Spatiotemporal feature integration shapes approximate numerical processing

Michele Fornaciai

Department of Psychological & Brain Sciences, University of Massachusetts, Amherst, MA, USA

Joonkoo Park

Department of Psychological & Brain Sciences and Commonwealth Honors College, University of Massachusetts, Amherst, MA, USA

Numerosity perception involves a complex cascade of processing stages comprising an early sensory representation stage followed by a later stage providing a conceptual representation of numerical magnitude. While much recent work has focused on understanding how nonnumerical spatial features (e.g., density, area) influence numerosity perception in this processing cascade, little is known about how the spatiotemporal properties of the stimuli affect numerosity processing. Whether numerosity information is integrated over space and time in the processing cascade is an important question as it can provide insights into how the system dedicated for numerosity interacts with other perceptual systems. To address these issues, in four independent experiments, we asked participants to judge the numerosities of various different kinds of dynamically presented dot arrays, such as dots randomly changing in their locations, moving in smooth trajectories, or flickering on and off. The results revealed a systematic overestimation of dynamically presented dot arrays, which implicates the existence of spatiotemporal integration mechanisms, both at the early sensory representation stage and the later conceptual representation stage. The results also revealed the influence of motion and color processing areas on numerosity processing. The findings thus provide empirical evidence that numerosity perception arises from a complex interaction between multiple perceptual mechanisms in the visual stream, and that it is shaped by the integration of spatiotemporal properties of visual stimuli.

Introduction

The ability to evaluate approximate numerosity is likely one of the most fundamental skills allowing animals to make successful choices for their survival. Such ability appears to be phylogenetically ancient, as it is widespread across species (Agrillo, Dadda, Serena, & Bisazza, 2008; Pepperberg, 2006; Piantadosi & Cantlon, 2017; Rugani, Vallortigara, Priftis, & Regolin, 2015), and ontogenetically innate, as it seems present from birth in human newborns (Izard, Sann, Spelke, & Streri, 2009; Xu, 2003; Xu & Spelke, 2000). Furthermore, it has been shown that approximate numerical abilities arise from a dedicated and genuinely perceptual brain mechanism (Anobile, Cicchini, & Burr, 2016a; Burr & Ross, 2008; Cicchini, Anobile, & Burr, 2016; Fornaciai, Cicchini, & Burr, 2016; but see Gebuis, Cohen Kadosh, & Gevers, 2016 for arguments against the existence of a dedicated mechanism for numerosity processing).

Much evidence in favor of an abstract (i.e., independent from the specific format and modality of sensory stimuli) sense of number comes from behavioral studies investigating cross-format (i.e., simultaneously and sequentially presented numerosities) and cross-modal (i.e., visual and auditory numerosities) adaptation effects, showing that the number sense is generalized across different formats and sensory modalities (Arrighi, Togoli, & Burr, 2014; Izard et al., 2009), and across both sensory and motor systems (Anobile, Arrighi, Togoli, & Burr, 2016b). However, at the neural level, much recent research has focused on the visual domain and has identified the parietal cortex to be critical for numerosity processing. For example, single-cell recording studies in nonhuman primates highlight the intraparietal sulcus during numerosity perception, showing the existence of neurons responding to the number of items with well-defined tuning curves (Nieder, 2016; but see Chen & Verguts, 2013, arguing the lack of evidence for the numerosity-selective tuning properties). The involvement of the intraparietal sulcus has been also demonstrated in humans by several neuroimaging studies (Castaldi, Aagten-Murphy, Tosetti, Burr, & Morrone, 2016;

Citation: Fornaciai, M., & Park, J. (2017). Spatiotemporal feature integration shapes approximate numerical processing. *Journal of Vision*, *17*(13):6, 1–15, doi:10.1167/17.13.6.

doi: 10.1167/17.13.6 Received July 12, 2017; published November 06, 2017 ISSN 1534-7362

Harvey, Klein, Petridou & Dumoulin, 2013; Piazza, Pinel, Le Bihan, & Dehaene, 2007).

