Fang [e,f,g,h,*]

[a] Center for Cognition and Brain Disorders, The Affiliated Hospital of Hangzhou Normal University, Hangzhou, 311121, China
[b] Institute of Psychological Sciences, Hangzhou Normal University, Hangzhou, 311121, China
[c] Key Laboratory for Biomedical Engineering of Ministry of Education, College of Biomedical Engineering and Instrument Sciences, Zhejiang University, Hangzhou, 311121, China
[d] Research Center for Advanced Artificial Intelligence Theory, Zhejiang Lab, Hangzhou, 311121, China
[e] School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing, 100871, China
[f] IDG/McGovern Institute for Brain Research, Peking University, Beijing, 100871, China
[g] Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing, 100871, China
[h] Peking-Tsinghua Center for Life Sciences, Peking University, Beijing, 100871, China

ARTICLE INFO

Keywords:
Ensemble size
Steady-state visual evoked potential (SSVEP)
Temporal response function (TRF)
Global interaction
Local interaction

ABSTRACT

To efficiently process complex visual scenes, the visual system often summarizes statistical information across individual items and represents them as an ensemble. However, due to the lack of techniques to disentangle the representation of the ensemble from that of the individual items constituting the ensemble, whether there exists a specialized neural mechanism for ensemble processing and how ensemble perception is computed in the brain remain unknown. To address these issues, we used a frequency-tagging EEG approach to track brain responses to periodically updated ensemble sizes. Neural responses tracking the ensemble size were detected in parieto-occipital electrodes, revealing a global and specialized neural mechanism of ensemble size perception. We then used the temporal response function to isolate neural responses to the individual sizes and their interactions. Notably, while the individual sizes and their local and global interactions were encoded in the EEG signals, only the global interaction contributed directly to the ensemble size perception. Finally, distributed attention to the global stimulus pattern enhanced the neural signature of the ensemble size, mainly by modulating the neural representation of the global interaction between all individual sizes. These findings advocate a specialized, global neural mechanism of ensemble size perception and suggest that global interaction between individual items contributes to ensemble perception.

1. Introduction

The visual system has limited processing capacity (Luck and Vogel, 1997; Palmer et al., 2011). One strategy to overcome the capacity limitation and optimize information processing is to summarize the complex and redundant information into ensemble coding (Alvarez, 2011; Parkes et al., 2001; Whitney and Yamanashi Leib, 2018). Our visual system is remarkably accurate in estimating ensemble properties (e.g., mean, variance) in multiple dimensions, including low-level visual features such as size (Ariely, 2001; Chong and Treisman, 2003) and orientation (Parkes et al., 2001), and high-level visual characteristics such as emotion, gender (Haberman and Whitney, 2007), face identities (Neumann et al., 2013), and biological motion (Sweeny et al., 2013). However, because individual items constitute the ensemble, it is difficult for traditional neuroimaging methods to dissociate the neural process of ensemble perception from that of individual item perception. As a result, although ensemble perception has been studied extensively at the behavioral level over the past two decades, how it is implemented in the brain remains controversial.

Two hypotheses have been proposed to explain how ensemble perception is achieved in the brain. The subsampling hypothesis (Myczek and Simons, 2008; Simons and Myczek, 2008; Solomon et al., 2011) proposes that ensemble perception does not recruit a global mechanism; instead, it can be achieved by sampling and summarizing a subset of items. Specifically, a small subset of items are randomly sampled by attention, and their properties are averaged to generate a mean

* Corresponding author. School of Psychological and Cognitive Sciences, Peking University, 52 Haidian Road, Beijing, 100087, China.
E-mail address: ffang@pku.edu.cn (F. Fang).

https://doi.org/10.1016/j.neuropsychologia.2022.108290
Received 28 December 2021; Received in revised form 2 June 2022; Accepted 7 June 2022
Available online 10 June 2022
0028-3932/© 2022 Elsevier Ltd. All rights reserved.

perception. Consequently, not all items in the stimulus set are processed in ensemble perception. The summary-statistic representation hypothesis (Ariely, 2001, 2008; Chong et al., 2008; Chong and Treisman, 2005a, 2005b), on the other hand, suggests a specialized global mechanism that computes summary-statistic representations over all displayed items. According to this hypothesis, the ensemble property of a stimulus set is processed in parallel with its individual items. All items are processed in the ensemble perception (Iakovlev and Utochkin, 2021), although their weights may vary (Choi and Chong, 2020). This claim is supported by EEG evidence that ensemble statistics can be available even before individual item properties (Epstein and Emmanouil, 2021). Despite that, no consensus has been reached thus far on whether ensemble perception recruits a global mechanism that processes all items in the display.

In addition to the controversy about the existence of a global mechanism, the computational process of summarization remains unknown. Taking ensemble size perception as an example, it is intuitively reasonable for the brain to initially respond to individual sizes and then take a linear average of neural responses to individual sizes to represent the ensemble size (Allik et al., 2013, 2014; Baek and Chong, 2020a). However, ensemble representation computation can be complicated because it is modulated by multiple factors, such as selective attention (Baek and Chong, 2020b; Choi and Chong, 2020), saliency of individual items (Iakovlev and Utochkin, 2021), and variance across items (Semizer and Boduroglu, 2021). Accordingly, weighted averaging and weighted variance computation models have been proposed for ensemble computation (Choi and Chong, 2020; de Gardelle and Summerfield, 2011; Jeong and Chong, 2021). Furthermore, in addition to the linear relationship between perception of individuals and ensemble, individuals may interact with each other and contribute to ensemble perception interdependently. Given extensive evidence in multisensory integration and Gestalt psychology (Angelaki et al., 2009; Wagemans et al., 2012), the brain often processes interactions between items. Therefore, interactions between individuals, especially between neighboring individuals and between all individuals, may be represented in the brain and contribute to ensemble perception.

To explore the neural signature and computational principle of ensemble perception, it is critical to separate the neural responses to individual items from that to the ensemble. To achieve this purpose, we combined the steady-state visual evoked potential (SSVEP) method with the temporal response function (TRF) method. SSVEP utilizes the fact that EEG activity tracks the rhythms of changes in visual stimulus, allowing neural representations to be "tagged" to visual features at a specific frequency. With SSVEP, a stimulus presented at a fixed frequency produces robust neural responses at the same frequency (Norcia et al., 2015). In our study, we manipulated the size of eight circles presented at fixed locations in the peripheral visual field such that the mean size of all circles varied periodically with a base frequency, while the sizes of individual circles varied randomly (Fig. 1b and c). Participants performed mean size-related or unrelated tasks (Fig. 1a). Under this manipulation, if the visual system is sensitive to the periodic change in the mean size of all displayed items, SSVEP at this base frequency should be observed (Ding et al., 2016; Norcia et al., 2015). In contrast, if the visual system processes only a small subset of items in the stimulus array, as suggested by the subsampling hypothesis, SSVEP at the base frequency would not be observed. TRF uses a linear forward modeling

