Context-dependent modulation of early visual cortical responses to numerical and non-numerical magnitudes

Joonkoo Park[1, 2*], Sonia Godbole[3], Marty G. Woldorff[3, 4], Elizabeth M. Brannon[5]

[1] Department of Psychological & Brain Sciences, University of Massachusetts

[2] Commonwealth Honors College, University of Massachusetts

[3] Center for Cognitive Neuroscience, Duke University

[4] Department of Psychiatry and Behavioral Sciences, Duke University

[5] Department of Psychology, University of Pennsylvania

* Correspondence should be addressed to:

Joonkoo Park, Ph.D.

Department of Psychological and Brain Sciences
University of Massachusetts
135 Hicks Way/Tobin Hall
Amherst, MA 01003. U. S. A.
Phone: (413) 545-0051
Email: joonkoo@umass.edu

Abstract

Whether and how the brain encodes discrete numerical magnitude differently from continuous non-numerical magnitude is hotly debated. In a previous set of studies, we orthogonally varied numerical (numerosity) and non-numerical (size and spacing) dimensions of dot arrays and demonstrated a strong modulation of early visual evoked potentials (VEPs) by numerosity and not by non-numerical dimensions. Although very little is known about the brain's response to systematic changes in continuous dimensions of a dot array, some authors intuit that the visual processing stream must be more sensitive to continuous magnitude information than to numerosity. To address this possibility, we measured VEPs of participants viewing dot arrays that changed exclusively in one non-numerical magnitude dimension at a time (size or spacing) while holding numerosity constant and compared this to a condition where numerosity was changed while holding size and spacing constant. We found reliable but small neural sensitivity to exclusive changes in size and spacing; however, changing numerosity elicited a much more robust modulation of the VEPs. Together with previous work, these findings suggest that sensitivity to magnitude dimensions in early visual cortex is context dependent: The brain is moderately sensitive to changes in size and spacing when numerosity is held constant, but sensitivity to these continuous variables diminishes to a negligible level when numerosity is allowed to vary at the same time. Neurophysiological explanations for the encoding and context dependency of numerical and non-numerical magnitudes are proposed within the framework of neuronal normalization.

Keywords

magnitude processing; numerosity; event-related potentials; neuronal normalization

Introduction

Humans possess a primitive, nonverbal ability to estimate discrete magnitude (henceforth, numerical magnitude or numerosity). While there has been considerable progress in understanding the cognitive and neural bases of this "number sense," it remains controversial as to whether our nervous system directly perceives discrete magnitude or extracts the numerical information from continuous, or non-numerical, magnitude information (e.g., see Burr & Ross, 2008; Durgin, 2008; Gebuis & Reynvoet, 2013). In a previous study (Park, DeWind, Woldorff, & Brannon, 2016), we used electrical recordings of brain activity (electroencephalogram; EEG) and a novel regression-based analytic method to quantify the unique contribution of different visual properties to the processing of dot array stimuli. We found very early visual event-related potentials (ERPs) to be considerably more sensitive to numerosity than to non-numerical magnitudes, suggesting a specific neural mechanism for the rapid and direct processing of discrete magnitude information (Park et al., 2016). These results have now been replicated in several subsequent studies (Fornaciai & Park, 2017; Fornaciai et al., 2017; DeWind et al., 2019).

There is a prevailing assumption in the literature that continuous magnitude information pervades the visual processing stream (for review, see Gebuis, Cohen Kadosh, & Gevers, 2016; Leibovich & Henik, 2013). While we did not find empirical evidence supporting such a view in our previous investigation (e.g., Park et al., 2016), the specifics of the experimental design may have led to a negative finding. In that previous study, dot arrays were constructed such that three orthogonal dimensions varied equally in magnitude: numerosity, size, and spacing. Numerosity refers to the number of dots in an array. Size refers to the dimension corresponding to the cumulative area of the dots that can be manipulated independently of spacing and numerosity. Spacing refers to the dimension corresponding to the inter-dot spacing of the array that can be manipulated independently of size and numerosity. [See **Section 2.2** below for the precise definitions of size and spacing]. The arrays were systematically sampled from a wide range of parameters for each of these dimensions and presented throughout the experiment. Thus, while the regression approach used in Park et al. (2016) allowed *statistical* assessment of the contribution of each dimension in explaining the variance in the ERPs, there was no condition that assessed the exclusive effects of one single dimension on the neural activity while the other two orthogonal dimensions were held constant. Therefore, it remained possible that a robust effect of numerosity within one stimulus set may have reduced sensitivity to size and spacing, leading to the observation of little effect of these continuous magnitude variables.

The present study was designed to assess the ERP effects of variations in one dimension at a time (orthogonal to the other two) (see **Fig. 1A**), particularly the effects of non-numerical magnitude (i.e., size and spacing) when numerosity was held constant. In Experiment 1, participants viewed dot arrays that changed exclusively in size while numerosity and spacing were held constant—i.e., we questioned whether the viewing of 16 smaller dots versus 16 larger dots would

elicit any difference in the visual ERPs. In Experiment 2, participants viewed dot arrays that changed exclusively in spacing while numerosity and size were held constant—i.e., we questioned whether the viewing of 16 dots that were widely scattered versus densely packed would show any difference in the visual ERPs. In both of these Experiments, blocks that varied exclusively in numerosity while holding size and spacing constant were also included, in order to compare the effects on the ERPs of varying size and spacing to the effects of varying numerosity.

Material and Methods

Participants

Across two experiments, a total of 65 participants from the Duke University subject pool participated in the study. Data from one participant could not be acquired due to equipment error, resulting in a final sample of 64 participants (age range of 18.1-22.6 years with mean of 19.6 years; 30 males). All participants were right-handed and had normal or corrected-to-normal vision. Participants provided written informed consent to a protocol approved by the Duke University Institutional Review Board.

Stimuli

Visual stimuli were white dot arrays presented on a black background. A unique dot array was generated for each trial using a custom algorithm that drew non-overlapping dots within an invisible circular field. The dot size was homogenous within each dot array.

A stimulus set was constructed that varied systematically in numerosity (*N*) and other continuous magnitudes, following the scheme used in our previous study (Experiment 2 of Park et al., 2016). Individual item perimeter (*IP*) refers to the perimeter encompassed by a single dot. Total perimeter (*TP*) refers to the total perimeter encompassed by the dots in an array and is simply $IP \times N$. Field area (*FA*) refers to the area of the invisible circle on the screen within which the dots were drawn. Sparsity (*Spar*) is defined by FA/N, thereby making it the inverse of density. As seen in **Figure 1**, logarithmic scaling of these parameters allowed a construction of two novel dimensions: size (*Sz*) and spacing (*Sp*).

Size (*Sz*) is the dimension that changes the overall perimeter of the dots when *N* is held constant, or mathematically expressed as:

$$\log(Sz) = \log(TP) + \log(IP).$$

This equation is derived from the relationship between *N*, *TP*, and *IP* ($N = TP/IP$; $\log(N) = \log(TP) - \log(IP)$), and since *TP* and *IP* are orthogonal, an orthogonal dimension to $\log(N)$ can be represented by $\log(TP) + \log(IP)$. Size (*Sz*) represents the dimension that changes the overall perimeter of the dots independent of *N*. That is, when numerosity (*N*) is held constant and *IP* is varied by some scaling factor, then *TP* must be varied by the same scaling factor.