While neuroimaging studies have mostly concerned a set of magnitude-specific brain areas across the parietal lobe, thought to underpin numerosity processing, computational models propose a more complex stream of processing stages (Dehaene & Changeux, 1993; Verguts & Fias, 2004). Particularly, according to Dehaene and Changeux's (1993) model, visual numerosity is processed in three fundamental computational steps: first, visually presented items are encoded in an object location map (OLM) while their size and shape are normalized; then, the overall activity in the OLM is summed and detected by number-sensitive neurons (summation coding); finally, the pattern of activity across number-sensitive neurons elicits the response of number-selective units, directly coding for the specific numerosity carried by sensory signals (but see also Stoianov & Zorzi, 2012, and Sengupta, Surampudi, & Melcher, 2014 for alternative models).

As these computational models predict, previous event-related potential studies from our lab demonstrated that a first stage of numerosity processing could be traced as early as 75 ms after stimulus onset, likely arising from early visual area such as V2 or V3 (Fornaciai, Brannon, Woldorff, & Park, 2017; Fornaciai & Park, 2017; Park, Dewind, Woldorff, & Brannon, 2016). However, while much research has been dedicated to characterizing the effect of nonnumerical spatial features on perceived numerosity and the functional and neural properties of the conceptual representation of number, little is known about the *spatiotemporal* integration properties of such multiple processing stages and how they interact with each other to give rise to the visual number sense.

In the present study, we examined to what extent the spatial and temporal properties of dynamically presented dot arrays become integrated during numerosity perception. In particular, we sought to understand how various spatiotemporal integrations occur in the earlier stage (in the putative OLM) or in the later stage (in the putative summation coding stage) of numerosity processing. In Experiment 1, we investigated whether temporal modulations of object-position information are integrated over time, by asking participants to compare static and dynamic dot arrays where a portion of the dots either randomly changed position at different frequencies or moved along smooth trajectories. In Experiment 2, we pitted various dynamic dot array stimuli against each other, as opposed to making participants compare dynamic versus static arrays, to directly measure the relative effects of spatiotemporal integration and to disentangle numerosity integration occurring at the early object location map stage or at the later summation coding stage. In Experiment 3, we then investigated whether spatiotemporal integration depends strictly on luminance contrast or generalizes to stimuli defined by color contrast, which would require the involvement of color-sensitive areas. Finally, in Experiment 4, we further tested whether fluctuations in the numerical magnitude of dot array stimuli gets integrated over time to define the final numerical representation.

Materials and methods

Participants

A total of 243 subjects took part in the study (187 females, mean age = 21.1 ± 1.6 years old) after signing a written informed consent, and were rewarded with course credits. All participants were naive to the aims of the experiment, had normal or corrected-to-normal vision, and reported no history of neurological, attentional, or psychiatric disorder. All the experimental procedures were approved by the Institutional Review Board of the University of Massachusetts at Amherst, and were in line with the Declaration of Helsinki.

Apparatus and stimuli

The experiment took place in a large room, with participants tested in groups of five to 16 subjects in each experimental session. Each participant sat in front of an individual monitor screen and performed the experiment independently. Stimuli were generated by means of the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on Matlab (version r2013b; The Mathworks, Inc., Natick, MA), and presented on a monitor screen running at 60 Hz, with a resolution of 1280 × 1024 pixels, and encompassing approximately 38 × 30 degrees of visual angle from a viewing distance of about 57 cm.

Stimuli were randomly generated dot arrays, presented either in black on a gray background (Experiments 1, 2, and 4) or in green on a red background (Experiment 3). Dot diameter was set to 0.4°, and the items were randomly positioned inside an area of $113^{\circ 2}$ (diameter = 12°). Since some of the dynamic stimuli used across the experiments (see below) necessarily involve overlap and crossover between dots, we did not include any constraints on the minimum interdot distance in any of the stimuli (including the static arrays) to keep the low-level statistics similar across all the stimuli. However, note that the percentage of arrays with overlapping dots was very low and that such overlapping could happen in both static and dynamic stimuli. Therefore, it is unlikely that such an overlap would have affected the results in any systematic way.