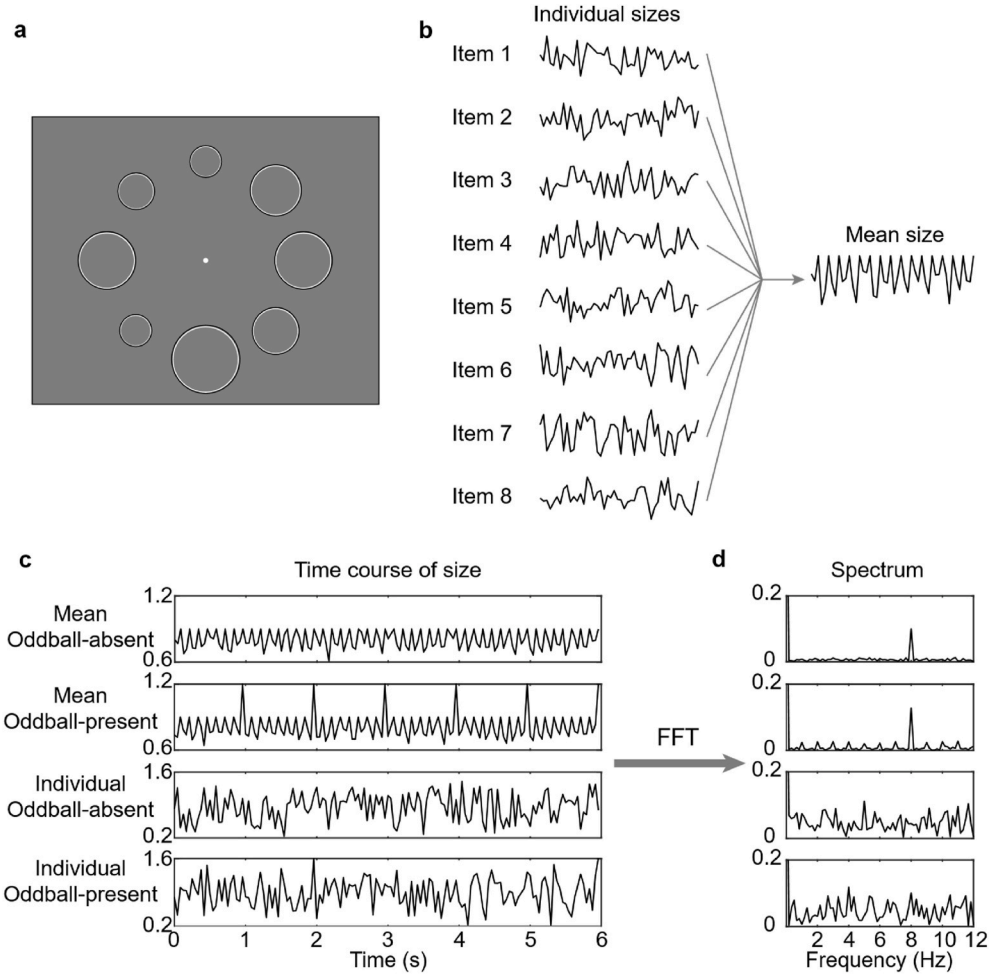


Fig. 1. Illustration of the stimulus display and examples of ensemble and individual circle size variations in Experiment 1. **a**. Illustration of the stimulus display. An array of circles was presented against a gray background at eight fixed locations. **b**. Illustration of the stimulus generation. The individual sizes varied randomly across time and were determined independently on the premise that their mean size changes periodically across time. **c**. The example variations of the ensemble size (top two rows) and individual circle size (bottom two rows) across time. The ensemble size varied periodically at 8 Hz, with an oddball appeared at 1 Hz in half of the trials. The sizes of individual circles varied randomly and independently in both oddball-absent and oddball-present trials. **d**. Fourier spectrums of the corresponding time series in **c**. The spectrums showed a clear 8 Hz response in all trials, and a clear 1 Hz response and its higher harmonics only in the oddball-present trials. The size variation of individual circles did not show any clear periodic component.

approach to extract EEG responses specific to certain sensory stimulus inputs from the overall EEG recordings (Ding and Simon, 2012; Lalor et al., 2006). Here, to explore the neural representations and contributions of individual sizes and their interactions, we defined three components: individual sizes, their local interactions, and their global interaction. Operationally, the local interactions were defined as the products of two neighboring individual sizes and the global interaction between all individuals was defined as the product over eight individual sizes (Kim et al., 2015; Werner and Noppeney, 2010). If the individual items interact with each other in the brain, the local and/or global interaction components should be represented in the brain and contribute to the ensemble perception.

Previous studies have demonstrated that ensemble representation is enhanced when attention is distributed over the global pattern as opposed to when attention is focused on a single item (Baek and Chong, 2020a; Chong and Treisman, 2005b; de Fockert and Marchant, 2008). It is unclear which components, individuals or their local or global interactions, are modulated by attentional distribution. Modulating the attentional distribution provides a chance to explore which components are critical to ensemble perception. Therefore, Experiment 2 further investigated modulation of spatial attention distribution on ensemble perception. Taken together, the present study aims to address a series of questions regarding the neural mechanism of ensemble perception, including (1) the neural signature of ensemble size perception, (2) whether the individual sizes and their local and global interactions are represented in the brain and contribute to ensemble size perception, and (3) whether the representations of individual sizes and their interactions are modulated by attentional distribution.

2. Materials and methods

2.1. Participants

A total of 45 human volunteers (23 female, age range: 18–24 years) participated in this study. Experiment 1 (n = 23) and Experiment 2 (n = 22) were performed at Peking University and Hangzhou Normal University, respectively. Our sample sizes were comparable to those in recent studies using similar TRF methods (Broderick et al., 2019; Jia et al., 2017; O'Sullivan et al., 2019), which were typically from 16 to 20. We decided to recruit 23 and 22 participants for the two experiments because our data analysis would estimate 17 independent neural responses (see below) from the overall EEG signals. All participants were right-handed, reported normal or corrected-to-normal vision, and had no known neurological or visual disorders. They gave written, informed consent in accordance with the procedures and protocols approved by the Human Subject Review Committee of Peking University or Hangzhou Normal University.

2.2. Experiment 1

Apparatus. We used MATLAB (version 9.5, The MathWorks) and Psychotoolbox-3 extensions (Brainard, 1997; Pelli, 1997) to generate and display visual stimuli and record behavioral responses. The visual stimuli were presented on a Display++ LCD monitor (Cambridge Research Systems) with a 1920 × 1080 spatial resolution and a 120 Hz refresh rate. Electroencephalography (EEG) recording was carried out in a dark room shielded from sound and electromagnetic signals. The participants were comfortably seated at 57 cm from the screen with their heads stabilized on a chin rest.

Stimuli. An array of circles were presented against a gray background (luminance: 54 cd/m^2) at eight fixed locations (Fig. 1a). The eccentricity of the center of each circle was 9° and the distance between adjacent circle centers was 6.9°. As depicted in Fig. 1, the edge of each circle consisted of an outer black line (luminance: 0.41 cd/m^2) and an inner white line (luminance: 107.49 cd/m^2), so that the mean luminance of each circle was equal to the background luminance.

A 6-s stimulus sequence consisted of 144 circle arrays, which were presented successively. Each circle array was presented for 41.7 ms (5 frame), so the sequence was updated at 24 Hz. The circle sizes in each array were determined independently using the following procedure (Fig. 1b). First, the radius of each of the eight circles was drawn from a uniform distribution between 0.5° and 1°. At the 3rd update of every 3 updates, the mean radius of the eight circles was set to 0.9°, so that the mean size changed periodically at 8 Hz (i.e., base frequency; Fig. 1b). In an oddball-present sequence, at the last update of every 24 updates, the mean radius of the eight circles was set to a larger size of 1.2° or 1.4°, so that the oddball array appeared periodically at 1 Hz. Second, to enlarge the size differences among the circles in each array, four reference sizes were randomly drawn from a uniform distribution between 0.1° and 0.4°. The radii of four randomly selected circles were increased by the reference sizes and the radii of the rest four circles were reduced by the reference sizes. This procedure ensured that, while the radius of each circle changed randomly, the mean radius of the eight circles changed periodically at 8 Hz, and the oddball array appeared periodically at 1 Hz (Fig. 1c and d).

Procedure. Each participant completed two blocks of 60 trials. In a trial, a 6-s stimulus sequence was presented, and participants were instructed to maintain fixation on a small point at the center of the display. The mean radius of the oddball array was 1.2° in one block and 1.4° in the other block. Half of the trials in each block were oddball-present trials, and the other half were oddball-absent trials. The order of oddball-present trials and oddball-absent trials was randomized in each block. The order of the two blocks was counterbalanced across participants. Participants were asked to judge whether oddball arrays were present or not by pressing one of two keys. Then, the next trial started after a 1–1.5s interval. Participants were prompted to rest after completing every 30 trials. Before the formal experiment, the participants completed 10 trials to get familiar with the experimental procedure. It took about 20 min to complete the whole experiment.