Likewise, spacing (*Sp*) is the dimension that changes the overall spacing of the dots when *N* is held constant, or mathematically expressed as:

$$\log(Sp) = \log(FA) + \log(Spar).$$

Again, this equation is derived from the relationship bewteen *N*, *FA*, and *Spar* (*N* = *FA*/*Spar*; log(*N*) = log(*FA*) – log(*Spar*)), and since *FA* and *Spar* are orthogonal, an orthogonal dimension to log(*N*) can be represented by log(*FA*) + log(*Spar*). Spacing (*Sp*) represents the dimension that changes the overall spacing of the dots independent of *N*. That is, when numerosity (*N*) is held constant and *FA* is varied by some scaling factor, then sparsity (*Spar*) must be varied by the same scaling factor.

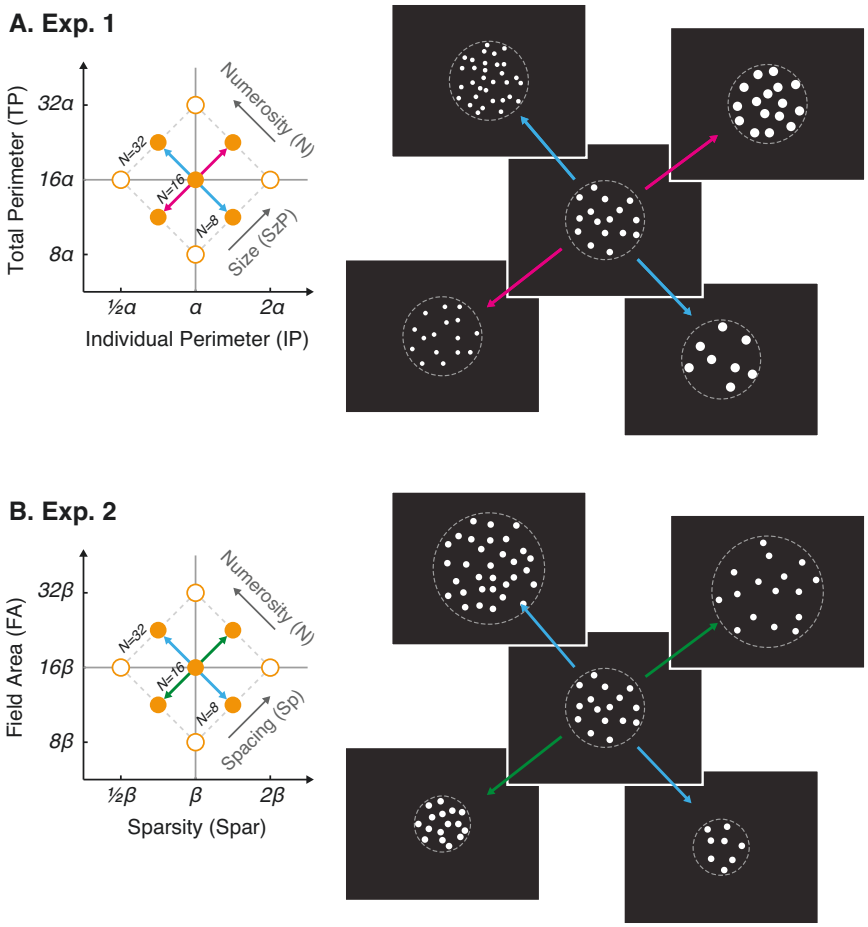

Figure 1*. Stimulus design and schematic illustrations of exemplar dot arrays in the two Experiments. A. In Experiment 1, participants were presented with dot arrays of different sizes (represented by red arrows) but constant numerosity and spacing, or they were presented with dot arrays of different numerosity (represented by blue arrows) but constant size and spacing, in alternating blocks. B. In Experiment 2, participants were presented with dot arrays of different spacing (represented by green arrows) but constant numerosity and size, or they were presented with dot arrays of different numerosity (represented by blue arrows) but constant size and spacing, in alternating blocks. Gray dotted circles surrounding the exemplar dot arrays represent invisible circular field within where dots were randomly drawn; these are for illustration purposes only and did not appear in the experiments. C. The three-dimensional parameter space defined by the logarithmic scales of N, Sz, and Sp as cardinal axes and the representation of other magnitude dimensions in this parameter space.*

N, *Sz*, and *Sp* together construct a three-dimensional parameter space, and other visual parameters of interest, such as *IP*, *TP*, *FA*, and *Spar*, can be derived from a linear combination of these three orthogonal dimensions. This parameter space also allows the construction of two other interpretable dimensions defined as a function of *Sz* and *Sp* (see DeWind et al., 2015). First, we define coverage as *TP*/*FA* (or equivalently *IP*/*Spar*), which translates into the equation log(*Coverage*) = ½ log(*Sz*) – ½ log(*Sp*). When *N* is constant, increasing *Sz* makes the dots become closer to each other proportionately to their increase in perimeter. Coverage is exactly the dimension quantifying this amount. Second, we defined apparent closeness as a dimension orthogonal to coverage, which is mathematically defined as log(*Apparent Closeness*) = ½ log(*Sz*) + ½ log(*Sp*). It captures the overall scaling of the stimulus and works as if one is zooming in and out into the image. **Table 1** lists mathematical definitions of all the dimensions described so far as well as the relations between those dimensions, and **Figure 1C** shows the graphical illustrations.

Table 1. Mathematical relations between various magnitude dimensions.

Dimension	As a function of n, r_d, r_f	Relationship with N, Sz, and Sp	Intuitive explanations
Individual perimeter (IP)	$2\pi r_d$	log(IP) = ½log(Sz) – ½log(N)	The perimeter of each dot.
Total perimeter (TP)	$n \times 2\pi r_d$	log(TP) = ½log(Sz) + ½log(N)	The cumulative perimeter of all the dots.
Field area (FA)	πr_f^2	log(FA) = ½log(Sp) + ½log(N)	The area of the invisible circle within which the dots are drawn.
Sparsity (Spar)	$\pi r_f^2/n$	log(Spar) = ½log(Sp) – ½log(N)	Inverse of density, which is defined by field area divided by the number of dots.
Individual area (IA)	πr_d^2	log(IA) = –log(4π) + log(Sz) – log(N)	The area of each dot.
Total area (TA)	$n \times \pi r_d^2$	log(TA) = –log(4π) + log(Sz)	The cumulative area of all the dots.

Coverage	$2n \times r_d / r_f^2$	log(Cov) = ½log(Sz) – ½log(Sp)	The amount measuring how much dots are closer to each other proportionately to their increase in perimeter.
Apparent Closeness	$2\pi^2 \times r_d \times r_f^2$	log(Closeness) = ½log(Sz) + ½log(Sp)	Overall scaling of the stimulus as if one is zooming in and out of the image.

Note: n = number; r_d = radius of individual dot; r_f = radius of the invisible circular field in which the dots are drawn; log(Sz) = log(TP) + log(IP); log(Sp) = log(*FA*) + log(*Spar*).

