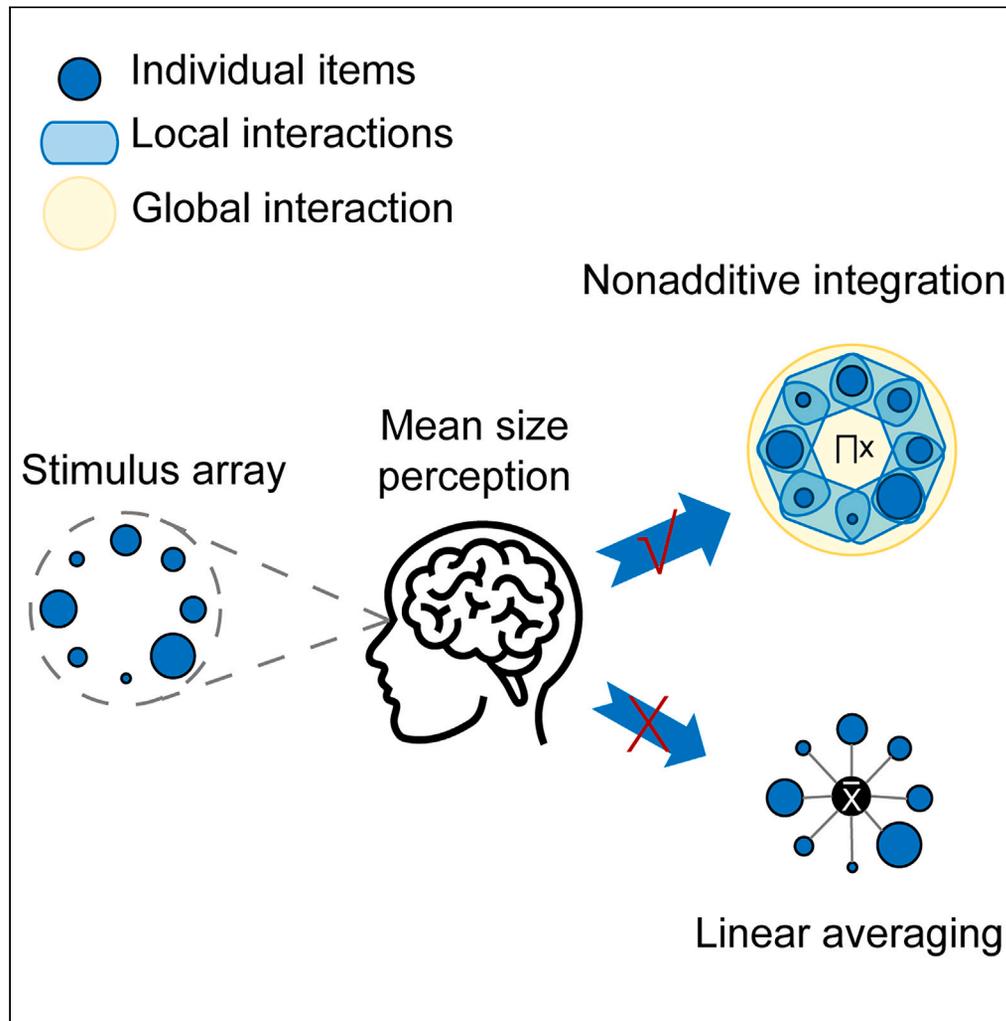


Article

# Nonadditive integration of visual information in ensemble processing



Tongyu Wang,  
 Yuqing Zhao,  
 Jianrong Jia

[jianrongjia@hznu.edu.cn](mailto:jianrongjia@hznu.edu.cn)

Highlights

The interactions between individual stimuli can elicit independent neural responses

Ensemble perception involves the global interaction of individual stimuli

Spreading attention enhances the neural representation of global interaction



## Article

## Nonadditive integration of visual information in ensemble processing

Tongyu Wang,<sup>1</sup> Yuqing Zhao,<sup>1</sup> and Jianrong Jia<sup>1,2,3,\*</sup>

## SUMMARY

**Statistically summarizing information from a stimulus array into an ensemble representation (e.g., the mean) improves the efficiency of visual processing. However, little is known about how the brain computes the ensemble statistics. Here, we propose that ensemble processing is realized by nonadditive integration, rather than linear averaging, of individual items. We used a linear regression model approach to extract EEG responses to three levels of information: the individual items, their local interactions, and their global interaction. The local and global interactions, representing nonadditive integration of individual items, elicited rapid and independent neural responses. Critically, only the neural representation of the global interaction predicted the precision of the ensemble perception at the behavioral level. Furthermore, spreading attention over the global pattern to enhance ensemble processing directly promoted rapid neural representation of the global interaction. Taken together, these findings advocate a global, nonadditive mechanism of ensemble processing in the brain.**

## INTRODUCTION

The visual system is constantly confronted with complex and cluttered information. Due to the limited capacity of the visual system,<sup>1,2</sup> processing detailed information from each item is mostly impossible. To facilitate information processing, the visual system would capture statistical regularities inherent in the visual experience, such as the mean value of visual features, and create compressed representations in the form of summary statistics.<sup>3–5</sup> Extensive work has shown that various visual features can be summarized at a glance, even faster than the processing of the visual features themselves.<sup>6,7</sup> The ability to represent ensemble visual information is a foundation of gist perception,<sup>3</sup> which may be realized by a specialized neural mechanism in the brain.<sup>8</sup> However, how the individual items are integrated to form the ensemble representation and its underlying neural mechanism are still unknown.

Intuitively, a mean representation can be generated by linearly averaging the representations of individual items,<sup>9–11</sup> either over all items<sup>4,5</sup> or over a subset of items,<sup>12,13</sup> where individual items were treated independently. On the other hand, ensemble perception is largely influenced by the internal structure of the stimulus set, such as the variance,<sup>14</sup> the Gestalt grouping,<sup>15</sup> and the number of items,<sup>16</sup> suggesting interdependencies of individual items during ensemble perception. Recent studies have further shown that individual items are not treated equally in ensemble perception. For example, items with extreme properties tend to be downweighted in the stimulus set, just as a statistician might exclude an outlying data point.<sup>17,18</sup> Weighting on individual items has been found in the mean<sup>19</sup> and variance<sup>20</sup> computations of ensemble perception. A critical logic behind these findings is that to be appropriately weighted, the individual items in the stimulus set would not be processed independently. Instead, it is necessary to capture the whole picture, the relationships between the individual items in the stimulus set. This raises the possibility that the interactions between individual items are involved in the computation of the ensemble representation.

Involving the interactions between individual items in the ensemble representation means that the individual items are integrated non-additively rather than linearly. Indeed, evidence of nonadditive integration of local information has been found in the brain for Gestalt psychology<sup>15,21</sup> and multisensory integration.<sup>22,23</sup> Given the shared characteristics of global processing in these processes, it is natural to ask whether nonadditive integration of individual items is represented in the brain and contributes to ensemble processing. To represent nonadditive integration, we built a linear regression model that included the information for the interactions between individual items (Figure 1A) and the information for the individual items as predictors. Specifically, in two experiments, we recorded electroencephalogram (EEG) signals from human participants and estimated the neural representations and contributions of three levels of information – the individual items, the local interactions between adjacent items, and the global interaction over all items – in mean size perception. The nonadditive integration was represented by two forms of interactions (Figure 1A): the local interaction, defined as the products of two adjacent items to capture the local horizontal connections,<sup>22,23</sup> and the global interaction, defined as the product over eight individual items to capture the recurrent neural connections across the neuronal populations representing all individual items.<sup>22,23</sup> Because multiplications of information from multiple items

<sup>1</sup>Department of Psychology, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China

