

1 **Early Numerosity Encoding in Visual Cortex is Not Sufficient for the Representation of**
2 **Numerical Magnitude**

3

4 Michele Fornaciai^{1,*} & Joonkoo Park^{1,2}

5 ¹ Department of Psychological and Brain Sciences, University of Massachusetts, Amherst MA, USA.

6 ² Commonwealth Honors College, University of Massachusetts, Amherst, MA, USA.

7

8 * Corresponding author.

9 Michele Fornaciai

10 (mfornaciai@umass.edu)

11 University of Massachusetts

12 Department of Psychological and Brain Sciences

13 135 Hicks Way, Amherst, 01003

14 USA

15

16

17

18

1 ABSTRACT

2 Recent studies have demonstrated that the numerosity of visually presented dot arrays are represented
3 in low-level visual cortex extremely early in latency. However, whether or not such an early neural
4 signature reflects the perceptual encoding of numerosity remains unknown. Alternatively, such a
5 signature may indicate the raw sensory representation of the dot-array stimulus prior to becoming the
6 perceived representation of numerosity. Here, we addressed this question by using the connectedness
7 illusion, whereby arrays with pairwise connected dots are perceived to be less numerous compared to
8 arrays containing isolated dots. Using EEG and fMRI, we measured neural responses to dot-array
9 stimuli comprising 16 or 32 dots, either isolated or pairwise connected. The effect of connectedness,
10 which reflects the segmentation of the visual stimulus into perceptual units, was observed in the neural
11 activity after 150 ms post stimulus onset and in area V3. In contrast, earlier neural activity before 100
12 ms and in area V2 was strictly modulated by numerosity regardless of connectedness, suggesting that
13 this early activity reflects the sensory representation of a dot array prior to perceptual segmentation.
14 Our findings thus demonstrate that the neural representation for numerosity in early visual cortex is
15 not sufficient for visual number perception, and suggest that the perceptual encoding of numerosity
16 occurs at or after the segmentation process that takes place later in area V3.

17

18

1 INTRODUCTION

2 Among the many properties of our rich visual environment, numerical magnitude is one of the crucial
3 dimensions needed to achieve a basic description of the external environment. A growing amount of
4 evidence suggests that humans (e.g. Gallistel & Gelman, 1992; Dehaene, 2011) and many other animal
5 species (e.g. Agrillo et al., 2008; Pepperberg, 2006; Piantadosi & Cantlon, 2017; Rugani et al., 2015)
6 are endowed with a primitive sense of number allowing a rapid estimation or comparison of
7 approximate numerical magnitudes. In this context, approximate numerical magnitude could indeed
8 be regarded as a fundamental perceptual dimension (Anobile et al., 2016b; Burr et al., 2018), processed
9 by a dedicated and largely format- and modality-independent mechanism (Arrighi et al., 2014; Anobile
10 et al., 2016a; Cicchini et al., 2016; but see Durgin, 2008, for an argument opposing to the existence of
11 a visual sense of number).

12
13 Despite the fundamental nature of numerosity processing, however, the key cortical mechanisms
14 allowing the extraction and representation of numerical magnitudes remain unclear. A plethora of
15 studies to date have implicated the parietal cortex as the core neural region involved in numerosity
16 perception. Specifically, neurons selectively tuned to numerosity have been highlighted in single-cell
17 recording studies (Nieder, 2011; but see Chen & Verguts, 2013, arguing against the existence of
18 number-selective responses). Likewise, tuning curves for numerosity and a topographic organization
19 of numerosity-selective responses have been reported in humans using functional magnetic resonance
20 imaging (fMRI) (Piazza et al., 2004; Harvey et al., 2013).

21
22 In contrast to these previous investigations, recent studies have increasingly highlighted the role of the
23 early visual cortex for numerosity processing. Particularly, a number of studies have now demonstrated
24 that numerical information is encoded even at extremely early latencies and across multiple stages
25 comprising early visual areas (Roggeman et al., 2011; Park et al., 2016; Fornaciai & Park, 2017b;
26 Fornaciai et al., 2017). For instance, previous event-related potential (ERP) studies using dot arrays

demonstrate that numerosity sensitive neural activity emerges as early as 75 ms after stimulus onset (Park et al., 2016), originating from early visual areas such as V2 and V3, with possible contributions even from V1 (Fornaciai et al., 2017). These recent studies thus suggest that the representation of numerosity emerges earlier in the visual stream than previously thought; however, the specific nature of early cortical representation of numerosity remain unclear.

In particular, one crucial remaining question is whether such early activity represents a sufficient correlate of numerosity perception. In other words, does the early visual cortical activity (V1, V2, and V3 around 75 ms) represent the content of the subjective perceptual experience? To define the properties of a “sufficient” numerical representation, the characteristics of numerosity perception should be taken into account. For instance, it has been shown that numerosity perception is based on perceptual units defined by the topological properties of visual images (He et al., 2015), which under specific circumstances could induce distortions in the perceived magnitude of a visual stimulus (He et al., 2009; Franconeri et al., 2009; Fornaciai et al., 2016). To be considered a sufficient correlate, a neural signature of numerical representation should then reflect such topological properties of visual images, and the segmentation of elements into perceptual units. Alternatively, such an early neural signal may indicate the initial raw sensory representation of the visual stimulus prior to segmentation, thus still coded in a format only weakly related to the emerging percept.

In the present study, we used electroencephalogram (EEG) and functional magnetic resonance imaging (fMRI), in combination with multivariate pattern analysis techniques, to address the role of early neurophysiological signals and early visual areas in numerosity representation. To do so, we exploited the connectedness illusion, which has been previously shown to highlight the perceptual segmentation processes leading to the representation of numerosity. Under this illusion, the numerical magnitude of an array of dots connected pairwise by task-irrelevant lines is robustly underestimated compared to an array with the same number of isolated dots (He et al., 2009; Franconeri et al., 2009). On the one hand,

1 if a processing stage in the visual stream encodes the numerical magnitude of the perceptual units
2 contained in a visual stimulus, the neural activity should be maximally sensitive to the connectedness
3 of the dot arrays. In turn, the effect of connectedness would reflect the processes giving rise to the
4 perceptual units underlying the perception of numerical magnitude. On the other hand, if a processing
5 stage in the visual stream encodes the raw sensory representation of the stimulus, the neural activity
6 should be maximally sensitive to the numerosity of the dot arrays regardless of their connectedness.
7 Such an effect would reflect processing stages prior to the segmentation of visual images into
8 perceptual units, which is not sufficient to explain the properties of numerosity perception observed at
9 the phenomenological/behavioral level. Our results demonstrate that the effect of connectedness is
10 traceable at around 150 ms after stimulus onset and in visual area V3. However, earlier neural signals
11 at or before 100 ms and in visual area V2 show no effect of connectedness and are strictly modulated
12 by numerosity regardless of its connectedness, suggesting that such a processing stage encodes only
13 the raw sensory representation of the visual stimulus that will later be further processed turning the
14 elements into perceptual units.

17 METHODS

18 *Participants.*

19 Twenty participants (13 females, age ranging from 18 to 27 years) took part in the EEG experiment.
20 Another group of 20 participants (16 females, age ranging from 18 to 27 years) took part in the fMRI
21 experiment, although one participant was excluded from analysis due to excessive movements inside
22 the scanner, leaving a total of 19 participants in the fMRI experiment. All the participants signed a
23 written informed consent before participating in the experiment and were rewarded for their time with
24 monetary compensation (\$10/hour in the EEG experiment; \$15/hour in the fMRI experiment). All the
25 subjects included in the study were naïve 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.

In both the EEG and fMRI experiments, visual stimuli were generated using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007) on MATLAB (version r2013b; The Mathworks, Inc.). In the EEG experiment, stimuli were presented on a monitor screen running at 144 Hz, encompassing approximately 34×19 degrees of visual angle from the viewing distance of 90 cm, and with a resolution of 1920 x 1080 pixel. In the fMRI experiment, stimuli were presented on an MRI-compatible monitor screen positioned behind the scanner (60 Hz, 1920 x 1080 resolution), made visible to the participants by means of a mirror located inside the scanner, and encompassing 28×16 deg from a visual distance of about 137 cm. In both experiments, stimuli were arrays of black dots presented on a grey background, that could be either pairwise connected by straight lines (henceforth named connected dots) or presented as isolated dots (henceforth named isolated dots). The dot arrays were generated off-line, and systematically constructed in order to span approximately equal ranges on three orthogonal dimensions of numerosity (N), size (Sz) and spacing (Sp) (as in DeWind et al., 2015; Park et al., 2016). More specifically, while numerosity indicates the number of dots contained in the array, size reflects the combination of the individual dot area (IA) and the total area covered by the dots (TA), so that $\log(\text{Sz}) = \log(\text{TA}) + \log(\text{IA})$. Conceptually, greater size means greater total dot area and individual dot area together while holding numerosity constant. Spacing reflects the combination of the area upon which the dots were drawn (field area, FA) and their sparsity (Spar; i.e. the inverse of density), so that $\log(\text{Sp}) = \log(\text{FA}) + \log(\text{Spar})$. Conceptually, greater spacing means greater field area and sparsity together while holding numerosity constant. Such stimulus design was used in this study for a systematic sampling of numerical and non-numerical attributes of a dot array, even though the comparison between numerical and non-numerical effects on neural activity was not the primary focus of the current study (for those studies see Park et al., 2016; Fornaciai & Park, 2017a;

Park, 2017; DeWind et al., 2015). For more details about this stimuli construction scheme, see DeWind et al. (2015) and Park et al. (2016).