2.3. Experiment 2

Apparatus. The visual stimuli were presented on a CRT monitor with a 1024 × 768 spatial resolution and an 85 Hz refresh rate. EEG recordings were carried out in a dark room shielded from sound noise. The participants were comfortably seated 70 cm from the screen with their heads stabilized on a chin rest.

Stimuli. An array of circles were presented at the same fixed locations on a gray background (luminance: 52 cd/m^2) with the same eccentricity and circle-to-circle distance as in Experiment 1 (Fig. 1A). Similar to Experiment 1, the edge of each circle consists of an outer black line (luminance: 0 cd/m^2) and an inner white line (luminance: 104 cd/m^2), so that the mean luminance of each circle was equal to the background luminance.

A 6-s stimulus sequence consisted of 102 circle arrays, which were presented successively. Each circle array was presented for 58.5 ms (5 frame), so the sequence was updated at 17 Hz. At the 3rd update of every 3 updates, the mean radius of the eight circles was set to 0.9°, so that the mean size of the circles changed periodically at 5.7 Hz. The circle sizes in each array were determined using the same procedure as in Experiment 1, ensuring that while the radius of each circle changed randomly, the mean radius of the eight circles changed periodically at 5.7 Hz.

Procedure. Each participant completed 3 blocks of 40 trials. In the three blocks, the participants were required to attend to the ensemble pattern, the fixation, and a single item, respectively. In the attend-to-ensemble block, in 10 randomly selected trials, all circles were set to be of the same size for 1 s at a random time between 1 and 6 s after trial onset. In the attend-to-fixation block, in 10 randomly selected trials, the luminance of the fixation point was reduced from 104 to 82 cd/m^2 for 0.35 s at a random time between 1 and 6 s after trial onset. In the attend-to-item block, in 10 randomly selected trials, the Michelson contrast of the circle at the left side of the fixation was reduced from 1 to 0.93 for

0.47 s at a random time between 1 and 6 s after trial onset. The participants were asked to detect whether the circles were of the same size (the attend-to-ensemble condition), the changes in luminance of the fixation (the attend-to-fixation condition), and changes in contrast of the left circle (the attend-to-item condition) in three blocks respectively, by pressing a key. In each trial, a 6-s stimulus sequence was presented, and participants were instructed to maintain fixation on a small point at the center of the display. There was a 1–1.5s interval between trials. The order of the 3 blocks was counterbalanced across participants. At the beginning of each block, the participants completed 5 trials to get familiar with the task. It took about 15 min to complete the whole experiment.

2.4. EEG data acquisition and processing

EEG signals were recorded continuously at 1000 Hz using two BrainAmp amplifiers and a 64-channel EasyCap (BrainProducts). FCz electrode was used as reference and electrode impedances were maintained below 5 kΩ during data acquisition. Horizontal and vertical electrooculograms were recorded by two additional electrodes around the subjects' eyes. EEG signals were preprocessed using the FieldTrip toolbox (Oostenveld et al., 2011). They were re-referenced to the average value of all channels except two electrooculograms and were offline band-pass filtered between 0.1 and 40 Hz using a Butterworth IIR filter with the order of 2. The signals were then downsampled to the same frequency as the screen refresh rate (i.e., 120 Hz in Experiment 1 and 85 Hz in Experiment 2) for temporal response function (TRF) estimation (Lalor et al., 2006). Independent component analysis was performed to remove eye-movement and artifact components, and the remaining components were back-projected onto the EEG electrode space. 6-s EEG epochs during stimulus sequence presentation were segmented for each trial and used for further analyses.

2.5. SSVEP analysis

Evoked activities were computed by averaging EEG epochs for each condition and each participant. FFT was applied to the evoked activities after applying a Hanning taper to calculate the power spectrum for each channel with a frequency resolution (the size of the frequency bins) of 0.167 Hz. A baseline-correction procedure was used to extract SSVEP responses from baseline noise across the frequency spectrum (Meigen and Bach, 2000). Specifically, the difference between the power in the bin of interest and the mean power in the six surrounding bins was computed. In Experiment 2, only the trials without any manipulation were included in the analysis to exclude the influence of target detection on SSVEP.

2.6. Predicting EEG responses using individual circle sizes and their interactions as predictors

We used a forward TRF approach to predict EEG responses using individual circle sizes and their interactions as predictors. TRF describes the brain's linear transformation of the stimulus input, S(t), to the neural response output, R(t), as R(t) = TRF * S(t), where * denotes the convolution operator (Jia et al., 2017, 2019; Lalor et al., 2006). TRF was defined as a 1 s length neural response to each unit change in a predictor and was computed by a regularized linear regression between the predictor value and EEG amplitude. A parameter λ was used to control overfitting in the ridge regression.

In this study, we used eight circle sizes ($I_{1\text{-}8}$, they are I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8), eight local interactions ($L_{1\text{-}8}$, they are $I_1 \times I_2$, $I_2 \times I_3$, $I_3 \times I_4$, $I_4 \times I_5$, $I_5 \times I_6$, $I_6 \times I_7$, $I_7 \times I_8$, $I_8 \times I_1$) and a global interaction (G, that is $I_1 \times I_2 \times I_3 \times I_4 \times I_5 \times I_6 \times I_7 \times I_8$) to predict EEG responses (Best and Wolf, 2015; Smith and Kutas, 2015; Werner and Noppeney, 2010):

$$EEG = \left(\sum_{i=1}^{8} TRF_{Ii} \ * \ I_i\right) + \left(\sum_{i=1}^{8} TRF_{Li} * L_i\right) + TRF_G * G$$

where I_i and TRF_{Ii} are the size of the i-th circle and the corresponding TRF, L_i and TRF_{Li} are the i-th local interaction and the corresponding TRF, and G and TRF_G are the global interaction and the corresponding TRF. The individual size predictors represented the information of the eight circles, the local-interaction predictors represented the information of interactions between two neighboring individuals, and the global-interaction predictor represented the information of the highest-order interaction over all circles (Smith and Kutas, 2015; Werner and Noppeney, 2010).

The TRF-based EEG prediction was performed using the multivariate temporal response function (mTRF) toolbox (Michael J. Crosse et al., 2016). The λ values in all models were set to 1 for all subjects in our experiments. Each predictor was converted to z score before model fitting to reduce structural multicollinearity (Frost, 2019).

We quantified how well the individual circle sizes and their local and global interactions were encoded in the EEG signals using a leave-one-trial-out cross-validation procedure. TRFs were trained on N-1 trials and convolved with the predictors of the left-out trial to predict the channel-specific EEG signals (Broderick et al., 2019; Ding and Simon, 2012). The squared Pearson correlation coefficient (Michael J. Crosse et al., 2016; Frost, 2019) between the predicted and the recorded EEG signals were used to quantify the predicted accuracy (i.e., predicted R^2). The advantage of using predicted accuracy for model evaluation is that it is sensitive to model overfitting. Because it is impossible to predict random noise, the predicted accuracy must drop for an overfit model that adds random noise to the model as predictors (Michael J. Crosse et al., 2016; Frost, 2019). If the individual circle sizes and their interactions were encoded in the EEG signals, the predicted accuracy should increase after adding these predictors to the TRF model. We defined a full model using the individual circle sizes, local interactions, and global interaction as the predictors of the TRF model to predict the EEG signals (the *ILG* model) (see above). We then defined three reduced models that used the individual circle sizes and the local interactions (the *IL* model), the individual circle sizes and the global interaction (the *IG* model), or the local interactions and the global interaction (the *LG* model) as the predictors to predict the EEG signals, respectively. The encoding of the individual circle sizes, the local interactions and the global interaction can be quantified as the predicted accuracy difference between the ILG and LG models, the ILG and IG models and the ILG and IL models, respectively. If the added predictor is encoded in the EEG signals, the predicted accuracy difference would be positive, and if not, negative or zero (Michael J. Crosse et al., 2016; Frost, 2019).