In this study, each of the three dimensions (*N*, *Sz*, and *Sp*) varied in three levels in the ratio of 1:2:4. Thus, on a $\log_2$-scale each level differed by 1 from its subsequent level. Specifically:

The three levels of *N* were 8, 16, and 32 dots (ratio of 1:2:4), and their $\log_2(N)$ values were 3, 4, and 5 (increment by 1).

The three levels of *Sz* were 1.55×10^5, 3.10×10^5, and 6.20×10^5 (ratio of 1:2:4). These values were computed from 2^log(*Sz*) = 2^(log(*TP*) + log(*IP*) = 2^(log(*IP* × *N*)+log(*IP*)), where *N*=16 and IP ranged from a minimum of 31.4 pixels (encompassing 0.21° visual angle [10 pixels] in diameter) to a maximum of 62.8 pixels (encompassing 0.42° visual angle [20 pixels] in diameter). Accordingly, the three levels of $\log_2(Sz)$ were 13.9, 14.9, and 15.9 (increment by 1).

The three levels of *Sp* were 0.809×10^8, 1.62×10^8, and 3.24×10^8 (ratio of 1:2:4). These values were computed from 2^log(*Sp*) = 2^(log(*FA*) + log(*Spar*)) = 2^(log(*FA*)+log(*FA* / *N*)), where *N*=16 and FA ranged from a minimum of 35,968 pixel^2 (encompassing 4.45° visual angle [214 pixels] in diameter) to a maximum of 71,631 pixel^2 (encompassing 6.27° visual angle [302 pixels] in diameter). Accordingly, the three levels of $\log_2(Sp)$ were 26.3, 27.3, and 28.3 (increment by 1).

Note that in our previous study (Park et al., 2016) two different definitions of size were used: size in area (*SzA*) in Experiment 1 and size in perimeter (*SzP*) in Experiment 2. In the present study, *Sz* refers to *SzP*. This is a more conservative treatment from the standpoint of the *N* and *Sp* dimensions because while *SzP* varies in the ratio of 1:2:4, *SzA* varies in the ratio of 1:4:16. In other words, while there was a four-fold difference in numerosity across stimuli in this experiment, the difference in total area or individual area was sixteen-folds.

Task and Procedure

Each participant completed four blocks, each of which consisted of the presentation of 400 unique arrays. In each block, participants viewed centrally presented dot arrays, each of 200 ms duration, with stimulus onset asynchronies (SOAs) varying between 700 and 900 ms. A fixation dot appeared at the center of the screen between stimuli. To ensure that participants paid attention to the stimuli, participants were asked to engage in an oddball detection task. Specifically, the participants were instructed to press a button when the dot array was displayed in red (5% of trials). Participants used their left index finger to respond in two blocks

and their right index finger to respond for the other two. The finger order was counterbalanced across participants. The red oddball trials were not included in the analyses.

In Experiment 1, participants (N=31) received experimental blocks in which dot arrays were modulated exclusively either in *N* or *Sz* (see **Fig. 1A** and **Table 2**)**:**

In the blocks where *N* was modulated, dot arrays were randomly drawn that differed only in *N* ($\log_2(N)$=2, 3, or 5; 8, 16, or 32 dots), but with constant *Sz* ($\log_2(Sz)$=14.9) and constant *Sp* ($\log_2(Sp)$=27.3).

In the blocks where *Sz* was modulated, dot arrays were randomly drawn that differed only in *Sz* ($\log_2(Sz)$=13.9, 14.9, or 15.9), but with constant *N* ($\log_2(N)$=4; 16 dots) and constant *Sp* ($\log_2(Sp)$=27.3).

The two different types of blocks alternated (e.g., *N*-varying block, *Sz*-varying block, *N*-varying block, *Sz*-varying block) while the type of the starting block was randomized across participants.

In Experiment 2, participants (N=33) received experimental blocks in which dot arrays were modulated exclusively either in *N* or *Sp* (see **Fig. 1B** and **Table 2**):

In the blocks where *N* was modulated, dot arrays were randomly drawn that differed only in *N* ($\log_2(N)$=2, 3, or 5; 8, 16, or 32 dots), but with constant *Sz* ($\log_2(Sz)$=14.9) and constant *Sp* ($\log_2(Sp)$=27.3).

In the blocks where *Sp* was modulated, dot arrays were randomly drawn that differed only in *Sp* ($\log_2(Sp)$=27.3, 26.3, or 28.3), but with constant *N* ($\log_2(N)$=4; 16 dots) and constant *Sz* ($\log_2(Sz)$=14.9).

The two different kinds of blocks alternated (e.g., *N*-varying block, *Sz*-varying block, *N*-varying block, *Sz*-varying block) while the identity of the starting block was randomized across participants.

Table 2. Numerical values of magnitude dimensions as a function of exclusive changes in N, Sz, and Sp.

	Exclusive change in N			Exclusive change in Sz			Exclusive change in Sp		
$\log_2$(N)	3	4	5	4	4	4	4	4	4
$\log_2$(Sz)	14.9	14.9	14.9	13.9	14.9	15.9	14.9	14.9	14.9
$\log_2$(Sp)	27.3	27.3	27.3	27.3	27.3	27.3	26.3	27.3	28.3
IP	20π = 62.83	14π = 43.98	10π = 31.42	10π = 31.42	14π = 43.98	20π = 62.83	14π = 43.98	14π = 43.98	14π = 43.98
TP	$20\pi \cdot 8$ = 5.03e2	$14\pi \cdot 16$ = 7.04e2	$10\pi \cdot 32$ =1.01e3	$10\pi \cdot 16$ =5.03e2	$14\pi \cdot 16$ = 7.04e2	$20\pi \cdot 16$ = 1.01e3	$14\pi \cdot 16$ = 7.04e2	$14\pi \cdot 16$ = 7.04e2	$14\pi \cdot 16$ = 7.04e2
FA	$107^2\pi$ = 3.60e4	$127^2\pi$ =5.07e4	$151^2\pi$ =7.16e4	$127^2\pi$ =5.07e4	$127^2\pi$ =5.07e4	$127^2\pi$ =5.07e4	$107^2\pi$ = 3.60e4	$127^2\pi$ =5.07e4	$151^2\pi$ =7.16e4
Spar	$107^2\pi/8$ = 4.50e3	$127^2\pi/16$ = 3.17e3	$151^2\pi/32$ = 2.24e3	$127^2\pi/16$ = 3.17e3	$127^2\pi/16$ = 3.17e3	$127^2\pi/16$ = 3.17e3	$151^2\pi/32$ = 2.24e3	$127^2\pi/16$ = 3.17e3	$107^2\pi/8$ = 4.50e3
Coverage	$20 \cdot 8/107^2$	$14 \cdot 16/127^2$	$10 \cdot 32/151^2$	$10 \cdot 16/127^2$	$14 \cdot 16/127^2$	$20 \cdot 16/127^2$	$14 \cdot 16/107^2$	$14 \cdot 16/127^2$	$14 \cdot 16/151^2$

	=0.0140	=0.0139	=0.0140	=0.0099	=0.0139	=0.0198	=0.0196	=0.0139	=0.0098
Apparent Closeness	$\pi^2 \cdot 20 \cdot 10 7^2$ =2.26e6	$\pi^2 \cdot 14 \cdot 12 7^2$ =2.23e6	$\pi^2 \cdot 10 \cdot 15 1^2$ =2.25e6	$\pi^2 \cdot 10 \cdot 10 7^2$ =1.13e6	$\pi^2 \cdot 14 \cdot 12 7^2$ =2.23e6	$\pi^2 \cdot 20 \cdot 15 1^2$ =4.50e6	$\pi^2 \cdot 14 \cdot 10 7^2$ =1.58e6	$\pi^2 \cdot 14 \cdot 12 7^2$ =2.23e6	$\pi^2 \cdot 14 \cdot 15 1^2$ =3.15e6

Note: Values that remain constant across the changes are grey shaded. Small numerical differences when identical values are expected is due to rounding errors, as pixels are integers.