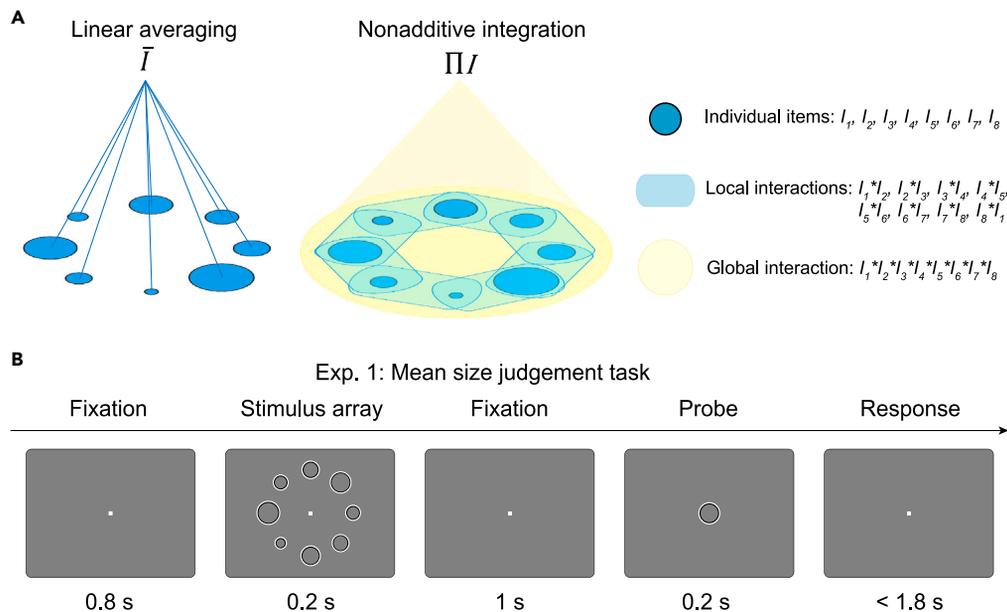
<sup>2</sup>Zhejiang Philosophy and Social Science Laboratory for Research in Early Development and Childcare, Hangzhou Normal University, Hangzhou 311121, Zhejiang, China

<sup>3</sup>Lead contact

\*Correspondence: [jianrongjia@hznu.edu.cn](mailto:jianrongjia@hznu.edu.cn)

<https://doi.org/10.1016/j.isci.2023.107988>





**Figure 1. Experimental hypotheses and the paradigm of Exp. 1**

(A) The linear averaging hypothesis (left) suggests that the ensemble representation is achieved by averaging the representations of individual items. The nonadditive integration hypothesis (middle) proposes that interactions (defined as multiplications) of items contribute to ensemble processing. Definitions of the individual items and their local and global interactions for the present study are illustrated (right).

(B) Each trial began with the presentation of a fixation point followed by a stimulus array. The stimulus array consisted of eight circles of different radii. A single probe circle was presented last. The participants responded by judging whether the size of the probe circle was larger than the average size of the stimulus array.

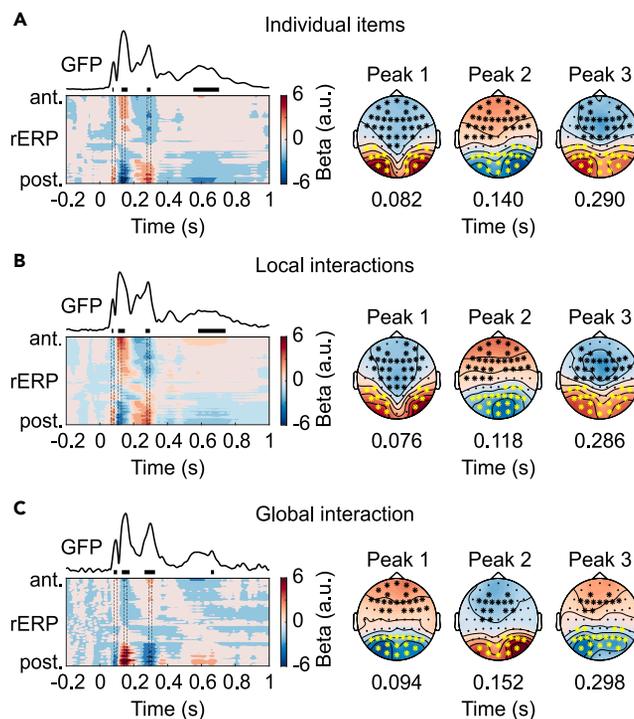
produce the neural representations of the interaction terms, they could not be decomposed into the representations of each individual item independently.

To disentangle the neural responses to each of the three levels of information from the overall EEG recordings, we used a regression-based event-related potentials (rERPs) framework.<sup>24–26</sup> In the regression framework, the temporal dynamics of the rERPs could reveal the dynamics of brain responses to a given level of visual information. We observed that the individual items and their local and global interactions were represented in the brain, and that each level of information contributed uniquely to the brain signals. Moreover, only the neural response to the global interaction was correlated with the precision of ensemble perception at the behavioral level, suggesting a critical role of the global interaction in ensemble processing. These findings were further supported by our manipulations on attention distributions within participants. Spreading attention, a process that enhances ensemble perception, modulated rapid neural responses to the global interaction which was specifically correlated with the ensemble perception rather than the item perception. Altogether, these results provide strong evidence for rapid neural representations of interactions between individual items during ensemble perception and suggest global, nonadditive integration of individual items in ensemble processing.

## RESULTS

### Neural representations of the individual items and their local and global interactions

In Experiment 1 ( $n = 24$ ), observers estimated the mean size of a circle array composed of eight circles of different radii. They judged whether the size of a probe circle was larger or smaller than the mean size (Figure 1B). Based on the previous findings that the visual system relies on the radius rather than the mathematical area<sup>27,28</sup> to determine the size of a circle, the radius was used as the measure of the circle size. Across trials, the mean radius of the circle array was uniformly distributed across 10 levels. In contrast, the radii of the individual circles in the array were randomly distributed. To explore the neural representations of the three levels of information (Figure 1A), we built a linear regression model. For each participant, three components representing different levels of information were included as predictors in the regression model to predict the amplitude of the recorded EEGs at each time point during the presentation of the circle array (–0.2 to 1 s relative to the onset of the circle array). Specifically, information for the individual items ( $I$ ) was represented by the radii of eight individual circles; information for the local interactions ( $L$ ) between individual items was represented by the eight products between adjacent circles; information for the global interaction, the highest order interaction over all circles,<sup>23,26</sup> was represented by the product of the eight circles ( $G$ ). The neural response weights generated by the linear regression models were obtained in each electrode, providing the rERPs for the seventeen predictors. The rERPs for the eight individual items and the eight local interactions were averaged to represent the neural representations of the individual items and the local interactions, respectively.



**Figure 2. rERPs to individual items and their local and global interactions in Experiment 1**

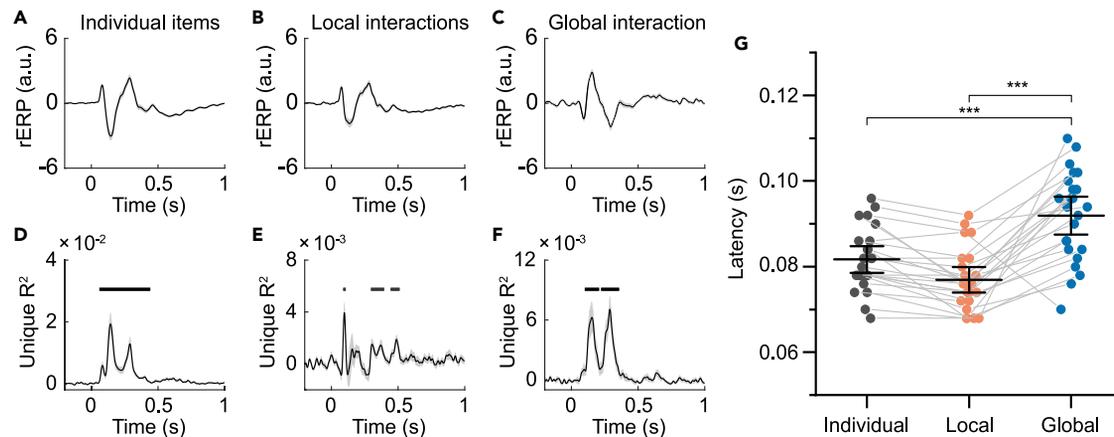
(A–C) (left) The rERP patterns (i.e., the beta values in all electrodes) and the corresponding global field powers (GFP) of the individual items (A), the local interactions (B), and the global interaction (C). Electrodes are arranged from the posterior to the anterior (bottom to top). The thick horizontal bars below the GFP curves mark the significant time points above the baseline. (right) The topographies at the first three peaks of the three predictors. The colors of the rERP patterns and the topographic maps indicate the sign and magnitude of the beta values. The black asterisks indicate frontoparietal clusters and the yellow asterisks indicate occipital clusters as revealed by the cluster-based permutation test. The a.u. indicates the arbitrary unit.

Figure 2 shows the rERP time courses for the three levels of information in an antero-posterior gradient from prefrontal to occipital electrodes. To quantify the overall activity at each time point, we calculated the global field power (GFP) of the rERPs across all valid electrodes.<sup>29</sup> As shown in the left panels of Figure 2, the GFP showed three peaks after the onset of the circle array in the three predictors. These peaks were significant for the individual items (Peak 1: 0.082 s; Peak 2: 0.14 s; Peak 3: 0.29 s), their local interactions (Peak 1: 0.076 s; Peak 2: 0.118 s; Peak 3: 0.286 s), and their global interaction (Peak 1: 0.094 s; Peak 2: 0.152 s; Peak 3: 0.298 s). Thus, the brain generated rapid neural responses to the individual items and their local and global interactions.