After obtaining the predicted EEG signals in the full model and the three reduced models, we performed the same SSVEP analysis on the predicted EEG signals as on the recorded EEG signals to measure the base frequency power. To investigate the contribution of the individual circle sizes, the local interactions, and the global interaction to the ensemble size representation, the base frequency power of the predicted EEG signals was compared between the full model and the three reduced models. Our logic is that if a certain predictor contributes to the ensemble size representation, adding the predictor to the TRF model should increase the base frequency power. Otherwise, the base frequency power will not increase (Frost, 2019).

2.7. Statistical tests

The statistical significance of SSVEP power and predicted accuracy against zero in single conditions were evaluated using nonparametric one-sample Wilcoxon signed-rank tests (two-tailed). The statistical significance between conditions was evaluated using nonparametric Wilcoxon signed-rank tests (two-tailed). All p values were Bonferroni corrected. Effect sizes (rank biserial correlation, r_{rb}) were provided in all

tests. The difference of response accuracy across the three attention conditions in Experiment 2 was tested using a repeated measures ANOVA.

3. Results

3.1. SSVEP to the ensemble size

The participants performed well in the oddball detection task in Experiment 1, with a response accuracy of 0.96 ± 0.01 (Mean ± SD). We first examined whether an SSVEP could be observed at the frequency of the ensemble size variation. In Experiment 1, while the sizes of the individual circles varied randomly and independently, the ensemble size varied at a base frequency of 8 Hz in all trials and at an oddball frequency of 1 Hz in the oddball-present trials. The SSVEP power was computed for each participant and condition and then averaged across participants. At parieto-occipital electrodes, we found a clear 8 Hz SSVEP in both the oddball-present (Fig. 2a, left) and the oddball-absent trials (Fig. 2a, right) and a clear 1 Hz SSVEP only in the oddball-present trials (Fig. 2b). POz exhibited the strongest power (i.e., the peak power electrode). Consistent with the observation, at POz, significant higher-than-baseline SSVEPs at 8 Hz were found in both the oddball-present and the oddball-absent trials (both $ps < 0.001$), whereas there was no significant difference between them ($W = 109.00$, $p = 0.393$, $r_{rb} = -0.21$) (Fig. 2c). This result reveals a neural signature of the ensemble size perception.

3.2. Contributions of individual circle sizes, local interactions, and global interaction to the ensemble size representation

Having demonstrated a neural signature of the ensemble size perception, we set out to explore the role of the local and global interactions between individual circles in the ensemble size representation. If the brain linearly averages the neural responses to all the individual sizes, the ensemble size representation should be solely based on the individual size representations. On the other hand, the brain may also process and take account of local and global interactions between individual sizes (Hock and Schöner, 2016; Wagemans et al., 2012).

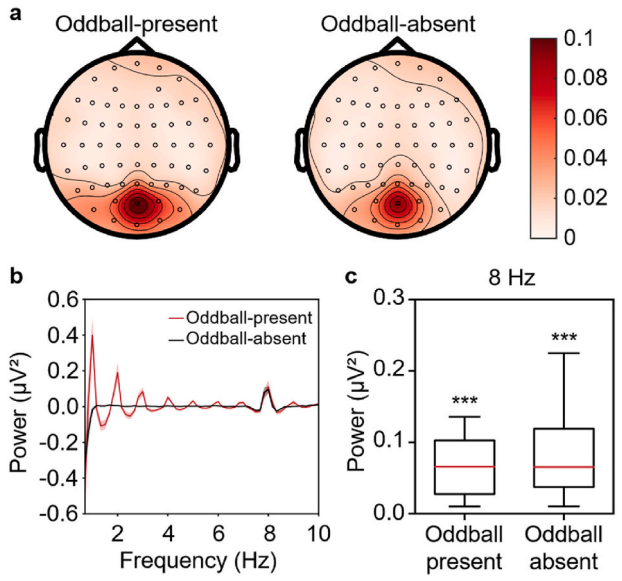


Fig. 2. SSVEP at the frequency of the ensemble size variation. **a**. SSVEP topographies at the frequency of the ensemble size variation in the oddball-present and the oddball-absent trials. **b**. Power spectrums of the EEG signals recorded at the peak-power electrode POz. Shaded areas around the curves denote one SEM (standard error of the mean) across participants. **c**. SSVEP power at the frequency of the ensemble size variation. Boxplot whiskers indicate minimum to maximum values, horizontal lines inside boxes indicate medians across participants, and boxes indicate values between the upper and lower data quartiles. ***$p < 0.001$.

To explore the two possibilities, we first evaluated the neural representation of the individual circle sizes and their local and global interactions in the recorded EEG signals at each electrode. Employing the TRF approach, we defined a full model (the ILG model) and three reduced models (the IL, IG, and LG models) to predict the EEG signals. In the full model, the individual circle sizes (I) and their local (L) and global (G) interactions were used as three predictors; in the reduced models, only two of the three predictors were used (see Methods). The neural representation of a predictor was quantified as the predicted accuracy (quantified by the squared Pearson correlation coefficient between the recorded and predicted EEG signals) difference between the full model and the reduced model without that predictor. For example, the neural representation of the individual sizes (I) was quantified as difference in the predicted accuracy between the ILG and LG models.

As shown in Fig. 3a, the individual circle sizes, the local interactions, and the global interaction were mainly encoded at parieto-occipital electrodes, and the global interaction was encoded at frontal electrodes additionally. We selected five electrodes, marked in Fig. 3a, with the best representation performance for the individual circle sizes, the local interactions, and the global interaction for further statistical analyses. The neural representation of the individual circle sizes, local interactions, and global interaction were significantly above zero (individuals circle sizes: $W = 276.00$, $p < 0.001$, $r_{rb} = 1.00$; local interactions: $W = 276.00$, $p < 0.001$, $r_{rb} = 1.00$; global interaction: $W = 276.00$, $p < 0.001$, $r_{rb} = 1.00$) (Fig. 3b). Note that the predicted R^2 are comparable to those in previous studies using the TRF approach (Broderick et al., 2019; M.J. Crosse et al., 2016; M.J. Crosse et al., 2016) which are typically between 0.01 and 0.1. Furthermore, in our data analysis, the overall predicted R^2 of the EEG signals was divided into three components, i.e., individual circle sizes, local interactions, and global interaction. The predicted R^2 of the individual sizes and the local interactions were larger than that of the global interaction. This is because both the individual sizes and the local interactions had eight terms in the regression model, while the global interaction had only one term. Despite that, the significant representations of the interactions, especially the global interaction, in the EEG signals demonstrate the existence of an interaction mechanism during ensemble size perception in the brain.

Next, we examined the contribution of the individual circle sizes, the local interactions, and the global interaction to the ensemble size representation. We calculated the 8 Hz SSVEP power of the predicted EEG signals in the full model and the three reduced models. Following the same logic as above, if a predictor contributed to the ensemble size perception, adding the predictor to the model should increase the SSVEP power. Otherwise, the SSVEP power would not change or even decrease. The same five electrodes for each predictor as above were selected for statistical analysis. As shown in Fig. 3c, the neural representation of the global interaction ($W = 216.00$, $p = 0.049$, $r_{rb} = 0.57$), rather than the individual circle sizes or the local interactions (individual circle sizes: $W = 83.00$, $p = 0.294$, $r_{rb} = -0.40$; local interactions: $W = 82.00$, $p = 0.275$, $r_{rb} = -0.41$), had a significant positive contribution to SSVEP predictive power. These results provide strong evidence for the key role of the global interaction in ensemble size perception.

