


The divisive normalization model of visual number sense: model predictions and experimental confirmation

Jenna Croteau¹, Michele Fornaciai², David E. Huber³, Joonkoo Park ^{1,4,*}

¹Department of Psychological and Brain Sciences, University of Massachusetts Amherst, 135 Hicks Way, Amherst, MA 01003, United States

²Institute for Research in Psychology (IPSY) and Institute of Neuroscience (IoNS), Université Catholique de Louvain, Place du Cardinal Mercier 10, Louvain-la-Neuve, 1348, Belgium

³Department of Psychology and Neuroscience, University of Colorado Boulder, Muenzinger D244, 345 UCB, Boulder, CO 80309, United States

⁴Commonwealth Honors College, University of Massachusetts Amherst, 157 Commonwealth Avenue, Amherst, MA 01003, United States

*Corresponding author: Joonkoo Park, Department of Psychological and Brain Sciences, University of Massachusetts Amherst, 135 Hicks Way, Amherst MA 01003, United States. Email: joonkoo@umass.edu

Our intuitive sense of number allows rapid estimation for the number of objects (numerosity) in a scene. How does the continuous nature of neural information processing create a discrete representation of number? A neurocomputational model with divisive normalization explains this process and existing data; however, a successful model should not only explain existing data but also generate novel predictions. Here, we experimentally test novel predictions of this model to evaluate its merit for explaining mechanisms of numerosity perception. We did so by consideration of the coherence illusion: the underestimation of number for arrays containing heterogeneous compared to homogeneous items. First, we established the existence of the coherence illusion for homogeneity manipulations of both area and orientation of items in an array. Second, despite the behavioral similarity, the divisive normalization model predicted that these two illusions should reflect activity in different stages of visual processing. Finally, visual evoked potentials from an electroencephalography experiment confirmed these predictions, showing that area and orientation coherence modulate brain responses at distinct latencies and topographies. These results demonstrate the utility of the divisive normalization model for explaining numerosity perception, according to which numerosity perception is a byproduct of canonical neurocomputations that exist throughout the visual pathway.

Key words: computational modeling; divisive normalization; numerosity perception; visual illusion.

Introduction

The ability to make a rapid estimation of the number of items in a scene, or numerosity, is a ubiquitous feature of both vertebrate and invertebrate animal behavior (Butterworth et al. 2018; Kobylkov et al. 2023). Like other perceptual abilities, this “number sense” conforms to Weber’s law (Anobile et al. 2016). However, unlike many other perceptual domains, the number sense is thought to underlie more analytical abilities such as addition, subtraction, and averaging (Feigenson et al. 2004; Dehaene 2011; Clarke and Beck 2021; Katzin et al. 2021; Togoli et al. 2021), and some authors suggest that it underlies symbolic numerical abilities (e.g. Halberda et al. 2008; Starr et al. 2013; but see Wilkey and Ansari 2019).

While the predominant view is that there exists an independent mechanism for the number sense (Burr and Ross 2008; Dehaene 2011), the precise computational mechanisms for how numerosity is represented within the visual system are not well understood. This lack of mechanistic explanations has caused some to challenge the traditional view. For instance, based on numerous findings that numerosity judgments can be influenced by other continuous magnitude dimensions (such as density and total area of the items) and vice versa (Gebuis and Reynvoet 2012a, 2012b; Leibovich and Henik 2014; DeWind et al. 2015; Anobile et al. 2016; Starr et al. 2017; Piazza et al. 2018), some authors have

proposed that other continuous dimensions are integrated and used as a proxy for number in later decision-making stages (Gebuis et al. 2016; Leibovich et al. 2017). Others have suggested that both numerosity and continuous dimensions are encoded and sustained interdependently and multidimensionally throughout visual processing (Lourenco and Aulet 2022). However, these alternative accounts still do not provide a detailed computational account of how other continuous dimensions are integrated to represent numerosity or how numerosity and other dimensions are encoded interdependently.

Computational modeling studies have provided some insights into the mechanisms of numerosity perception. Dehaene and Changeux first proposed that the number sense is extracted from feedforward visual processing stages in which objects on the retina are coded by their position regardless of the size of each object (Dehaene and Changeux 1993). Recent computational models following this foundational framework of feedforward mechanisms have shown that some units in a deep convolutional neural network can be sensitive to numerosity, indicating that some combination of convolution and pooling of neural activity across multiple layers gives rise to numerosity representation (Stoianov and Zorzi 2012; Nasr et al. 2019; Kim et al. 2021). There are distinct advantages to some of these learning models, such as their capacity to explain the development of a visual

number sense in newborn infants and in animal species like rhesus monkeys and honeybees (Testolin et al. 2020). However, successful application of these models is only a first step. Although learning or nonlinear computations across the multiple layers of these models can produce network units sensitive to numerosity among other magnitude dimensions, additional investigations are needed to determine which neurocomputational mechanisms are necessary and which are not (Bowers et al. 2023; Kobylkov et al. 2023).

In our prior modeling work, we demonstrated that divisive normalization applied to template filters is the crucial neurocomputational mechanism for the visual number sense (Park and Huber 2022). To simulate number sense for dot arrays, this divisive normalization mechanism was implemented using center-surround filters, at differing spatial scales, within a single-layer convolutional network. The output of these retinotopic filters was then normalized, consistent with research on divisive normalization within the animal visual system (Carandini and Heeger 2012). We then quantified the output of the model in response to images of dot arrays varying in number (the number of dots in an array), size (the surface area of a dot array while holding number constant), and spacing (the overall inter-dot distance of the array while holding number constant). Importantly, all three dimensions can be manipulated independent of each other and various continuous dimensions such as the density, convex hull, total surface area, and individual dot area of the array can be expressed as a linear combination of the three independent dimensions (see DeWind et al. 2015). We found that the output activity of the divisive normalization model was strongly modulated by the number of dots in an array but not by size and spacing, and it was subject to various visual illusions (as described in detail below). Most importantly, this neurocomputational model provided a mechanistic explanation of the manner in which the nervous system is sensitive to numerosity information (Park and Huber 2022).

Center-surround filters were used in modeling the perception of dot arrays because they are the ideal filter for dots. However, that does not mean that numerosity representations arise only from center-surround neural responses. Rather, the proposed computational account suggests that divisive normalization is the neural mechanism that gives rise to the number sense. Because divisive normalization exists at all levels of visual processing, displays of objects other than dots may rely on behavioral “read out” from other brain regions. In other words, this model suggests that there is not just one processing stage for numerosity, but rather many stages.

To test the hypothesis that divisive normalization underlies sensitivity to numerosity, we took advantage of the coherence illusion in numerosity perception, where participants systematically underestimate heterogeneous arrays compared to homogeneous arrays (Lee et al. 2016; DeWind et al. 2020; Qu et al. 2022). One previous study showed that arrays in which items are oriented randomly are underestimated compared to arrays in which all the items are oriented identically (DeWind et al. 2020). Likewise, another previous study hinted at the possibility that arrays with heterogeneous item areas are underestimated relative to arrays with homogeneous item areas (Lee et al. 2016). While both illusory effects are considered the same coherence illusion, as will be explained, the divisive normalization model suggests different underlying neural activity for these two effects.

Previous electroencephalography (EEG) studies of numerosity perception show that brain activity becomes sensitive to numerosity extremely early in the visual cortical hierarchy. When

participants view dot-array images that vary systematically in number, size, and spacing, the amplitude of their visual evoked potentials become modulated primarily by number, but much less so—if any—by size or spacing, at the latency of 75 to 100 ms at central occipital scalp locations (likely arising from V1, V2, and V3) and again around 200 to 300 ms at bilateral occipital scalp locations (likely arising from V3 and beyond) (Park et al. 2016, 2021; Fornaciai et al. 2017; Fornaciai and Park 2018a, 2018b). The sensory, perceptual, or cognitive processes underlying these stages are unknown but could be determined by proposing and testing theories about the manner in which numerosity perception unfolds across different temporal stages of visual processing.

In this study, we use computational modeling to develop a theory of numerosity perception and use EEG to test empirical predictions of the theoretical framework. Specifically, we hypothesize that divisive normalization underlies numerosity perception, as demonstrated by the success of our original computational model (Park and Huber 2022). However, a successful model should not only explain existing data but also make novel, testable predictions, which lead to new experiments. To achieve this, we conducted a psychophysical experiment, designed and implemented a computational model to explain our psychophysical results and to elicit novel predictions, and then performed an EEG experiment to test the prediction of the model. First, we test for and confirm the behavioral effects of the coherence illusion in area (e.g. rectangles of same or different areas in an array) and orientation (e.g. rectangles of same or different orientations in an array). Second, we demonstrate that the divisive normalization model simulates the area coherence illusion with a first layer of center-surround filters but that it only simulates the orientation coherence illusion after augmenting our prior model with a second layer of Gabor filters of different spatial frequencies. This model simulation predicts that the orientation coherence illusion arises from a different, later visual processing stage than the area coherence illusion. Third, we test this prediction using the high temporal resolution of the EEG. The analysis of the visual evoked potentials confirms the prediction that these two coherence illusions occur in different stages of visual processing as determined by the latency and topography of the neural activity.

Materials and methods

Participants

Twenty-seven subjects ($n = 27$) recruited from the departmental research participation system participated in the behavioral psychophysical experiment. They included 4 men and 23 women with the age range from 18.2 to 22.9 years with the mean of 20.2 years. Self-reported ethnic and racial category were as follows: Hispanic or Latino (3) and Not Hispanic or Latino (24); White or Caucasian (17), Asian (8), and Black or African American (2).

Another group of 45 subjects from the same pool participated in the EEG experiment. These participants did not overlap with those who participated in the psychophysical experiment. Data from two subjects were corrupted due to equipment failure and one subject kept falling asleep; therefore, data from these three subjects were removed. Therefore, the final sample size was $n = 42$. This sample included 20 men and 22 women with the age range from 18.4 to 24.4 years with the mean of 20.4 years. Self-reported ethnic and racial category were as follows: Hispanic or Latino (2) and Not Hispanic or Latino (40); White or Caucasian (31), Asian (9), and Black or African American (2).

All the participants signed a written informed consent before participating in the experiment and were compensated with course credit. All the participants included in the study were naive to the purpose of the experiment and had normal or corrected-to-normal vision. All the experimental procedures were approved by the University of Massachusetts institutional review board and were in line with the Declaration of Helsinki.

Stimuli

A large set of arrays containing rectangular elements (see Fig. 1A) was created for the psychophysical experiment. All the visual stimuli were generated by adapting the previously published code for generating dot-array stimuli for magnitude perception studies (<https://osf.io/s7xer/>) in MATLAB version R2022b (The MathWorks, Inc.). In the psychophysical and EEG experiments, the stimuli were presented on a monitor screen running at 144 Hz with a resolution of 1920 × 1080, encompassing $\sim 34^\circ \times 19^\circ$ of visual angle from the viewing distance of 90 cm.