Electrophysiological Recording and Preprocessing

Electrophysiological recording and analysis were similar to our previous study (Park et al., 2016). The electroencephalogram (EEG) was recorded continuously from 64 channels mounted in a customized, elastic electrocap (Duke64Waveguard cap layout, Advanced Neuro Technology, the Netherlands) at a sampling rate of 512 Hz, a low-pass filter with a high-frequency cutoff at 138 Hz, and an online averaged reference. Our custom cap is designed such that the electrodes are equally spaced across the cap, while also providing extended coverage of the head from just above the eyebrows anteriorly to below the inion posteriorly (Woldorff et al., 2002). The ground electrode was placed on the left collarbone, and the electrooculogram (EOG) was monitored with electrodes below the left eye, and slightly lateral to each external canthus. Electrode impedances were maintained below 10 kΩ for the EOG channels and 5 kΩ for all other channels.

The continuous EEG data were first band-pass filtered offline from 0.01–100 Hz in asalabTM (www.ant-neuro.com). The rest of the event-related potential (ERP) analyses were conducted using the EEGLAB software package (Delorme & Makeig, 2004) and the associated ERPLAB toolbox (Lopez-Calderon & Luck, 2014) in Matlab R2013a. EEG epochs time-locked to the onset of the dot arrays were extracted from 200 ms before to 600 ms after stimulus onset, to which a prestimulus (-200 to 0 ms) baseline correction was applied. A step-like artifact rejection tool in EEGLAB (moving window width = 400 ms; window step = 20 ms; threshold = 30 μV) was used to identify any epochs contaminated by eye movements or blinks, which were then removed prior to selective averaging. The average artifact rejection rate across participants was 25.1%. Finally, the selectively-averaged ERPs were low-pass filtered at 30 Hz in each participant, after which statistical analyses and grand averaging of the ERPs across participants were performed for each experiment.

Electrophysiological analysis

In our previous study, a regression-based analysis resulted in modulations of ERPs as a function of the three orthogonal regressors—numerosity, size, and spacing—primarily in the medial (Oz') and bilateral (PO7i and PO8i) occipital channels (see Fig. 3 and 6B in Park et al., 2016). Both in that previous work and in this study, the locations of PO7i and PO8i in the montage layout were slightly (~0.14 radians) inferior to PO7 and PO8 in the standard 10–20 system, and the location of Oz' was equivalent to the location of Oz in the standard 10-20 system. Thus, we restricted our primary statistical analysis to those three channels of interest. Note

that the current results did indeed suggest that the peak effects of numerosity, size, and spacing are largely located on those three channels throughout the time course of a trial (see **Figures 2 & 3**).

The effects of magnitude dimension on the ERPs in each of the channels of interest were assessed using a cluster-based nonparametric method (Maris & Oostenveld, 2007). This approach first identified consecutive time points that show significant linear contrast effects of a magnitude dimension. Then, a maximum cluster-level test statistic (sum of the F-statistic within a cluster exceeding a height threshold of $p < .05$) was evaluated against the null distribution, which was generated from random permutations of the waveforms by randomly assigning condition category labels to each waveform. Considering that six tests (3 channels × 2 conditions) were performed in each Experiment, a Bonferroni correction was further applied so that contiguous time regions with α values (false positive rate of the cluster size) below 0.0083 (=0.05/6) were considered statistically significant.

Results

Color oddball detection performance

Across the two experiments, the hit rate for detection of the red oddball target was 97.1% with a median reaction time of 405 ms, indicating that the participants were alert and attentively engaged during the experiment.

Experiment 1: effects of size

We first examined the extent to which ERPs were modulated by the variation of size of a dot array while holding numerosity and spacing constant. As depicted in **Figure 2**, size modulated the ERPs in various latency points at sites Oz' and PO8i. The sensitivity to size first emerged around [49 84] ms in PO8i, and again around [135 271] ms in Oz' and [174 219] ms in PO8i. As a post hoc analysis, we computed Cohen's d for the mean ERP amplitude within a 20-ms bin centering the local peak of the contrast wave. These measures were d=0.840 (PO8i at 62 ms), 1.236 (Oz' at 189 ms), and 0.716 (PO8i at 189 ms), for the three time windows. These results demonstrate that early visual cortex is sensitive to changes in the size of a dot array when other magnitude properties were held constant.

In the other experimental block of this experiment, in which numerosity varied while size and spacing was held constant, the effects of varying numerosity were considerably more robust. The ERPs were sensitive to numerosity in three latency windows in Oz' ([57 119] ms, Cohen's d=-0.992 at 90 ms; [174 229] ms, d=0.847 at 207 ms; [313 367] ms, d=-0.908 at 332 ms), around [115 238] ms in PO7i (d=1.122 at 199 ms), and around [115 262] ms in PO8i (d=1.446 at 203 ms). These results replicate the previous finding where the effects of numerosity were observed to occur very early in the medial occipital channels and somewhat later in the bilateral occipital channels. In comparison to the case where size was exclusively varied, these results suggest that the effects of numerosity were indeed more robust than the effects of size.