General procedure

All the experiments included a first baseline numerosity discrimination task, comprising two blocks of 81 trials. Each individual condition of Experiments 1, 3, and 4 comprised four blocks of 162 trials, while Experiment 2 comprised four blocks of 81 trials.

In the baseline condition, while participants were asked to keep their gaze on a central fixation cross, probe (with variable numerosity) and reference stimuli were presented on the right and left of the fixation point (horizontal eccentricity = 10°), with their positions (right or left) randomly chosen for each trial. The reference stimulus always contained 20 dots, while the probe numerosity was randomly chosen on each trial (8, 11, 14, 17, 23, 27, 31, 35, or 40, with each level tested an equal number of times). The stimuli were presented for 500 ms, after which the central fixation cross turned red signaling participants to give a response by pressing the appropriate key on a keyboard, indicating which stimulus seemed to contain either the larger ("judge more" task) or the smaller ("judge less" task) amount of dots. After providing a response, the next trial started automatically after 350 ms. In the conditions following the baseline measurement, either the reference stimulus (Experiments 1, 3, and 4), or both probe and reference stimuli (Experiment 2), underwent spatiotemporal modulations according to the aim of the specific experiment (see Results). With the exception of the spatiotemporal modulation of the stimuli, the experimental procedures were identical for all the experiments. Note that different tasks (judge more and judge less) were tested separately on different groups of participants, in order to control for response biases induced by dynamic stimuli (see Results). Each block took about 5 min, and participants were free to take a break between different blocks. An entire experimental session took 45–50 min.

Data analysis

Subjects' responses as a function of the stimuli presented in each trial were pooled together, and each condition was analyzed separately. The proportion of responses as a function of probe numerosities was fitted with a cumulative Gaussian function, following the maximum likelihood method described by Watson (1979). As a measure of subject accuracy in the task, we took the point of subjective equality (PSE), defined as the median of the best-fitting cumulative Gaussian curve to all the data of a given condition. PSEs represent the probe numerosity perceptually matching (i.e., indistinguishable from) the reference stimulus, thus providing a measure of the perceived numerosity of the reference stimulus under different stimulation conditions (i.e., spatiotemporal modulations). As a measure of precision, we used the Weber fraction (WF), calculated according to the Weber-Fechner law as the ratio between the minimum discriminable increment (just noticeable difference; the standard deviation of the underlying Gaussian function) and the perceived magnitude of the reference stimulus (PSE).

In order to assess the changes in PSE caused by experimental manipulations, we calculated a perceived numerosity change index based on the difference in PSEs between a given test condition and the baseline condition:

$$\text{Perceived numerosity change index} = \left([PSE_{test} - PSE_{baseline}] \;/ PSE_{baseline} \right) \times 100 \tag{1}$$

where PSE_{test} refers to the PSE obtained in the conditions in which we manipulated the stimuli to test specific predictions (see below for more details), and $PSE_{baseline}$ refers to the PSE obtained in the baseline condition. Such an index represents the net effect of experimental manipulations on perceived numerosity.

Exclusion criteria were based on WF and the perceived numerosity change index. Namely, a given participant was excluded from further analysis if we observed either WF > 0.75 or perceived numerosity change index $> 100\%$. This procedure led to the exclusion of 69 individual conditions (out of 674, which was the total number of conditions tested across all the participants and experiments). Additionally, 20 more individual conditions were excluded from data analysis due to the inability of performing the judge less task (i.e., participants who selected the more numerous stimulus even in the judge less task). Exclusion criteria were applied independently to different conditions, so that the same participant could be excluded from one condition but kept in another. Doing so, statistical analyses were performed in a between-subjects fashion. The number of participants reported in the Results represents the actual number of subjects included in the analyses. Note that such relatively high number of excluded cases (about 10%) may have been driven by the little close supervision of the participants as we ran the experiments in large groups, which in turn may have resulted in occasional unmotivated participants (e.g., pressing keys at random). However, poor overall performance should not be related to our hypotheses of interest in any way.