First, a set of “probe” arrays were created whose dimensions of number (N), size (S_z), and spacing (S_p) were systematically determined following previous work (Fornaciai and Park 2018a, 2021). This stimulus construction scheme allows the quantification of the effects of numerical and non-numerical dimensions on a dependent measure (DeWind et al. 2015; Park 2022). This scheme was used here to provide systematic variations in non-numerical dimensions while the subjects were asked to make a judgment on number. The different levels of non-numerical dimensions, however, were collapsed during data analysis because the primary goal of the study was on coherence illusory effects, not on the influence of non-numerical dimensions on numerosity judgments.

These probe arrays consisted of black (50%) and white (50%) rectangles on a gray background of size 600 × 600 pixels. When there was an odd number of rectangles, either the set of black or white rectangles (randomly chosen) outnumbered the other set by one. The size of the rectangles within an array was identical in these probe arrays, and the height of each rectangle was twice as large as its width (d). The minimum center-to-center distance between any two rectangles was at least twice as d . The number of rectangles (N) in a probe array ranged from 12 to 48 in seven equidistant levels on a logarithmic scale (i.e. $n=12, 15, 19, 24, 30, 38, 48$). The width (d) ranged from 7 to 14 pixels likewise in seven equidistant levels on a logarithmic scale. These rectangles were drawn in random positions within an invisible circle with the field radius (r_f) that ranged from 150 to 300 pixels in seven equidistant levels on a logarithmic scale. This configuration generated 91 unique sets of parameters (N , S_z , and S_p) across the three-dimensional parameter space defined by N , S_z , and S_p . A total of 455 unique probe arrays were pre-generated for the area coherence condition.

These arrays were used as the probe images in the “area coherence” condition. Note that the area of each rectangle was homogeneous within a probe array. The probe images in the “orientation coherence” condition were generated by rotating all the rectangles in each array in one of five different orientations: 12° , 84° , 156° , 228° , and 300° . Thus, while the orientations of the rectangles varied across arrays, they were homogeneous within a probe array. These orientations were chosen to avoid cardinal (0° and 90°) and oblique (45° and 135°) angles. As in the area coherence condition, a total of 455 unique probe arrays were pre-generated for the orientation coherence condition.

Second, separately for the area coherence and orientation coherence conditions, a set of “reference” arrays was created,

which contained both homogeneous and heterogeneous rectangles within each array. The reference arrays contained 25 rectangles, which was one more than the median number of rectangles ($n=24$) across all the probe arrays. This was intended as we expected the manipulation of coherence to cause an underestimation. That is, we anticipated that the perceived numerosity of the homogeneous reference array to be slightly greater than the median numerosity of the probe arrays, while the perceived numerosity of the heterogeneous reference array to be slightly smaller than the median numerosity of the probe arrays.

In the area coherence condition, homogeneous reference arrays were constructed by creating 25 rectangles using the median values of r_f ($=212$ pixels) and d ($=10$ pixels) from the probe arrays. Heterogeneous reference arrays were constructed likewise but by creating five rectangles with $d=4$, five rectangles with $d=6$, five rectangles with $d=9$, five rectangles with $d=12$, and five rectangles with $d=15$. This ensured that the area of the rectangles varied within each array (thus, creating heterogeneity) while holding the total area equivalent to that of the homogeneous reference arrays. A total of 175 unique homogeneous reference arrays and 175 unique heterogeneous reference arrays were pre-generated.

In the orientation coherence condition, homogeneous reference arrays were constructed by creating 25 rectangles using the median values of r_f ($=212$ pixels) and d ($=10$ pixels) from the probe arrays, in one of five orientations (i.e., 12° , 84° , 156° , 228° , 300°). Heterogeneous reference arrays were constructed likewise but by creating five rectangles with 12° , five rectangles with 84° , five rectangles with 156° , five rectangles with 228° , and five rectangles with 300° within each array. A total of 175 unique homogeneous reference arrays and 175 unique heterogeneous reference arrays were pre-generated.

In the EEG experiment, the aforementioned homogeneous and heterogeneous reference arrays—designed and used in the psychophysical experiment—were used. In half of the trials (see *Experimental paradigms*), a rectangle array was immediately followed by a backward mask. The backward mask images were created by randomly translating (up to -120 pixels and $+120$ pixels horizontally and vertically) each of 64 randomly chosen reference arrays and overlapping the images successively with an exclusive disjunction (xor) operation on each pixel, after which a circular aperture with a radius of 256 pixels cropped the image. This procedure of creating mask images was to ensure that the visual characteristic of the mask was similar to that of the reference array images. A total of 64 masks for the area coherence condition was created using the reference arrays from that condition; another 64 masks for the orientation coherence condition was created likewise.

In simulations using the computational model, the aforementioned homogeneous and heterogeneous reference arrays as well as the probe arrays were used. However, the size of the image was down sampled to 128×128 pixels for computational efficiency and the shade of the images was converted so that all the rectangles were white on a black background.

Data acquisition

All experiments were performed on individual subjects in a quiet room. The psychophysical experiment was built and presented using PsychoPy version 2022 February 4 (Peirce et al. 2019). The EEG experiment was built and presented using Psychtoolbox 3 on MATLAB versions R2013a. The EEG data were continuously recorded using an active electrode amplifier (actiCHamp; Brain Products GmbH) from 64 channels distributed in an extended

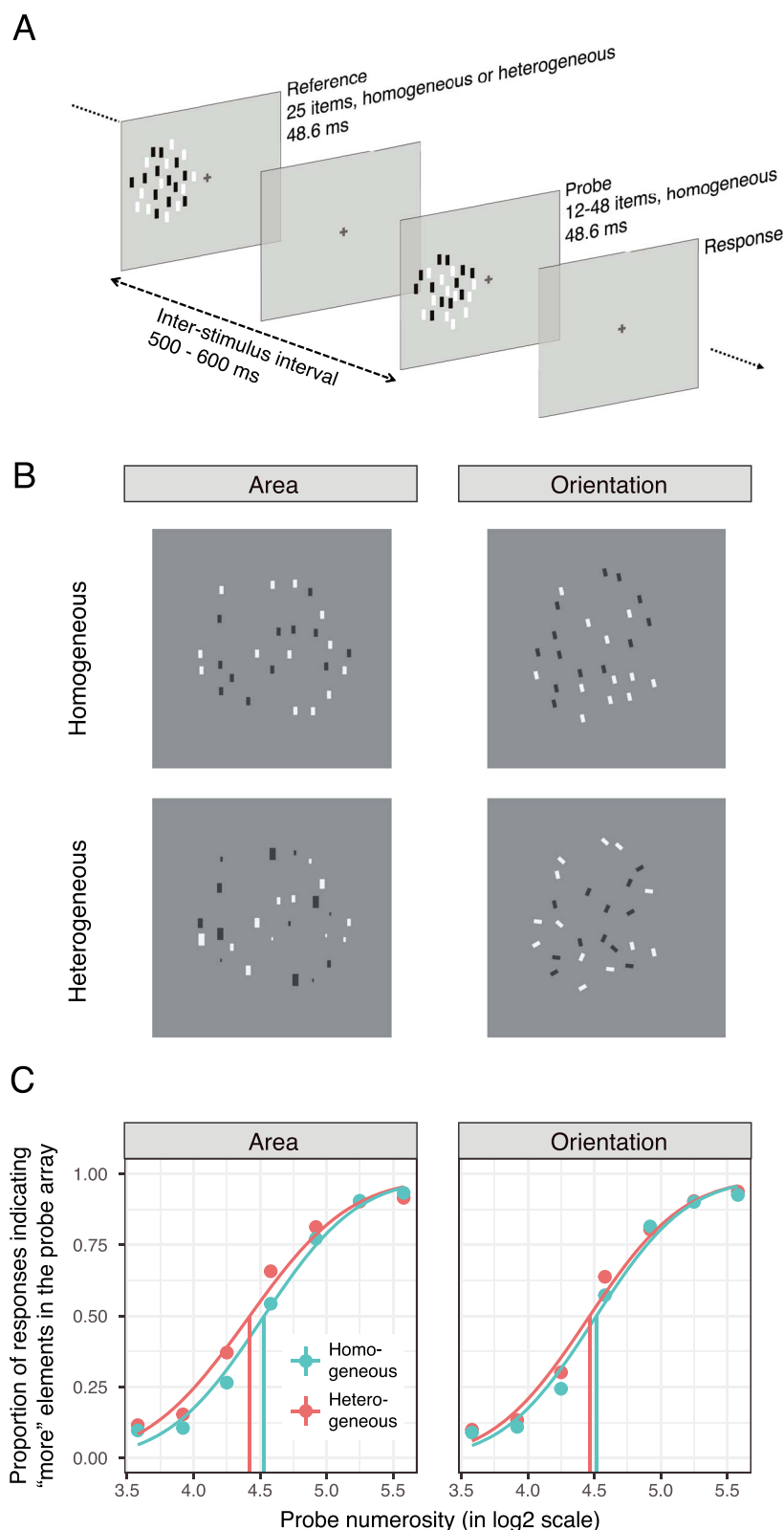


Fig. 1. Experimental procedure and results of the psychophysical experiment. A) Structure of an example trial. Participants judged which of the two arrays sequentially presented in the same position (randomly on either left or right of the fixation point) contained more items. The reference array always contained 25 homogeneous or heterogeneous items, and the probe array contained from 12 to 48 homogeneous items. The images are not drawn to scale. B) Example stimuli for the area and the orientation conditions. C) Psychometric curves computed from aggregate data across all participants describing how the probe stimuli were perceptually matched with either homogeneous or heterogeneous item arrays.

coverage, triangulated, equidistance cap (M10; EasyCap GmbH), with a sampling rate of 1000 Hz. During the recording, all the channels were referenced to the vertex (Cz). One vertical electrooculogram (EOG) electrode was placed below the left eye and two EOG electrodes were placed lateral to the left and right canthi. Channel impedances were kept below 15 k Ω most of the time, although up to 35 k Ω was tolerated.

Experimental paradigms

Psychophysical experiment

In the psychophysical experiment, each participant was given the area coherence condition (4 blocks of 98 trials) and the orientation coherence condition (4 blocks of 98 trials) in a counterbalanced order. The two conditions were identical in their experimental paradigm, except for the stimuli used for numerosity judgment. The paradigm was consistent with our previous experiments on the connectedness illusion (Fornaciai and Park 2018a, 2021). On each trial, a reference array consisting of 25 items was presented for 48.6 ms (7 screen flips under the refresh rate of 144 Hz) on either the left or the right side of a central fixation cross (eccentricity of 5.3°). The reference array always contained exactly 25 items that were either homogeneous or heterogeneous in nature (see Fig. 1A). For example, in the area condition either 25 homogeneously sized or 25 heterogeneously sized rectangles appeared on the left or right side. The reference array was followed by a variable interstimulus interval (ISI; 500 to 600 ms from a uniform distribution) where only the central fixation cross remained on the screen. After the ISI, a probe array was presented for 48.6 ms in the same spatial position as the reference array. The probe array contained from 12 to 48 homogeneous items to be compared against the reference array. The participant then judged which array contained more elements by pressing the “1” key (first/reference array contained more) or the “2” key (second/probe array contained more) on a standard keyboard. After the response, the following trial started after a variable intertrial interval of 950 to 1150 ms. The very brief stimulus presentation (~50 ms) and the peripheral spatial positions were specifically chosen to maximize the coherence bias. Biases in numerosity perception, such as the connectedness illusion and the serial dependence effect, are indeed usually stronger at short stimulus presentations and at peripheral locations (Franconeri et al. 2009; He et al. 2009; Fischer and Whitney 2014; Fornaciai and Park 2018b).