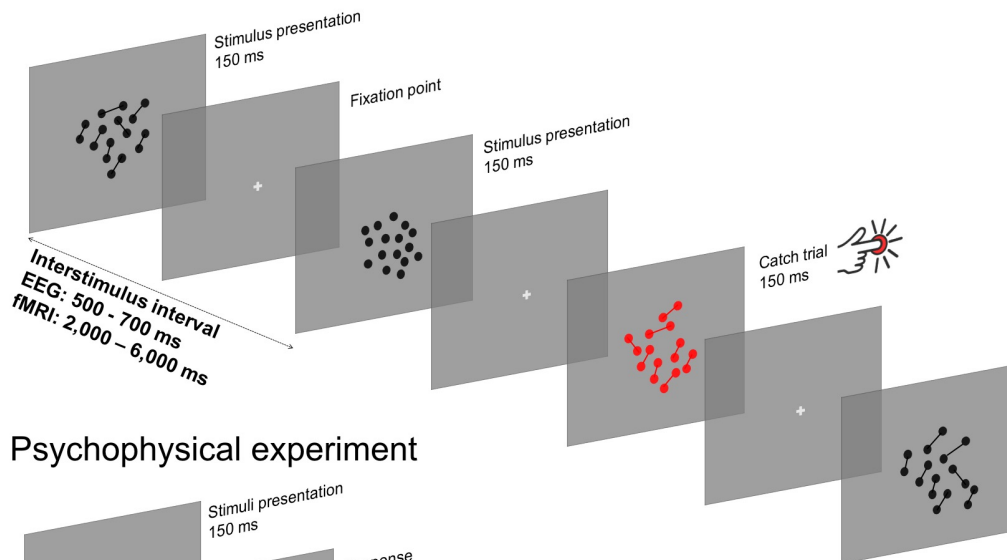
In the EEG and fMRI experiments (see *Task and Procedure*), the dot arrays were generated from one of eight combinations of parameters, comprising two levels of numerosity (16 and 32 dots) by two levels of size by two levels of spacing. In terms of the specific parameters, arrays of 16 dots were drawn with individual dot area of 0.067 or 0.093 deg² and orthogonally with field area of 5.25 or 7.44 deg². Arrays of 32 dots were drawn with individual dot area of 0.046 or 0.067 deg² and orthogonally with field area of 7.44 and 10.44 deg². Furthermore, dot arrays drawn from each of the eight parameter combinations were either pairwise connected or isolated (see Fig. 1), raising the total to 16 unique conditions. All the stimuli were generated using the same off-line routine, which iteratively calculates the coordinates of each dot in order to keep the items separated by at least 0.75 deg, while calculating the possible connections between pairs of dots in order to avoid lines crossing each other or other dots. Then, during the experiment, lines could be either presented or not on the exact same configuration of dots in order to display a connected or an isolated dot-array. This procedure ensured that in both cases the dot arrays were generated according to the same constraints, so keeping the low-level statistics of the stimuli (i.e. regularity) similar. The same set of stimuli was used for both the EEG and fMRI experiment.

In the psychophysical part of the study (which was identical in both the EEG and fMRI experiments), where participants judged which of two dot arrays contained more dots (see *Task and Procedure*), the dot arrays generated by the aforementioned parameters served as one of the two dot arrays (the “reference” stimulus) presented on each trial. The other dot array (the “probe” stimulus) varied widely in numerosity (8-28 dots or 16-56 dots, when paired with a 16-dot or 32-dot reference stimulus, respectively) and was generated to match the reference stimulus (1) in total dot area and in sparsity, (2) in individual dot area and in sparsity, (3) in total dot area and in field area, or (4) in individual dot

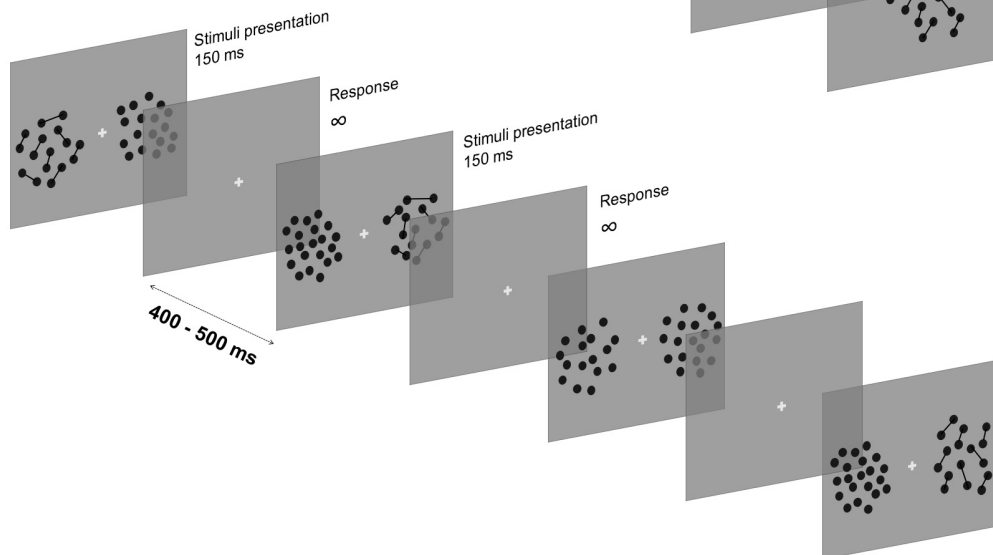
1 area and in field area, with equal probability. In the analysis of behavioral data, such different types of
 2 probe stimuli were collapsed together.

3

A EEG and fMRI experiment



B Psychophysical experiment



4

5 *FIGURE 1. Experimental procedures. (A) Sequence of stimuli in the EEG and fMRI experiments. Dot*
 6 *arrays were presented sequentially and displayed for 150 ms, with variable inter-stimulus intervals.*
 7 *At unpredictable times during each block, the stimulus could appear in red, and in such cases the*
 8 *participants were instructed to press a specific button on a joypad as quickly as they could. (B)*
 9 *Sequence of trials in the psychophysical part of the study. On each trial, two arrays of dots were*
 10 *presented on the screen: a reference stimulus, comprising 16 or 32 dots, either isolated or connected,*
 11 *and a probe stimulus, comprising a variable number of isolated dots. The stimuli were displayed for*

150 ms, and after the presentation the participants were instructed to indicate the stimulus containing more dots. After the response, the next trial started automatically after 400-500 ms. Stimuli are not depicted in scale.

Task and procedure.

EEG. The experiment took place in a quiet and dimly illuminated room. Participants sat in front of a monitor screen and were instructed to keep their gaze on a central fixation cross, which was briefly presented before the presentation of the first stimulus and during the inter-stimulus interval (ISI) between consecutive stimuli. Dot arrays were presented serially at the center of the screen, for a duration of 150 ms and with a variable ISI drawn from a uniform distribution between 500 and 700 ms. The experiment mostly required the participants to passively view the stream of stimuli. However, to ensure that participants kept their attention on the screen, they were required to perform a simple color oddball detection task. Namely, at unpredictable times during each block, the dot array was presented in red, and participants were instructed to press a button on a joypad as soon as possible when they detected a red stimulus. Each of the 16 unique stimulus types were presented 25 times in each block. Of the total of 400 stimuli presented in each block, 16 of them were presented as an oddball which were separated by a number of standard trials randomly chosen between 9 and 19. No other instructions were given to the participants, and nothing about number or magnitude was mentioned since participant recruitment until when participants completed this EEG part of the study. The average reaction time in the color oddball task was $412 \text{ ms} \pm 48 \text{ ms}$, and the average hit rate was $98\% \pm 2.3\%$. Each participant completed six blocks of 400 trials

fMRI. The procedure of the fMRI experiment was largely identical to the EEG experiment. While lying down within the scanner and watching the screen by means of a mirror, participants were instructed to keep their gaze on the central fixation point and to pay attention to the stimuli. The experiment

comprised six runs. During each run, each of the 16 unique stimulus types were presented 6 times. A total of 96 dot array stimuli with 48 null events (fixation cross that make the trial indistinguishable from an ISI) were presented sequentially (duration = 150 ms) with an ISI of 2,000 ms. Null events were limited to a maximum of two consecutive events. Thus, considering the null events, the effective ISI ranged between 2,000 to 6,000 ms. Participant performed a color-oddball detection task as in the EEG experiment, responding to red stimuli by pressing a button on a response device. In each run, 8 stimuli were presented as an oddball. Beside the oddball task, no other instructions were given to the participants, and nothing about number or magnitude was mentioned until the completion of the fMRI scan. The average reaction time in the color oddball task was $411 \text{ ms} \pm 52 \text{ ms}$, and the average hit rate was $98\% \pm 4.6\%$.