Interestingly, the rERP showed distinct spatial distributions between the global interaction and the other two components in each GFP peak. To better reveal the topography of the neural responses, we averaged the rERPs within a 20-ms window around each GFP peak and identified the significant electrodes using a cluster-based permutation test.<sup>30</sup> At the first GFP peak, the individual items (Figure 2A right) and their local interactions (Figure 2B right) showed two significant positive clusters of electrodes in the bilateral occipital area and one significant negative cluster of electrodes in the frontoparietal area. However, only two significant clusters were observed in the global interaction, one in the frontoparietal area and one in the occipital area across the three GFP peaks (Figure 2C right). For all the three neural representations, the spatial pattern reversed at the second GFP peak and reversed again at the third GFP peak.

### Unique contributions of the individual items and their local and global interactions to electroencephalogram signals

Due to the multiplicative nature of the local and global interactions, the three components of our model were inevitably correlated. Therefore, it is reasonable to assume that the neural representations for the local and global interactions across individual items are derived from the neural representation of the individual items. To examine this possibility, we tested whether the individuals, the local interactions, and the global interaction contributed independently to the EEG signals.<sup>31</sup> We constructed a full model that included the individual items, the local interactions, and the global integration as predictors (the *ILG* model). We then defined three reduced models using the local interactions and the global interaction (the *LG* model), the individuals and the global interaction (the *IG* model), and the individuals and the local interactions (the *IL* model) as predictors, respectively. The  $R^2$  differences between the full model and the three reduced models could reveal the unique contribution of the individuals (*ILG* vs. *LG*), the local interactions (*ILG* vs. *IG*), and the global interaction (*ILG* vs. *IL*). To compare the rERPs and the unique contributions to the EEG signals between the individual items and their local and global interactions, we restricted the analysis to six electrodes (P5, P7, PO7, P6, P8, and PO8) from the occipital region that commonly showed significant responses to all three levels of information (Figure 2 right). Figures 3A–3C showed the averaged rERPs across the six electrodes for the three predictors estimated in the full



**Figure 3. The unique contributions and response latencies of the individual items, the local interactions, and the global interaction**

(A–C) The rERPs to the individual items (A), the local interactions (B), and the global interaction (C) in selected occipital electrodes.

(D–F) The unique  $R^2$  of the three predictors across time. Horizontal bars indicate the significant time points above the baseline. Shaded areas indicate standard error of the mean (SEM) across participants.

(G) The latencies of the first rERP peak to the three predictors. Error bars indicate the 95% confidence interval. \*\*\* indicate  $p < 0.001$ . The a.u. indicates the arbitrary unit.

model. The unique  $R^2$ , as shown in Figures 3D–3F, were significant for all the three predictors shortly after the stimulus onset. Significant clusters were observed for the individual items (0.058–0.440 s), the local interactions (0.088–0.108 s, 0.298–0.400 s, and 0.446–0.512 s), and the global interaction (0.094–0.210 s and 0.218–0.360 s). Therefore, the local and global interactions of individual circles contribute independently to the EEG signals.

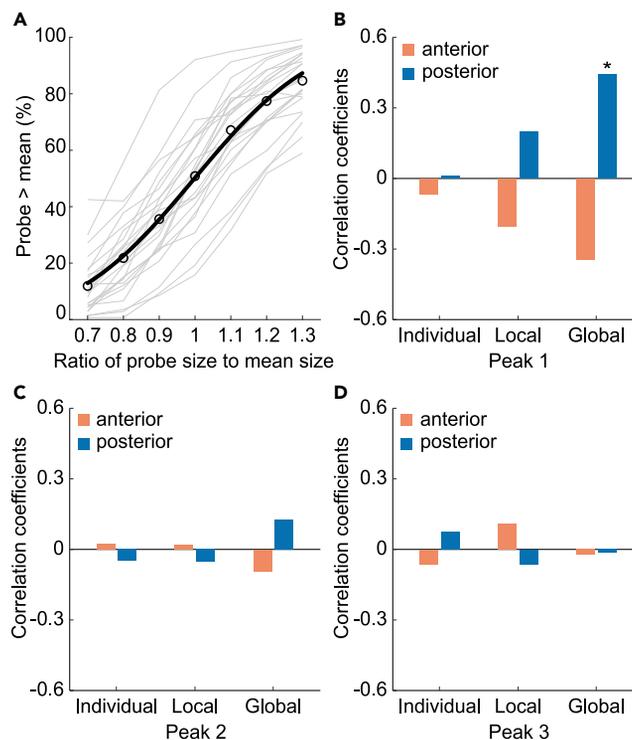
### The processing hierarchy of the individual items, the local interactions, and the global interaction

We examined the response latencies of the rERPs in the six selected occipital electrodes (Figures 3A–3C) to the individual items and their local and global interactions. The response latencies for the three predictors were extracted from the first rERP peaks (time range: 0.06–0.12 s) in each participant (Figure 3G). A repeated measures ANOVA showed a significant difference between the three latencies ( $F(1.29, 29.56) = 32.01$ ,  $p < 0.001$ , partial  $\eta^2 = 0.58$ ; Greenhouse-Geisser corrected). The Bonferroni-corrected post hoc analysis showed that the latency of the global interaction (0.092 s) was significantly later than that of the individual items (0.082 s,  $t(23) = 5.35$ ,  $p < 0.001$ , Cohen's  $d = 1.22$ ) and the local interactions (0.077 s,  $t(23) = 7.83$ ,  $p < 0.001$ , Cohen's  $d = 1.78$ ). The latency of the local interactions was not significantly different from that of the individual items ( $t(23) = 2.48$ ,  $p = 0.051$ , Cohen's  $d = 0.56$ ). The later peak latency suggested that the global interaction was processed at a higher level of the information processing hierarchy.<sup>32</sup>

### Behavioral correlates of the early neural response to the global interaction

While the neural representations of the three predictors have been demonstrated in the brain, it is unclear whether they are directly related to ensemble perception at the behavioral level. To answer this question, we estimated the correlation between the neural responses of the three predictors and the precision of ensemble perception across participants. In Experiment 1, the participants performed a size judgment task in which they judged whether the probe circle size was larger than the average size of the circle array. The overall response accuracy was  $76.69 \pm 5.04\%$  (mean  $\pm$  SD). As shown in Figure 4A, the proportion of responses for “probe is larger” monotonically as the ratio of the probe size to the mean size increased. A cumulative Gaussian function was fitted to the curve. The sigma value of the fitted function indicated the perceptual sensitivity to the difference between the probe size and the mean size. The smaller the sigma value, the higher the sensitivity, and the larger the sigma value, the lower the sensitivity.

The rERPs were extracted from the electrodes in the statistically significant clusters in the frontoparietal and occipital regions of Figure 2. For each predictor, the peak rERP values were calculated by first averaging the rERP amplitudes within a 20-ms window (peak  $\pm$  10 ms) around the three peak times respectively and then averaging the peak amplitudes across the electrodes in each cluster. Pearson correlations were performed between the sigma and peak rERP values in each cluster across participants. The results (Figure 4B) showed that, for the first peak, neither the rERPs of the individual items (Figure 4B left, anterior cluster:  $r = -0.07$ ,  $p = 0.753$ ; posterior cluster:  $r = 0.003$ ,  $p = 0.990$ ) nor those of the local interactions (Figure 4B middle, anterior cluster:  $r = -0.20$ ,  $p = 0.338$ ; posterior cluster:  $r = 0.20$ ,  $p = 0.348$ ) was correlated with the perceptual sensitivity of the mean size. The first peak of the global interaction in the occipital area was significantly and positively correlated with the mean size perception precision (Figure 4B, right;  $r = 0.44$ ,  $p = 0.031$ ); the correlation was not significant for electrodes in the frontal area (Figure 4B, right;  $r = -0.35$ ,  $p = 0.095$ ). The neural responses of the two clusters at the second (Figure 4C) and third (Figure 4D) peaks were not correlated with the behavioral performance ( $ps > 0.56$ ). Therefore, the sensitivity in ensemble size perception is only reflected in the early



**Figure 4. Group-level behavioral performance and the behavioral correlates of the rERPs of the individuals, the local interactions, and the global interaction**

(A–D) (A) Behavioral performance in Exp. 1. The ratio of the probe radius to the mean radius of the circle array (X axis) is plotted against the proportion of “probe is larger” responses (Y axis). A cumulative Gaussian function was fitted to the curve for each participant (gray line) and the group average (black line). The correlation coefficients between the sigma value of the fitted cumulative Gaussian function and the rERPs at the first (B), second (C), and third (D) peaks of the individuals, the local interactions, and the global interaction. The rERPs for the posterior and anterior clusters are tested separately. \* indicates correlations with  $p < 0.05$ .

neural response to the global interaction. This finding further suggests that the nonadditive integration of individual items in the stimulus array supported the ensemble size perception.