The reference and probe arrays were randomly chosen from the pre-generated images as described in the Stimuli section. The reference array always contained 25 rectangles. Importantly, a random half of the trials presented a homogeneous reference array, and the other half presented a heterogeneous reference array. The probe array contained one of the seven number of rectangles ($n=12, 15, 19, 24, 30, 38, 48$). The probability of the appearance of 12, 24, and 48 rectangles to the probability of the appearance of 15, 19, 30, and 38 rectangles was in a 3:4 ratio. This distribution was to ensure that slightly more trials are allocated for more challenging judgments and relatively less trials are allocated to trials that are the easiest or that need to be largely guessed.

Since the stimuli were presented always in the same order (reference first, followed by the probe array), the overall accuracy of numerosity judgments might be influenced by time-order errors (Hellström 1985; Hellström and Rammsayer 2015), namely, a systematic under- or overestimation of the first stimulus in the sequence, depending on the properties of the stimuli and the type of judgment. However, the present study focused on the relative difference in perceived numerosity due to the manipulation

of stimulus homogeneity and not on the absolute accuracy of numerical judgments. Even if possibly present, time-order errors are thus not expected to affect the interpretation of the results.

EEG experiment

In the EEG experiment, four different types of the reference arrays were presented at the center of the screen in random order. The four types were area homogeneous arrays, area heterogeneous arrays, orientation homogeneous arrays, and orientation heterogeneous arrays. Each reference array was presented for a duration of 48.5 ms (7 screen flips), which was followed by a variable intertrial interval (663 to 1063 ms from a uniform distribution). A total of 288 arrays were presented in each block, and a total of 8 blocks were given to each participant. A black fixation cross was presented at the center of the screen throughout the trial and across all the trials. Occasionally at random times, the central fixation cross turned red. Participants were instructed to press a button under their right or left index finger (counterbalanced), at which point the fixation cross turned yellow to provide button response feedback to the participants. The purpose of this color oddball detection task was to ensure that participants kept their attention on the center of the screen, and it was orthogonal to the purpose of the experiment. Of the 288 arrays in each block, 24 were presented as an oddball, and these oddball trials were excluded from the analysis. The average (\pm SD) of the median response time (RT) across subjects in the color oddball task was 438 ± 44 ms, and the average hit rate was 95.8%. In random half of the trials, the reference array was immediately followed by a very short (13.9 ms) fixation cross which was followed by a backward mask image described in Stimuli for 48.5 ms. The purpose of this manipulation was to assess the impact of backward masking on the visual evoked potentials for the design and interpretation of future studies. These trials are not the focus of the current work and therefore were not reported in this paper.

Data analysis

Psychophysical data

Behavioral response from the psychophysical experiment was analyzed to obtain individual subject's accuracy and precision in the numerosity judgment task. A cumulative Gaussian function was fitted to each individual subject's choice data under four conditions (area homogeneous, area heterogeneous, orientation homogeneous, orientation heterogeneous) using the *quickpsy* toolbox implemented in R, with the lapse rate set to 0.02 and all other parameters set to their default values (Linares and López-Moliner 2017). The median point of the best-fitting curve for each condition was taken as the point of subjective equality (PSE) for that condition, which was our primary measure of underestimation effect expected due to coherence illusion. In addition to this primary measure, the difference in numerosity between the median point and the 75% correct response from the fitted curve was taken as the just noticeable difference (JND). The Weber fraction was derived from JND divided by PSE. JND and Weber fraction were interpreted as the precision in the numerosity judgment task.

EEG data

The EEG data were analyzed off-line using the functions provided by the EEGLAB software package and the ERPLAB toolbox (Delorme and Makeig 2004; Lopez-Calderon and Luck 2014). As part of preprocessing, the raw continuous EEG data were high-pass filtered at 0.01 Hz and referenced to the average value of all the 64 channels. The continuous data were then segmented into epochs from 100 ms before to 500 ms after stimulus onset, with

a baseline correction using the pre-stimulus interval. We then performed an independent component analysis (ICA) to remove components identified as eye blinks. After ICA, epochs with large signal amplitudes (signal going outside the range of $[-150, 150]$ μV) were further identified and rejected using an artifact rejection tool in ERPLAB. This procedure led to the median rejection rate of 3.8%. One of the subjects had the rejection rate $>50\%$, and this subject was removed from further analysis, making $n=41$. Finally, the epochs were selectively averaged for each of the four conditions (area homogeneous, area heterogeneous, orientation homogeneous, orientation heterogeneous), followed by a low-pass filter at 30 Hz before computing the grand average of the event-related potentials (ERPs). Based on the previous results showing early ERP sensitivity to numerosity in the medial occipital channel (Park et al. 2016; Fornaciai and Park 2017, 2018a), our primary analysis was focused on the channel Oz. In order to test our hypothesis concerning the neural effects of coherence illusion in the case of area manipulation and in the case of orientation manipulation, the ERPs evoked by homogeneous arrays were contrasted with the ERPs evoked by heterogeneous arrays across each time point in Oz separately for the area conditions and for the orientation conditions.

Computational modeling

The previously published neural network model for magnitude perception served as the starting point for the computational model in this study (Park and Huber 2022). Therefore, some of the descriptions below are a reiteration of what was explained in that previous work. Images fed into the model were the four different types of reference arrays (area homogeneous, area heterogeneous, orientation homogeneous, and orientation heterogeneous) as in the EEG experiment and the probe arrays used in the psychophysical experiment, except the shades of the images were converted so that white rectangles were displayed on a black background and the size of the images was rescaled to 128×128 pixels for computational efficiency.

Difference-of-Gaussians layers

The first part of the neural network model consisted of a set of convolutional layers with five different sizes of difference-of-Gaussians (DoG) filters. This part of the architecture hence contained $128 \times 128 \times 5$ simulated neurons. The DoG filters were formally defined as the following:

$$\Gamma(x, y) = I \cdot \left(\frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} - \frac{1}{2\pi K^2\sigma^2} e^{-\frac{x^2+y^2}{2K^2\sigma^2}} \right), \quad (1)$$

I is the input image, σ^2 is the spatial variance of the narrower Gaussian, and K is the scaling factor between the two variances. $K=1.6$ was used to achieve balanced bandwidth and sensitivity of the filters (Marr and Hildreth 1980). Considering that the input values range $[0, 1]$, the DoG filters were reweighted so that the sum of the positive portion equals to 1 and the sum of the negative portion equals to -1 , making the summation across all domains 0. This reweighting ensured that the response is maximized when the input matches the DoG filter regardless of filter size and that the filter produces a response of value 0 if the input is constant across a region regardless of filter size. Finally, the output of this convolution process was followed by half-wave rectification at each simulated neuron, where negative responses were replaced by zero (Heeger 1991). This stipulation sets the “firing threshold” of the network such that the simulated neurons would not fire if the input does not match its DoG filter.

Five different σ values were used ($\sigma_k=1, 2, 4, 8, 16$ for filter size k , respectively) that covered spatially fine features like the edge of the smallest rectangles to spatially broad features like the landscape of the rectangles of the image. The activity of each simulated neuron, i , in filter size k following this convolution procedure is referred to as the driving input, $D_{i,k}$.

The activity of each simulated neuron was represented by the driving input of that neuron normalized by the driving inputs of other neighboring neurons. This divisive normalization procedure is modeled as the following (Carandini and Heeger 2012):

$$R_{i,k} = \frac{D_{i,k}^\gamma}{c + \sum_{j,k} \eta_{(i,j,k)} D_{j,k}^\gamma}, \quad (2)$$

where the neighborhood weight η is defined as

$$\eta_{(i,j,k)} = e^{-\frac{d_{(i,j)}}{r_k}}. \quad (3)$$

D_i is the driving input of neuron i (i.e., the output of the convolution procedure described above), $d_{(i,j)}$ is the Euclidean distance between neuron i and neuron j in any filter size, and c is a constant that prevents division by zero. The denominator minus this constant, which was set to 1, is referred to as the normalization factor. The parameter r_k , or the scaling factor, serves to scale between local and global normalization. As r_k gets larger, activities from a broader set of neurons constitute the normalization factor, and it was defined as a function of σ_k so that neurons with larger filter sizes have their normalization factor computed from a broader pool of neighboring neurons. The parameter γ determines the degree of amplification of individual inputs and serves to scale between winner-take-all and linear normalization. Following our previous work, r_k was set to $2\sigma_k$ and γ was set to 2. $R_{i,k}$ represents the normalized response of neuron i in filter size k (Park and Huber 2022).

The normalized responses were summed across the five layers of different filter sizes, $\sum_k R_{i,k}$, which resulted in a map of 128×128 simulated activations. This map of neural representation was fed into the next part of the neural network architecture (see Gabor layers below). The summed value across all neurons and all filter sizes, $\sum_{i,k} R_{i,k}$, was interpreted as the summarized neural activity in response to the given input image.

Gabor layers

In our investigation of the orientation coherence illusion, we assumed another set of layers capable of capturing orientation representations. Thus, a second part of the neural network model consisted of a set of convolutional layers with Gabor filters of four wavelengths (4, 8, 16, 32 pixels/cycle) and six orientations ($0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ$) as implemented in MATLAB version R2022b. This part of the architecture hence contained $128 \times 128 \times 4 \times 6$ simulated neurons. To keep the range of the simulated neural activity comparable across different stages of the neural network architecture, the output map of the DoG layers, R_i , was linearly rescaled before being fed into the set of Gabor filters. Specifically, a constant scaling factor was computed so that one of the output maps of the DoG layers would be rescaled to $[0, 1]$, and all other output maps would be rescaled using that constant scaling factor. Otherwise, the simulated values become infinitesimally small as the signal feeds through subsequent layers due to the immense divisive operations.

Convolution of the input activity with Gabor filters resulted in the driving input, $D_{i,\theta}$, in each simulated neuron i , in wavelength

λ , in orientation θ . As in the DoG layers, the driving input was normalized by the principle of divisive normalization (Eq. 2). The summation of normalized responses across all the wavelengths, orientations, neurons $\Sigma_{i,\lambda,\theta} R_{i,\lambda,\theta}$, was interpreted as the summarized neural activity after this second part of the model.