Behavioral. After either the EEG or the fMRI experiment, participants performed a numerosity discrimination task, comprising 4 blocks of 128 trials. The primary purpose of this second experiment was to confirm previous results demonstrating underestimation of numerosity when dots are pairwise connected (He et al., 2009; Franconeri et al., 2009; Fornaciai et al., 2016), but with our unique set of dot array stimuli. Participants were instructed to compare two dot arrays, one comprising 16 or 32 either pair-wise connected or isolated dots (reference stimulus), with a probe stimulus comprising a variable number of unconnected dots (randomly chosen on each trial, ranging from -0.30 logarithmic units (LU) to +0.24 LU of difference with respect to the reference numerosity; note that we used an asymmetric range with larger values at the lower hand according to the expected underestimation of connected stimuli). Probe and reference stimuli were presented simultaneously for 150 ms on the right and on the left of a central fixation point (with their position randomly chosen on each trial). Participants were asked to discriminate which one contained more dots by pressing the appropriate key on a standard keyboard. There were no time limitations for the response, and after a key press the subsequent trial was initiated with a random inter-trial interval ranging from 400 to 500 ms. All

different combinations of numerosity and connectedness were randomly intermixed within each block and were presented with equal probability.

Behavioral data analysis

Data from the numerosity discrimination experiment were analyzed to obtain measures of participants' accuracy and precision in the task. Cumulative Gaussian curves were fitted to the data according to the Maximum Likelihood method (Watson, 1979). The final estimate of the point of subjective equality (PSE), representing the accuracy of subjects' discrimination performance, was taken as the median of the best-fitting cumulative Gaussian curve to all the data of each participant in each condition. As a measure of precision, we used the just-noticeable difference (JND) taken as the difference in numerosity between chance level responses and 75% correct responses. Furthermore, to achieve a direct measure of the effect induced by connectedness, we calculated an index of perceived numerosity change, indicating how much the perceived numerosity of connected stimuli differs from the perceived numerosity of isolated stimuli. The perceived numerosity change index was calculated as follows:

$$\text{Perceived numerosity change} = ([\text{PSE}_{\text{connected}} - \text{PSE}_{\text{isolated}}] / \text{PSE}_{\text{isolated}}) \times 100; \quad (1)$$

where $\text{PSE}_{\text{connected}}$ refers to the PSE obtained with the connected dot arrays used as reference stimuli (either 16 or 32 dots), and $\text{PSE}_{\text{isolated}}$ refers to the PSE obtained with isolated dot arrays used as reference (either 16 and 32).

EEG electrophysiological recording and analysis

Data collection. EEG was continuously recorded for the entire duration of the EEG experiment (actiCHamp, Brain Products, GmbH), using a 64-channel, extended 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). In order to monitor artifacts due to blinks or eye

movements, the electro-oculogram (EOG) was monitored by means of electrodes positioned below the left eye and lateral to the left and right canthi. Channel impedances were usually kept below 15 k Ω . However, on some occasions, impedances up to 35 k Ω were tolerated.

Preprocessing. The EEG data were analyzed offline in Matlab (version R2013b), using the functions provided by the EEGLAB software package (Delorme & Makeig, 2004) and the ERPLAB toolbox (Lopez-Calderon & Luck, 2014). During the offline analysis, the EEG signals were high-pass filtered (0.1 Hz) and were re-referenced to the average value of all the 64 channels. The continuous EEG data were then segmented into epochs from -100 ms before to 400 ms after stimulus onset, with a baseline correction using the pre-stimulus interval. We then excluded trials containing eye-blink artifacts, by applying the step-like artifact rejection tool provided by EEGLAB. Trials were rejected whenever activity from the eye-channels exceeded a threshold equal to 30 μ V (in a time window spanning 400 ms, with 20 ms steps). This procedure led to an average ($\pm$ standard deviation) rejection rate of 7.97% $\pm$ 6.40%. Finally, the epochs were selectively averaged for each of the 16 stimulus types, followed by a low-pass filter (30 Hz) prior to computing the grand average.

Regions of interest. Based on previous results showing magnitude-sensitive ERP peaks (Park et al., 2016; Fornaciai & Park, 2017a), our primary analysis was focused on a specific channel of interest and a specific latency window. In particular, in line with those previous studies, we chose the occipital channel Oz, and a latency window spanning from 70 ms to 100 ms after stimulus onset, which encompasses the numerosity-sensitive peaks found in previous studies (75 ms in Park et al., 2016 and 88 ms in Fornaciai & Park, 2017a). Visual evoked potentials corresponding to the different class of stimuli, collapsed according to numerosity and connectedness (i.e. 16 isolated, 32 isolated, 16 connected, 32 connected) in that channel and time window were averaged. The distribution of responses across the group of participants were then tested using a two-way repeated measure ANOVA with factors numerosity (16 vs. 32) and connectedness (isolated vs. connected).

1
2 *Regression analysis.* In order to assess the effects of numerosity, connectedness, and the interaction
3 between the two outside the aforementioned predefined region of interest, a regression model was run
4 on the entire epoch using a moving-window approach. ERPs corresponding to each stimulus type were
5 entered as responses, with their amplitude averaged over shorter time windows (window width = 10
6 ms) spanning the entire duration of a trial (window step = 10 ms, from -100 to 400 ms). The analysis
7 was run individually for each subject to allow testing for the significance of the distribution of beta
8 values obtained for the different regressors. To do so, we used a cluster-based nonparametric test
9 (Maris & Oostenveld, 2007), comparing the clusters of significant time windows emerging from the
10 actual data with the extent of significant time windows emerging from a null distribution of beta
11 estimates (height threshold for significance corresponding to $p < 0.001$). The null distribution was
12 computed by randomly permuting the design matrix, and repeating this procedure for 10,000 iterations.
13 The final estimate of the beta values at the group level was computed as the average of the individual
14 beta estimates corresponding to different regressors ($\beta_{\text{Connectedness}}$, $\beta_{\text{Numerosity}}$, $\beta_{\text{Interaction}}$).
15

16 *Multivariate pattern analysis in the time domain.* To achieve a better description of the temporal
17 dynamics of the neural activity pattern, we applied a multivariate pattern analysis method (King &
18 Dehaene, 2014), using the Neural Decoding Toolbox (Meyers, 2013). This method allowed us to
19 evaluate how neural activity patterns coming from multiple sensors differ between different
20 experimental conditions and how well such differences can be generalized across the time domain.
21 More specifically, this neural decoding analysis involved the training of a support vector machine
22 (SVM) classifier on a subset of the data corresponding to specific conditions (i.e. the presentation of a
23 specific stimulus type), and to make predictions about which stimulus was presented in the remaining
24 subset of data and across all the time-course of activity. All the different stimuli presented across the
25 experiment were collapsed in four categories (16 isolated, 32 isolated, 16 connected, 32 connected),
26 and we compared pairs of them in order to assess the degree to which neural activity patterns for

1 different numerosity (i.e., 16 isolated vs 32 isolated and 16 connected vs 32 connected) or
2 connectedness (i.e., 16 isolated vs 16 connected and 32 isolated vs 32 connected) can be distinguished.
3 Each comparison was tested separately, training the SVM classifier with the responses (i.e. activity
4 recorded at all the channels with the exception of the EOG channels) to the two class of stimuli at hand
5 and testing it on another subset of trials not used in the training phase (using a leave-one-trial-out cross
6 validation). We followed a practice suggested in Grootswagers et al. (2017) to account for high noise
7 in single trial EEG data. First, we created pseudo trials by averaging randomly selected groups of 16
8 trials to improve signal to noise ratio. Second, to avoid overfitting, the number of features (i.e.
9 channels) included in the analysis was limited to the five most significant features computed from a
10 univariate ANOVA. Third, this decoding procedure was repeated 40 times for each participant using
11 different subsets of data for training and testing (as well as different ways to generate pseudo trials),
12 and the average of the 40 runs was taken as the final estimate of the decoding performance. The
13 outcome of the decoding procedure was a temporal generalization plot showing the performance of
14 the classifier at each time point, with classification accuracy (CA) reflecting how well the pattern
15 classifier can discriminate two conditions (see Fig. 5).

16

17 *fMRI data acquisition and analysis.*

18 *Image acquisition.* In the fMRI experiment, brain images were recorded using a Siemens Magnetom
19 Skyra 3T MRI scanner housed in the Human Magnetic Resonance Center at the University of
20 Massachusetts Amherst. The functional images comprised 34 axial slices and were acquired with an
21 echo-planar imaging (EPI) pulse sequence measuring the blood oxygen level dependent (BOLD) T2*
22 contrast. The following scanning parameters were used: TR = 2,000 ms, TE = 30 ms, slice thickness
23 = 3 mm, field of view = 204 mm, acquisition matrix = 68 x 68, flip angle = 79 deg, and parallel
24 acceleration = 2. After the first four volumes in each of the six functional runs were discarded to allow
25 scanner equilibrium, a total of 144 volumes were acquired in each run. High quality T1-weighted
26 structural images were acquired after half of the functional runs, allowing the participant to take a

break from the task. Structural images were acquired with a MPAGE sequence using the following parameters: TR = 1,800 ms, TE = 2.13 ms, slice thickness = 1 mm, field of view = 256 mm, flip angle = 9 deg.