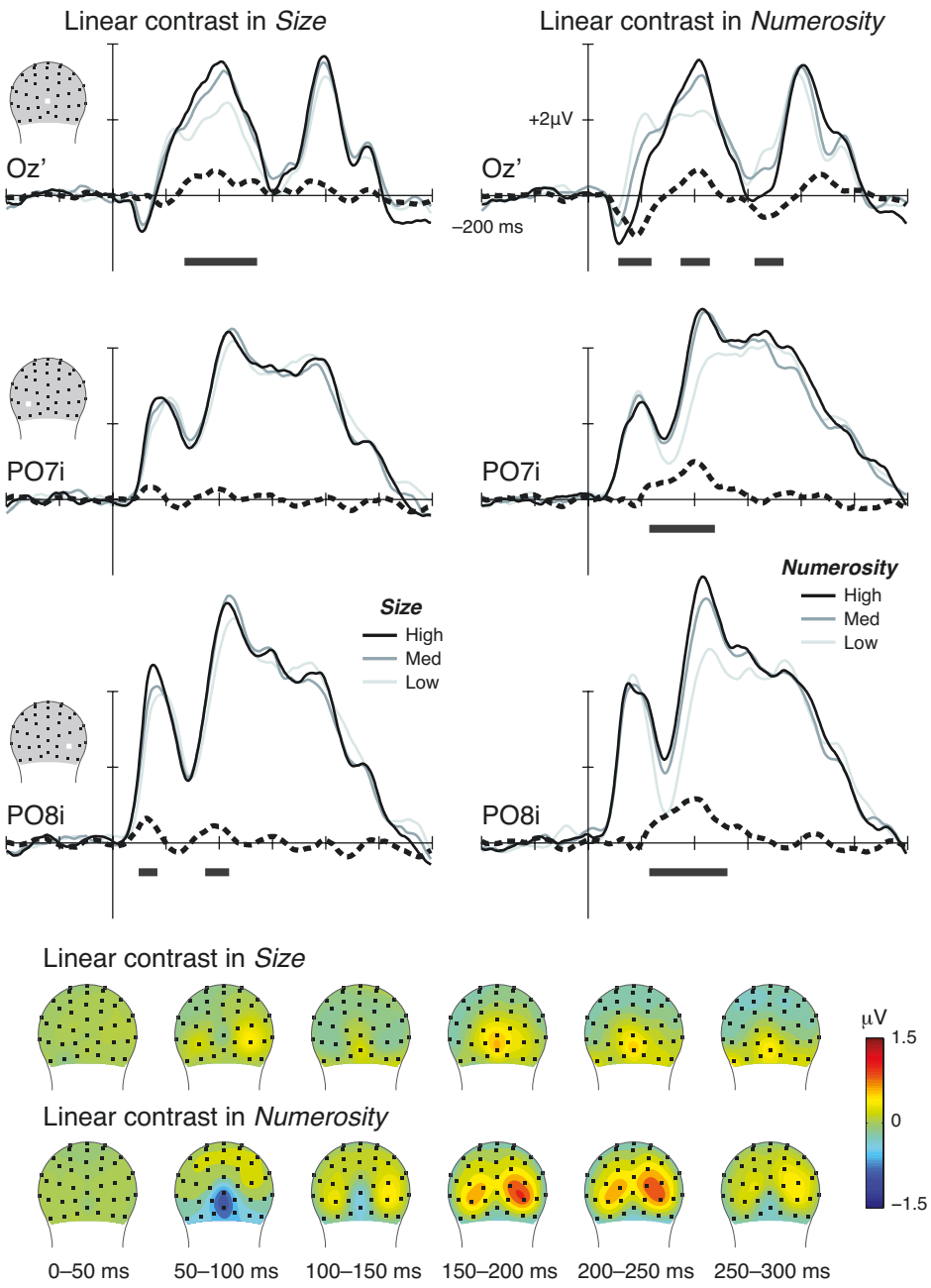

Figure 2. Brainwaves (top) and topographic maps (bottom) from Experiment 1. Different levels of size in Exp. 1 represent the manipulation along the size dimension while holding numerosity and spacing constant, as indicated by the red arrows in Figure 1A. Different levels of numerosity represent the manipulation along the numerosity dimension while holding size and spacing constant. Dotted line illustrates the linear contrast of the three waveforms. Gray bar indicates the latency window reaching Bonferroni-corrected statistical significance. The topographic map illustrates the posterior view of the grand-averaged ERPs for the linear contrasts. These maps show averaged ERPs within a 50-ms time window as indicated.

Experiment 2: effects of spacing

Similarly, we examined the extent to which ERPs were modulated by the variation of the spacing of a dot array while holding numerosity and size constant. Unlike the case of the size manipulation in Experiment 1 (*Section 3.2*), spacing had little effects on the ERPs, as depicted in **Figure 3**. The only significant effect was observed around [145 238] ms in PO8i (d=-0.841 at 203 ms).

On the other hand, as in Experiment 1, there was again a strong effect of numerosity when numerosity varied while other orthogonal dimensions were held constant: the ERPs were sensitive to numerosity in two latency windows in Oz' ([55 119] ms, d=-1.368 at 90 ms; [168 230] ms, d=0.899 at 201 ms), around [146 240] ms in PO7i (d=1.253 at 199 ms), and around [111 262] ms in PO8i (d=1.270 at 191 ms). Thus, again, the effects of numerosity were much more robust than the effects of spacing on the ERPs. Moreover, the pattern of those effects was virtually identical that of Experiment 1.