An analysis of variance (ANOVA) was used to compare the dependent measures across multiple experimental conditions, and a one-sample t test was used to assess whether the distribution of numerosity change index is significantly greater than a hypothesized population mean that equals zero, representing the case where the experimental manipulation has no effect on numerosity perception. When needed (i.e., failure of normality test), a nonparametric test was used.

Results

Experiment 1

We first tested whether and how the spatiotemporal properties of dot-array stimuli get integrated and concur in shaping numerical representations.

In the random position change condition, the reference stimulus comprised a portion of dots randomly changing position at different frequencies—that is, disappearing and reappearing in new positions—but keeping the total number of dots constant throughout the duration of the presentation. In a 2×3 experimental design, we tested two different proportions of randomly changing stimuli (25% and 50%), and three different temporal frequencies (2, 6, or 8 Hz). For example, a condition of 25% and 6 Hz for a dot array containing 20 dots meant that each of the three times during the 500-ms presentation time, five of the 20 dots disappeared from their locations but five other dots appeared in new locations (Supplementary Movie 1). The average shift in position underwent by the randomly changing dots was about 5.4° ($SD = 0.57$; calculated across a simulation of 10,000 repetitions of the stimuli), which ensured that no apparent motion can be elicited by those shifts in position (i.e., Gepshtein, Tyukin, & Kubovy, 2007; Grossberg & Rudd, 1992). The lack of any motion energy due to the large spatial extent and instantaneous shifts in position assured that randomly changing stimuli represent a distinct class of stimuli compared to the smooth motion condition (see below), and could not be interpreted as extremely fast motion.

In the smooth motion condition, we tested whether motion integration is involved in this stage of analysis. There, static dot array images were compared with a dot arrays comprising a portion of moving dots (50% or 100%) with different speeds (3°/s, 6°/s, or 9°/s), in a 2×3 design similar to the random position change condition described above. For example, a condition of 50% and 3°/s for a dot array containing 20 dots meant that 10 of the 20 dots moved in random linear trajectories at a speed of 3°/s. Moving dots were contained within an invisible circle where the dots are drawn; the dots bounced back when they reached the boundary of that invisible circle (Supplementary Movie 2). Our main prediction about the effect of motion integration in the smooth motion condition concerns the possibility that it might prevent some of the effect provided by the spatiotemporal modulation. Namely, integrating the different positions of the dots across the motion paths might reduce any effect due to the shifts in position of the items. According to this prediction, the parameters of the moving dots were chosen to allow a more conservative comparison with the random position change condition. First, in the smooth motion condition we used larger proportions of moving dots. Furthermore, considering temporal frequency = speed / spatial frequency, and taking two times the size of an individual dot as a measure of spatial frequency (0.8°), the putative temporal frequency of the smooth motion stimuli is slightly higher than the corresponding conditions of the random position change stimuli (3.75, 7.50, or 11.25 Hz, respectively for 3, 6, or 9°/s).

Both conditions were performed separately in blocks and in a random order. In each condition, different combinations of proportion and temporal frequency/speed were randomly intermingled within each block and tested an equal number of times. However, since the dynamic dot arrays are intrinsically more salient compared to the static arrays, response biases could confound the effect of dynamic visual displays (e.g., one could argue that participants just responded to whichever side with more changes). To address this issue, we tested two independent groups of participants with different tasks: While one group was instructed to choose the stimulus containing more dots (judge more task; $N = 46$ and 47, respectively for the random position change and the smooth motion conditions), the other group was instructed to choose the stimulus containing the smaller amount of dots (judge less task; $N = 42$ and 44).