Image preprocessing. Functional and structural data were preprocessed and analyzed using SPM8 (<http://www.fil.ion.ucl.ac.uk/spm/>) on Matlab (R2015b). First, functional images were slice-time corrected and realigned to the first volume acquired in the first run, and the structural scan was coregistered to the mean functional image of the time series. The structural image was then segmented into white and gray matter, and the gray matter was normalized into the standardized MNI space. The normalization parameters were applied to the realigned functional images, with a spatial resolution of 3 mm x 3 mm x 3 mm. Finally, normalized functional images were smoothed with a FWHM Gaussian kernel (width = 6 mm). If not indicated otherwise, SPM8 default parameters were used.

Regions of interest. To focus the analysis on a subset of brain areas relevant for our experimental question, we limited our primary analysis to four visual regions of interest (ROIs): V1, V2, V3, and IPS. These visual ROIs were defined using the probabilistic atlas provided by Wang et al. (2014), with a probability cut-off of 33%. Each of the early visual ROIs collapsed all four quadrants (dorsa/ventral and left/right). The IPS ROI collapsed six subdivisions (IPS0-IPS5) in both hemispheres in the original atlas. To avoid including the same voxels in multiple ROIs, we assigned overlapping voxels to the ROIs where they show the highest probability. The four visual ROIs are depicted in Fig. 6B.

Activation analysis. To achieve a measure of neural activity corresponding to the different class of stimuli presented during the experiment, activation was evaluated with a general linear model (GLM), including separate regressors for each of the 16 stimulus classes, which were convolved with a canonical hemodynamic response function. Additionally, motion parameters (head translation and rotation), 1-TR history of the motion parameters, and the square of both of them were added as

nuisance covariates in the GLM. Neural response was evaluated by taking the mean parameter estimate values (for specific contrasts) across voxels within each ROI, for each participant. The distribution of average parameter estimate value was then tested using a 2 x 2 ANOVA full factorial design for each of the ROIs, comprising the two levels of numerosity (16 and 32 dots) and the two levels of connectedness (isolated and connected dots).

Multivoxel pattern analysis. As in the EEG analysis, we employed a multivariate approach (multivoxel pattern analysis; MVPA) to evaluate the discriminability of brain activity patterns across different classes of stimuli. We used the SVM classifier provided by the libsvm package on Matlab (R2015b) (Chang & Lin, 2011). The MVPA analysis involved training the classifier on the pattern of voxel activity corresponding to different classes of stimuli (i.e. 16 isolated vs. 32 isolated; 16 isolated vs. 16 connected), and then testing whether the classifier can successfully predict which stimulus corresponds to novel instances of voxel activity not used for training. More specifically, each instance of a pattern for a specific stimulus class was defined by the corresponding activity in each run. This approach thus resulted in 6 instances for each stimulus class in each subject, as there were 6 runs in the experiment. Also, to avoid overfitting, we only included in the analysis 10% of the voxels, based on their discriminative power (using a univariate t-test) tested before entering the data into the MVPA routine. A leave-two-out cross-validation procedure was employed, training the classifier (using the cost parameter, $C = 1$) on ten instances corresponding to two classes of stimuli (5 instances each), and then testing it on the two remaining instances. This procedure was run individually for each subject and was repeated 6 times in order to cover different combinations of training and test sets. The average across all repetitions was taken as the final estimate of classification accuracy for each subject and for each comparison, indicating how well the pattern of voxel activity within each ROI allowed the classifier to successfully predict which stimulus corresponded to the instances included in the test set. The distribution of classification accuracies across the group was then tested with one-sample t-tests against the null hypothesis of chance level classification accuracy (i.e. classification accuracy = 0.5).

RESULTS

Psychophysical results

To confirm that pairwise connections between dots can successfully change perceived numerosity as reported in previous studies (He et al., 2009; Franconeri et al., 2009; Fornaciai et al., 2016), we analyzed the magnitude of the effect induced by connectedness, measured with a psychophysical test performed by participants after completing either the EEG or fMRI experiment. In this behavioral paradigm, participants were asked to discriminate the numerosity of two simultaneously presented dot-arrays: a reference containing either 16 or 32 dots, either isolated or pairwise connected, and a probe containing a variable number of dots (Fig. 1B). Note that besides the different levels of numerosity, other non-numerical continuous magnitudes were also systematically manipulated, which can be summarized by the orthogonal dimensions of *size* and *spacing* (see Methods; DeWind et al., 2015; Park et al., 2016). Figure 2A shows the average psychometric curves obtained by pooling the data of the two groups of participants tested in the EEG and fMRI experiments. Curves describing performances with connected dots were strongly shifted toward the left, indicating a robust underestimation compared to arrays of isolated dots. Figure 2B depicts the difference in the point of subjective equality (PSE) induced by connectedness across all participants. The perceived numerosity of connected items was substantially lower compared to isolated items in both conditions (16-dot reference; numerosity change: $-19.2\% \pm 2.1$; Wilcoxon Signed Rank test, $Z = -5.39$, $p < 0.001$; 32-dot reference; numerosity change: $-25.5\% \pm 2.2$; $Z = -5.44$, $p < 0.001$).

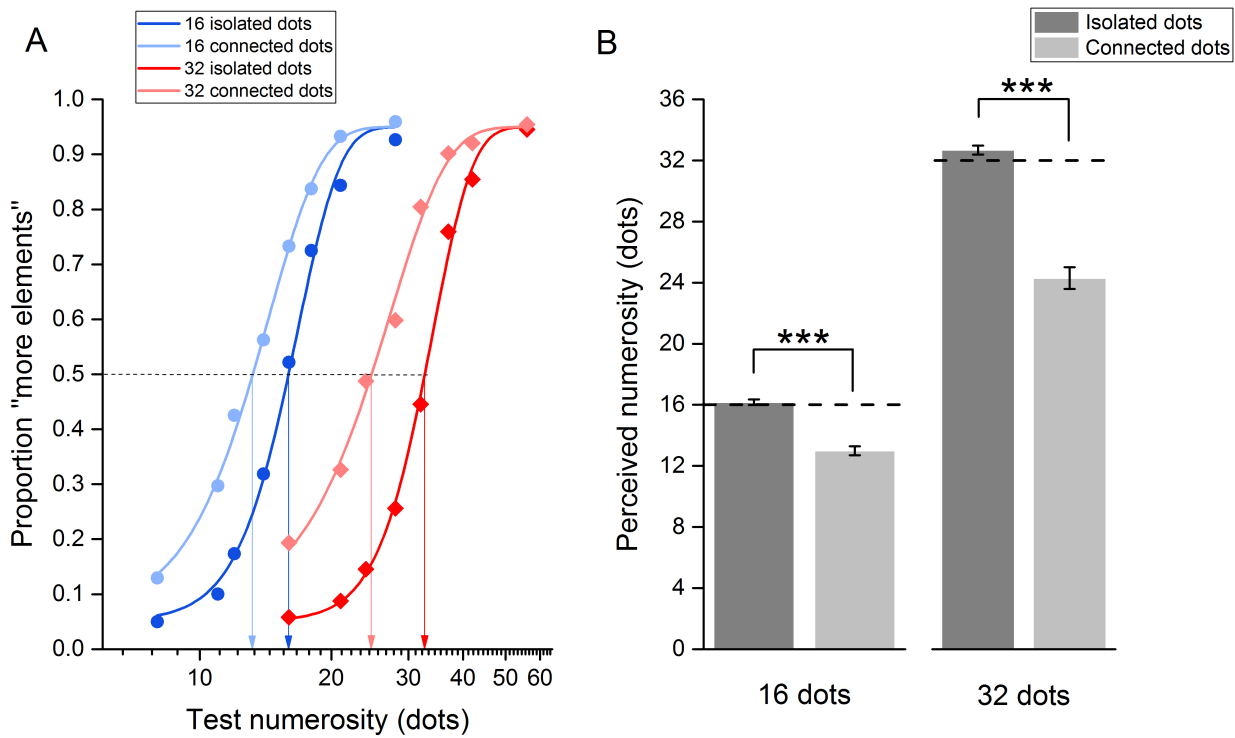
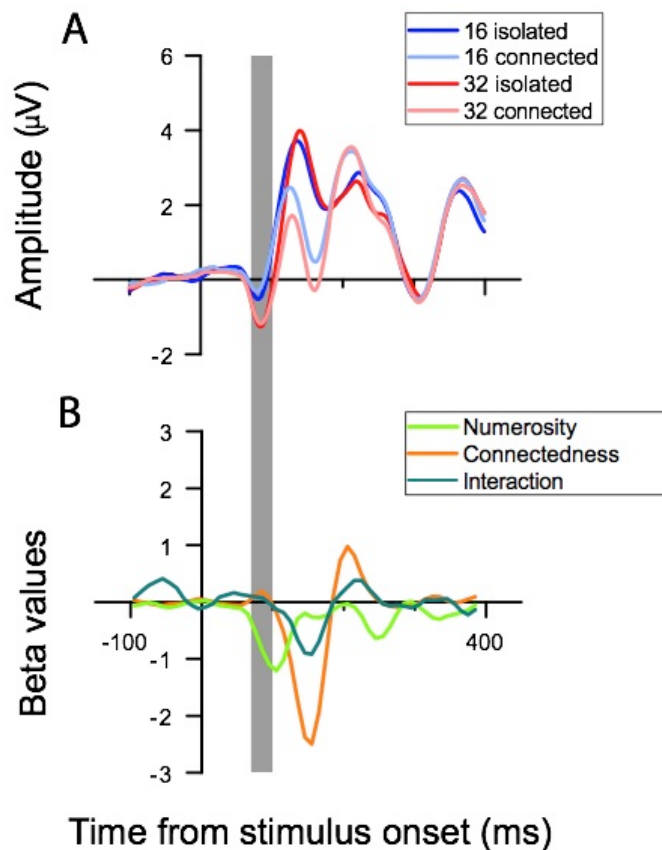


FIGURE 2. Behavioral results. (A) Average psychometric curves in the psychophysical task, obtained by pooling the data of all the participants tested in the EEG and fMRI experiments. The leftward shift of the curves representing connected dots (light blue and light red) suggests a systematic underestimation of connected items compared to the isolated ones. (B) Average PSEs in the different conditions of the psychophysical experiment. Dashed lines indicate the veridical numerosity of the reference stimulus in two different trial types (16-dot reference and 32-dot reference). Asterisks refer to the results of statistical tests comparing PSEs across the different conditions: *** $p < 0.001$. Error bars represent SEM.