Unlike center-surround normalization, orientation normalization depends on the similarity between orientations (Hubel and Wiesel 1962). To capture the difference between neighboring objects of the same orientation versus neighboring objects of a different orientation, the scaling factor, r , for neighborhood weight η (Eq. 3) was assumed to vary as a function the difference between the orientation of the filter for one neuron θ_i and the orientation of the filter for one neuron θ_j . Initially, we set these parameters so that the model can capture cross-orientation suppression effects for overlapping stimuli: when $|\theta_i - \theta_j| = 0^\circ$, $r = 2\lambda$; when $|\theta_i - \theta_j| = 30^\circ$ or 150° , $r = 4\lambda$; when $|\theta_i - \theta_j| = 60^\circ$ or 120° , $r = 8\lambda$; when $|\theta_i - \theta_j| = 90^\circ$, $r = 16\lambda$. In this setting, different orientations are constantly more suppressed than similar orientations (e.g. Brouwer and Heeger 2011). However, this assumption runs counter to other studies reporting the opposite effect, in that presentation of identical or similar orientations in the surround produces more suppression in V1 cells (e.g. Meese et al. 2007; Coen-Cagli et al. 2015). Thus, we also performed simulations under the iso-orientation suppression assumption in which similar orientations are constantly more suppressed than different orientations: when $|\theta_i - \theta_j| = 0^\circ$, $r = 16\lambda$; when $|\theta_i - \theta_j| = 30^\circ$ or 150° , $r = 8\lambda$; when $|\theta_i - \theta_j| = 60^\circ$ or 120° , $r = 4\lambda$; when $|\theta_i - \theta_j| = 90^\circ$, $r = 2\lambda$. Finally, we also simulated the case where the scaling factor stayed constant, at $r = 2\lambda$, with no dependency in orientation.

The MATLAB code used to implement the model can be found in the following public repository: <https://osf.io/4rwjs/>.

Results

Numerosity of a heterogeneous array is underestimated compared to a homogeneous array

We first verified behavioral effects of coherence illusion in the case of area and orientation. While coherence illusion for orientation has been studied systematically (DeWind et al. 2020), empirical results concerning coherence illusion for area are less well documented. A psychophysical study on the perception of area, numerosity, and mean size suggested the possibility of underestimation of numerosity for arrays with heterogeneous individual areas, but a systematic investigation of homogeneous versus heterogeneous arrays was not its focus (Lee et al. 2016).

In a numerosity comparison task, participants indicated which of the two successively presented arrays of rectangles contained more items (Fig. 1A). In both the area and the orientation conditions, participants compared an array with homogeneous items (ranging in number from 8 to 32) to an array with 25 heterogeneous items or an array with 25 homogeneous items (Fig. 1B; see *Materials and methods* for the explanation of systematic manipulations of visual features of these arrays). To investigate the underestimation effects, a mixed-effects linear analysis was conducted on the point of subjective equality (PSE) using the lmer package (Bates et al. 2015). We included the heterogeneity of the array as well as the condition (i.e. area or orientation) as fixed effects and the participant as a random effect. There was a moderate underestimation of numerosity with heterogeneous displays in both conditions (Fig. 1C). In the area condition, the PSE of the array with 25 heterogeneous rectangles was 21.1 ($=2^{4.399}$) on average, which was significantly lower than the PSE of the array with

25 homogeneous rectangles, which was 22.9 ($=2^{4.515}$) on average ($t(22) = 5.197$, $p < 0.001$). In the orientation condition, the mean PSE of the heterogeneous array was 21.9 ($=2^{4.454}$) and the mean PSE of the homogeneous array was 22.6 ($=2^{4.499}$) ($t(22) = 2.868$, $p = 0.009$). The interaction between heterogeneity and condition was significant ($t(66) = -2.173$, $p = 0.033$), indicating that underestimation due to coherence illusion was stronger in the area condition than in the orientation condition. It is worth noting that the strength of the coherence illusion observed in our experiment was largely consistent with previous observations, which is reported to be ~5% to 7% of underestimation (DeWind et al. 2020; Qu et al. 2022), although a quantitative comparison between the effect sizes needs caution.

Divisive normalization model explains sensitivity to numerosity and insensitivity to size and spacing

In a previous computational modeling work, we proposed the number sense divisive normalization model using a single layer of center-surround filters, testing for the effects of number, size, and spacing (Fig. 2B) of circular dot arrays on the network unit activity (Park and Huber 2022). Here, we apply the same model in its original form to the rectangle array stimuli used in the psychophysical experiment.

The divisive normalization model contains a single convolutional layer with difference-of-Gaussians (DoG) filters of five different sizes (Fig. 2A). The driving input D for each network unit is determined by convolving the stimulus image with a DoG filter of a particular size at a particular position (Eq. 1). The driving input is then normalized by the weighted summation of neighboring neurons and filter sizes (Eq. 2), resulting in a normalized response R for each network unit. As applied to the arrays with rectangular items, the total normalized response across all network units, ΣR , was strongly modulated by the number of rectangles whereas any effect of size or spacing of the array was minimal. This was true for the stimuli used in both the area and orientation conditions (Fig. 2C). A regression was performed with the simulated summed normalized responses as the dependent measure and the three dimensions (N , Sz , Sp) as the independent variables to assess the relative size of the effects of N , Sz , Sp on the output activity. This analysis revealed a much larger coefficient estimate for N ($b = 19.47$ for area and $b = 19.47$ for orientation) than for Sz ($b = 2.70$ for area and $b = 2.64$ for orientation) and for Sp ($b = 1.64$ for area and $b = 1.70$ for orientation). The R-squared value for area was 0.965, and the R-squared value for orientation was 0.966.

Divisive normalization model predicts underestimation of numerosity for heterogeneity of area and orientation for different reasons

We then assessed the outputs of the divisive normalization model in response to rectangle arrays with heterogeneity manipulations in area and orientation separately. The visual arrays used in the psychophysical experiment were fed into the original DoG center-surround divisive normalization model. We compared the results for arrays with 25 rectangles that were homogeneous in area versus heterogeneous in area (identical orientation in all cases). The same comparison was done with 25 rectangles that were homogeneous versus heterogeneous in orientation (identical area in all cases). The model produced 9.75% smaller summed normalized responses, ΣR , for area heterogeneous arrays compared to area homogeneous arrays (Cohen's $d = 6.85$), but there was no reduction in the normalized responses in the case of orientation arrays (Fig. 3A).

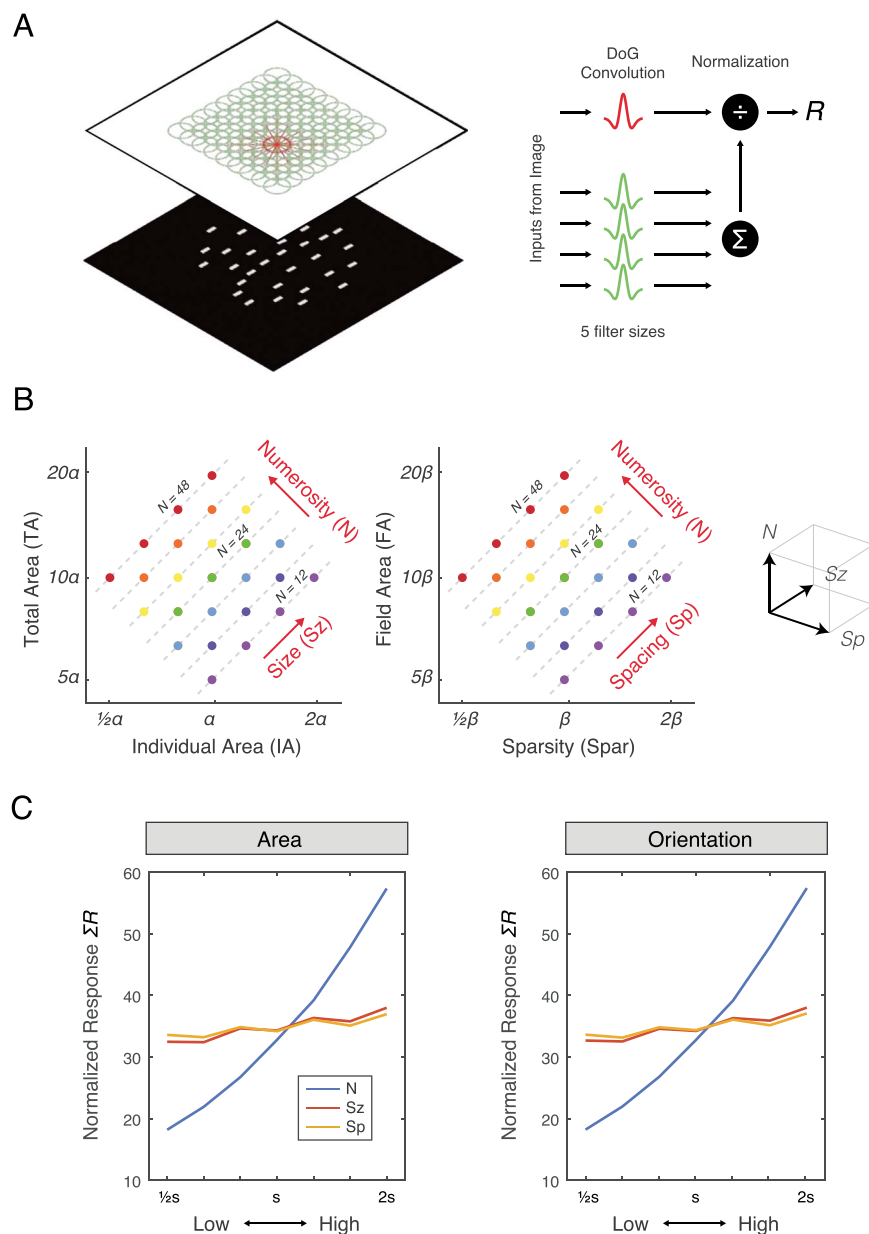


Fig. 2. Illustration of the computational model and simulation results of the effect of number (N), size (Sz), and spacing (Sp) on the normalized responses as applied to the arrays of rectangles. A) Bitmap images of item arrays were fed into a convolutional layer with difference-of-Gaussian (DoG) filters with divisive normalization (for details, see computational modeling in *Methods*). The resulting normalized responses were considered the output of the neural network. B) The probe arrays were constructed systematically across the log-scaled dimensions of N , Sz , and Sp following previous studies (DeWind et al. 2015; Park and Huber 2022). N refers to the number of items in an array. Sz refers to the dimension independent of N that is associated with both individual item area and total item area. In other words, it is the dimension that changes with individual item area or total item area while holding N constant (because they necessarily change together under the constant N). Sp refers to another dimension independent of N that is associated with sparsity (inverse of density) and field area. It is the dimension that changes with sparsity and field area while holding N constant (again because they necessarily change together under the constant N). Sz and Sp are independent of each other as well. This architecture allows an easy visualization of stimulus parameters and a systematic analysis of the effects of various magnitude dimensions on choice behavior or neural data. For more details, see DeWind et al. (2015) and Park and Huber (2022). C) Normalized responses from the neural network were simulated from the probe arrays across all the stimulus parameters. The summed normalized responses, ΣR , were strongly modulated by N and much less so, if any, by Sz and Sp . The value s on the horizontal axis indicates the median value for each dimension.