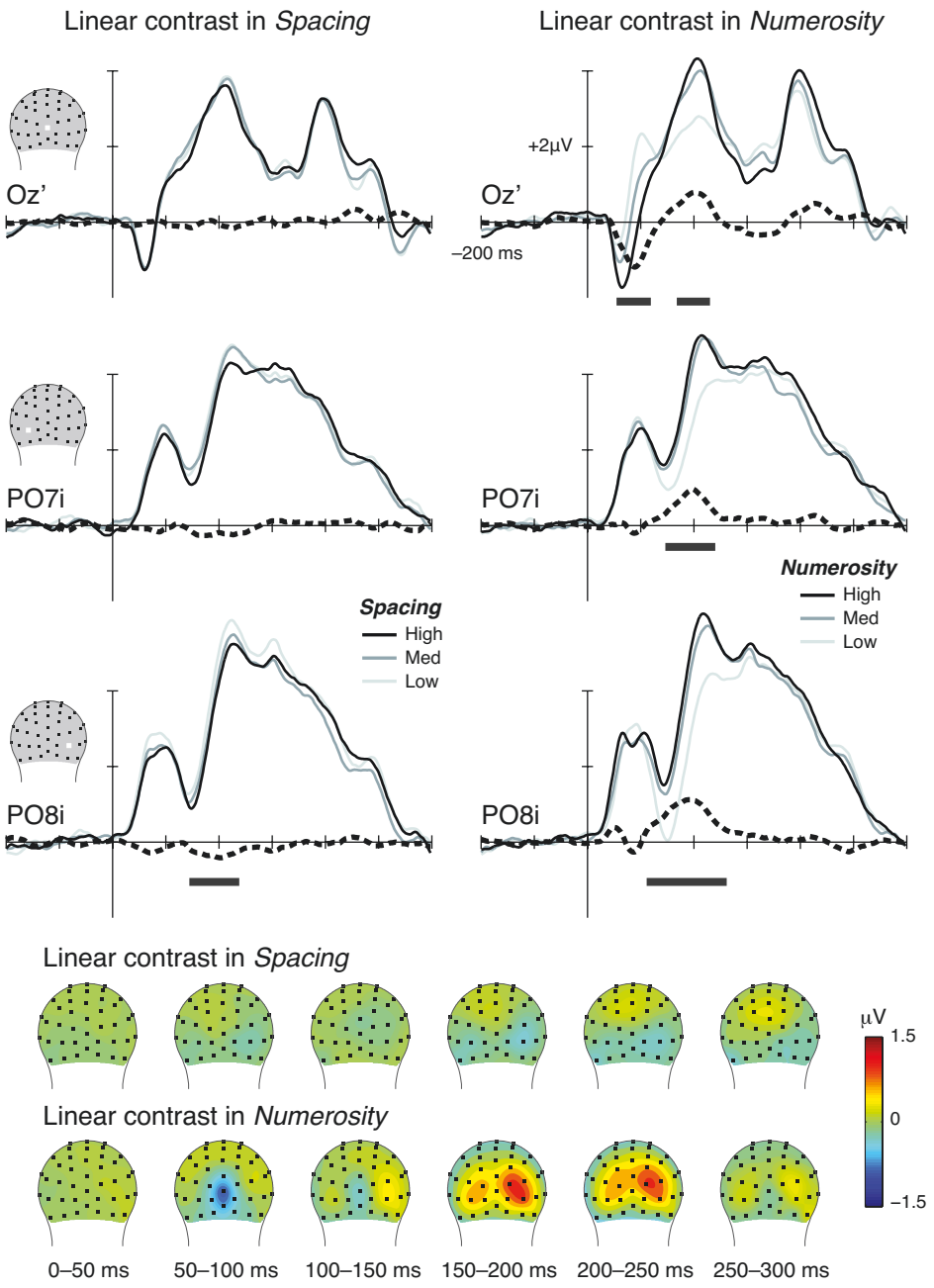

Figure 3. Brainwaves (top) and topographic maps (bottom) from Experiment 2. Different levels of spacing in Exp. 2 represent the manipulation along the spacing dimension while holding numerosity and size constant, as indicated by the green arrows in Figure 1A. Different levels of numerosity represent the manipulation along the numerosity dimension while holding size and spacing constant. Dotted line illustrates the linear contrast of the three waveforms. Gray bar indicates the latency window reaching Bonferroni-corrected statistical significance. The topographic map illustrates the posterior view of the grand-averaged ERPs for the linear contrasts. These maps show averaged ERPs within a 50-ms time window as indicated.

Discussion

In this study, participants viewed dot arrays that changed in one of two dimensions that were orthogonal to numerosity while performing a color oddball detection task. In Experiment 1, the size of the dots changed while the numerosity and the spacing of the array were held constant in one set of blocks, and the numerosity changed while the other two dimensions were held constant in another set of blocks. In Experiment 2, the spacing of the dots changed while the numerosity and size were held constant in one set of blocks, and again the numerosity changed while the other two dimensions were held constant in another set of blocks. Unlike our previous studies (Park et al., 2016; Fornaciai & Park, 2017; Fornaciai et al., 2017; DeWind et al., 2019), the current experimental design allowed us to directly assess the effects of continuous, non-numerical magnitudes (size and spacing) on the evoked neural activity while holding discrete, numerical magnitude constant.

Effects of numerosity, size, and spacing under exclusive manipulations

Overall, significant effects of size and spacing were identified in the ERPs. The effect of size was relatively strong very early in latency bilaterally and around [135 271] ms medially. Such a strong effect of size somewhat resembles our previous finding (Exp. 2 of Park et al., 2016) that demonstrated a medial occipital effect of total perimeter (see Fig. 8 of Park et al., 2016). This could mean that the visual cortex is quite sensitive to perimeter, when size was exclusively manipulated. However, it is important to consider that, in that previous work, no significant effect of size was observed when size was defined in terms of area (*SzA*) (in Exp. 1 of Park et al., 2016). Thus, an alternative explanation is that the effect of size when defined in terms of perimeter could be larger simply because there was a greater amount of difference in area than in perimeter over the course of the experiment, in that a 4-fold range of perimeter designed in the current study translates to a 16-fold range of area. Regardless of these explanations, the fact that the effect of size starts more bilaterally and then proceeds medially (which is opposite of a common expectation that information transfers from the most medial striate cortex to extrastriate cortices) is an interesting pattern that requires future investigations.

While exclusive manipulations of size and spacing resulted in ERP modulations, the manipulation of numerosity while holding those orthogonally-defined, non-numerical dimensions constant resulted in a substantially greater modulation of the ERPs, especially starting early in the time course at the medial occipital channel (prior to 100 ms). The effect of numerosity in the current study resembles the pattern observed in our previous study in terms of both latency and site (Park et al., 2016). Moreover, our other previous studies show that early ERP sensitivity to numerosity at or before 100 ms reflects initial sensory representations from feedforward activity arising from V2 and V3 (Fornaciai et al., 2017; Fornaciai & Park, 2018). The current findings, thus, again support the idea that numerosity is one of the most salient cues processed early in the visual stream.

Our findings, nevertheless, elicits two important questions. First, what is the neurophysiological explanation for the robust effect of numerosity (in this and

aforementioned previous studies) and the negligible effects of non-numerical magnitude in early visual cortex (at least when all the magnitude dimensions are varied simultaneously within a stimulus set)? Second, how do we explain the significant effects of non-numerical cues when they are manipulated in isolation, as newly tested in this study? Below we provide explanations for these two questions using the neurocomputational framework of neuronal normalization.

Neuronal normalization across space may explain the absence of size and spacing effects

A classic computational model of numerosity processing assumes multiple layers of neuronal representation of an item array (Dehaene & Changeux, 1993): from an *input* layer, where stimuli are directly registered on to a representative model-version of a "retina," to an *intermediate* layer, where item location and size are normalized, and then to a *summation* layer that represents the sum of all outputs from the intermediate layer (see also Verguts & Fias, 2004). The intermediate layer is most critical for our discussion here. This layer contains a two-dimensional sheet of neuronal clusters that code the location of objects normalized for their size. Each cluster of this sheet functions as a difference-of-Gaussian filter, and one dimension of this sheet represents the positions for the center of the filter, while the other dimension represents the width for the filters. Lateral inhibition between clusters across different filter widths in the same or neighboring locations imposing a winner-take-all competition allows only one cluster whose filter width best matches with the size of the raw input to be represented in this sheet of clusters. This procedure is thus called object normalization.

While Dehaene & Changeux's model provided a proof of concept, it may be overly simplistic. For instance, in the intermediate layer a single large dot on a background (thus, a greater luminance difference) would be coded exactly the same as a single small dot, which is implausible as the visual cortical activity generally increases monotonically as a function of local contrast until it saturates. We propose that a more neurophysiologically plausible explanation for numerosity encoding can be derived from the principle of neuronal normalization.

Neuronal normalization is a canonical computational principle that has been proposed to operate in the brain more generally (for review, see Carandini & Heeger, 2012). Under this principle, the response of a neuronal pool is a result of the pool's driving input (numerator) divided by the summation of the response from a larger pool elsewhere in the visual field (the normalization factor in the denominator). A simple analogy would be the phenomenon of surround suppression widely observed in V1, whereby a neuron's responses to stimuli inside the receptive field are suppressed by additional stimuli surrounding that field (Cavanaugh, Bair, & Movshon, 2002a, 2002b; Webb, Dhruv, Solomon, Tailby, & Lennie, 2005), which is thought to be achieved by lateral inhibition (Carandini & Heeger, 2012). More generally, a growing body of literature suggests that visual information over wide regions of the visual space gets integrated to form a global, coherent visual percept, even in the earliest processing stages of the visual stream (for review see Allman, Miezin, & McGuinness, 1985; Kapadia, Westheimer, & Gilbert, 1999; Sceniak,

Ringach, Hawken, & Shapley, 1999), and divisive normalization may play a key role in that process (Carandini & Heeger, 2012).