Figure 1 shows the results of the different conditions tested in Experiment1, both with the judge more (Figure 1A) and judge less (Figure 1B) tasks. First, the random position change stimuli were significantly overestimated compared to the static ones, in all the conditions tested (one-sample t tests; $t[45] = 7.2$, $t[45] = 9.8$, $t[45] = 10.9$, $t[45] = 9.54$, $t[40] = 12.57$, $t[45] = 15.07$, respectively for the six conditions reported in the leftmost panel of Figure 1A; all p values < 0.001). A two-way repeated measure ANOVA with factors proportion (25% and 50%) and temporal frequency (2, 6, 8 Hz) showed a statistically significant main effect of both factors, $F(1, 274) = 62.75$, $p < 0.001$, and $F(2, 274) = 31.25$, $p < 0.001$, respectively for proportion and temporal frequency, although with a significant interaction between the two factors, $F(3, 274) = 6.98$, $p = 0.002$. On the other hand, a post hoc multiple comparisons procedure (Holm-Sidak multiple comparisons) showed that overall the effect increased both as a function of proportion (25% vs. 50%, $t[274] = 7.92$, $p < 0.001$) and temporal frequency (2 Hz vs. 6 Hz:

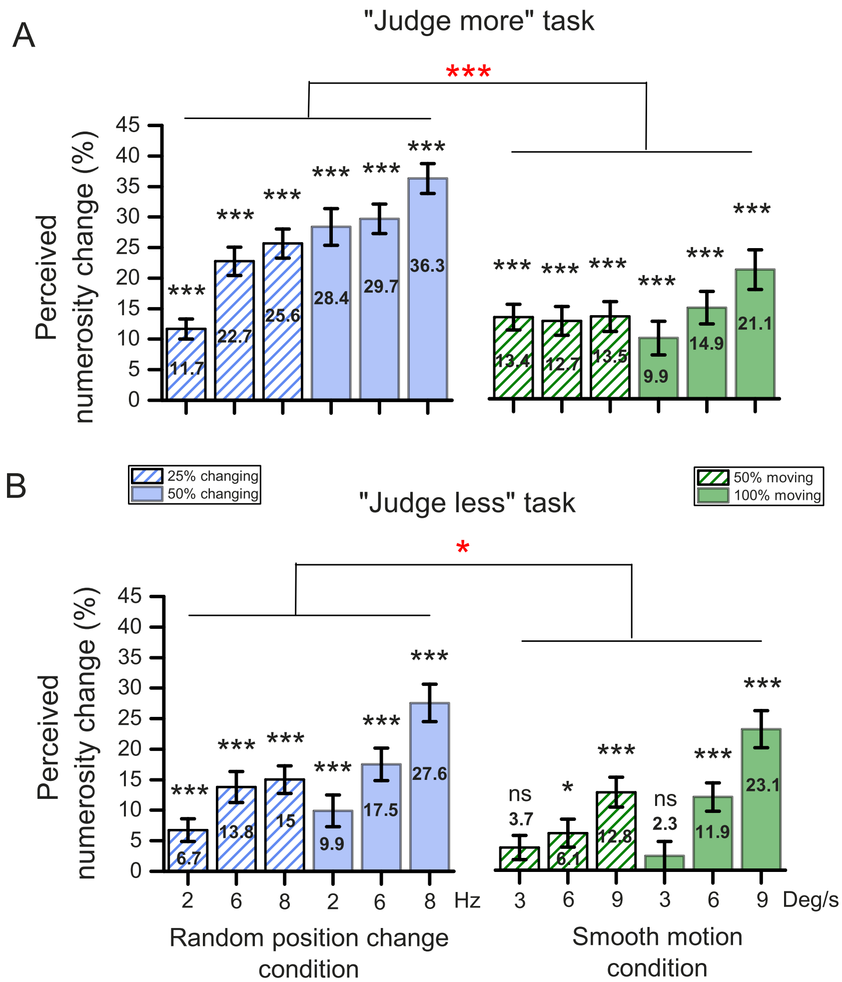

Figure 1. Results of Experiment 1. (A) Effects of dynamic stimuli (random position change and smooth motion) on perceived numerosity in Experiment 1, under the judge more task. While both conditions revealed substantial effects on perceived numerosity, the random position change condition resulted in greater overestimation than the smooth motion condition. (B) Effects of dynamic stimuli in Experiment 1 under the judge less task. The overall smaller effects compared to the judge more task suggests that response biases can affect numerosity discrimination, but accounting only for a limited proportion of the effect. Asterisks above each single bar represent the results of one-sample tests against a null hypothesis of zero perceived numerosity change. Red asterisks indicate the significance of the statistical tests used to compare the overall effect of the different classes of stimuli. Numbers within the bars or above them represent the numerosity change effect in the corresponding condition. Error bars represent *SEM*. ns = not significant, *$p < 0.05$, ***$p < 0.001$.