3.3. Attentional effect on the SSVEP to the ensemble size

Studies have demonstrated that the ensemble representation is enhanced when attention is distributed over the global pattern, as opposed to when attention is focused on a single item (Baek and Chong, 2020a; Chong and Treisman, 2005b; de Fockert and Marchant, 2008). Therefore, if the SSVEP in Experiment 1 is a neural signature of the ensemble size perception, we predict that the SSVEP should be enhanced when attention is paid to the global pattern. To explore this, Experiment

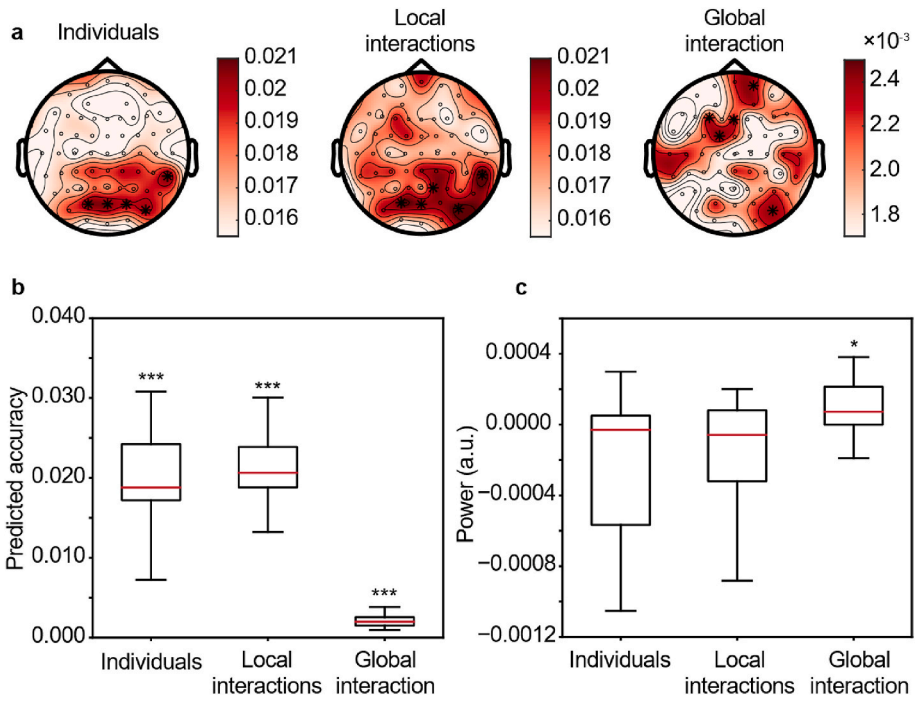


Fig. 3. Neural representation of the individual circle sizes, the local interactions, and the global interaction and their contributions to the SSVEP to the ensemble size. **a**. Topographies of predicted accuracy differences between the full and the three reduced models for quantifying the neural representation of the three predictors. Five electrodes with the best representation performance for each predictor, marked with asterisk, were selected for statistical analyses. **b**. Predicted accuracy differences between the full and the three reduced models, which were ascribed to the three predictors. **c**. Contributions of the three predictors to the SSVEP power. ***$p < 0.001$, *$p < 0.05$.

2 included three attention conditions: attend-to-ensemble, attend--to-fixation, and attend-to-item.

The response accuracies were 0.78 ± 0.04, 0.84 ± 0.01, and 0.76 ± 0.03 for the attend-to-ensemble, attend-to-fixation, and attend-to-item conditions, respectively. There was no significant difference across the three conditions ($F(2,42) = 2.19$, $p = 0.12$, partial $\eta^2 = 0.10$).

Replicating the findings in Experiment 1, the largest SSVEP power was observed also at the POz electrode (Fig. 4) and significant SSVEPs at the base frequency were found in all the three conditions (attend-to-ensemble: $W = 252$, $p < 0.001$, $r_{rb} = 0.99$; attend-to-fixation: $W = 253$, $p < 0.001$, $r_{rb} = 1.00$; attend-to-item: $W = 250$, $p < 0.001$, $r_{rb} = 0.98$).

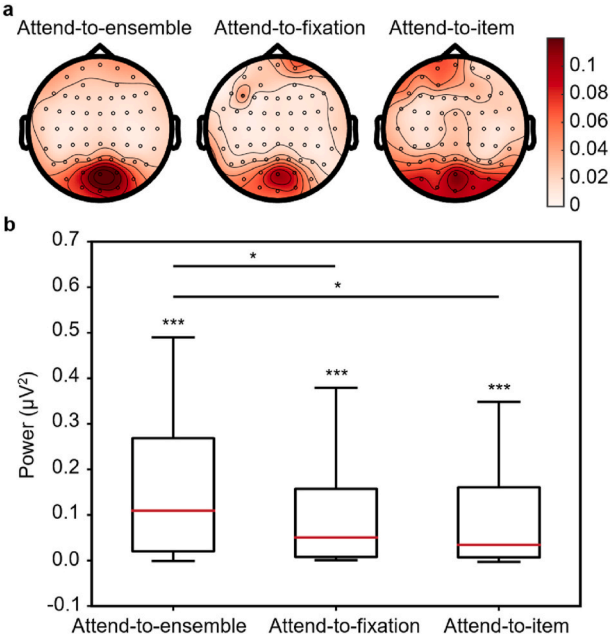


Fig. 4. Attentional effect on the SSVEP to the ensemble size. **a**. Topographies of the SSVEP power in the three attention conditions. **b**. SSVEP power at POz in the three attention conditions. ***$p < 0.001$, *$p < 0.05$.

More importantly, the SSVEP power was significantly greater in the attend-to-ensemble condition than in the attend-to-fixation ($W = 210$, $p = 0.016$, $r_{rb} = 0.66$) and attend-to-item ($W = 213$, $p = 0.011$, $r_{rb} = 0.68$) conditions. No significant difference was found between the attend-to-fixation and the attend-to-item conditions ($W = 170$, $p = 0.498$, $r_{rb} = 0.34$). These results are consistent with the attentional effect on ensemble perception at the behavioral level (Baek and Chong, 2020a; Chong and Treisman, 2005b; de Fockert and Marchant, 2008) and demonstrate that distributed attention over the global pattern enhances the specialized brain response to the ensemble size.

3.4. Attentional effects on the contributions of individual circle sizes, local interactions, and global interaction to the ensemble size perception

The manipulation of attention also allowed to examine how attentional distribution over the circle array affects the neural representations of the individual circle sizes, the local interactions, and the global interaction, as well as their contributions to the SSVEP. As in Experiment 1, we calculated the predicted accuracy differences between the full model and the three reduced models. The same five electrodes as in Experiment 1 were selected for statistical analyses. We found that the neural representation of the individual sizes was significant in all three attention conditions (Fig. 5a, all *ps* < 0.001), with no cross-condition differences (all *ps* > 0.05). Similarly, the neural representation of the local interactions was significant in all three conditions (Fig. 5b, all *ps* < 0.01), with no cross-condition differences (all *ps* > 0.05). However, the neural representation of the global interaction exhibited a different pattern (Fig. 5c). While the neural representation of the global interaction was significantly above zero in the attend-to-ensemble condition ($W = 253.00$, $p < 0.001$, $r_{rb} = 1.00$), it was not different from zero in the attend-to-fixation condition ($W = 89.00$, $p = 0.704$, $r_{rb} = -0.30$) and even significantly below zero in the attend-to-item condition ($W = 18.00$, $p < 0.001$, $r_{rb} = -0.86$). Importantly, the neural representation of the global interaction in the attend-to-ensemble condition was significantly greater than those in the attend-to-fixation ($W = 239.00$, $p < 0.001$, $r_{rb} = 0.90$) and the attend-to-item ($W = 252.00$, $p < 0.001$, $r_{rb} = 0.99$) conditions, and there was no significant difference between the attend-to-fixation and the attend-to-item conditions ($W = 179.00$, $p =$