Temporal dynamics: regression analysis of EEG data.

In the EEG experiment, participant watched a stream of dot-array stimuli comprising different numerosities (16 or 32 dots), either isolated or pairwise connected, and performed an oddball detection task on a dimension (i.e. color) irrelevant to the dimension of interest (i.e. numerosity). Figure 3A shows the brainwaves over the midline occipital channel Oz, our primary channel of interest given previous findings (Park et al., 2016; Fornaciai et al., 2017), for the four stimulus categories. As can be

1 seen from the brainwaves, the ERPs evoked by different absolute numbers of dots (16 vs. 32) were
2 dissociable early in the time course (< 100 ms) regardless of whether the stimuli were connected or
3 not. In stark contrast, at later latencies (100-200 ms), the ERPs differed more based on whether the
4 dots were isolated or connected than on the absolute number of dots.
5



6
7 *FIGURE 3. Brainwave plots and results of the regression analysis. (A) ERPs recorded at channel Oz.*
8 *(B) Beta values obtained in the regression analysis for channels Oz. The same results in other posterior*
9 *channels of interest are reported in Table 1. The position of channel Oz is indicated in red in Fig. 4A-*
10 *C. The shaded area highlights our time window of interest (70-100 ms).*

11
12 To address our central research question concerning the extremely early visual cortical activity, we
13 assessed the effects of numerosity and connectedness using a two-way repeated measures ANOVA on
14 the ERPs at channel Oz from the latency window 70-100 ms, again defined based on previous findings

(Park et al., 2016; Fornaciai et al., 2017; for details see *Regions of interest* in *Methods*). Specifically, all the stimuli presented during the experiment were collapsed into four classes (i.e. 16 isolated, 32 isolated, 16 connected, 32 connected), and ERPs corresponding to these different classes of stimuli were averaged across a target latency window defined in order include the peaks of early numerosity-sensitive activity found in previous studies (Park et al., 2016; Fornaciai et al., 2017). The results showed a significant main effect of numerosity ($F(1,19) = 23.04$, $p < 0.001$), but no main effect of connectedness ($F(1,19) = 1.26$, $p = 0.28$) or interaction ($F(1,19) = 1.89$, $p = 0.18$). They demonstrate that early cortical activity is exclusively modulated by the numerosity of the stimuli, with very little, if any, effect of the connectedness illusion.

We then performed a point-by-point regression analysis to provide a more comprehensive picture of the observed effect across the entire epoch (Fig. 3B). This model included numerosity (16 or 32 dots), connectedness (0 or 1 categorical coding), and the interaction between the two as regressors to explain the variance in the ERPs. The distribution of beta values obtained with the regression analysis was tested by means of a cluster-based nonparametric test, performed across the entire time course of activity. At channel Oz, a significant effect of numerosity emerged early in the time course, peaking at around 105 ms post-stimulus (75-125 ms, peak $\beta_N = -1.21$; $p < 0.0001$, average adjusted $R^2 = 0.51 \pm 0.28$) – consistently with the significant effect found within our primary latency window of interest. The effect of numerosity was followed by a significant effect of connectedness (115-175 ms, peak $\beta_C = -2.50$; $p < 0.0001$, average adjusted $R^2 = 0.81 \pm 0.10$) and a significant interaction (135-165 ms, peak $\beta_I = -0.91$; $p = 0.0001$, average adjusted $R^2 = 0.81 \pm 0.10$), both peaking at around 150 ms. In addition, a later positive peak of the connectedness effect was evident at about 205 ms post-stimulus (195-215 ms, peak $\beta_C = 0.97$; $p = 0.0004$, average adjusted $R^2 = 0.43 \pm 0.24$) followed by a later peak of the numerosity effect around 245 ms (245-256 ms, peak $\beta_N = -0.67$; $p < 0.0001$, average adjusted $R^2 = 0.47 \pm 0.25$). These results show that modulations of numerosity and connectedness both have a strong effect on neural responses, although with different timing.

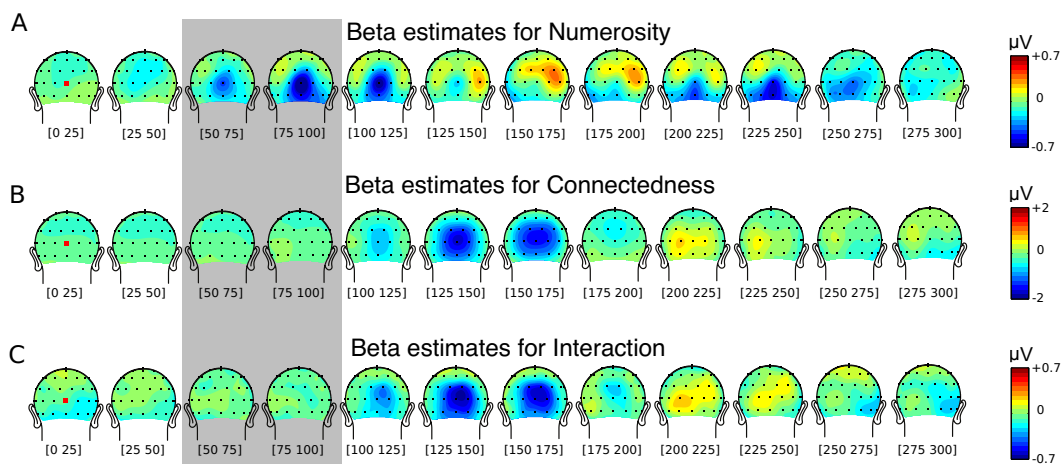
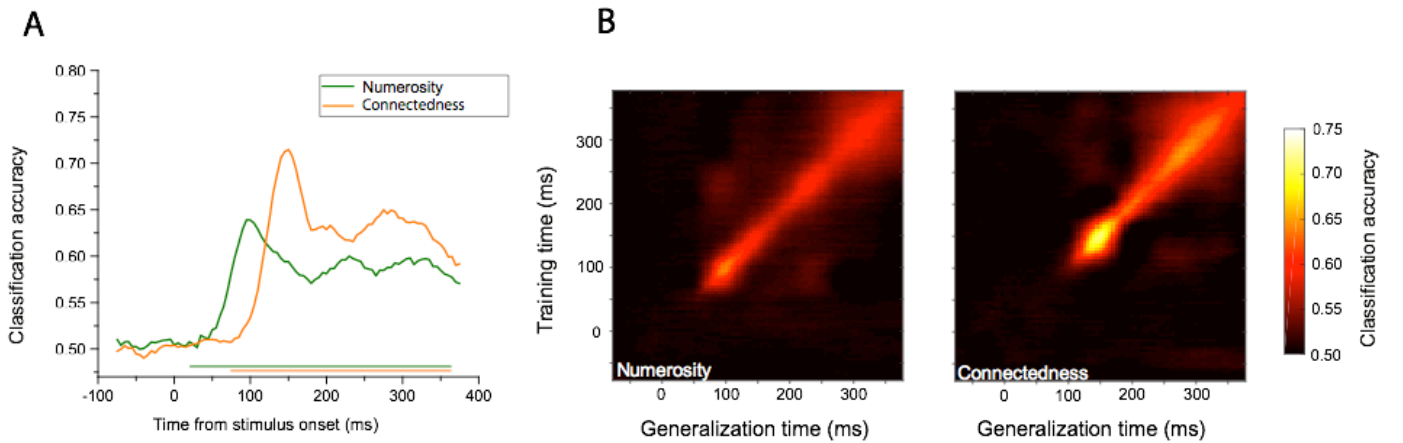


FIGURE 4. Topographic maps of beta estimates. (A-C) Posterior-perspective topographic maps of beta estimates averaged across different time windows (indicated in square brackets below each plot) for numerosity (C), connectedness (D), and the interaction between numerosity and connectedness (E). Shaded area indicates topographic maps encompassing our latency window of interest.