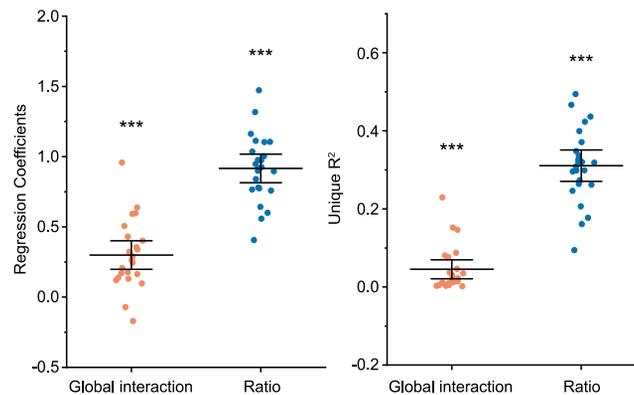
### Global interaction between items influences ensemble perception

Based on the previous results, we found that only the neural responses to the global interaction were correlated with ensemble perception. It can be inferred that if the cognitive processing of the global interaction leads to ensemble perception, then the global interaction of the stimulus can also be used to predict ensemble perception. To test this hypothesis, we quantified the influence of the global interaction on the ensemble perception within participants and across trials. The influences were calculated using a multivariate logistic regression of the behavioral choice (i.e., whether the probe size was larger than the mean size) on a linear combination of the global interaction and the ratio of the probe size to the mean size (i.e., the probe-to-mean size ratio). To reveal the unique contribution of each predictor, we then compared the adjusted  $R^2$  between this full model with two reduced models using either the global interaction or the probe-to-mean size ratio as the predictor.

As shown in Figure 5, the regression coefficients of the global interaction and the probe-to-mean size ratio were both significantly above zero (the global interaction:  $t(23) = 6.06$ ,  $p < 0.001$ , Cohen’s  $d = 1.24$ ; the probe-to-mean ratio:  $t(23) = 18.70$ ,  $p < 0.001$ , Cohen’s  $d = 3.82$ ), indicating substantial influences of the global interaction in addition to the probe-to-mean size ratio on the ensemble perception across trials. The unique  $R^2$  of the global interaction and the probe-to-mean ratio were both positively significant (the global interaction:  $t(23) = 3.83$ ,  $p < 0.001$ , Cohen’s  $d = 0.78$ ; the probe-to-mean ratio:  $t(23) = 15.96$ ,  $p < 0.001$ , Cohen’s  $d = 3.26$ ), indicating independent contributions of the global interaction and the probe-to-mean size ratio to the ensemble perception. These results strengthened the previous results that the global interaction between items is represented in the brain and mediates ensemble size perception.

### Spread attention enhances early neural responses to the global interaction

The unique contribution and the behavioral correlates of the global interaction component indicate that the global nonadditive integration in the brain supported the ensemble perception. To further explore this hypothesis, Experiment 2 ( $n = 20$ ) manipulated the attentional distribution over items in the ensemble. Studies have widely demonstrated that ensemble perception is enhanced when attention spreads over the global pattern compared to when attention is focused on a single item.<sup>33–35</sup> Based on the nonadditive integration hypothesis, if the global



**Figure 5. Regression coefficients (left) and the unique  $R^2$  (right) of the global interaction and the probe-to-mean size ratio**

Both predictors predicted the probability of the “probe being larger” responses significantly (left) and had unique contributions to the ensemble perception (right). The error bars indicate the 95% confidence interval. \*\*\* indicate  $p < 0.001$ .

interaction component represents the ensemble perception, we expect that the spread of attention over the global pattern would enhance the neural representation of the global interaction component. In contrast, the neural representation for the individual items and their local interactions would not be influenced.

To manipulate the attentional distribution, participants were instructed to either spread attention over the entire circle array (the attend-to-ensemble condition) or to concentrate attention on a single circle in the array (the attend-to-item condition). They performed a probe vs. mean size comparison task or a probe vs. single size comparison task accordingly (Figure 6A). We estimated the GFPs for the individual items (Figure 6B), the local interactions (Figure 6C), and the global interaction (Figure 6D) separately for the two conditions. Regardless of whether attention was spread over the global pattern or concentrated on a single item, significant GFPs were observed in the early and late phases after the circle array onset for all three predictors. However, the influence of attention distribution was found mainly in global interaction. Specifically, attending to the ensemble induced a larger GFP for the global interaction component than focusing attention on a single circle at 0.070–0.094 and 0.338–0.358 s after the circle array onset. In the global attention condition, it was also found that the individual item component had a larger GFP at a later time of 0.418–0.448 s. This could be due to a feedback effect from the global pattern to the individual items.

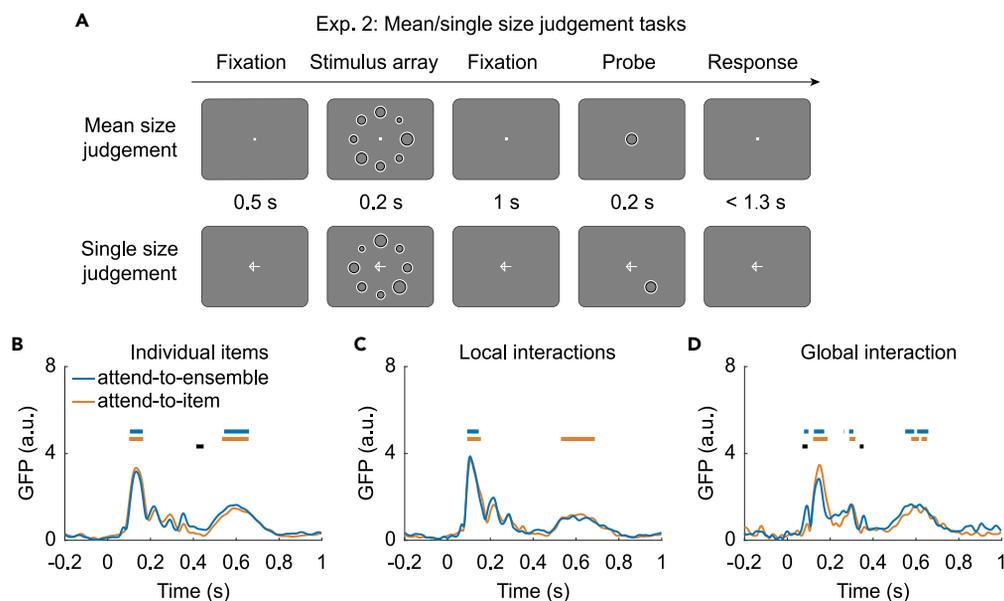
### The early neural response to the global interaction specifically supports ensemble perception

To further explore the functional significance of the global interaction, as in Experiment 1, we computed the Pearson correlation between the Peak 1 amplitude of the rERP to the global interaction and the ensemble size perception sensitivity and the single size perception sensitivity in the two attention conditions, respectively. As shown in Figure 7, the Peak 1 amplitude significantly correlated with the mean size sensitivity ( $r = 0.49$ ,  $p = 0.029$ ) but not with the single size sensitivity ( $r = 0.06$ ,  $p = 0.814$ ). Therefore, the neural response to the global interaction is specifically correlated with ensemble perception. These results replicated the findings of Experiment 1 and confirmed that global nonadditive integration supports ensemble representation.

## DISCUSSION

Summarizing visual information into an ensemble representation requires the visual system to integrate information from the individual items. Here we hypothesized that rather than linearly averaging individual items’ properties (e.g., sizes), the ensemble processing may be realized through nonadditive integration of individual items. We performed a regression-based ERP estimation, which revealed significant and unique neural representations of the individual items and their local and global interactions. Critically, only the neural representation of the global interaction, the highest-order interaction over all circles,<sup>22,25</sup> predicted the behavioral precision in ensemble perception within and across participants. Furthermore, manipulating attentional distribution modulated neural changes mainly in the global interaction, at both early and late phases. Finally, the behavioral correlates of the global interaction are specific to the ensemble perception but not the item perception. These findings concomitantly reveal a global nonadditive integration of visual information during ensemble perception and suggest a global mechanism of ensemble perception in the brain.