By this account, the absence of size and spacing effects, especially in early visual areas, can be explained by the cancellation of the neuronal pools' driving input arising from local contrast by the normalization factor. For instance, imagine a neuronal pool with its receptive field on a particular area of a visually-presented dot array. An array with greater size may result in a greater driving input for this particular neuronal pool encoding the local contrast of the specific visual space. However, because all the other dots also increase in size, there will be greater local contrasts elsewhere in the visual space, which in turn could increase the normalization factor and cancel out the increased stimulus driving effect for this neuronal pool. As another example, a dot array with greater spacing will elicit responses from neuronal pools that are responsive to stimuli at greater eccentricity (because local contrasts exist in those regions of the visual space), thereby increasing the stimulus drive; however, increased sparsity elsewhere in the visual space may result in decreased neuronal responses because local contrasts will be weaker on average, thereby again cancelling out the effect of increased stimulus drive.

According to this neurophysiological explanation, then, the absence of size and spacing effects may be a byproduct of a neuronal property in the visual cortex representing an almost complete normalization (i.e., increase/decrease in the neuronal pool's driving input approximately matching the increase/decrease in the normalization factor). At the same time, a change in the neuronal pool's driving input may be associated with a non-linear change in the normalization factor. For example, when the dots of an array are very large or very densely populated, the normalization factor for that metric may saturate (i.e., denominator not as large as numerator) resulting in an incomplete normalization, which in turn would result in nontrivial effect of size or spacing. In contrast, the robust neuronal modulation by numerosity may represent the case in which the increase/decrease in the driving input differs substantially from the increase/decrease in the normalization factor. By this logic, the encoding of numerosity is a result of the lack of neuronal normalization for local contrasts.

It should be noted that our proposed model assumes a local neural computation at the level of initial sensory representations in early visual cortex, likely at the feedforward stage within V1, V2, and V3, as demonstrated in a recent study (Fornaciai & Park, 2018). Therefore, this mechanism is fundamentally different from other claims that numerosity is extracted via other non-numerical continuous magnitudes (Gebuis et al., 2016; Leibovich & Henik, 2013), which lack some precise mechanistic explanations (e.g., see criticisms in Open Peer Commentary in Leibovich et al., 2017). Note also that our idea is distinct from a model based on a deep network in which numerosity and other non-numerical dimensions (such as total area) emerge as a statistical property of an image (Stoianov & Zorzi, 2012).

Finally, we note that our stimulus design and ERP results are not discussed in terms of spatial frequency, a well-established currency of the visual system since Hubel and Wiesel (1959, 1968). In a set of recent studies, Adriano, Girelli, and Rinaldi (2021a; 2021b) used a clever manipulation to equalize the overall structure of spatial frequency content and luminance between two images of dot arrays. They demonstrated that participants' judgment of numerosity was not affected when spatial frequency was equated indicating that numerosity perception is independent of the overall spatial frequency structure of an image. Spatial frequency cannot be quantified by a single measure, and therefore does not directly map onto our stimulus parameter framework, originally developed by DeWind and colleagues (2015). Spatial frequency and the dimensional metrics used in the current study provide different levels of explanations, and we suggest it would be valuable for future studies to investigate how those two levels of explanations can be reconciled. Our proposed neurophysiological explanation for magnitude processing could provide an important starting point to this investigation. For instance, it may be that a neuronal pool's driving input quantified based on the local contrast of a specific visual field could be translated into a metric of spatial frequency. At the same time, the normalization factor computed from the rest of the visual field may also be contextualized in terms of measures of spatial frequency.

Neuronal normalization across time may explain context-dependent modulations

A novel finding in the current study is the non-negligible effect of size and spacing when they are manipulated exclusively in isolation. This phenomenon may best be explained by the idea that the sensory cortex becomes sensitive to statistical regularities in the stimuli over time (Barlow, 2001), which again reflects the principle of neuronal normalization. Perceptual adaptation, in which the sensory system re-distributes the representation of the stimulus, is another example of neuronal normalization in effect. In this case, the driving input at a certain time point is normalized by the summation of the neural response working over a period of recent time. This, in the end, alters neuronal sensitivity to match prevailing stimulus regularities (Kohn, 2007; Wark, Lundstrom, & Fairhall, 2007).

Light adaptation occurring at the level of retina and contrast adaptation occurring in low-level visual cortex are two well-known examples of visual adaptation. At the computational level, these visual adaptations are explained by a horizontal shift in the response function of a neuronal pool (Kohn, 2007; Carandini & Heeger, 2012). The same computational mechanism may have played a role in the current study. For instance, when size (or spacing) was exclusively manipulated within an experimental block, the response function of the visual system may have shifted so that the sensory system was recalibrated to detect even subtle changes in that dimension. As described above, this may be achieved through altered saturation point for a normalization factor for that metric. In contrary, such an increased sensitivity to size (or spacing) would be overridden when all the other dimensions vary simultaneously, as shown in our previous work (Park et al., 2016; Fornaciai & Park, 2017; Fornaciai et al., 2017; Fornaciai & Park, 2018; DeWind et al., 2019). This is because the size (or spacing) is no longer the only prevailing stimulus

regularity and because the visual system becomes sensitive also to changes in other dimensions over time. Under this scenario, exclusive changes in numerosity while holding constant the mathematically-defined orthogonal dimensions (i.e., size and spacing) may also have resulted in a shift in the response function to better match variations in numerosity. Nevertheless, our data show that the visual system's sensitivity to numerosity is so strong that further increase in its sensitivity due to adaptation makes little difference in the overall neural activity patterns. Thus, these results, again, suggest that numerosity may be the most prominent cue processed by the visual stream.

Summary

We assessed the sensitivity of ERP measures of early visual sensory processing activity to numerical and non-numerical (size and spacing) magnitudes by selectively varying one magnitude dimension at a time. The results demonstrate that the visual cortex is sensitive to exclusive manipulations in size and spacing that are orthogonal to the dimension of numerosity, although it is much more sensitive to exclusive manipulations in numerosity. Together with previous work, these findings suggest that sensitivity to magnitude dimensions in early visual cortex is context dependent but is most responsive to numerosity when multiple magnitude dimensions of the stimulus vary simultaneously. We propose that the encoding of numerical magnitude information may be best explained by (the absence of) neuronal normalization in effect across visual space. Context-dependent modulation of neural activities may also be explained by neuronal normalization, but in this case working across time, as the visual system becomes sensitive to the regularities of the presented stimuli.

Acknowledgements

We thank Pawan Mathew, Chandra Swanson, and Rachel Roberts for their assistance in data collection. We appreciate Drs. David Huber and Michele Fornaciai for helpful discussions. This work was supported by the National Science Foundation CAREER Award BCS1654089 to J.P.