To obtain a comprehensive view of the observed effects across the channels, we plotted topographic maps of beta estimates. As shown in Fig. 4A, the effect of numerosity showed two main peaks: one around 75-125 ms over medial occipital channels and another around 150-175 ms over occipito-parietal scalp sites. This pattern is consistent with earlier findings demonstrating two separate peaks of numerosity-sensitive activity at about 75 ms and 180 ms post-stimulus (Park et al., 2016). In contrast, the effect of connectedness (Fig. 4B) was mostly evident around 125-175 ms over medial occipital scalp sites. Finally, the topographic distribution of the interaction effect was much alike that of the connectedness effect (Fig. 4C). It should be noted that top-perspective topographic maps (not shown in figure) show no noteworthy effect of numerosity, connectedness, or interaction, confirming that the effects of experimental manipulations are maximally distributed in occipital regions.

Temporal generalization analysis in the time domain.

1 A multivariate neural decoding analysis was performed to characterize the dynamics of neural
2 representations across all the recording channels (see Methods). In this decoding analysis, a support
3 vector machine (SVM) classifier was used to test how well two classes of neural activity patterns are
4 distinguishable (e.g. 16 isolated vs. 32 isolated or 16 isolated vs. 16 connected), using classification
5 accuracy as a metric for that distinctiveness. The comparison of 16 isolated versus 32 isolated
6 conditions together with the comparison of 16 connected and 32 connected conditions allowed us to
7 infer about how the neural activity patterns for two different numerosities differed. The comparison of
8 16 isolated versus 16 connected conditions together with the comparison of 32 isolated and 32
9 connected conditions allowed us to infer about how the neural activity patterns for isolated and
10 connected items differed.



11
12 *FIGURE 5. Results of the multivariate neural decoding analysis across all the recording channels*
13 *under the contrast of numerosity (16 isolated versus 32 isolated, 16 connected versus 32 connected)*
14 *and the contrast of the connectedness effect (16 isolated versus 16 connected, 32 isolated versus 32*
15 *connected). (A) Classification accuracies obtained by training and testing the classifier at the same*
16 *time point. The two successive peaks of the numerosity effect (~100 ms) and connectedness effect (~150*
17 *ms) clearly illustrate the temporal offset between the types of neural representations. The colored bars*
18 *at the bottom of panel A indicate the significant time windows corresponding to the different*
19 *conditions. (B) Temporal generalization matrices obtained by training the classifier at one time point*
20 *and testing it at all the other time points.*

Fig. 5A shows the pattern of classification accuracy obtained by training and testing the classifier at the same time point. In the case of the numerosity contrast (i.e., different number of dots regardless of their connectedness), classification accuracy showed a first peak at about 95 ms (average peak accuracy = 0.64 ± 0.067), consistent with the results of the regression analysis, followed by a sustained stream of above-chance classification accuracy until later in the time course. In contrast, the neural decoding pattern corresponding to the connectedness effect (i.e., same number of isolated vs. connected dots) peaked slightly later in the time course at 150 ms (average peak accuracy = 0.71 ± 0.09), again consistent with the results of the regression analysis. The classification accuracy was generally greater in distinguishing the effect of connectedness than in distinguishing the effect of numerosity, suggesting that the processes triggered by the connectedness illusion are much more robust compared to the raw numerical representation.

To assess the extent to which patterns of activity generalize across multiple time windows, we also obtained temporal generalization plots by training the classifier at one time point and testing it across all the time points (Fig. 5B). While training and testing the classifier at the same time point (which represents the diagonal of the temporal generalization plots shown in Fig. 5B) provides information about the dynamics of brain activity patterns across the analyzed time-course, the temporal generalization provides further information about potential recurring patterns and sustained stages of activation over time (King & Dehaene, 2014). The temporal generalization plots (Fig. 5B) first confirmed the stronger effect in the connectedness comparison. One additional noticeable point here is that at earlier latencies, above-chance decoding was mostly distributed along the diagonal in both the numerosity and connectedness cases, suggesting that these dot-array stimuli are likely represented by a series of processing stages with unique patterns of activity through the visual stream. On the other hand, the pattern of classification accuracy showed a larger generalization across time points later in the time course (~300 ms), likely representing the involvement of a relatively sustained processing

stage (King & Dehaene, 2014). Note that the effects represented in the diagonal of a temporal generalization plot in Fig. 5B reflect the decoding results shown in Fig. 5A.

Anatomical correlates: fMRI results.

Overall, the EEG results showed that the early-latency ERPs are exclusively modulated by numerosity irrespective of connectedness, suggesting that the segmentation processes necessary to turn the raw visual input into perceptual units emerge only later in the processing stream. Indeed, as shown by both the regression and decoding analyses, the effect of connectedness emerged at later latencies (150 ms), well beyond the early numerosity-sensitive signals highlighted in previous studies (Park et al., 2016; Fornaciai et al., 2017). Previous results (Fornaciai et al., 2017) pinpointed such early activity as arising from early visual areas such as V2 or V3. To identify the roles of different visual areas in representing the raw numerical information versus the perceptual units necessary for numerosity perception, we further investigated the anatomical correlates of the connectedness illusion. To this aim, we performed an independent fMRI experiment using the same paradigm and the same set of stimuli as in the EEG experiment to address this question.

In the fMRI experiment, participants performed the same task viewing the same dot-array stimuli as in the EEG experiment. The only difference was the longer interstimulus interval ranges that accommodated the sluggish hemodynamic response (see Methods). Participant's blood-oxygen-level-dependent (BOLD) signal associated with each of the four experimental conditions (i.e., 16 isolated, 16 connected, 32 isolated, 32 connected) was modeled using the general linear model (GLM), and the regression (beta) coefficients were interpreted as a proxy for neural activity. We then assessed the effects of numerosity and connectedness in anatomically defined early visual regions of interest (ROIs) including areas V1, V2, V3, and the intraparietal sulcus (IPS) (Fig. 6B). These ROIs were chosen given the medial occipital effects obtained in the current EEG experiment (Fig. 4) and the recent evidence for numerosity processing primarily in V2 and V3 (Fornaciai et al., 2017). Besides these

primary regions of interest, however, we also examined the effects in the intraparietal sulcus (IPS) for its relevance in numerosity processing (Piazza et al., 2004; Piazza et al., 2007; Harvey et al., 2013).

Neural responses representing the encoding of numerosity were obtained by extracting the average regression coefficients (from the GLM in fMRI analysis) for different levels of numerosity (i.e. 16 and 32 dots), and for different levels of connectedness (i.e. isolated and connected), within each of the selected ROIs (V1, V2, V3, IPS). We then assessed the neural responses to the different class of stimuli using a 2 x 2 repeated measures ANOVA, separately for each ROI. However, none of the ROIs showed any significant effect of either numerosity ($F(1,18) < 1.26$, p -values > 0.27) or connectedness ($F(1,18) < 3.89$, p -values > 0.064), and no interaction between the two factors ($F(1,18) < 2.61$, p -values > 0.12). These weak effects suggest that the average activity across the different ROIs might not be sensitive enough to pick up a relatively small effect caused by the parametric modulation of our dot-array stimuli.

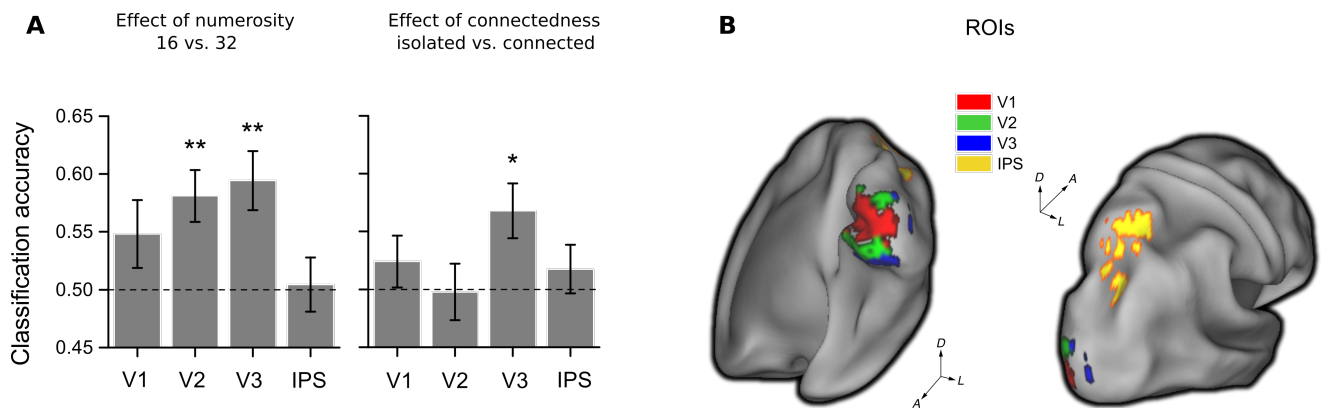


FIGURE 6. fMRI results. (A) Average classification accuracies obtained with the MVPA analysis across the ROIs, corresponding to the effect of numerosity (left panel) and the effect of connectedness (right panel). While the effect of numerosity was observed in both area V2 and V3, the effect of connectedness was observed only in area V3. This result suggests that V3 is the key area underpinning the segmentation processes turning the raw numerical information into the perceptual units necessary

for numerosity perception. Error bars are SEM. * $p < 0.025$, ** $p < 0.01$. (B) Illustration of the anatomically-defined visual ROIs visualized on an inflated right hemisphere. Early visual cortex ROIs (V1, V2, V3) included both left and right, and dorsal and ventral sub-regions. The IPS ROI included 6 sub-regions (IPS0-IPS5) in both hemispheres. Axes show the orientation of the brain along the anterior-posterior, dorsal-ventral, and medial-lateral directions: A-anterior; D-dorsal; L-lateral.