Whether a global mechanism exists for ensemble perception has triggered a long-standing debate.<sup>4,12,36</sup> Yet research on this issue is limited. One reason is the lack of techniques to disentangle the neural processing of the ensemble from the processing of individual items, as the ensemble and individuals are spatially inseparable. The present study employed a regression-based ERP estimation to extract the independent neural responses to three levels of information: the individuals, the local interactions, and the global interaction. This method revealed a rapid neural response to the global interaction and confirmed its unique contribution to EEG signals after controlling for the contribution of the other two components. Furthermore, the global interaction was specifically correlated with the precision of the ensemble perception across two experiments, implying its functional role in ensemble representation. Therefore, the nonadditive integration of all individual items may serve as a global mechanism in ensemble perception.



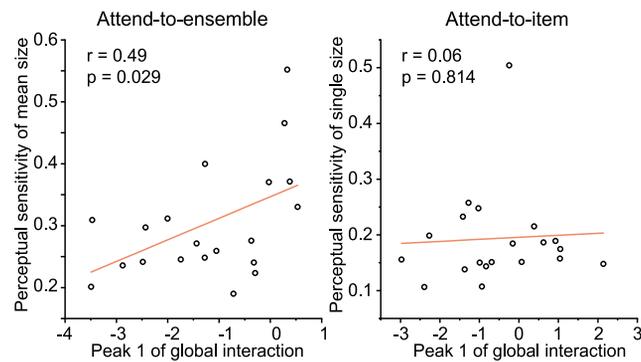
**Figure 6. Paradigm and GFPs of Experiment 2**

(A–D) (A) Procedures of the attend-to-ensemble (top) and attend-to-item (bottom) conditions in Experiment 2. During the attend-to-ensemble trials, participants judged whether the size of the probe circle was larger than the average size of the circle array or not. During the attend-to-item trials, participants judged whether the size of the probe circle was larger than the size of the cued single circle in the circle array. In the attend-to-item task, the probe was randomly placed in one of the eight locations of the stimulus array. The GFPs of were calculated for the individuals (B), the local interactions (C), and the global interaction (D) in the attend-to-ensemble and attend-to-item conditions. The blue horizontal bar illustrates the significant time points for the attend-to-ensemble condition; the orange horizontal bar illustrates the significant time points for the attend-to-item condition; the black horizontal bar illustrates significant differences between the two attention distribution conditions. The a.u. indicates the arbitrary unit.

So far, studies have focused on the computational principle underlying linear, additive integration of individuals for ensemble representation. Recent studies have suggested that the mean perception of an ensemble is not a simple averaging of individuals. Instead, individual items are weighted by their relative physical saliency or task relevance before linear averaging.<sup>17–20,37</sup> A critical logic behind the weighted averaging account is that the individual items are interdependent. Namely, assigning proper weights over any individual item is unlikely without a global representation of all individuals. This interdependency across individuals was further revealed in our finding with the nonadditive integration of individual items. Because the multiplication of all items represents the global interaction component, its contribution cannot be decomposed onto any individual item. In this regard, the interdependency term should be examined in addition to the individual items in ensemble representation. Notably, because the weighted averaging only considers linear relations between individuals, the previous finding that weighted averaging contributes to ensemble perception does not exclude the contribution of the global interaction in the present results. Future studies could explore the contribution of global interaction over individuals after considering their weights.

The abstract global integration of the individual sizes may carry functional significance relative to the linear averaging of the individual sizes. Although the mean and product are closely related, the global integration mechanism can be more efficient than the linear averaging mechanism in reflecting the contribution of the individual items. For instance, when two items are averaged, the small item might modulate the mean much less than the large item, making the smaller item difficult to decode. In contrast, when the two items are multiplied, the small item can modulate the product as effectively as the large item, allowing both items to be easily decoded.<sup>32</sup> We suggest that the multiplicative nature of the global interaction supports the ensemble perception to represent the accurate shape of the distribution of items.<sup>38</sup>

Our results that ensemble processing takes into account the interactions between individual items are consistent with the normalization theory.<sup>39</sup> When multiple stimuli are spatially presented, the canonical neural computation, such as normalization, is likely to be involved. Given that normalization is responsible for neural interactions among cells, the normalization is intrinsically nonlinear.<sup>39</sup> Serving a role in nonlinear pooling, normalization can also be implemented in high-level visual processing which spans a large visual field as in the ensemble processing in the present study. Empirical evidence has been found for suppression in ventral stream area V4<sup>40</sup> and the inferotemporal cortex<sup>41</sup> when a non-preferred stimulus (or object) is presented with a preferred stimulus, supporting that normalization can be implemented in large visual field integration. Furthermore, the idea of pooling in large receptive fields is also common in broader theories of visual perception. For example, ensemble statistics have been viewed as a form of gist representation,<sup>42</sup> the rough impression of the whole scene without knowing the details. A prominent “reverse hierarchy theory”<sup>43</sup> suggests that the gist percept arises at the top of the fast feedforward stream, where neurons with large receptive fields respond to large portions of the visual field. In contrast, detailed vision requires slow feedback propagation to lower-level neurons with small receptive fields.



**Figure 7. Behavioral correlates of the global interaction in the two attention distribution conditions**

The Pearson correlation between the Peak 1 amplitude of the rERP to the global interaction and the perceptual sensitivity (i.e., the fitted sigma value) was significant in the attend-to-ensemble condition (left) but non-significant in the attend-to-item condition (right).

Although it has not been demonstrated in ensemble perception, the nonadditive integration of local information is likely to be a basic computational principle of global information processing in the brain.<sup>21,44,45</sup> For instance, nonadditive integration of local information has been observed in Gestalt psychology<sup>15,21</sup> and multisensory integration.<sup>22,23</sup> Such global information processing, including ensemble representation, can be built by pooling together the activity of local low-level feature detectors across the visual field. A global receptive field can then be generated as a combination of local receptive fields to build a holistic representation of a visual scene in high-level visual processing.<sup>46–48</sup> In line with this classical feedforward model, the representation latency for the global integration is later than that for the individual items.<sup>48</sup> A nonadditively integrated ensemble representation, which contains information about individual items and the relations between individuals, is more consistent with the context-dependent properties of visual processing.<sup>45</sup>

Notably, the global interaction term, computed as the multiplication of eight individual sizes, is mathematically related to the geometric mean, which is the linear averaging of the logarithmic compressed individual sizes. Treating the global interaction as the geometric mean makes the individual sizes independent in the ensemble mean size processing. While this is an interesting possibility, there has been no evidence for logarithmic compression in size perception or ensemble processing in the brain. Furthermore, if the visual system follows a linear averaging of individual sizes, the individuals do not interact with each other for ensemble processing. Then the ensemble perception process would not be influenced by the variance or saliency of the individual ensemble stimuli. This conflicts with the previous findings on the influence of variance and saliency on the ensemble perception.<sup>14,49</sup> Nevertheless, both the nonadditive integration and the geometric mean explanations are worth exploring in future studies. Furthermore, this study only explores how individuals are integrated in the ensemble computation, and it cannot answer the question of the neural representation of ensemble size in the brain.<sup>50</sup> The relationship between ensemble representation and nonadditive integration of individuals is another promising direction for future research.

In our results, relative to attending to a single item, attending to the global pattern increased the neural response to the global interaction in the early and late phases, yet global attention increased the neural response to the individual items only at a later phase. The early attentional enhancement of global interaction indicates that the rapid global integration is not a simple fast sweep in the early visual area, which is supposed not to be modulated by attention.<sup>51</sup> Instead, the global integration is a feedback processing,<sup>48,51</sup> and the attention modulation is transmitted to the individual items in a late object recognition stage that scrutinizes individual items.<sup>47,48</sup> These results suggest that the perception of the ensemble and its local information follows a reverse hierarchy.<sup>43,48</sup>

In summary, the present study shows that the visual system responds rapidly to the global interaction between individual items during ensemble perception. The global interaction, which specifically predicts the precision of ensemble perception within and across participants, reveals a global nonadditive integration of individual items supporting ensemble representation. Ensemble perception may be realized through a global mechanism with nonadditive integration of individual items.