As explained in our previous work, the underestimation of area with heterogeneous arrays arises from the nonlinear saturation of the normalized response as a function of item area (Park and Huber 2022). That is, the summed normalized response of an individual item increases logarithmically (not linearly) as the area of the item increases. Note that heterogeneity in area is achieved by making some items smaller and other items larger while keeping the total area constant. Because of the saturation

effect, the increase in the normalized response from making some items larger is not as great as the decrease in the normalized response from making some items smaller by the same amount. Thus, when individual item areas are heterogeneous, the overall normalized response becomes smaller than when all the items have the same area. The current simulation demonstrates that this effect holds regardless of whether the items are circular dots or rectangles, suggesting that the area coherence

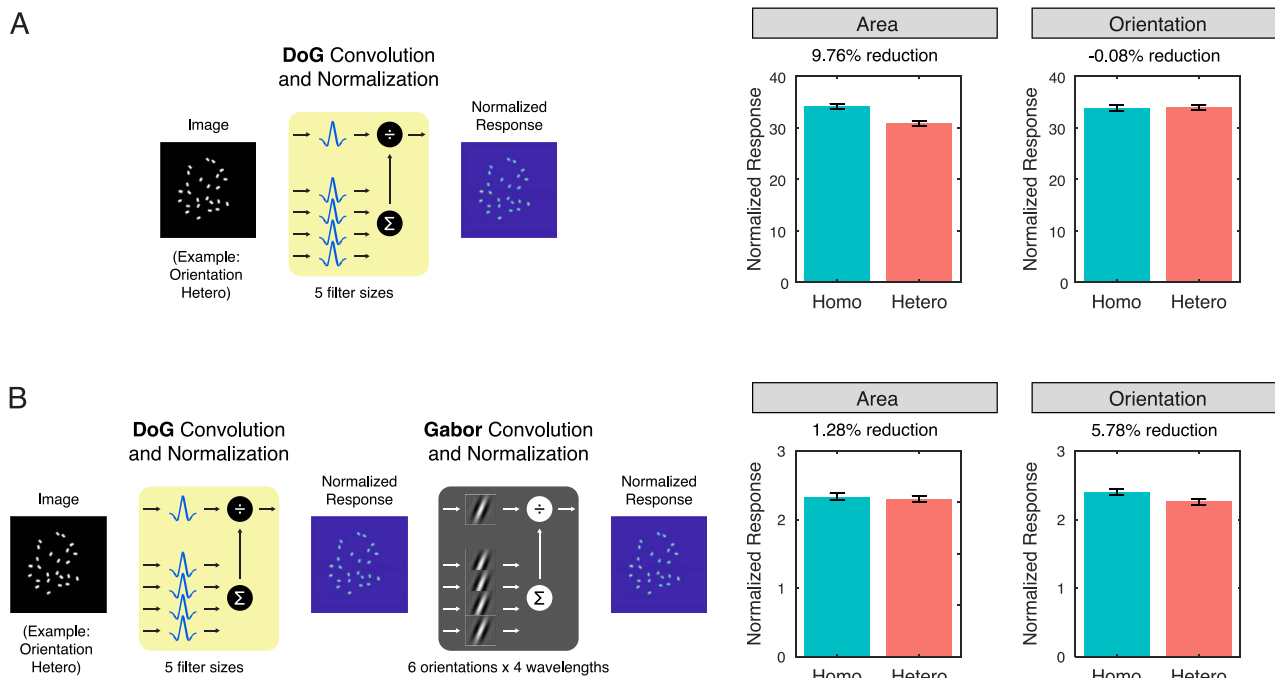


Fig. 3. Simulation of area and orientation coherence illusions from the divisive normalization model. A) A neural network with one layer of center-surround difference-of-Gaussian (DoG) convolution and divisive normalization. Compared to the homogeneous item arrays, heterogeneous arrays resulted in reduced normalized responses (by 9.76%) in the area condition. This reduction was not observed in the orientation condition, indicating that the neural activity underlying the coherence illusion is different for the two heterogeneity illusions. B) A neural network with a layer of DoG convolution and divisive normalization followed by another layer of Gabor convolution and divisive normalization. After this additional layer in the neural network, the heterogeneous arrays resulted in reduced normalized responses (by 5.78%) compared to the homogeneous arrays in the orientation condition. This reduction was much smaller in the area condition. The simulations all together suggest that the orientation coherence illusion arises from a downstream brain region as compared to the area coherence illusion.

illusion arises from a very early visual process (i.e., center-surround filters).

As might be expected considering that the center-surround model has no capacity to represent orientation, the original, unaltered divisive normalization model was completely unable to explain underestimation for heterogeneity of orientation. This failure, despite success at explaining the heterogeneity of area effect, suggests that the brain region producing the area coherence illusion is different than the brain region producing the orientation coherence illusion. To further explore the differences between these two different heterogeneity illusions, we added the capacity to represent orientation to the divisive normalization model by adding a second layer. If successful in explaining the orientation heterogeneity effect, the addition of this second layer predicts that the neural substrate underlying the heterogeneity effect resides in a different neural process at a longer delay from stimulus onset than the heterogeneity effect of area.

To simulate the orientation coherence illusion, a second convolutional layer that contained oriented Gabor filters rather than center-surround filters was added to the model (Fig. 3B). In this augmented model, the normalized responses from the center-surround DoG filters were fed into 24 different types of Gabor filters representing six different orientations of four different sizes (i.e., different spatial frequencies as captured by adjusting the wavelength parameter), with orientation-dependent normalizations (see *Materials and methods*). The dependency on the orientation difference resulted in a convolutional layer that attempts to determine the dominant orientation at each location, consistent with classic physiological findings (Hubel and Wiesel 1962; Klímová et al. 2021), and consistent with the “pooling” process that

is often used in machine vision convolutional neural networks (Krizhevsky et al. 2012).

Images of rectangle arrays with homogeneous and heterogeneous orientations were submitted to this augmented model, with cross-orientation suppression. The output of the model showed smaller summed normalized responses, ΣR , for heterogeneous arrays compared to homogeneous arrays by 5.78% (Cohen’s $d=3.10$) (Fig. 3B). These results demonstrate that the “underestimation” of heterogeneous arrays in the orientation condition can arise from this additional convolutional layer of Gabor filters with divisive normalization where different orientations are constantly more suppressed than similar orientations (Brouwer and Heeger 2011).

The greater normalization between filters of different orientations is likely the cause of the underestimation of heterogeneous arrays in the augmented divisive normalization model. When the rectangles have the same orientation, they will not suppress each other as much within a local region, but when there are of differing orientations within a local region, there will be greater normalization for that region. The simulations resulted in a different pattern when the augmented model assumed iso-orientation suppression; compared to homogeneous arrays, the summed normalized responses in heterogeneous arrays were greater by 2.60% ($d=-1.26$). When no dependency in orientation was assumed, the responses in heterogeneous arrays were smaller by 0.90% ($d=0.30$). Orientation-dependent normalization in the brain is likely to be a complex mixture of these different cases depending on the anatomical location and the context of the visual parameters. However, regardless of whether the simulated Gabor layer corresponds to area V1 versus a later “pooling” region

(Krizhevsky et al. 2012) that is attempting to identify orientation regardless of location, the simulation results reveal that the center-surround layer cannot capture the orientation coherence illusion, and, furthermore, that this effect can be captured by some form of normalization in a downstream brain region that is sensitive to orientation.

The model does not contain feedback between the layers, so the inclusion of the Gabor layer does not alter the area coherence effect in the DoG layer. In the area coherence condition, the objects did not differ in orientation. Thus, it is sensible that behavior reflects the summed activity of the DoG in the area coherence condition because participants focused on blob rather than orientation detection. However, to assess the specificity of the two coherence effects, we also evaluated the summed activity of the Gabor layer for an area heterogeneity effect. The area heterogeneity effect was much reduced in the Gabor layer as compared to the DoG layer (Fig. 3B). In other words, the two different kinds of coherence effects were predicted to reflect divisive normalization in different brain regions. This is consistent with the idea that numerosity perception does not arise from a single processing stage. Instead, numerosity perception arises from multiple processing stages, with each stage reflecting divisive normalization in the brain region that is best suited to identify the elements that make up a scene (i.e., if assessing identically oriented rectangles, center-surround is best, but if assessing object that differ in orientation, Gabor filters are needed). Thus, the orientation coherence effect primarily lives in the Gabor layer and the area coherence effect primarily lives in the DoG layer.

In summary, the divisive normalization model explains the area coherence illusion as arising from nonlinear saturating responses in the center-surround filter layer, whereas the orientation coherence illusion primarily reflects competition between orientations in the oriented Gabor filter layer. Thus, the two effects are predicted to reside in different brain regions and to occur at different delays from onset of the image.

Visual evoked potentials show an earlier effect of area coherence and a later effect of orientation coherence as predicted by the computational model

We used EEG to test the predictions that area and orientation coherence illusions arise at different time points as expected if the origin of orientation coherence is downstream from area coherence. In the EEG experiment, participants saw the “reference” arrays from the psychophysical experiment flashed briefly on the screen. These arrays contained rectangular items with identical values of number, size, and spacing but were either homogeneous or heterogeneous in area, or were either homogeneous or heterogeneous in orientation (see Fig. 1A). Prior studies demonstrated that visual evoked potentials in the midline occipital channel are sensitive to numerosity as early as 75 ms (Park et al. 2016; Fornaciai et al. 2017). Based on this finding, the primary channel of interest was Oz. As expected, the visual evoked potential was modulated by heterogeneity, as computed by the brainwaves of the homogeneous conditions minus those of heterogeneous condition (Fig. 4A). A point-by-point t-test with multiple comparisons correction (false discovery rate, FDR $q < 0.05$) was performed, which indicated a significant effect of area heterogeneity at [88, 112] ms (Cohen’s $d = -0.641$) and a significant effect of orientation heterogeneity at [188, 252] ms ($d = -0.753$) and at [302, 326] ms ($d = 0.511$). Confirming model predictions, the effect of area heterogeneity was earlier than the effect of orientation heterogeneity.

The topographic pattern across the electrodes was then examined to assess whether the two effects were reflected in different brain regions. The posterior-view topographic maps of the two heterogeneity effects (Fig. 4B) indicate different electrode sources for the two effects. The area heterogeneity effect was observed early in the midline occipital site (75 to 125 ms) likely indicating V1, V2, and V3 as the source (Fornaciai et al. 2017), while the orientation heterogeneity effect was observed later in bilateral occipital sites (175 to 275 ms) from V3 and beyond, which became focal again in the midline occipital site after 275 ms, with inverted polarity, possibly indicating feedback signals into V1, V2, and V3. The observed bilateral effects motivated us to conduct a post hoc analysis on the four channels left and right of Oz (channel locations equivalent to PPO7h, POO7h, POO8h, and PPO8h in the standard 10–20 system) (Fig. 4C). A point-by-point t-test revealed a significant effect of orientation heterogeneity at [110, 128] ms ($d = -0.515$) and [176, 278] ms ($d = -1.054$).

In summary, the earliest effect of area heterogeneity arises from the medial occipital site (at 88 ms), followed by the earliest effect of orientation heterogeneity in the bilateral occipital sites (at 110 ms), consistent with a downstream cortical source for the orientation heterogeneity effect as compared to the area heterogeneity effect predicted by our computational model.