As in the EEG experiment, we adopted a multivariate approach to achieve better sensitivity to the brain activity elicited by different experimental conditions. In this multi-voxel pattern analysis (MVPA), an SVM classifier was used to assess the discriminability of two classes of brain response patterns within each ROI. The discriminability of neural activity patterns based on numerosity (i.e. 32 isolated vs. 16 isolated combined with 32 connected vs. 16 connected) was significantly above chance in V2 and V3 (Left panel of Fig. 6A; CA = 0.58 ± 0.02 , and 0.59 ± 0.02 ; one-sample t-test: $t(18) = 3.62$, $p = 0.002$, and $t(18) = 3.70$, $p = 0.0016$, respectively for V2 and V3), indicating that the pattern of activation across voxels within these areas can successfully discriminate between different absolute number of dots, regardless of their connectedness. The discriminability of neural activity pattern based on the effect of connectedness (i.e. 16 isolated vs. 16 connected combined with 32 isolated vs. 32 connected) was significantly above chance in V3 only (Right panel of Fig. 6A; CA = 0.57 ± 0.02 , $t(18) = 2.87$, $p = 0.01$). Except for these cases, other ROIs including the IPS failed to discriminate neural activity patterns for numerosity or connectedness (p -values > 0.11).

DISCUSSION

Numerical magnitude is one of the fundamental attributes of the visual world, yet a comprehensive account of the neural pathway leading to the representation of numerosity is still lacking. Recent studies demonstrate robust neurophysiological correlates of numerosity perception, with numerosity-sensitive brain signal emerging at multiple stages during the time-course of activity. Importantly, it has been shown that such numerosity-sensitive activity emerges extremely early after stimulus onset,

1 as early as 75 ms (Park et al., 2016). Later studies identified this activity as arising from early visual
2 areas, such as V2 and V3, with possible contributions even from V1 (Fornaciai et al., 2017). However,
3 it has been demonstrated that numerosity perception is not simply based on the physical amount of
4 items present in the visual scene, but rather on segmentation processes defining a set of “perceptual
5 units” according to the topological properties of the visual images (i.e. as in the case of items connected
6 by lines, or items enclosed within the same boundary; Franconeri et al., 2009; He et al., 2009; He et
7 al., 2015; Fornaciai et al., 2016). It is thus unclear whether the extremely early activity highlighted in
8 previous studies actually reflects such perceptual units forming the basis of numerosity perception, or
9 a raw sensory representation of the visual stimuli prior to segmentation. Indeed, only in the former
10 case a neural signature could be regarded as a sufficient correlate of numerosity perception.

11
12 To investigate whether early neural activity for numerosity in the low-level visual cortex represents a
13 sufficient correlate of numerosity perception, we exploited the connectedness illusion (He et al., 2009;
14 Franconeri et al., 2009; He et al., 2015; Fornaciai et al., 2016), which is based on one of the most basic
15 principles of perceptual organization (i.e. uniform connectedness; Wertheimer, 1912; Palmer & Rock,
16 1994). Numerosity perception and its connectedness illusion provide an ideal condition to address this
17 issue, because there is a clear distinction between raw sensory information and the perceptual units
18 giving rise to the connectedness illusion, and because the effect of this illusion can be manipulated
19 parametrically.

20
21 The results of our EEG experiment from both the regression (Fig. 3 and Fig. 4) and multivariate
22 decoding analyses (Fig. 5) show a striking dissociation between neural activity reflecting raw
23 numerosity (e.g., illustrated in the contrast 16 versus 32 isolated dots) and the neural activity reflecting
24 the connectedness illusion (e.g., illustrated by the contrast 16 isolated versus 16 connected dots). The
25 neural activity pattern representing numerosity emerges early in the visual stream (~100 ms) followed
26 by an effect of connectedness at a later latency (~150 ms), with largely overlapping topographic

distributions. The results from our fMRI experiment further demonstrate that while numerosity is represented in both visual area V2 and V3, the effect of connectedness emerges only in V3. Thus, collectively, our results suggest that early visual signals (< 100 ms), arising from areas V2 and V3, do *not* represent a sufficient correlate of numerosity perception. Instead, the segmentation processes creating the perceptual units necessary for numerosity perception, as highlighted by the connectedness illusion, emerge only later in the visual stream, around or after 150 ms post stimulus onset.

One plausible explanation for this pattern of results is that numerosity information first processed in V2/V3 in a feed-forward manner (at or before 100 ms) gets processed again within that same cortical region (V3) after re-entrant feedback from other areas (around 150 ms), such as higher-level parietal or frontal regions. However, given the tight timing between the two processing stages (just ~ 50 ms), a more plausible explanation is that the re-entrant feedback arises from the interaction between mid-level visual areas along the dorsal (i.e. V3A, MT) and ventral (i.e., V4, LOC) stream, which process different visual features. This idea is consistent with recent behavioral observations suggesting the involvement of multiple visual areas in the early levels of numerosity processing (Fornaciai & Park, 2017b), and with the fact that visual area V3 represents a crucial node in the visual pathway, with the peculiar features of neurons in this area (i.e. sensitivity to multiple features of the stimuli like color and motion) making it an intersection between the dorsal and ventral stream (Gegenfurtner et al., 1997). Thus, according to this idea, the re-entrant feedback signals from a network of mid-level visual areas to V3 mediates the object segmentation processes providing the source of the perceptual units underlying numerosity perception.

How does this re-entrant processing fit into the numerosity processing stream highlighted by previous studies? For instance, Park and colleagues (2016) demonstrated two distinct stages of numerosity perception, first around 75 ms after stimulus onset followed by a later stage around 185 ms, while Fornaciai et al. (2017) showed that the first of these two stages occurs in early visual areas such as V2

1 or V3. In this context, the effect of connectedness emerging at around 150 ms may lie on the
2 information transformation pathway between the two numerosity processing stages observed in
3 previous studies (i.e. 75-88 ms and 185-215 ms; Park et al., 2016; Fornaciai et al., 2017). Interestingly,
4 while the later processing stage has been previously interpreted as a signature of activity in parietal
5 cortex – consistently with results showing activity in the IPS (Piazza et al., 2004; Harvey et al., 2013)
6 – we did not observe any significant effect of numerosity or connectedness in the IPS. This lack of
7 parietal activation in the context of our passive-viewing paradigm is consistent with a recent study
8 showing little or no IPS activation in the absence of an explicit task (Cavdaroglu et al., 2015; DeWind
9 et al., Under review), and raises the possibility that a decision process is necessary for the involvement
10 of the IPS (also see Göbel et al., 2004).

11
12 While the current findings reject the hypothesis that early visual cortical activity represents a sufficient
13 correlate of numerosity perception and rather illustrate that the neural signature for perceptual
14 segmentation occurs only at or after 150 ms in area V3, does this mean that this later activity is
15 sufficient to give rise to the subjective perceptual experience of numerosity? Unfortunately, our results
16 alone cannot determine whether those representations are sufficient to give rise to subjective
17 perceptual experience, because the passive viewing nature of our paradigm prevents us from assessing
18 whether numerosity-sensitive neural responses at different levels are correlated with the actual
19 numerosity experienced by participants. Further processing stages may be required to reach visual
20 awareness. Nevertheless, considering previous studies addressing the neurophysiological correlates of
21 visual awareness (see Koivisto & Revonsuo, 2010, for a review), there is evidence to argue that neural
22 representations in V3 around 150 ms may form sufficient basis to support the subjective conscious
23 experience of numerosity. For instance, the most widely acknowledged ERP correlate of visual
24 awareness emerging from studies comparing seen versus unseen stimuli, is the visual awareness
25 negativity (VAN), peaking between 150 ms and 200 ms over occipito-parietal sites (although the
26 timing and scalp topography are highly variable across different studies; Railo et al., 2011; Pitts et al.,

2014), and the timing of which is relatively consistent with the peak of the connectedness effect (150 ms). Similarly, studies investigating illusory percepts show the involvement of mid-level vision processing stages (145-155 ms) in representing illusory information such as illusory contours (Halgren et al., 2003), again consistent with the timing of the connectedness illusion observed in our results. However, since our results do not allow us to draw strong conclusions about whether signals in V3 may underlie the emergence of numerosity information into visual awareness, this remains an intriguing question for future studies. Regardless of this possibility, we can conclude that relatively early processing occurring in visual area V3 is the earliest processing stage where we can trace a sufficient correlate of numerosity perception, reflecting the perceptual units setting the bases for numerical perception.

Overall, we have demonstrated that numerosity perception undergoes at least two distinct stages: first, the extraction of raw sensory information of a dot-array stimulus early in the visual stream (< 100 ms in V2/V3) which does not yet form a sufficient basis for numerosity perception and, second, the segmentation of the stimulus into perceptual units slightly later in the visual stream (150 ms in V3) that are likely to serve as a foundation for perceived representation of numerosity. These results thus also highlight the active and constructive nature of perception, whereby numerical information passes through different processing stages in early visual cortex before being turned into a suitable format for numerosity perception.