### Limitations of the study

Although our study provides strong evidence that ensemble processing is nonadditive rather than linear for the integration of individual items, and that spreading attention can enhance the nonadditive integration of individual items, there are still some unanswered questions in this study that deserve further investigation. First, although interactions between individual items are involved in ensemble processing, the specific neurocomputational mechanisms of ensemble processing remain unclear. Future computational modeling studies could investigate how interactions between individual items contribute to ensemble processing. Second, the EEG technique used in this study does not provide information about the brain regions where individual items are integrated. Future fMRI- and MEG-based studies could investigate in which brain regions individual items are integrated linearly and nonlinearly. Third, it is unclear how ensemble processing and individual item processing interact with each other. Future studies investigating these questions in more detail may shed more light on the cognitive neural mechanisms of ensemble processing.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- [KEY RESOURCES TABLE](#)
- [RESOURCE AVAILABILITY](#)
  - Lead contact
  - Materials availability
  - Data and code availability
- [EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS](#)
  - Participants
- [METHOD DETAILS](#)
  - Stimuli and apparatus
  - General procedure and design
  - Behavioral measurement of the mean size judgment
  - EEG acquisition and preprocessing
- [QUANTIFICATION AND STATISTICAL ANALYSIS](#)
  - rERP estimation
  - Unique explained variance of each predictor
  - Influence of global interaction on ensemble perception
  - Cluster-based permutation testing
  - Global field power (GFP)
  - GFP testing

## ACKNOWLEDGMENTS

Funding: This work was supported by the Natural Science Foundation of Zhejiang Province (Grant LY23C090001 to J.J.), National Natural Science Foundation of China (Grant 32000735 to J.J.), and Major Project of Philosophy and Social Science Research of the Ministry of Education of China (22JZD044).

## AUTHOR CONTRIBUTIONS

Conceptualization: J.J.

Methodology: J.J., T.W.

Investigation: T.W., Y.Z., J.J.

Visualization: T.W., J.J.

Supervision: J.J.

Writing—original draft: T.W., J.J.

Writing—review & editing: J.J., T.W.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: November 2, 2022

Revised: September 3, 2023

Accepted: September 16, 2023

Published: September 19, 2023

REFERENCES

- Palmer, J. (1990). Attentional Limits on the Perception and Memory of Visual Information. *J. Exp. Psychol. Hum. Percept. Perform.* 16, 332–350.
- Luck, S.J., and Vogel, E.K. (1997). The capacity of visual working memory for features and conjunctions. *Nature* 390, 279–281.
- 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>.
- Arieli, D. (2001). Seeing Sets: Representation by Statistical Properties. *Psychol. Sci.* 12, 157–162. <https://doi.org/10.1111/1467-9280.00327>.
- Chong, S.C., and Treisman, A. (2003). Representation of statistical properties. *Vis. Res.* 43, 393–404.
- Whitney, D., and Yamanashi Leib, A. (2018). Ensemble Perception. *Annu. Rev. Psychol.* 69, 105–129. <https://doi.org/10.1146/annurev-psych-010416-044232>.
- Epstein, M.L., and 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](https://doi.org/10.1162/jocn_a_01704).
- Jia, J., Wang, T., Chen, S., Ding, N., and Fang, F. (2022). Ensemble size perception: Its neural signature and the role of global interaction over individual items. *Neuropsychologia* 173, 108290. <https://doi.org/10.1016/j.neuropsychologia.2022.108290>.
- Allik, J., Toom, M., Raidvee, A., Averin, K., and 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>.
- Lee, J., and Chong, S.C. (2021). Quality of average representation can be enhanced by refined individual items. *Atten. Percept. Psychophys.* 83, 970–981. <https://doi.org/10.3758/s13414-020-02139-3>.
- Parke, L., Lund, J., Angelucci, A., Solomon, J.A., and Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nat. Neurosci.* 4, 739–744. <https://doi.org/10.1038/89532>.
- Myczek, K., and 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>.
- Simons, D.J., and 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>.
- Semizer, Y., and 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>.
- Corbett, J.E. (2017). The Whole Warps the Sum of Its Parts: Gestalt-Defined-Group Mean Size Biases Memory for Individual Objects. *Psychol. Sci.* 28, 12–22. <https://doi.org/10.1177/0956797616671524>.
- Kanaya, S., Hayashi, M.J., and Whitney, D. (2018). Exaggerated groups: amplification in ensemble coding of temporal and spatial features. *Proc. Biol. Sci.* 285, 20172770. <https://doi.org/10.1098/rspb.2017.2770>.
- de Gardelle, V., and Summerfield, C. (2011). Robust averaging during perceptual judgment. *Proc. Natl. Acad. Sci. USA* 108, 13341–13346. <https://doi.org/10.1073/pnas.1104517108>.
- Li, V., Herce Castañón, S., Solomon, J.A., Vandormael, H., and Summerfield, C. (2017). Robust averaging protects decisions from noise in neural computations. *PLoS Comput. Biol.* 13, e1005723. <https://doi.org/10.1371/journal.pcbi.1005723>.
- Choi, Y.M., and Chong, S.C. (2020). Effects of Selective Attention on Mean-Size Computation: Weighted Averaging and Perceptual Enlargement. *Psychol. Sci.* 31, 1261–1271. <https://doi.org/10.1177/0956797620943834>.
- Jeong, J., and 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>.
- Wagemans, J., Elder, J.H., Kubovy, M., Palmer, S.E., Peterson, M.A., Singh, M., and 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>.
- Kim, S.S., Gomez-Ramirez, M., Thakur, P.H., and 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>.
- Werner, S., and 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>.
- Lalor, E.C., Pearlmutter, B.A., Reilly, R.B., McDarby, G., and 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>.
- Smith, N.J., and 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>.
- Smith, N.J., and Kutas, M. (2015). Regression-based estimation of ERP waveforms: I. The rERP framework. *Psychophysiology* 52, 157–168. <https://doi.org/10.1111/psyp.12317>.
- Raidvee, A., Toom, M., Averin, K., and Allik, J. (2020). Perception of means, sums, and areas. *Atten. Percept. Psychophys.* 82, 865–876. <https://doi.org/10.3758/s13414-019-01938-7>.
- Yousif, S.R., and Keil, F.C. (2021). How We See Area and Why It Matters. *Trends Cognit. Sci.* 25, 554–557. <https://doi.org/10.1016/j.tics.2021.03.017>.
- Skrandies, W. (1990). Global field power and topographic similarity. *Brain Topogr.* 3, 137–141. <https://doi.org/10.1007/BF01128870>.
- Maris, E., and Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *J. Neurosci. Methods* 164, 177–190. <https://doi.org/10.1016/j.jneumeth.2007.03.024>.
- Groen, I.I.A., Ghebreab, S., Prins, H., Lamme, V.A.F., and Scholte, H.S. (2013). From Image Statistics to Scene Gist: Evoked Neural Activity Reveals Transition from Low-Level Natural Image Structure to Scene Category. *J. Neurosci.* 33, 18814–18824. <https://doi.org/10.1523/JNEUROSCI.3128-13.2013>.
- Ratan Murty, N.A., and Arun, S.P. (2018). Multiplicative mixing of object identity and image attributes in single inferior temporal neurons. *Proc. Natl. Acad. Sci. USA* 115, E3276–E3285. <https://doi.org/10.1073/pnas.1714287115>.
- Baek, J., and Chong, S.C. (2020). 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>.
- Chong, S.C., and Treisman, A. (2005). Attentional spread in the statistical processing of visual displays. *Percept. Psychophys.* 67, 1–13. <https://doi.org/10.3758/BF03195009>.
- de Fockert, J.W., and 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>.
- Zhao, Y., Zeng, T., Wang, T., Fang, F., Pan, Y., and Jia, J. (2023). Subcortical encoding of summary statistics in humans. *Cognition* 234, 105384. <https://doi.org/10.1016/j.cognition.2023.105384>.
- Utchkin, I.S., Choi, J., and Chong, S.C. (2022). A Population Response Model of Ensemble Coding. *Neuroscience*. <https://doi.org/10.1101/2022.01.19.476871>.
- Chetverikov, A., Campana, G., and Kristjánsson, Á. (2017). Representing Color Ensembles. *Psychol. Sci.* 28, 1510–1517. <https://doi.org/10.1177/0956797617713787>.
- Carandini, M., and Heeger, D.J. (2011). Normalization as a canonical neural computation. *Nat. Rev. Neurosci.* 13, 51–62. <https://doi.org/10.1038/nrn3136>.
- Reynolds, J.H., and Desimone, R. (2003). Interacting Roles of Attention and Visual Saliency in V4. *Neuron* 37, 853–863. [https://doi.org/10.1016/S0896-6273\(03\)00097-7](https://doi.org/10.1016/S0896-6273(03)00097-7).
- Zoccolan, D., Cox, D.D., and DiCarlo, J.J. (2005). Multiple Object Response Normalization in Monkey Inferotemporal Cortex. *J. Neurosci.* 25, 8150–8164. <https://doi.org/10.1523/JNEUROSCI.2058-05.2005>.
- Cohen, M.A., Dennett, D.C., and Kanwisher, N. (2016). What is the Bandwidth of Perceptual Experience? *Trends Cognit. Sci.* 20, 324–335. <https://doi.org/10.1016/j.tics.2016.03.006>.
- Hochstein, S., and Ahissar, M. (2002). View from the top: Hierarchies and reverse hierarchies in the visual system. *Neuron* 36, 791–804.
- Zhang, P., Jamison, K., Engel, S., He, B., and He, S. (2011). Binocular Rivalry Requires Visual Attention. *Neuron* 71, 362–369. <https://doi.org/10.1016/j.neuron.2011.05.035>.
- Albright, T.D., and Stoner, G.R. (2002). Contextual Influences on Visual Processing. *Annu. Rev. Neurosci.* 25, 339–379. <https://doi.org/10.1146/annurev.neuro.25.112701.142900>.
- Oliva, A., and Torralba, A. (2006). Chapter 2 Building the gist of a scene: the role of global image features in recognition. In *Progress in Brain Research* (Elsevier), pp. 23–36. [https://doi.org/10.1016/S0079-6123\(06\)55002-2](https://doi.org/10.1016/S0079-6123(06)55002-2).
- Oliva, A., and Torralba, A. (2001). Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope.