Discussion

The results presented here demonstrate divisive normalization as a crucial mechanism underlying the visual number sense. In its broadest terms, the model produces sensitivity to the number of items (but not to size or spacing of the array) owing to local regions of suppression (normalization) within convolutional filter layers (e.g., center-surround or oriented Gabor). The computational model explains various illusions of numerosity perception, including the area coherence illusion and the orientation coherence illusion. These two illusions appear to describe the same phenomenon; an array with less coherent (i.e., heterogeneous) items is underestimated compared to an array with more coherent (i.e., homogeneous) items. However, the model suggested otherwise, and this prediction was confirmed by the EEG data.

In the calculation of divisive normalization, there is both self-normalization (i.e., self from a single item’s perspective) as well as local neighborhood normalization. These two aspects of divisive normalization underlie different kinds of underestimation for arrays of heterogeneous items. First, self-normalization of center-surround filters underlies the area coherence effect. Heterogeneous area arrays were constructed such that some items were smaller and other items were larger than the items in the homogeneous array, while maintaining the same total area across the arrays. Thus, because of the nonlinear saturation effect with self-normalization, the resulting increase in output to larger items was smaller than the decrease in output to smaller items, producing a net reduction of the total response across the items.

In contrast, our simulation results suggest that the orientation coherence illusion reflects local neighborhood normalization rather than self-normalization. Unsurprisingly, the original center-surround divisive normalization model was unable to simulate the underestimation of heterogeneously oriented arrays. However, in the layer of oriented Gabor filters, neurons representing different orientations within the same retinotopic region are modeled to be competitive (i.e., cross-orientation suppression), which resulted in greater normalization for arrays of heterogeneously oriented rectangles than homogeneously oriented rectangles. It should be noted that other assumptions about

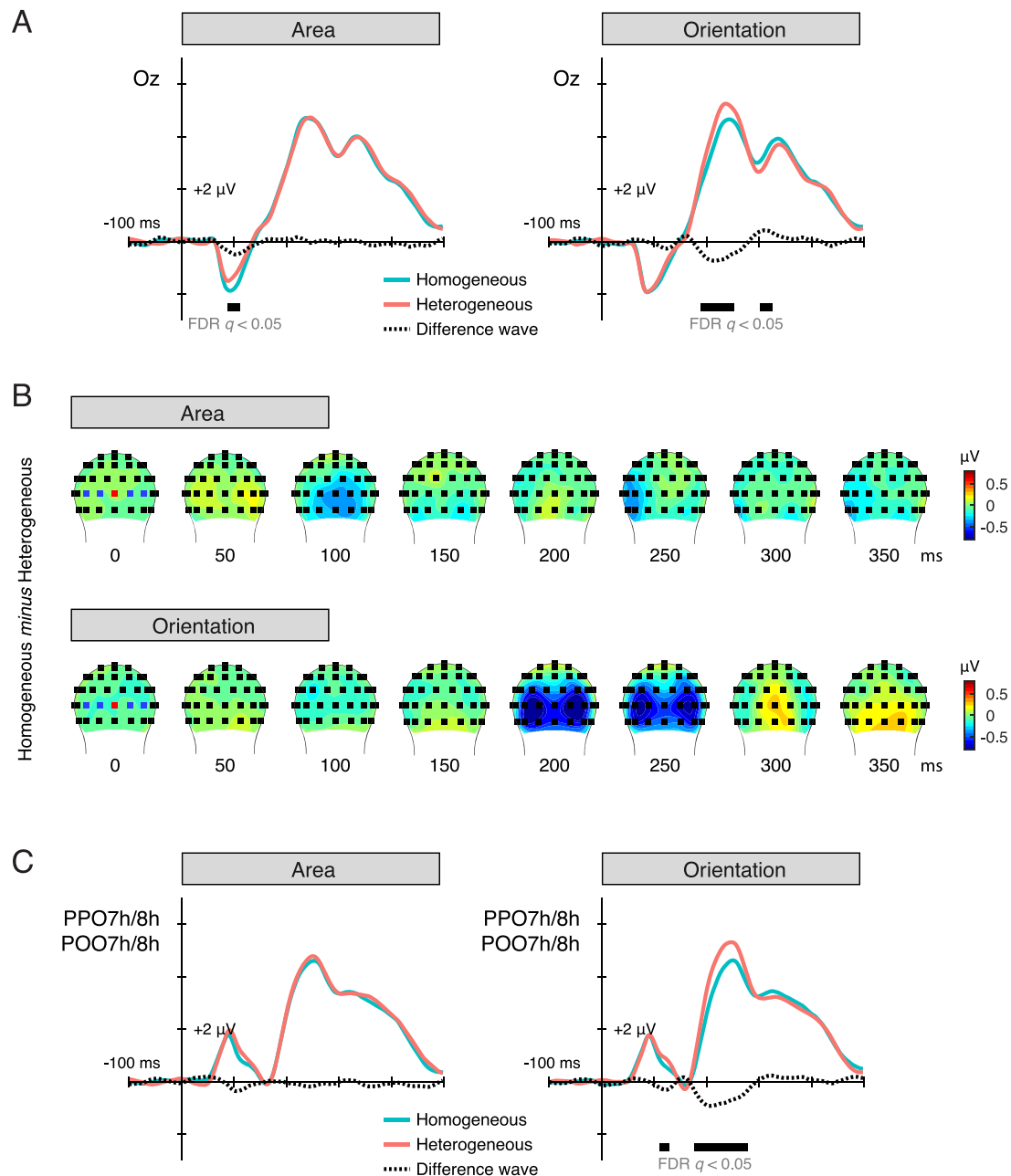


Fig. 4. Neural activity in response to homogeneous and heterogeneous arrays in the area and the orientation conditions. A) Brain waves from the medial occipital channel (Oz) indicate that the effect of coherence (homogeneous versus heterogeneous) is present at an earlier time point in the area condition (first observed at 88 ms) than in the orientation condition (first observed at 188 ms). The difference waves (dotted lines) are computed from homogeneous minus heterogeneous waves, and the black horizontal lines indicate the time points when the difference wave is significantly different from zero (corrected for multiple comparisons across time points using FDR). B) Posterior-view topographic maps of difference waves across time bins with a width of 50 ms (e.g. time 0 ms includes data from -25 to 25 ms, and time 25 ms includes data from 25 to 75 ms). These topomaps demonstrate the earlier and centralized effect of the area coherence effect and the later and lateralized effect of the orientation coherence effect. C) Identical to panel A, except the brain waves illustrate the average signal from the four bilateral channels from Oz (PPO7h, PPO8h, POO7h, and POO8h). The effect of orientation heterogeneity in these bilateral channels was first observed at 110 ms.

orientation-dependent normalization, unsurprisingly, resulted in different patterns of results. The significance of our simulations, however, is that the center-surround contrast filters are not capable of capturing any differences in the neural activities in response to homogeneously versus heterogeneously oriented arrays. Thus, our computational model predicts that orientation coherence illusion must be captured by another form of normalization after the center-surround layer. Within the scope of our current work, our focus was not to test the specifics of the orientation-tuned normalization. Nevertheless, the current results invite exciting

future research on the modeling of orientation-tuned normalization that seems to exhibit, according to the literature, both cross-orientation suppression and iso-orientation suppression depending on the brain region and the extent of overlap between receptive fields (Hubel and Wiesel 1962; Meese et al. 2007; Brouwer and Heeger 2011; Coen-Cagli et al. 2015).

The result from the augmented divisive normalization model suggests that the orientation coherence illusion arises in a downstream brain region, predicting that the effect of orientation heterogeneity on number perception should occur with a

longer delay and likely with a different topography than area heterogeneity. Confirming this novel prediction with EEG, the modulatory effect of orientation heterogeneity was found in bilateral occipital sites later than the modulatory effect of area heterogeneity in medial occipital sites. Such differences in timing and topography for the two different illusory effects provide strong credence to the divisive normalization model as a mechanistic explanation for numerosity perception.

Several of our previous works have demonstrated that visually evoked potentials in response to dot arrays ranging systematically across number, size, and spacing are uniquely sensitive to number starting from as early as 75 to 100 ms in the medial occipital channel and again from 175 to 250 ms in the bilateral occipital channels (Park et al. 2016; Fornaciai et al. 2017). The “earlier” activity has been suggested to reflect part of the C1, a component thought to be generated from feedforward processing in early visual areas (Jeffreys and Axford 1972; Clark et al. 1994; Di Russo et al. 2003; Ales et al. 2010, 2013; Kelly et al. 2013a, 2013b). An EEG experiment with a manipulation in the visual field of stimulus presentation and an fMRI study indeed suggested that the anatomical source of this early numerosity-sensitive activity is in V2 and V3 (Fornaciai et al. 2017; Fornaciai and Park 2018a). The second activity pattern resembles the P2p, which has traditionally been associated with numerical thinking broadly sourced in the bilateral middle occipital gyri or the inferior parietal lobule (Dehaene 1996; Pinel et al. 2001; Libertus et al. 2007; Hyde and Spelke 2012).

The timing and topography of the two heterogeneity effects in the current study resemble these previous findings. The heterogeneity effect of area at 88 ms in the medial occipital channel coincides with the previously documented early neural sensitivity to numerosity at 75 to 100 ms. This suggests that the mechanism which initially encodes numerosity is the same mechanism that drives underestimation in area coherence illusion. Given the current computational model, this explanation makes sense because the encoding of numerosity is explained by normalization within local regions and the heterogeneity effect of area is explained also by normalization within local regions, although a much narrower scope (i.e., self-normalization). On the other hand, the previously documented later neural sensitivity to numerosity at 175 to 250 ms coincides with the heterogeneity effect of orientation. This suggests that the involvement of downstream cortical patches could distort the representation of numerosity due to normalization across activity of cells with different types of receptive fields.

It should be noted that in a previous study with both fMRI and EEG, we identified V3 at a “later” latency (180 ms) as the source of the connectedness illusion (Franconeri et al. 2009; He et al. 2009; He et al. 2015; Fornaciai et al. 2016), and thus argued that the “earlier” latency activity does not directly give rise to the perception of numerosity (Fornaciai and Park 2018a). However, unlike that previous claim of ours, the current results do indicate that the earliest visual representation of numerosity can be behaviorally relevant in the case of the area coherence illusion. At the same time, the illusory effect of orientation coherence was associated with a later time point in the visual stream.

As noted in the introduction, divisive normalization exists in many brain regions, leading to the hypothesis that encoding of numerosity information may happen in multiple brain regions depending on how well the region is suited to identify the elements that make up the numerosity, and that the strength of encoding depends on the contextual information of the stimuli (Park et al. 2021). If the objects vary in orientation, orientation-selective brain regions may be engaged. In contrast, if the objects

vary in size, then center-surround brain regions may be best suited to the task. However, one limitation of the current model is that it does not specify the nature of this “read out” process. Nevertheless, behavioral studies indicate that numerosity estimation is influenced by a variety of different sensory factors reflecting different levels of visual processing, such as spatial frequency (Bonn and Odic 2024), color and pitch (Grasso et al. 2022), and motion (Fornaciai and Park 2017). These findings suggest that multiple stages in the visual hierarchy can contribute to numerosity representations and can remain susceptible to illusory effects. Future work could begin to study how participants use attention or learning to read out the appropriate neural response from different brain regions, such as by examining the influence of previous trials in an experiment in which different trials use different kinds of displays and judgment tasks.