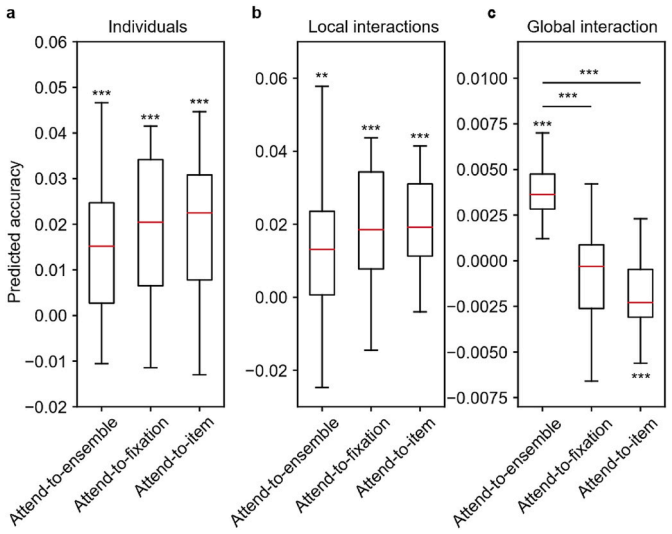


Fig. 5. Attentional effects on the neural representation of (**a**) the individual circle sizes, (**b**) the local interactions, and (**c**) the global interaction in the three attention conditions. ***$p < 0.001$, **$p < 0.01$.

0.275, $r_{rb} = 0.42$). These results demonstrated that the neural representations of the individual circle sizes and the local interactions were not modulated by the attentional distribution. In contrast, the neural representation of the global interaction could be enhanced by distributed attention on the global pattern.

Next, we directly examined how the contributions of the neural representation of the individual circle sizes, the local interactions, and the global interaction to the ensemble size perception were modulated by attention. We compared the SSVEP power of the predicted EEG signal in the full model with those in the three reduced models. As shown in Fig. 6, the neural representation of the individual sizes and the local interactions did not contribute significantly to the SSVEP power in all three attention conditions (all $ps > 0.05$). In contrast, the neural representation of the global interaction contributed significantly to the SSVEP power in the attend-to-ensemble ($W = 235.00$, $p < 0.001$, $r_{rb} = 0.86$) and attend-to-fixation conditions ($W = 211.00$, $p = 0.014$, $r_{rb} = 0.67$), but not in the attend-to-item condition ($W = 117.00$, $p = 1$, $r_{rb} = -0.08$). The contribution of the global interaction was significantly

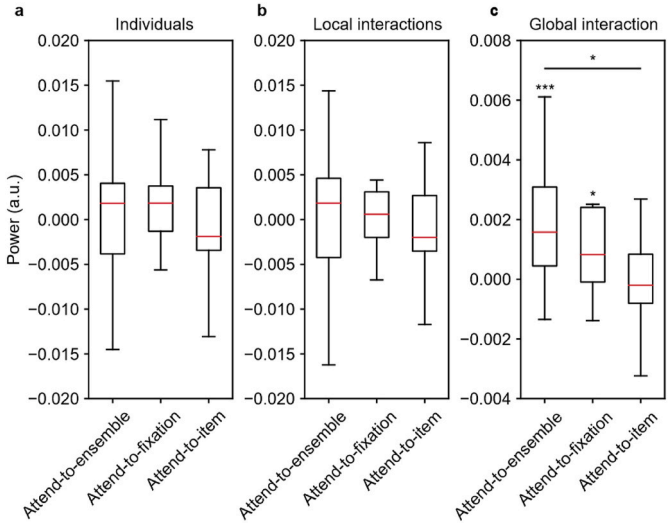


Fig. 6. Attentional effects on the contributions of (**a**) the individual circle sizes, (**b**) the local interactions, and (**c**) the global interaction to the SSVEP on the ensemble size in the three attention conditions. ***$p < 0.001$, *$p < 0.05$.

greater in the attend-to-ensemble condition than in the attend-to-item condition ($W = 204.00$, $p = 0.031$, $r_{rb} = 0.61$), demonstrating that attention to the ensemble enhanced the contribution of the global interaction to the ensemble size perception.

4. Discussion

We used a frequency-tagging technique, steady-state visual evoked potential (SSVEP), in combination with the temporal response function (TRF) technique, to study the neural signature and computational principle of ensemble size representation. Our findings provided evidence for a global mechanism of ensemble perception. Specifically, first, SSVEP showed clear electrophysiological responses that were synchronized with the frequency of the mean size changes, revealing that the human brain has a specialized neural response to ensemble size perception; second, using the TRF approach to predict EEG responses to individual sizes and their local and global interactions, we identified that the global interaction of all items in the display was encoded in EEG signals and contributed directly and significantly to ensemble size perception; finally, we identified, for the first time, that the attentional enhancement effects on ensemble perception were accompanied with an increased contribution only from the global interaction component, suggesting that the attentional enhancement on ensemble size perception derives from the effect of attention on the global interaction processing. Together, our findings support a specialized and global neural mechanism for ensemble size perception and suggest that the global interaction over all individuals contributes to ensemble size perception.

The existence of a specialized and global neural mechanism for ensemble perception is much debated. The subsampling hypothesis argues that ensemble size can be accurately estimated by randomly sampling and linearly averaging a few items strategically (Myczek and Simons, 2008; Solomon et al., 2011). However, this would not predict the EEG signals synchronized with the base frequency, which results from the variation of the mean size rather than combinations of a subset of individual items. Furthermore, the subsampling hypothesis could not explain the global interaction result, i.e., the interaction of all items contributed to the ensemble representation and was modulated by attention distribution. Together, these findings argue against the sub-sampling hypothesis. The present results thus support the summary-statistic representation hypothesis (Ariely, 2001; Chong and Treisman, 2003; Leib et al., 2016; Oh et al., 2019) and promote the existence of a global and specialized neural mechanism for ensemble perception.

Studies have shown that estimate of ensemble sizes can be more accurate than that of individual sizes. A linear averaging principle explains this phenomenon (Allik et al., 2013, 2014; Alvarez, 2011; Baek and Chong, 2020a; Sun and Chong, 2020). Linear averaging of multiple items tends to cancel out random errors in the processing of individuals, resulting in higher precision in ensemble size estimate, although individual size estimate is less accurate. However, this principle does not accommodate the attentional limitations, which restricts the visual processing to a limited number of individual sizes. Alternatively, the higher precision in the estimation of ensemble size can be achieved by utilizing interactions between individual items (Hock and Schöner, 2016; Wagemans et al., 2012). Taking advantage of the TRF method, the present study permits examination of these theoretical hypotheses directly with EEG responses. In addition to the neural representations of the individual circles, we consistently observed neural representations of the local and global interactions in EEG responses during ensemble perception. Crucially, the global interaction directly contributed to ensemble size perception and the contribution was enhanced by global attention. These results concomitantly indicate that ensemble size perception utilizes the high-order interaction information among individual sizes.

While interdependency among individual items has not been discussed in previous studies, researchers have argued that simple

linearization (e.g., simple averaging) could not generate efficient ensemble perception. Recent studies have demonstrated that the contributions of individuals are weighted in their mean (Choi and Chong, 2020) and variance (Jeong and Chong, 2021) computations for ensemble perception. Thus, these models and ours commonly suggest that ensemble perception is not achieved by simple averaging. However, because the present study did not manipulate the weights of individual circles and our TRF method could not compute the weights of individual circles in single trials, the current results could not examine the contribution of weighted averaging in ensemble perception. The global interaction component does not exclude the effect of weighted averaging either.