References

Adriano, A., Girelli, L., & Rinaldi, L. (2021). Non-symbolic numerosity encoding escapes spatial frequency equalization. Psychological Research. https://doi.org/10.1007/s00426-020-01458-2

Adriano, A., Girelli, L., & Rinaldi, L. (2021). The ratio effect in visual numerosity comparisons is preserved despite spatial frequency equalisation. Vision Research. https://doi.org/10.1016/j.visres.2021.01.011

Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. *Annual review of neuroscience, 8*(1), 407-430. https://doi.org/10.1146/annurev.neuro.8.1.407

Barlow, H. (2001). The exploitation of regularities in the environment by the brain. *Behavioral and Brain Sciences*, *24*(4), 602–607. https://doi.org/10.1017/S0140525X01000024

Burr, D., & Ross, J. (2008). A Visual Sense of Number. *Current Biology*, *18*(6), 425–428. https://doi.org/10.1016/j.cub.2008.02.052

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*(1), 51–62. https://doi.org/10.1038/nrn3136

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002a). Nature and Interaction of Signals From the Receptive Field Center and Surround in Macaque V1 Neurons. *Journal of Neurophysiology*, *88*(5), 2530–2546. https://doi.org/10.1152/jn.00692.2001

Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002b). Selectivity and Spatial Distribution of Signals From the Receptive Field Surround in Macaque V1 Neurons. *Journal of Neurophysiology*, *88*(5), 2547–2556. https://doi.org/10.1152/jn.00693.2001

Clark, V. P., Fan, S., & Hillyard, S. A. (1994). Identification of early visual evoked potential generators by retinotopic and topographic analyses. *Human Brain Mapping*, *2*(3), 170–187. https://doi.org/10.1002/hbm.460020306

Dehaene, S. (1996). The Organization of Brain Activations in Number Comparison: Event-Related Potentials and the Additive-Factors Method. *Journal of Cognitive Neuroscience*, *8*(1), 47–68. https://doi.org/10.1162/jocn.1996.8.1.47

Dehaene, S., & Changeux, J. P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*. https://doi.org/10.1162/jocn.1993.5.4.390

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal*

of Neuroscience Methods, *134*(1), 9–21. https://doi.org/10.1016/j.jneumeth.2003.10.009

DeWind, N. K., Park, J., Woldorff, M. G., & Brannon, E. M. (2019). Numerical encoding in early visual cortex. *Cortex*, *114*, 76–89. https://doi.org/10.1016/j.cortex.2018.03.027

Di Russo, F., Martínez, A., Sereno, M. I., Pitzalis, S., & Hillyard, S. A. (2002). Cortical sources of the early components of the visual evoked potential. *Human Brain Mapping*, *15*(2), 95–111. https://doi.org/10.1002/hbm.10010

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*. https://doi.org/10.1016/j.cub.2008.07.053

Fornaciai, M., Brannon, E. M., Woldorff, M. G., & Park, J. (2017). Numerosity processing in early visual cortex. *NeuroImage*, *157*, 429–438. https://doi.org/10.1016/j.neuroimage.2017.05.069

Fornaciai, M., & Park, J. (2017). Distinct Neural Signatures for Very Small and Very Large Numerosities. *Frontiers in Human Neuroscience*, *11*. https://doi.org/10.3389/fnhum.2017.00021

Fornaciai, M., & Park, J. (2018). Early Numerosity Encoding in Visual Cortex Is Not Sufficient for the Representation of Numerical Magnitude. *Journal of Cognitive Neuroscience*, *30*(12), 1788–1802. https://doi.org/10.1162/jocn_a_01320

Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than approximate number system underlies numerosity processing: A critical review. *Acta Psychologica*, *171*, 17–35. https://doi.org/10.1016/j.actpsy.2016.09.003

Gebuis, T., & Reynvoet, B. (2013). The neural mechanisms underlying passive and active processing of numerosity. *NeuroImage*, *70*, 301–307. https://doi.org/10.1016/j.neuroimage.2012.12.048

Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. The Journal of Physiology, 195(1), 215–243. https://doi.org/10.1113/jphysiol.1968.sp008455

Hubel, D. H., & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. The Journal of Physiology, 148(3), 574–591. https://doi.org/10.1113/jphysiol.1959.sp006308

Hyde, D. C., & Spelke, E. S. (2012). Spatiotemporal dynamics of processing nonsymbolic number: An event-related potential source localization study. *Human Brain Mapping*, *33*(9), 2189–2203. https://doi.org/10.1002/hbm.21352

Jeffreys, D. A., & Axford, J. G. (1972). Source locations of pattern-specific components of human visual evoked potentials. I. Component of striate cortical origin. *Experimental Brain Research*, *16*(1). https://doi.org/10.1007/BF00233371

Johannes, S., Münte, T. F., Heinze, H. J., & Mangun, G. R. (1995). Luminance and spatial attention effects on early visual processing. *Cognitive Brain Research*, *2*(3), 189–205. https://doi.org/10.1016/0926-6410(95)90008-X

Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (1999). Dynamics of spatial summation in primary visual cortex of alert monkeys. *Proceedings of the National Academy of Sciences*, *96*(21), 12073–12078. https://doi.org/10.1073/pnas.96.21.12073

Kohn, A. (2007). Visual Adaptation: Physiology, Mechanisms, and Functional Benefits. *Journal of Neurophysiology*, *97*(5), 3155–3164. https://doi.org/10.1152/jn.00086.2007

Leibovich, T., & Henik, A. (2013). Magnitude processing in non-symbolic stimuli. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2013.00375

Lopez-Calderon, J., & Luck, S. J. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, *8*. https://doi.org/10.3389/fnhum.2014.00213

Maris, E., & Oostenveld, R. (2007). Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, *164*(1), 177–190. https://doi.org/10.1016/j.jneumeth.2007.03.024

Normann, R. A., & Perlman, I. (1979). The effects of background illumination on the photoresponses of red and green cones. *The Journal of Physiology*, *286*(1), 491–507. https://doi.org/10.1113/jphysiol.1979.sp012633

Park, J., Dewind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and Direct Encoding of Numerosity in the Visual Stream. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhv017

Picton, T. ., Woods, D. ., & Proulx, G. . (1978). Human auditory sustained potentials. II. Stimulus relationships. *Electroencephalography and Clinical Neurophysiology*, *45*(2), 198–210. https://doi.org/10.1016/0013-4694(78)90004-4

Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience*, *2*(8), 733–739. https://doi.org/10.1038/11197

Stoianov, I., & Zorzi, M. (2012). Emergence of a "visual number sense" in hierarchical generative models. Nature Neuroscience, 15(2), 194–196. https://doi.org/10.1038/nn.2996

Verguts, T., & Fias, W. (2004). Representation of Number in Animals and Humans: A Neural Model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504. https://doi.org/10.1162/0898929042568497

Wark, B., Lundstrom, B. N., & Fairhall, A. (2007). Sensory adaptation. *Current Opinion in Neurobiology*, *17*(4), 423–429. https://doi.org/10.1016/j.conb.2007.07.001

Webb, B. S. (2005). Early and Late Mechanisms of Surround Suppression in Striate Cortex of Macaque. *Journal of Neuroscience*, *25*(50), 11666–11675. https://doi.org/10.1523/JNEUROSCI.3414-05.2005

Woldorff, M. ., Liotti, M., Seabolt, M., Busse, L., Lancaster, J. ., & Fox, P. . (2002). The temporal dynamics of the effects in occipital cortex of visual-spatial selective attention. *Cognitive Brain Research*, *15*(1), 1–15. https://doi.org/10.1016/S0926-6410(02)00212-4