$t[274] = 4.46$, $p < 0.001$; 2 Hz vs. 8 Hz: $t[274] = 7.88$, $p < 0.001$; 6 Hz vs. 8 Hz: $t[274] = 3.42$, $p < 0.001$).

Smooth motion stimuli were similarly effective, resulting in a significant overestimation of the moving stimulus compared to the static one, $t(46) = 6.5$, $t(46) = 5.5$, $t(46) = 5.6$, $t(46) = 3.6$, $t(46) = 5.6$, $t(46) = 6.6$, all p values < 0.001. A two-way repeated measures ANOVA with factors proportion (50% and 100%) and speed (3°/s, 6°/s and 9°/s) further showed a significant main effect of speed, $F(2, 281) = 11.31$, $p < 0.001$, and no main effect of proportion, $F(1, 281) = 3.52$, $p = 0.067$, although again with a significant interaction between the two factors, $F(3, 281) = 11.06$, $p < 0.001$. A multiple comparisons procedure confirmed that the overall effect does not significantly increase with the proportion of moving dots (50% vs. 100%, $t[281] = 1.88$, $p = 0.067$), and showed that while the effect does not significantly increase with speed with the lower proportion (50%) of moving dots (3°/s vs. 6°/s: $t[281] = 2.95$, $p = 0.004$; 3°/s vs. 9°/s: $t[281] = 6.66$, $p < 0.001$; 6°/s vs. 9°/s: $t[281] = 3.71$, $p < 0.001$), it does so with the larger proportion (100%) of moving dots (3°/s vs. 6°/s: $t[281] = 0.37$, $p = 0.915$; 3°/s vs. 9°/s: $t[281] = 0.05$, $p = 0.96$; 6°/s vs. 9°/s: $t[281] = 0.43$, $p = 0.96$).

In order to compare the magnitude of the effects between the random position change and smooth motion conditions, we averaged the effects of all the different combination of temporal frequency/speed and proportion in each subject, and directly compared the distributions of average effects in the two experimental conditions. This test confirmed the stronger overestimation effect in the random position change condition compared to the smooth motion condition (two-sample *t* test, $t[91] = 3.78$, one-tailed *p* value < 0.001, Cohen's $d = 1.10$). The relatively smaller overestimation effect of the smooth motion condition indicates that the integration of spatial and temporal information is dampened by motion processing, suggesting that the motion processing system directly interacts with the numerosity processing system in the brain. However, as the overestimation effect provided by spatiotemporal modulation strongly depends on the parameters chosen for the two conditions—which can be only roughly compared—caution is needed when interpreting the comparison between the two classes of stimuli.