- Forensic Sci. Int. 120, 145–154. <https://doi.org/10.1023/A:1011139631724>.
48. Hochstein, S., Pavlovskaya, M., Bonneh, Y.S., and Soroker, N. (2015). Global statistics are not neglected. *J. Vis.* 15, 7. <https://doi.org/10.1167/15.4.7>.
  49. Iakovlev, A.U., and 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>.
  50. Oh, B.-I., Kim, Y.-J., and Kang, M.-S. (2019). Ensemble representations reveal distinct neural coding of visual working memory. *Nat. Commun.* 10, 5665. <https://doi.org/10.1038/s41467-019-13592-6>.
  51. Lamme, V.A., and 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](https://doi.org/10.1016/S0166-2236(00)01657-X).
  52. Brainard, D.H. (1997). The Psychophysics Toolbox. *Spatial Vis.* 10, 433–436.
  53. Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
  54. Ehinger, B.V., and Dimigen, O. (2019). Unfold: an integrated toolbox for overlap correction, non-linear modeling, and regression-based EEG analysis. *PeerJ* 7, e7838. <https://doi.org/10.7717/peerj.7838>.
  55. Baek, J., and Chong, S.C. (2020). Distributed attention model of perceptual averaging. *Atten. Percept. Psychophys.* 82, 63–79. <https://doi.org/10.3758/s13414-019-01827-z>.
  56. H. Best, and C. Wolf, eds. (2015). *The SAGE handbook of regression analysis and causal inference* (SAGE Reference).
  57. Jia, J., Fang, F., and 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>.
  58. Jia, J., Liu, L., Fang, F., and Luo, H. (2017). Sequential sampling of visual objects during sustained attention. *PLoS Biol.* 15, e2001903. <https://doi.org/10.1371/journal.pbio.2001903>.
  59. Frost, J. (2019). *Regression Analysis: An Intuitive Guide for Using and Interpreting Linear Models* (James D. Frost).
  60. C.M. Michel, T. Koenig, D. Brandeis, L.R.R. Gianotti, and J. Wackermann, eds. (2009). *Electrical Neuroimaging* (Cambridge University Press). <https://doi.org/10.1017/CBO9780511596889>.
  61. Benjamini, Y., and Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. Roy. Stat. Soc. B* 57, 289–300.

## STAR★METHODS

### KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Experimental data and analysis code (Matlab)	This paper	<a href="https://doi.org/10.17632/vtsx2694p7.1">https://doi.org/10.17632/vtsx2694p7.1</a>
Software and algorithms		
MATLAB	MathWorks	<a href="https://www.mathworks.com/">https://www.mathworks.com/</a>
Psychtoolbox-3	Brainard, D.H., 1997 <sup>52</sup>	<a href="http://psychtoolbox.org/">http://psychtoolbox.org/</a>
EEGLAB	Delorme, A. et al., 2004 <sup>53</sup>	<a href="https://scn.ucsd.edu/eeglab/index.php">https://scn.ucsd.edu/eeglab/index.php</a>
Unfold Toolbox	Ehinger BV et al., 2019 <sup>54</sup>	<a href="https://www.unfoldtoolbox.org/">https://www.unfoldtoolbox.org/</a>

### RESOURCE AVAILABILITY

#### Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Jianrong Jia ([jianrongjia@hznu.edu.cn](mailto:jianrongjia@hznu.edu.cn)).

#### Materials availability

This study did not generate new unique reagents or other materials.

#### Data and code availability

- The data has been deposited at Mendeley data repository and is publicly available as of the date of publication. DOI is listed in the [key resources table](#).
- The code has been deposited at Mendeley data repository and is publicly available as of the date of publication. DOI is listed in the [key resources table](#).
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

### Participants

Twenty-four (15 females, mean age 20.54 years, range 19-24 years, Mongoloid ethnicity) and twenty (18 females, mean age 19.10 years, range 18-22 years, Mongoloid ethnicity) healthy participants with normal or corrected-to-normal vision took part in Experiment 1 and Experiment 2, respectively. All participants were naive to the purpose of the experiments. They provided written informed consent and received monetary compensation for participation. The experimental procedure was approved by the Departmental Ethical Committee of Hangzhou Normal University.

## METHOD DETAILS

### Stimuli and apparatus

The stimuli were generated and controlled with MATLAB (the MathWorks) and Psychophysics Toolbox (Psychtoolbox-3).<sup>52</sup> The stimuli were presented on a 21-inch CRT screen with a resolution of 1024×768 pixels and a refresh rate of 100 Hz. The participants were comfortably seated at 70 cm from the screen with their heads stabilized on a chin rest. Both experiments were conducted in a dimly lit room.

### General procedure and design

#### Experiment 1

Each trial started with the presentation of a white fixation point (84.53 cd/m<sup>2</sup>, diameter is 0.16 ° visual angle) at the center of a gray screen (42.80 cd/m<sup>2</sup>) (see [Figure 1](#)). After 800 ms, a stimulus array consisting of eight circles was presented for 200 ms. At 1000 ms after the offset of the stimulus array, a probe circle was presented at the center of the screen for 200 ms. Participants judged whether the probe's size was larger or smaller than the mean size of the circles in the stimulus array by pressing one of two keys on the keyboard. The response should be made within 1800 ms after the probe offset. Participants were instructed to fixate at the fixation point during the whole trial.

The eight circles in the stimulus array were presented in eight fixed locations evenly distributed on an imaginary circle centered on the fixation point with a 279-pixel (9° in visual angle) radius. Across trials, the average radius of the stimulus array varied at 10 levels, which were 35, 40, 45, 50, 55, 60, 65, 70, 75, and 80 pixels. We adopted a procedure to determine each circle's size (i.e., radius) while keeping the circle's mean size at a given level. First, at a given level of the mean size, four reference radii were randomly drawn from a uniform distribution between 0.2 and 0.8 times the average radius. Then, each of the four reference radii was added to and subtracted from the average radius, producing eight circle radii. Finally, the eight radii were randomly assigned to circles at the eight fixed locations. The radius of the probe circle was 0.7, 0.8, 0.9, 1, 1.1, 1.2, or 1.3 times the average radius of the stimulus array in different trials. The circle edge was composed of an outer white line (luminance: 84.53 cd/m<sup>2</sup>) and an inner black line (luminance: 0 cd/m<sup>2</sup>), which ensured the mean luminance of each circle was roughly equal to the background luminance. Each participant completed 980 trials in about 2 hrs. A forced break was inserted after every 140 trials.

### Experiment 2

Experiment 2 had two blocks, one for the attend-to-ensemble condition and one for the attend-to-item condition. The attend-to-ensemble block used the same stimuli and procedure as Experiment 1, except that the behavioral response should be made within 1300 ms after the probe offset. The subsequent trial started after a 500-ms interval following the response (upper row, Figure 6A). The attend-to-item block used a similar procedure. Still, the participants made a judgment on the relative size of the probe to the size of an individual circle (the left circle) in the stimulus array. The probe circle was presented on a random location of the stimulus array (lower row, Figure 6A). To match the task difficulty, the radius of the probe circle was 0.82, 0.88, 0.94, 1, 1.06, 1.12, or 1.18 times that of the cued circle. The order of the two blocks was counterbalanced across participants. Each block contained 980 trials. A forced break was inserted after every 140 trials.