Intuitively, the ensemble size can be computed in a hierarchically connected, feedforward neural network. In this framework, low-level processing (i.e., the processing of individuals) determines high-level processing (i.e., the processing of the ensemble), and the representations of individual sizes are independent of each other. However, this pure feedforward framework does not explain the present results that the contribution of the representations of individuals to the ensemble size perception is not significant. It does not accommodate previous findings that the representations of individuals were degraded in ensemble perception (Allik et al., 2013, 2014; Ariely, 2001), either. The present results advocate recurrent computations (Edelman and Gally, 2013; Jastrzębowska et al., 2021; Lamme and Roelfsema, 2000; Singer, 2021) in ensemble size perception. The recurrent processing in the brain integrates the local visual elements and forms interactions between them (Roelfsema, 2006; Singer, 2021). Feedback connections from high-level visual areas may generate global interactions between all individual items, serving as a global mechanism for ensemble size perception. In contrast, the horizontal connections within each visual area could generate the local interactions. The function of interaction is thus a joint contribution of individuals which cannot be broken down to single ones.

To date, only a limited number of studies have investigated the neural substrates involved in ensemble statistics processing. Cant and Xu (2012) reported that the anterior-medial ventral visual cortex responded to ensemble statistics, whereas the lateral occipital area responded to individual objects. Im et al. (2017) proposed that the dorsal and ventral visual streams were involved in the ensemble emotion perception and individual face expression perception, respectively. Tark et al. (2021) found a gradual increase of selective responses to the mean orientation of multiple stimuli along the visual hierarchy when the mean orientation was task-relevant. Future research should explore the neural substrates involved in representing the individual items, and their local and global interactions.

The attention manipulation in Experiment 2 extends our knowledge of the influence of attention on ensemble size perception from behavioral performance (Baek and Chong, 2020a; Chong and Treisman, 2005b) to neural responses. Critically, our results confirm that the enhanced ensemble perception is accompanied with an increased contribution from the global interaction component rather than the individual item component. This finding suggests that the effect of attention on ensemble size perception derives from the effect of attention on the global interaction process. While it is a new finding in ensemble perception, attention improving interaction representation has been found widely at different visual hierarchies (Gordon et al., 2019; Zhang et al., 2011). We speculate that distributed attention can directly facilitate visual systems to integrate information from multiple items and send feedback information from higher areas to multiple lower areas, promoting global interactions between local elements.

In summary, the present study demonstrates a specialized and global neural mechanism for ensemble size perception in the brain. Specifically, ensemble size perception utilizes the representation of the global interaction between all individual sizes, which is enhanced by distributed attention over the global pattern.

Author contributions

J.J. and F.F. designed research; J.J., T.W., and S.C. performed research; J.J. and N.D. analyzed data; J.J. and F.F. wrote the paper.

Declaration of competing interest

The authors declare no conflicts of interest.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant 32000735 to J.J., Grant 31930053 to F.F.); Beijing Academy of Artificial Intelligence (BAAI) to F.F.; National Science and Technology Innovation 2030 Major Program 2022ZD0204802 to F.F.

References

Allik, J., Toom, M., Raidvee, A., Averin, K., Kreegipuu, K., 2014. Obligatory averaging in mean size perception. Vis. Res. 101, 34–40. https://doi.org/10.1016/j.visres.2014.05.003.

Allik, J., Toom, M., Raidvee, A., Averin, K., Kreegipuu, K., 2013. An almost general theory of mean size perception. Vis. Res. 83, 25–39. https://doi.org/10.1016/j.visres.2013.02.018.

Alvarez, G.A., 2011. Representing multiple objects as an ensemble enhances visual cognition. Trends Cognit. Sci. 15, 122–131. https://doi.org/10.1016/j.tics.2011.01.003.

Angelaki, D.E., Gu, Y., DeAngelis, G.C., 2009. Multisensory integration: psychophysics, neurophysiology, and computation. Curr. Opin. Neurobiol. 19, 452–458. https://doi.org/10.1016/j.conb.2009.06.008.

Ariely, D., 2008. Better than average? When can we say that subsampling of items is better than statistical summary representations? Percept. Psychophys. 70, 1325–1326. https://doi.org/10.3758/PP.70.7.1325.

Ariely, D., 2001. Seeing sets: representation by statistical properties. Psychol. Sci. 12, 157–162. https://doi.org/10.1111/1467-9280.00327.

Baek, J., Chong, S.C., 2020a. Ensemble perception and focused attention: two different modes of visual processing to cope with limited capacity. Psychon. Bull. Rev. https://doi.org/10.3758/s13423-020-01718-7.

Baek, J., Chong, S.C., 2020b. Distributed attention model of perceptual averaging. Atten. Percept. Psychophys. 82, 63–79. https://doi.org/10.3758/s13414-019-01827-z.

Best, H., Wolf, C. (Eds.), 2015. The SAGE Handbook of Regression Analysis and Causal Inference. SAGE Reference, Los Angeles [Calif.

Brainard, D.H., 1997. The psychophysics toolbox. Spatial Vis. 10, 433–436.

Broderick, M.P., Anderson, A.J., Lalor, E.C., 2019. Semantic context enhances the early auditory encoding of natural speech. J. Neurosci. 39, 7564–7575. https://doi.org/10.1523/JNEUROSCI.0584-19.2019.

Cant, J.S., Xu, Y., 2012. Object ensemble processing in human anterior-medial ventral visual cortex. J. Neurosci. 32, 7685–7700. https://doi.org/10.1523/JNEUROSCI.3325-11.2012.

Choi, Y.M., Chong, S.C., 2020. Effects of selective attention on mean-size computation: weighted averaging and perceptual enlargement. Psychol. Sci. 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. Percept. Psychophys. 70, 1327–1334. https://doi.org/10.3758/PP.70.7.1327.

Chong, S.C., Treisman, A., 2005a. Statistical processing: computing the average size in perceptual groups. Vis. Res. 45, 891–900. https://doi.org/10.1016/j.visres.2004.10.004.

Chong, S.C., Treisman, A., 2005b. Attentional spread in the statistical processing of visual displays. Percept. Psychophys. 67, 1–13. https://doi.org/10.3758/BF03195009.

Chong, S.C., Treisman, A., 2003. Representation of statistical properties. Vis. Res. 43, 393–404.

Crosse, Michael J., Di Liberto, G.M., Bednar, A., Lalor, E.C., 2016. The multivariate temporal response function (mTRF) toolbox: a MATLAB toolbox for relating neural signals to continuous stimuli. Front. Hum. Neurosci. 10, 604. https://doi.org/10.3389/fnhum.2016.00604.

Crosse, M.J., Di Liberto, G.M., Lalor, E.C., 2016. Eye can hear clearly now: inverse effectiveness in natural audiovisual speech processing relies on long-term crossmodal temporal integration. J. Neurosci. 36, 9888–9895. https://doi.org/10.1523/JNEUROSCI.1396-16.2016.

de Fockert, J.W., Marchant, A.P., 2008. Attention modulates set representation by statistical properties. Percept. Psychophys. 70, 789–794. https://doi.org/10.3758/PP.70.5.789.

de Gardelle, V., Summerfield, C., 2011. Robust averaging during perceptual judgment. Proc. Natl. Acad. Sci. USA 108, 13341–13346. https://doi.org/10.1073/pnas.1104517108.

Ding, N., Melloni, L., Zhang, H., Tian, X., Poeppel, D., 2016. Cortical tracking of hierarchical linguistic structures in connected speech. Nat. Neurosci. 19, 158–164. https://doi.org/10.1038/nn.4186.