One of the most puzzling problems in numerosity perception is how the brain encodes discrete magnitude in a rapid manner when the primary means of neural computation, such as firing rates and population codes, is continuous in nature. Theoretically, this can be achieved via a mechanism for serially accumulating continuous information with some kind of discretization process (Leslie et al. 2008). A thematically relevant idea has been suggested from an empirical study that demonstrated a positive relationship between the number of visual fixations during numerosity estimation and the performance in the task (Cheyette and Piantadosi 2019). However, our model assumes instantaneous perception from a single fixation. Moreover, the duration of stimulus presentation was <50 ms in our psychophysics experiment, which is assumed as the time for perceiving several dots in parallel within the parafoveal region (Cheyette and Piantadosi 2019). Thus, within a single fixation, parallel accumulation of the elements in an array may underlie numerosity even if additional information may be accumulated across fixations.

The success of the divisive normalization model for numerosity perception provides a firm framework for understanding how numerosity is encoded and represented in the brain. According to the current study and our prior work (Park and Huber 2022), neural sensitivity to numerosity arises as a byproduct of the canonical neural structure (i.e., convolutional filter layers) and computation (i.e., divisive normalization). The model is not “counting” the number of dots or projecting each individual items as one unit of activation and summing them up (Dehaene and Changeux 1993; Leslie et al. 2008). The model simply normalizes size and spacing information, which happens to make its output activity most correlated with the number of items in the array. Thus, strictly speaking, this process results in a representation of something like a number (i.e., a sensitivity to numerosity) but cannot be said to produce a representation of precise number consistent with modern human cultures (although see Clarke and Beck 2021 for an alternative view).

The current work has implications for neuroscience research broadly and suggests new exciting directions. The proposed mechanism represents a biologically plausible framework for vertebrate and invertebrate visual number perception, given the ubiquity of center-surround cells and divisive normalization in visual systems (Allman et al. 1985; Carandini and Heeger 2012). Importantly, our model only demonstrates a computational mechanism for extracting numerical features in the environment; across the phylogenetic tree, there may be many different “neural codes,” or neural representations, of number (Kobylykov et al. 2023). Comparing and contrasting features of numerosity perception across various species in light of our model may further generate and test new hypotheses about mechanisms for magnitude

perception and neural encoding. Beyond visual number processing, the divisive normalization model may also be able to account for perceived number or magnitude in other sensory modalities. Center-surround receptive fields can be found in the earliest stages of auditory and somatosensory processing (Waite 1973; Knudsen and Konishi 1978; Krupa et al. 1999). If divisive normalization is indeed a canonical neural computation, then it may be possible to use the current model with some modifications to explain number perception in other sensory modalities, or even mechanisms for an amodal representation of numerosity and magnitude (Feigenson et al. 2004; Togoli and Arrighi 2021).

The current findings demonstrate that a seemingly identical perceptual phenomenon can arise from normalization in different brain regions. Previous work on coherence illusions in the visual number sense has demonstrated that participants systematically underestimate heterogeneously sized, oriented, and colored arrays compared to homogeneous ones (Lee et al. 2016; DeWind et al. 2020; Qu et al. 2022). However, our evidence suggests that these coherence illusions cannot be isolated to one single processing stage. Instead, self-normalization at early visual processing stages appears to be responsible for the area coherence illusion, while neighborhood normalization at later visual processing stages appears to be responsible for the orientation coherence illusion. These findings emphasize the importance of computational modeling for generating and confirming novel testable predictions for advancing knowledge in neuroscience.

In sum, these results provide additional support for the hypothesis that divisive normalization within visual brain regions underlies the number sense. More specifically, the current findings demonstrate that seemingly identical perceptual phenomenon can arise from normalization in different brain regions. Divisive normalization is a general mechanism of neural processing, and it exists throughout the brain. Thus, this hypothesis predicts that numerosity can be represented in multiple processing stages depending on the nature of the visual display (e.g., the visual properties that vary in number) and task demands (e.g., the aspect of numerosity that is needed for the task). While previous work on coherence illusion has demonstrated that participants systematically underestimate heterogeneously sized, oriented, and colored arrays compared to homogeneous ones (Lee et al. 2016; DeWind et al. 2020; Qu et al. 2022), our evidence suggests that these coherence illusions cannot be isolated to one single processing stage. Instead, self-normalization at early visual processing stages appears to be responsible for the area coherence illusion, while neighborhood normalization at later visual processing stages appears to be responsible for the orientation coherence illusion. In both cases, neural sensitivity to numerosity is achieved as divisive normalization systematically removes effects of size and spacing. Together, these findings emphasize the importance of computational modeling for generating and confirming novel testable predictions for advancing knowledge in neuroscience.

Acknowledgments

The authors thank Matt Gillespie, Nathalie Valentin, Marina Morcos, Victoria Schille, Victoria Devian, and Diana Mendel for assistance in data collection.

Author contributions

Jenna Croteau (Conceptualization, Formal analysis, Methodology, Writing—original draft, Writing—review & editing), Michele Fornaciai (Methodology, Writing—original draft, Writing—review

& editing), David E. Huber (Formal analysis, Writing—review & editing), and Joonkoo Park (Conceptualization, Formal analysis, Methodology, Software, Writing—review & editing).

Funding

M.F. was supported by the European Union's Horizon Europe research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 101103020 ("PreVis").

Conflict of interest statement: None declared.

Data availability

The MATLAB code for the divisive normalization model of numerosity perception can be found in the following public repository: <https://osf.io/4rwjs/>.

References

- Ales JM, Yates JL, Norcia AM. On determining the intracranial sources of visual evoked potentials from scalp topography: a reply to Kelly et al. (this issue). *NeuroImage*. 2013;64:703–711. <https://doi.org/10.1016/j.neuroimage.2012.09.009>.
- Ales JM, Yates JL, Norcia AM. V1 is not uniquely identified by polarity reversals of responses to upper and lower visual field stimuli. *NeuroImage*. 2010;52:1401–1409. <https://doi.org/10.1016/j.neuroimage.2010.05.016>.
- Allman J, Miezin F, Mcguinness E. Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annu Rev Neurosci*. 1985;8:407–437. <https://doi.org/10.1146/annurev.ne.08.030185.002203>.
- Anobile G, Cicchini GM, Burr DC. Number As a primary perceptual attribute: a review. *Perception*. 2016;45:5–31. <https://doi.org/10.1177/0301006615602599>.
- Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. *J Stat Softw*. 2015;67:1–48. <https://doi.org/10.18637/jss.v067.i01>.
- Bonn D, Odic D. Effects of spatial frequency cross-adaption on the visual number sense. *Atten Percept Psychophys*. 2024;86:248–262. <https://doi.org/10.3758/s13414-023-02798-y>.
- Bowers JS, Malhotra G, Dujmović M, Llera Montero M, Tsvetkov C, Biscione V, Puebla G, Adolphi F, Hummel JE, Heaton RF, Evans BD, Mitchell J, Blything R. Deep problems with neural network models of human vision. *Behav Brain Sci*. 2023;46:e385, e385. <https://doi.org/10.1017/S0140525X22002813>.
- Brouwer GJ, Heeger D. Cross-orientation suppression in human visual cortex. *J Neurophysiol*. 2011;106:2108–2119. <https://doi.org/10.1152/jn.00540.2011>.
- Burr D, Ross J. Report a visual sense of number. *Curr Biol*. 2008;18:425–428. <https://doi.org/10.1016/j.cub.2008.02.052>.
- Butterworth B, Gallistel CR, Vallortigara G. Introduction: The origins of numerical abilities. *Philos Trans R Soc B Biol Sci*. 2018;373:1740. <https://doi.org/10.1098/rstb.2016.0507>.
- Carandini M, Heeger DJ. Normalization as a canonical neural computation. *Nat Rev Neurosci*. 2012;13:51–62. <https://doi.org/10.1038/nrn3136>.
- Cheyette SJ, Piantadosi ST. A primarily serial, foveal accumulator underlies approximate numerical estimation. *Proc Natl Acad Sci USA*. 2019;116:17729–17734. <https://doi.org/10.1073/pnas.1819956116>.