AUTHOR CONTRIBUTIONS

M.F. and J.P. designed the study. M.F. collected the data. M.F. and J.P. analyzed the data, interpreted the results, wrote the manuscript and revised it.

1 Conflict of interest

2 The authors declare no conflict of interest.

3

4 Acknowledgements

5 We thank Brynn Boutin, Courtney McMahon, and Bri Southwick for their assistance in EEG data
6 collection and Dr. Kwan-Jin Jung for his advice and assistance in fMRI data collection. We also thank
7 Drs. Kyle Cave and Marty Woldorff for helpful comments and discussions. This study was supported
8 by the National Science Foundation CAREER Award BCS1654089 to J. P.

9

10 REFERENCES

11 Anobile, G., Arrighi, R., Togoli, I., & Burr, D. C. (2016). A shared numerical representation for
12 action and perception. *eLife*, 5. <https://doi.org/10.7554/eLife.16161>

13 Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number As a Primary Perceptual Attribute: A
14 Review. *Perception*, 45(1–2), 5–31. <https://doi.org/10.1177/0301006615602599>

15 Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proceedings of the*
16 *Royal Society B: Biological Sciences*, 281(1797), 20141791.

17 <https://doi.org/10.1098/rspb.2014.1791>

18 Burr, D. C., Anobile, G., Arrighi, R., & Burr, D. C. (2017). Psychophysical evidence for the number
19 sense. *Philosophical Transactions of the Royal Society B-Biological Sciences*, 373.

20 <https://doi.org/10.1098/rstb.2017.0045>

21 Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, 10(4), 433–436.

22 <https://doi.org/10.1163/156856897X00357>

23 Cavdaroglu, S., Katz, C., & Knops, A. (2015). Dissociating estimation from comparison and
24 response eliminates parietal involvement in sequential numerosity perception. *NeuroImage*,

25 116, 135–148. <https://doi.org/10.1016/j.neuroimage.2015.04.019>

1 Chang, C., & Lin, C. (2013). LIBSVM : A Library for Support Vector Machines. *ACM Transactions*
2 *on Intelligent Systems and Technology (TIST)*, 2, 1–39.
3 <https://doi.org/10.1145/1961189.1961199>

4 Cicchini, G. M., Anobile, G., & Burr, D. C. (2016). Spontaneous perception of numerosity in
5 humans. *Nature Communications*, 7, 12536. <https://doi.org/10.1038/ncomms12536>

6 Dehaene, S. (2011). *The number sense: How the mind creates mathematics*. New York: Oxford
7 University Press.

8 Delorme, A., & Makeig, S. (2004). EEGLAB: An open source toolbox for analysis of single-trial
9 EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*,
10 134(1), 9–21. <https://doi.org/10.1016/j.jneumeth.2003.10.009>

11 DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate
12 number system to quantify the contribution of visual stimulus features. *Cognition*, 142, 247–65.
13 <https://doi.org/10.1016/j.cognition.2015.05.016>

14 DeWind, N. K., Park, J., Woldorff, M. G., Brannon, E. M. (Under review). Numerical encoding in
15 early visual cortex.

16 Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*,
17 18(18), R855–R856. <https://doi.org/10.1016/j.cub.2008.07.053>

18 Fornaciai, M., Cicchini, G. M., & Burr, D. C. (2016). Adaptation to number operates on perceived
19 rather than physical numerosity. *Cognition*, 151, 63–67.
20 <https://doi.org/10.1016/j.cognition.2016.03.006>

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

23 Fornaciai, M., & Park, J. (2017a). Distinct Neural Signatures for Very Small and Very Large
24 Numerosities. *Frontiers in Human Neuroscience*, 11(January), 1–14.
25 <https://doi.org/10.3389/fnhum.2017.00021>

1 Fornaciai, M., & Park, J. (2017b). Spatiotemporal feature integration shapes approximate numerical
2 processing. *Journal of Vision*, 17(13), 6. <https://doi.org/10.1167/17.13.6>

3 Franconeri, S. L., Bemis, D. K., & Alvarez, G. A. (2009). Number estimation relies on a set of
4 segmented objects. *Cognition*, 113, 1–13. <https://doi.org/10.1016/j.cognition.2009.07.002>

5 Gegenfurtner, K. R., Kiper, D. C., & Levitt, J. B. (1997). Functional properties of neurons in
6 macaque area V3. *Journal of Neurophysiology*, 77(4), 1906–1923.

7 Göbel, S. M., Johansen-Berg, H., Behrens, T., & Rushworth, M. F. S. (2004). Response-selection-
8 related parietal activation during number comparison. *Journal of Cognitive Neuroscience*,
9 16(9), 1536–1551. <https://doi.org/10.1162/0898929042568442>

10 Halgren, E., Mendola, J., Chong, C. D. R., & Dale, A. M. (2003). Cortical activation to illusory
11 shapes as measured with magnetoencephalography. *NeuroImage*, 18(4), 1001–1009.
12 [https://doi.org/10.1016/S1053-8119\(03\)00045-4](https://doi.org/10.1016/S1053-8119(03)00045-4)

13 Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of
14 numerosity in the human parietal cortex. *Science*, 341(September), 1123–1126.
15 <https://doi.org/10.1126/science.1239052>

16 He, L., Zhang, J., Zhou, T., & Chen, L. (2009). Connectedness affects dot numerosity judgment:
17 implications for configural processing. *Psychonomic Bulletin & Review*, 16(3), 509–517.
18 <https://doi.org/10.3758/PBR.16.3.509>

19 He, L., Zhou, K., Zhou, T., He, S., & Chen, L. (2015). Topology-defined units in numerosity
20 perception. *Proceedings of the National Academy of Sciences*, 112(41), E5647–E5655.
21 <https://doi.org/10.1073/pnas.1512408112>

22 King, J.-R., & Dehaene, S. (2014). Characterizing the dynamics of mental representations: the
23 temporal generalization method. *Trends in Cognitive Sciences*, 18(4), 203–210.
24 <https://doi.org/10.1016/j.tics.2014.01.002>

1 Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in
2 Psychtoolbox-3? *Perception ECVF 2007 Abstract Supplement*, 14.
3 <https://doi.org/10.1068/v070821>

4 Koivisto, M., & Revonsuo, A. (2010). Event-related brain potential correlates of visual awareness.
5 *Neuroscience & Biobehavioral Reviews*, 34(6), 922–934.
6 <https://doi.org/10.1016/j.neubiorev.2009.12.002>

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

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

13 Meyers, E. M. (2013). The Neural Decoding Toolbox. *Frontiers in Neuroinformatics*, 7(May), 8.
14 <https://doi.org/10.3389/fninf.2013.00008>

15 Palmer, S., & Rock, I. (1994). Rethinking perceptual organization: The role of uniform
16 connectedness. *Psychonomic Bulletin & Review*, 1(1), 29–55.
17 <https://doi.org/10.3758/BF03200760>

18 Park, J. (2017). A neural basis for the visual sense of number and its development: A steady-state
19 visual evoked potential study in children and adults. *Developmental Cognitive Neuroscience*.
20 <https://doi.org/10.1016/j.dcn.2017.02.011>

21 Park, J., Dewind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and Direct Encoding of
22 Numerosity in the Visual Stream. *Cerebral Cortex*, 26(2), 748–763.
23 <https://doi.org/10.1093/cercor/bhv017>

24 Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: transforming numbers
25 into movies. *Spatial Vision*, 10(4), 437–442. <https://doi.org/10.1163/156856897X00366>

- 1 Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate
2 numerosity in the human intraparietal sulcus. *Neuron*, 44(3), 547–555.
3 <https://doi.org/10.1016/j.neuron.2004.10.014>
- 4 Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A Magnitude Code Common to
5 Numerosities and Number Symbols in Human Intraparietal Cortex. *Neuron*, 53(2), 293–305.
6 <https://doi.org/10.1016/j.neuron.2006.11.022>
- 7 Pitts, M. A., Metzler, S., & Hillyard, S. A. (2014). Isolating neural correlates of conscious perception
8 from neural correlates of reporting one's perception. *Frontiers in Psychology*, 5(SEP).
9 <https://doi.org/10.3389/fpsyg.2014.01078>
- 10 Railo, H., Koivisto, M., & Revonsuo, A. (2011). Tracking the processes behind conscious
11 perception: A review of event-related potential correlates of visual consciousness.
12 *Consciousness and Cognition*, 20(3), 972–983. <https://doi.org/10.1016/j.concog.2011.03.019>
- 13 Roggeman, C., Santens, S., Fias, W., & Verguts, T. (2011). Stages of nonsymbolic number
14 processing in occipitoparietal cortex disentangled by fMRI adaptation. *The Journal of*
15 *Neuroscience : The Official Journal of the Society for Neuroscience*, 31(19), 7168–73.
16 <https://doi.org/10.1523/JNEUROSCI.4503-10.2011>
- 17 Watson, A. B. (1979). Probability summation over time. *Vision Research*, 19(5), 515–522.
18 [https://doi.org/10.1016/0042-6989\(79\)90136-6](https://doi.org/10.1016/0042-6989(79)90136-6)
- 19 Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift Für*
20 *Psychologie Und Physiologie Der Sinnesorgane*, 61(1), 161–265.