Ding, N., Simon, J.Z., 2012. Neural coding of continuous speech in auditory cortex during monaural and dichotic listening. J. Neurophysiol. 107, 78–89. https://doi.org/10.1152/jn.00297.2011.
Edelman, G.M., Gally, J.A., 2013. Reentry: a key mechanism for integration of brain function. Front. Integr. Neurosci. 7 https://doi.org/10.3389/fnint.2013.00063.
Epstein, M.L., Emmanouil, T.A., 2021. Ensemble statistics can Be available before individual item properties: electroencephalography evidence using the oddball paradigm. J. Cognit. Neurosci. 33, 1056–1068. https://doi.org/10.1162/jocn_a_01704.
Frost, J., 2019. Regression Analysis: an Intuitive Guide for Using and Interpreting Linear Models. James D. Frost.
Gordon, N., Hohwy, J., Davidson, M.J., van Boxtel, J.J.A., Tsuchiya, N., 2019. From intermodulation components to visual perception and cognition-a review. Neuroimage 199, 480–494. https://doi.org/10.1016/j.neuroimage.2019.06.008.
Haberman, J., Whitney, D., 2007. Rapid extraction of mean emotion and gender from sets of faces. Curr. Biol. 17, R751–R753. https://doi.org/10.1016/j.cub.2007.06.039.
Hock, H.S., Schöner, G., 2016. Nonlinear dynamics in the perceptual grouping of connected surfaces. Vis. Res. 126, 80–96. https://doi.org/10.1016/j.visres.2015.06.006.
Iakovlev, A.U., Utochkin, I.S., 2021. Roles of saliency and set size in ensemble averaging. Atten. Percept. Psychophys. 83, 1251–1262. https://doi.org/10.3758/s13414-020-02089-w.
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. Nat. Human Behav. 1, 828–842. https://doi.org/10.1038/s41562-017-0225-z.
Jastrzębowska, M.A., Chicherov, V., Draganski, B., Herzog, M.H., 2021. Unraveling brain interactions in vision: the example of crowding. Neuroimage 240, 118390. https://doi.org/10.1016/j.neuroimage.2021.118390.
Jeong, J., Chong, S.C., 2021. Perceived variability reflects the reliability of individual items. Vis. Res. 183, 91–105. https://doi.org/10.1016/j.visres.2021.02.008.
Jia, J., Fang, F., Luo, H., 2019. Selective spatial attention involves two alpha-band components associated with distinct spatiotemporal and functional characteristics. Neuroimage 199, 228–236. https://doi.org/10.1016/j.neuroimage.2019.05.079.
Jia, J., Liu, L., Fang, F., Luo, H., 2017. Sequential sampling of visual objects during sustained attention. PLoS Biol. 15, e2001903 https://doi.org/10.1371/journal.pbio.2001903.
Kim, S.S., Gomez-Ramirez, M., Thakur, P.H., Hsiao, S.S., 2015. Multimodal interactions between proprioceptive and cutaneous signals in primary somatosensory cortex. Neuron 86, 555–566. https://doi.org/10.1016/j.neuron.2015.03.020.
Lalor, E.C., Pearlmutter, B.A., Reilly, R.B., McDarby, G., Foxe, J.J., 2006. The VESPA: a method for the rapid estimation of a visual evoked potential. Neuroimage 32, 1549–1561. https://doi.org/10.1016/j.neuroimage.2006.05.054.
Lamme, V.A.F., Roelfsema, P.R., 2000. The distinct modes of vision offered by feedforward and recurrent processing. Trends Neurosci. 23, 571–579. https://doi.org/10.1016/S0166-2236(00)01657-X.
Leib, A.Y., Kosovicheva, A., Whitney, D., 2016. Fast ensemble representations for abstract visual impressions. Nat. Commun. 7, 13186 https://doi.org/10.1038/ncomms13186.
Luck, S.J., Vogel, E.K., 1997. The capacity of visual working memory for features and conjunctions. Nature 390, 279–281.
Meigen, T., Bach, M., 2000. On the statistical significance of electrophysiological steady-state responses. Doc. Ophthalmol. 98, 207–232.
Myczek, K., Simons, D.J., 2008. Better than average: alternatives to statistical summary representations for rapid judgments of average size. Percept. Psychophys. 70, 772–788. https://doi.org/10.3758/PP.70.5.772.
Neumann, M.F., Schweinberger, S.R., Burton, A.M., 2013. Viewers extract mean and individual identity from sets of famous faces. Cognition 128, 56–63. https://doi.org/10.1016/j.cognition.2013.03.006.
Norcia, A.M., Appelbaum, L.G., Ales, J.M., Cottereau, B.R., Rossion, B., 2015. The steady-state visual evoked potential in vision research: a review. J. Vis. 15, 4. https://doi.org/10.1167/15.6.4.
Oh, B.-I., Kim, Y.-J., Kang, M.-S., 2019. Ensemble representations reveal distinct neural coding of visual working memory. Nat. Commun. 10 https://doi.org/10.1038/s41467-019-13592-6.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.-M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 1–9. https://doi.org/10.1155/2011/156869, 2011.
O'Sullivan, A.E., Lim, C.Y., Lalor, E.C., 2019. Look at me when I'm talking to you: selective attention at a multisensory cocktail party can be decoded using stimulus reconstruction and alpha power modulations. Eur. J. Neurosci. 50, 3282–3295. https://doi.org/10.1111/ejn.14425.
Palmer, E.M., Fencsik, D.E., Flusberg, S.J., Horowitz, T.S., Wolfe, J.M., 2011. Signal detection evidence for limited capacity in visual search. Atten. Percept. Psychophys. 73, 2413–2424. https://doi.org/10.3758/s13414-011-0199-2.
Parkes, L., Lund, J., Angelucci, A., Solomon, J.A., Morgan, M., 2001. Compulsory averaging of crowded orientation signals in human vision. Nat. Neurosci. 4, 739–744. https://doi.org/10.1038/89532.
Pelli, D.G., 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spatial Vis. 10, 437–442.
Roelfsema, P.R., 2006. Cortical algorithms for perceptual grouping. Annu. Rev. Neurosci. 29, 203–227.
Semizer, Y., Boduroglu, A., 2021. Variability leads to overestimation of mean summaries. Atten. Percept. Psychophys. 83, 1129–1140. https://doi.org/10.3758/s13414-021-02269-2.
Simons, D.J., Myczek, K., 2008. Average size perception and the allure of a new mechanism. Percept. Psychophys. 70, 1335–1336. https://doi.org/10.3758/PP.70.7.1335.
Singer, W., 2021. Recurrent dynamics in the cerebral cortex: integration of sensory evidence with stored knowledge. Proc. Natl. Acad. Sci. USA 118, e2101043118. https://doi.org/10.1073/pnas.2101043118.
Smith, N.J., Kutas, M., 2015. Regression-based estimation of ERP waveforms: II. Nonlinear effects, overlap correction, and practical considerations. Psychophysiology 52, 169–181. https://doi.org/10.1111/psyp.12320.
Solomon, J.A., Morgan, M., Chubb, C., 2011. Efficiencies for the statistics of size discrimination. J. Vis. 11 https://doi.org/10.1167/11.12.13, 13–13.
Sun, J., Chong, S.C., 2020. Power of averaging: noise reduction by ensemble coding of multiple faces. J. Exp. Psychol. Gen. 149, 550–563. https://doi.org/10.1037/xge0000667.
Sweeny, T.D., Haroz, S., Whitney, D., 2013. Perceiving group behavior: sensitive ensemble coding mechanisms for biological motion of human crowds. J. Exp. Psychol. Hum. Percept. Perform. 39, 329–337. https://doi.org/10.1037/a0028712.
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, 118680. https://doi.org/10.1016/j.neuroimage.2021.118680.
Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M., von der Heydt, R., 2012. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure–ground organization. Psychol. Bull. 138, 1172–1217. https://doi.org/10.1037/a0029333.
Werner, S., Noppeney, U., 2010. Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. Cerebr. Cortex 20, 1829–1842. https://doi.org/10.1093/cercor/bhp248.
Whitney, D., Yamanashi Leib, A., 2018. Ensemble perception. Annu. Rev. Psychol. 69, 105–129. https://doi.org/10.1146/annurev-psych-010416-044232.
Zhang, P., Jamison, K., Engel, S., He, B., He, S., 2011. Binocular rivalry requires visual attention. Neuron 71, 362–369. https://doi.org/10.1016/j.neuron.2011.05.035.