- Clark VP, Fan S, Hillyard SA. Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Hum Brain Mapp*. 1994;2:170–187. <https://doi.org/10.1002/hbm.460020306>.
- Clarke S, Beck J. The number sense represents (rational) numbers. *Behav Brain Sci*. 2021;44:e178, e178. <https://doi.org/10.1017/S0140525X21000571>.
- Coen-Cagli R, Kohn A, Schwartz O. 2015. Flexible gating of contextual influences in natural vision. *Nat Neurosci*. 2015;18:1648–1655. <https://doi.org/10.1038/nn.4128>.
- Dehaene S. The organization of brain activations in number comparison: event-related potentials and the additive-factors method. *J Cogn Neurosci*. 1996;8:47–68. <https://doi.org/10.1162/jocn.1996.8.1.47>.
- Dehaene S. *The number sense: how the mind creates mathematics*. USA: Oxford University Press; 2011.
- Dehaene S, Changeux JP. Development of elementary numerical abilities: a neuronal model. *J Cogn Neurosci*. 1993;5:390–407. <https://doi.org/10.1162/jocn.1993.5.4.390>.
- Delorme A, Makeig S. EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods*. 2004;134:9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>.
- DeWind NK, Adams GK, Platt ML, Brannon EM. Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*. 2015;142:247–265. <https://doi.org/10.1016/j.cognition.2015.05.016>.
- DeWind NK, Bonner MF, Brannon EM. Similarly oriented objects appear more numerous. *J Vis*. 2020;20:4–4. <https://doi.org/10.1167/jov.20.4.4>.
- Di Russo F, Martinez A, Hillyard SA. Source analysis of event-related cortical activity during visuo-spatial attention. *Cereb Cortex*. 2003;13:486–499. <https://doi.org/10.1093/cercor/13.5.486>.
- Feigenson L, Dehaene S, Spelke E. Core systems of number. *Trends Cogn Sci*. 2004;8:307–314. <https://doi.org/10.1016/j.tics.2004.05.002>.
- Fischer J, Whitney D. 2014. Serial dependence in visual perception. *Nat Neurosci*. 2014;17:5:17:738–743.
- Fornaciai M, Park J. Distinct neural signatures for very small and very large numerosities. *Front Hum Neurosci*. 2017;11:21. <https://doi.org/10.3389/fnhum.2017.00021>.
- Fornaciai M, Park J. Early Numerosity encoding in visual cortex is not sufficient for the representation of numerical magnitude. *J Cogn Neurosci*. 2018a;30:1788–1802. https://doi.org/10.1162/jocn_a_01320.
- Fornaciai M, Park J. Serial dependence in numerosity perception. *J Vis*. 2018b;18:15. <https://doi.org/10.1167/18.9.15>.
- Fornaciai M, Park J. Disentangling feedforward versus feedback processing in numerosity representation. *Cortex*. 2021;135:255–267. <https://doi.org/10.1016/j.cortex.2020.11.013>.
- Fornaciai M, Cicchini GM, Burr DC. Adaptation to number operates on perceived rather than physical numerosity. *Cognition*. 2016;151:63–67. <https://doi.org/10.1016/j.cognition.2016.03.006>.
- Fornaciai M, Brannon EM, Woldorff MG, Park J. Numerosity processing in early visual cortex. *NeuroImage*. 2017;157:429–438. <https://doi.org/10.1016/j.neuroimage.2017.05.069>.
- Franconeri SL, Bemis DK, Alvarez GA. Number estimation relies on a set of segmented objects. *Cognition*. 2009;113:1–13. <https://doi.org/10.1016/j.cognition.2009.07.002>.
- Gebuis T, Reynvoet B. The interplay between nonsymbolic number and its continuous visual properties. *J Exp Psychol Gen*. 2012a;141:642–648. <https://doi.org/10.1037/a0026218>.
- Gebuis T, Reynvoet B. Continuous visual properties explain neural responses to nonsymbolic number. *Psychophysiology*. 2012b;49:1649–1659. <https://doi.org/10.1111/j.1469-8986.2012.01461.x>.
- Gebuis T, Cohen Kadosh R, Gevers W. Sensory-integration system rather than approximate number system underlies numerosity processing: a critical review. *Acta Psychol*. 2016;171:17–35. <https://doi.org/10.1016/j.actpsy.2016.09.003>.
- Grasso PA, Anobile G, Arrighi R, Burr DC, Cicchini GM. Numerosity perception is tuned to salient environmental features. *iScience*. 2022;25:1–13. <https://doi.org/10.1016/j.isci.2022.104104>.
- Halberda J, Mazocco MMM, Feigenson L. Individual differences in non-verbal number acuity correlate with maths achievement. *Nature* 2008 455:7213. 2008;455:665–668. <https://doi.org/10.1016/j.isci.2022.104104>.
- He L, Zhang J, Zhou T, Chen L. Connectedness affects dot numerosity judgment: implications for configural processing. *Psychon Bull Rev*. 2009;16:509–517. <https://doi.org/10.3758/PBR.16.3.509>.
- He L, Zhou K, Zhou T, He S, Chen L. Topology-defined units in numerosity perception. *Proc Natl Acad Sci USA*. 2015;112:E5647–E5655. <https://doi.org/10.1073/pnas.1512408112>.
- Heeger DJ. Nonlinear model of neural responses in cat visual cortex. In: Landy MS, Movshon JA, editors. *Computational models of visual processing*. Cambridge, MA: MIT Press; 1991. pp. 119–133.
- Hellström Å. The time-order error and its relatives. *Mirrors of cognitive processes in comparing*. *Psychol Bull*. 1985;97:35–61. <https://doi.org/10.1037/0033-2909.97.1.35>.
- Hellström Å, Rammsayer TH. Time-order errors and standard-position effects in duration discrimination: an experimental study and an analysis by the sensation-weighting model. *Atten Percept Psychophys*. 2015;77:2409–2423. <https://doi.org/10.3758/s13414-015-0946-x>.
- Hubel DH, Wiesel TN. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*. 1962;160:106. <https://doi.org/10.1113/jphysiol.1962.sp006837>.
- Hyde DC, Spelke ES. Spatiotemporal dynamics of processing non-symbolic number: an event-related potential source localization study. *Hum Brain Mapp*. 2012;33:2189–2203. <https://doi.org/10.1002/hbm.21352>.
- Jeffreys DA, Axford JG. Source locations of pattern-specific components of human visual evoked potentials. I. Component of striate cortical origin. *Exp Brain Res*. 1972;16:1–21.
- Katzin N, Rosenbaum D, Usher M. The averaging of numerosities: a psychometric investigation of the mental line. *Atten Percept Psychophys*. 2021;83:1152–1168. <https://doi.org/10.3758/s13414-020-02140-w>.
- Kelly SP, Schroeder CE, Lalor EC. What does polarity inversion of extrastriate activity tell us about striate contributions to the early VEP? A comment on Ales et al. (2010). *NeuroImage*. 2013a;76:442–445. <https://doi.org/10.1016/j.neuroimage.2012.03.081>.
- Kelly SP, Vanegas MI, Schroeder CE, Lalor EC. The cruciform model of striate generation of the early VEP, re-illustrated, not revoked: a reply to Ales et al. (2013). *NeuroImage*. 2013b;82:154–159. <https://doi.org/10.1016/j.neuroimage.2013.05.112>.
- Kim G, Jang J, Baek S, Song M, Paik SB. Visual number sense in untrained deep neural networks. *Sci Adv*. 2021;7:1–9. <https://doi.org/10.1126/sciadv.abd6127>.
- Klímová M, Bloem IM, Ling S. The specificity of orientation-tuned normalization within human early visual cortex. *J Neurophysiol*. 2021;126:1536–1546. <https://doi.org/10.1152/jn.00203.2021>.
- Knudsen EI, Konishi M. Center-surround organization of auditory receptive fields in the owl. *Science* (1979). 1978;202:778–780. <https://doi.org/10.1126/jn.00203.2021>.

- Kobylov D, Zanon M, Perrino M, Vallortigara G. Neural coding of numerosness. *Biosystems*. 2023;232:104999, 104999. <https://doi.org/10.1016/j.biosystems.2023.104999>.
- Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst*. 2012;25:1–9. <https://doi.org/10.1016/j.biosystems.2023.104999>.
- Krupa DJ, Ghazanfar AA, Nicolelis MAL. Immediate thalamic sensory plasticity depends on corticothalamic feedback. *Proc Natl Acad Sci*. 1999;96:8200–8205. <https://doi.org/10.1073/pnas.96.14.8200>.
- Lee H, Baek J, Chong SC. Perceived magnitude of visual displays: area, numerosity, and mean size. *J Vis*. 2016;16:12. <https://doi.org/10.1167/16.3.12>.
- Leibovich T, Henik A. Comparing performance in discrete and continuous comparison tasks. *Q J Exp Psychol*. 2014;67:899–917. <https://doi.org/10.1080/17470218.2013.837940>.
- Leibovich T, Katzin N, Harel M, Henik A. From “sense of number” to “sense of magnitude”: the role of continuous magnitudes in numerical cognition. *Behav Brain Sci*. 2017;40:e164, e164. <https://doi.org/10.1017/S0140525X16000960>.
- Leslie AM, Gelman R, Gallistel CR. The generative basis of natural number concepts. *Trends Cogn Sci*. 2008;12:213–218. <https://doi.org/10.1016/j.tics.2008.03.004>.
- Libertus ME, Woldorff MG, Brannon EM. Electrophysiological evidence for notation independence in numerical processing. *Behav Brain Funct*. 2007;3:1–15. <https://doi.org/10.1186/1744-9081-3-1>.
- Linares D, López-Moliner J. 2017. quickpsy: An R package to fit psychometric functions for multiple groups.
- Lopez-Calderon J, Luck SJ. ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Front Hum Neurosci*. 2014;8:75729. <https://doi.org/10.3389/fnhum.2014.00213>.
- Lourenco SF, Aulet LS. A theory of perceptual number encoding. *Psychol Rev*. 2022;130:155–182. <https://doi.org/10.1037/rev0000380>.
- Marr D, Hildreth E. Theory of edge detection. *Proc R Soc Lond B Biol Sci*. 1980;207:187–217.
- Meese TS, Summers RJ, Holmes DJ, Wallis SA. Contextual modulation involves suppression and facilitation from the center and the surround. *J Vis*. 2007;7:7. <https://doi.org/10.1167/7.4.7>.
- Nasz K, Viswanathan P, Nieder A. Number detectors spontaneously emerge in a deep neural network designed for visual object recognition. *Sci Adv*. 2019;5:1–10. <https://doi.org/10.1126/sciadv.aav7903>.
- Park J. Flawed stimulus design in additive-area heuristic studies. *Cognition*. 2022;229:104919, 104919. <https://doi.org/10.1016/j.cognition.2021.104919>.
- Park J, Huber DE. A visual sense of number emerges from divisive normalization in a simple center-surround convolutional network. *elife*. 2022;11:1–16. <https://doi.org/10.7554/eLife.80990>.
- Park J, Dewind NK, Woldorff MG, Brannon EM. Rapid and direct encoding of numerosity in the visual stream. *Cereb Cortex*. 2016;26:748–763. <https://doi.org/10.1093/cercor/bhv017>.
- Park J, Godbole S, Woldorff MG, Brannon EM. Context-dependent modulation of early visual cortical responses to numerical and nonnumerical magnitudes. *J Cogn Neurosci*. 2021;33:2536–2547. https://doi.org/10.1162/jocn_a_01774.
- Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, et al. PsychoPy2: Experiments in behavior made easy. *Behav Res Methods*. 2019;51:195–203.
- Piazza M, De Feo V, Panzeri S, Dehaene S. Learning to focus on number. *Cognition*. 2018;181:35–45. <https://doi.org/10.1016/j.cognition.2018.07.011>.
- Pinel P, Dehaene S, Rivière D, LeBihan D. Modulation of parietal activation by semantic distance in a number comparison task. *NeuroImage*. 2001;14:1013–1026. <https://doi.org/10.1006/nimg.2001.0913>.
- Qu C, DeWind NK, Brannon EM. Increasing entropy reduces perceived numerosity throughout the lifespan. *Cognition*. 2022;225:105096, 105096. <https://doi.org/10.1016/j.cognition.2022.105096>.
- Starr A, Libertus ME, Brannon EM. Number sense in infancy predicts mathematical abilities in childhood. *Proc Natl Acad Sci USA*. 2013;110:18116–18120. <https://doi.org/10.1073/pnas.1302751110>.
- Starr A, DeWind NK, Brannon EM. The contributions of numerical acuity and non-numerical stimulus features to the development of the number sense and symbolic math achievement. *Cognition*. 2017;168:222–233. <https://doi.org/10.1016/j.cognition.2017.07.004>.
- Stoianov I, Zorzi M. Emergence of a “visual number sense” in hierarchical generative models. *Nature Neurosci*. 2012;15:194–196. <https://doi.org/10.1038/nn.2996>.
- Testolin A, Zou WY, McClelland JL. Numerosity discrimination in deep neural networks: initial competence, developmental refinement and experience statistics. *Dev Sci*. 2020;23:e12940, e12940. <https://doi.org/10.1111/desc.12940>.
- Togoli I, Arrighi R. Evidence for an A-modal number sense: numerosity adaptation generalizes across visual, auditory, and tactile stimuli. *Front Hum Neurosci*. 2021;15:713565, 713565. <https://doi.org/10.3389/fnhum.2021.713565>.
- Togoli I, Fornaciai M, Buetti D. The specious interaction of time and numerosity perception. *Proc R Soc B*. 2021;288:20211577. <https://doi.org/10.1098/rspb.2021.1577>.
- Waite PME. Somatotopic organization of vibrissal responses in the ventro-basal complex of the rat thalamus. *J Physiol*. 1973;228:527–540. <https://doi.org/10.1113/jphysiol.1973.sp010098>.
- Wilkey ED, Ansari D. Challenging the neurobiological link between number sense and symbolic numerical abilities. *Ann N Y Acad Sci*. 2019;1464:76–98. <https://doi.org/10.1111/nyas.14225>.