### Behavioral measurement of the mean size judgment

The ratio of the probe circle's size to the circle array's mean size constituted seven steps in both experiments. First, we calculated the proportion of "larger" responses in each step across different sizes. Next, we fitted a cumulative Gaussian function to the proportion of "larger" responses against the actual ratio of the probe size to the mean size for each participant. The fitted function's standard deviation (i.e., the sigma value) was calculated to represent the precision of mean size judgement,<sup>55</sup> with a lower sigma value indicating higher sensitivity and a larger sigma value indicating lower sensitivity in mean size perception.

### EEG acquisition and preprocessing

EEG data were acquired at a sampling rate of 500 Hz using a 64-channel Easycap and two BrainAmp amplifiers (Brain Products). The horizontal electrooculogram (EOG) was recorded by an additional electrode placed near the outer canthus of the left eye; an additional electrode below the right eye recorded the vertical EOG. FCz was used as the reference electrode during recording. The impedance of all electrodes were kept below 5 kΩ. Offline preprocessing was carried out using EEGLAB.<sup>53</sup> The EEG data were re-referenced to the average value across all electrodes (except the HEOG and VEOG) and band-pass filtered between 1 and 30 Hz. Independent component analysis (ICA) was performed to remove eye-movement and other artifactual components. On average, 8 ICA components were removed for each subject. No further trial rejection was executed. Data were epoched from -200 ms to 1000 ms relative to the onset of the circle array.

## QUANTIFICATION AND STATISTICAL ANALYSIS

### rERP estimation

A linear regression model was constructed to predict the EEG signals during ensemble size perception. Because our goal was to examine whether the interactions between individuals would be represented in the brain, in addition to the individuals, we have only included the simplest 2-item local interactions and the highest-order 8-item global interaction in the model.<sup>23,26,56</sup> Specifically, the individual circle sizes ( $I_{1-8}$ :  $I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8$ ), the local interactions between adjacent circles ( $L_{1-8}$ :  $I_1 \times I_2, I_2 \times I_3, I_3 \times I_4, I_4 \times I_5, I_5 \times I_6, I_6 \times I_7, I_7 \times I_8, I_8 \times I_1$ ), and the global interaction across all eight circles ( $G$ :  $I_1 \times I_2 \times I_3 \times I_4 \times I_5 \times I_6 \times I_7 \times I_8$ )<sup>23,26</sup> were included in the regression model. Because the variance of EEG data is limited, not exhausting all potential interaction terms, such as 3- or 4-item interactions, also ensured a high signal-to-noise ratio in fitting the regression model. Previous studies demonstrated that the visual system relies on the radius to determine the size of a circle rather than the mathematical area<sup>27,28</sup>; therefore, we use the radius as the measure of size. The model was shown as follows:

$$EEG = \left( \sum_{i=1}^8 \beta_{I_i} \times I_i \right) + \left( \sum_{i=1}^8 \beta_{L_i} \times L_i \right) + \beta_G \times G$$

where  $I_i$  and  $\beta_{I_i}$  are the radius of the  $i$ -th circle and its neural response weight,  $L_i$  and  $\beta_{L_i}$  are the  $i$ -th local interaction and its neural response weight, and  $G$  and  $\beta_G$  are the global interaction and its neural response weight.

Ridge regression was used to estimate the neural response weights for each predictor.<sup>57,58</sup> The Unfold toolbox<sup>54</sup> was used for model fitting. A parameter  $\lambda$ , set through cross-validation in each participant separately, was included to control for model overfitting. Values in each predictor were converted to z scores before model fitting to reduce structural multicollinearity.<sup>59</sup> The neural response weight was estimated at each electrode and time point separately.

### Unique explained variance of each predictor

We adopted a stepwise regression logic to investigate the unique contribution of each level of information (i.e., the individual items, the local interactions, and the global interaction) to the EEG signals. First, the model above was used as a full model, in which the individual sizes, the local interactions, and the global interaction were simultaneously included as predictors for the EEG signals (the *ILG* model). Then, we defined three reduced models that used the individual sizes and the local interactions (the *IL* model), the individual sizes and the global interaction (the *IG* model), and the local interactions and the global interaction (the *LG* model) as predictors, respectively. The unique contribution of each predictor can be quantified as the model fit (explained variance or  $R^2$ ) difference between the full model and the reduced model without that specific predictor. Specifically, the individual items, the local interactions, the global interaction can be quantified as the explained variance difference between the *ILG* and *LG* models, the *ILG* and *IG* models, the *ILG* and *IL* models, respectively. Six electrodes (P5, P7, PO7, P6, P8, and PO8) on the bilateral occipital areas were selected for statistical analysis.

### Influence of global interaction on ensemble perception

The influences of the global interaction between the individual items and the ratio of the probe size to the mean size on the ensemble perception were calculated via a multivariate logistic regression of behavioral choice (i.e., whether the probe size was larger than the mean size) across trials and within participants. The regression model was a linear combination of the global interaction and the probe-to-mean size ratio.

$$P(\text{large}) = \Phi[\beta_G \times G + \beta_R \times R]$$

where  $P(\text{large})$  is the probability of judging the probe size being larger than the mean size of the circle array;  $\Phi[\cdot]$  is the cumulative normal density function;  $G$  and  $\beta_G$  are the global interaction and its regression coefficient, respectively;  $R$  and  $\beta_R$  are the probe-to-mean size ratio and its regression coefficient, respectively.

The adjusted  $R^2$  between the multivariate logistic regression model above was further compared with two univariate models using either the global size interaction or the probe-to-mean size ratio as the predictor to reveal the unique contribution of each predictor.

### Cluster-based permutation testing

To determine whether the unique explained variances were higher than the baseline we used a nonparametric sign-permutation test at each time point to test the significance.<sup>30</sup> The unique explained variances were first baseline-corrected, then their signs were randomly flipped 1000 times with a probability of 50%, generating a null distribution. The percentage of the unique explained variance obtained from the null distribution being equal to or larger than the observed unique explained variance was calculated and used as the significance level. This procedure was repeated for each time point. A cluster-based permutation test was then used to correct for multiple comparisons over time using a cluster significance threshold of  $p < 0.05$ . The mean accuracy before the stimulus onset (-200 to 0 ms) was used as the baseline.

The cluster-based permutation test<sup>30</sup> was used to examine the topographic distribution of the beta values (Figure 2, right). A one-sample  $t$ -test was performed on each sensor's beta value against zero separately. Sensors showing a significant effect ( $p < 0.025$ ) were clustered based on spatial adjacency, with a minimum of two adjacent sensors to form a cluster. The sign of the beta value was randomly flipped 1000 times with a probability of 50%, generating a null distribution. The clusters were further thresholded at  $p < 0.01$ .

### Global field power (GFP)

We computed the global field power (GFP) of the neural response to each predictor: the GFP quantifies the amount of activity at each time point considering the data from all recording electrodes simultaneously, which results in a reference-independent descriptor of the potential field.<sup>29</sup> We used the GFP to describe the overall responses of the multichannel recordings. Based on the GFP results, we further analyzed the topographic distribution of responses and the rERPs on the electrodes of the largest responses.

$$GFP = \sqrt{\frac{\sum_{i=1}^n (u_i - \bar{u})^2}{n}}$$

where  $u_i$  is the rERP amplitude of the electrode  $i$ ,  $\bar{u}$  is the average amplitude of all electrodes, and  $n$  is the number of electrodes. Before GFP computation, the estimated rERPs were converted to  $z$  scores across all channels and times.

### GFP testing

We used a nonparametric permutation test to investigate the significance of the observed GFP values.<sup>60</sup> Because the grand-averaged GFP across participants at one moment depends on the GFP of the individual ERPs and the consistency of the topography over participants, a series of steps were taken. For a given time point: (1) the GFP of the grand mean of rERPs across participants was computed; (2) the measurements across channels were randomly shuffled for each participant; (3) the GFP of the grand mean over the randomly shuffled data were computed and retained as one instance of the GFP under the null hypothesis; (4) steps (2) to (3) were repeated for 5000 times; (5) the percentage of cases where the GFP obtained after randomization was equal to or larger than the GFP obtained in the observed data was calculated. This percentage was treated as the significance of the GFP at the given time point. Multiple comparisons across the time were corrected by FDR correction.<sup>61</sup>

We used a nonparametric permutation test to examine the GFP difference between the two attention conditions in Experiment 2. For a given time, the GFPs from the two conditions of all participants were mixed with shuffled labels. The GFP difference was then computed on the label-shuffled data. Repeating this procedure 5000 times resulted in a permutation distribution for the difference between the two conditions. The percentage of the GFP difference obtained from the label-shuffled data being equal to or larger than the GFP difference obtained from the observed data was calculated and used as the significance level of the GFP difference at the given time point. Multiple comparisons across the time were corrected by FDR correction.<sup>61</sup>