In the judge less task (Figure 1B), we found a similar pattern of results, with robust and significant changes in perceived numerosity induced by random position change stimuli (one-sample signed rank tests, $Z = 3.5$, $Z = 4.8$, $Z = 5.0$, $Z = 3.6$, $Z = 5.2$, $Z = 5.6$; all *p* values < 0.001). A two-way repeated measures ANOVA showed a significant main effect of both proportion, $F(1, 251) = 24.60$, $p < 0.001$, and temporal frequency, $F(2, 251) = 25.11$, $p < 0.001$, and a significant interaction between the two factors, $F(3, 251) = 11.23$, $p < 0.001$. A multiple comparisons procedure confirmed that the overestimation effect overall increased both as a function of proportion (25% vs. 50%, $t[274] = 4.96$, $p < 0.001$) and temporal frequency (2 Hz vs. 6 Hz: $t[251] = 4.18$, $p < 0.001$; 2 Hz vs. 8 Hz: $t[251] = 7.04$, $p < 0.001$; 6 Hz vs. 8 Hz: $t[251] = 2.86$, $p = 0.005$; although within the lower proportion, the comparison 6 Hz vs. 8 Hz did not reach significance: $t[251] = 0.25$, $p = 0.80$). Smooth motion stimuli caused weaker but systematic overestimation effects in most of the conditions, with the exception of the lower speed (3°/s) conditions, which were the only two conditions where there was no significant effect, $Z = 1.5$, $p = 0.12$; $Z = 2.2$, $p = 0.025$; $Z = 4.5$, $p < 0.001$; $Z = 0.27$, $p < 0.79$; $Z = 4.5$, $p < 0.001$; $Z = 5.5$, $p < 0.001$. A two-way repeated measures ANOVA again showed significant main effects of both proportion, $F(1, 263) = 15.24$, $p < 0.001$, and speed, $F(2, 263) = 54.56$, $p < 0.001$, and a significant interaction between the two factors, $F(3, 263) = 13.41$, $p < 0.001$, while multiple comparisons, differently from the results of the smooth motion condition in the judge more task, showed that the effect overall increased both as a function of proportion (50% vs. 100%, $t[263] = 3.90$, $p < 0.001$) and speed (3°/s vs. 6°/s: $t[263] = 4.12$, $p < 0.001$; 3°/s vs. 9°/s: $t[263] = 10.37$, $p < 0.001$; 6°/s vs. 9°/s: $t[263] = 6.25$, $p < 0.001$; although the comparison 6 Hz vs. 8 Hz within the lower proportion did not reach significance: $t[263] = 1.03$, $p = 0.30$).

Similarly to the judge more task, we directly compared the random position change and smooth motion condition by averaging in each subject the effects across the six combinations of parameters within each condition. Again, we found a statistically significant difference between the two conditions, $t(84) = 1.93$, one-tailed *p* value $= 0.029$, Cohen's $d = 0.59$, demonstrating that random position change stimuli caused a significantly stronger overestimation effect than the smooth motion condition.

Finally, we did not observe any substantial difference in precision (WF) across the different tasks and conditions, suggesting that the effect does not critically depend on an increased difficulty in performing the task. On the other hand, the results showed significantly stronger effects in the judge more task, suggesting that response biases partially contributed to the magnitude of the effect (see Supplementary Materials), also potentially providing ceiling effects in some of the conditions (i.e., the three conditions corresponding to the lower proportion of moving dots in the smooth motion condition with the judge more task, where the effect seems identical independently of speed).

Experiment 2

While in Experiment 1 we observed fairly strong effects of the various dynamic visual displays used, the intrinsic properties of the dot array stimuli leave some important questions unanswered.

First, is the effect of the random position change condition driven by the changes in the dot position (i.e., spatial modulation) or by the number of repositioning events in a given time frame (i.e., temporal modulation)? An overestimation driven by spatial modulation must reflect the integration of information over time at the putative OLM (because the new positions would increase the total amount of activity across the map) rather than at the putative summation coding stage (because the total number of dots was constant at any time, so that a summation of the activity representing these dots would remain constant). In contrast, an overestimation driven by temporal modulation could reflect the integration of information at the putative summation coding stage, where the perceived numerosity of dynamic dot arrays might be biased by the concurrent accumulation of sequential events. Since in Experiment 1 the rate of position changes coincided with the rate of temporal modulation, the two effects could not be disentangled. Therefore, in Experiment 2 we asked participants to directly compare a flickering dot array (stationary flicker condition), which provides a purely temporal modulation that might bias numerosity via an increased number of events, against a flickering dot array also modulated in position as in the random position change ("flicker with random position change" condition).

Second, is the difference between the random position change and the smooth motion condition due to qualitative differences between the two classes of stimuli? Particularly, while moving stimuli smoothly follow linear trajectories, randomly changing stimuli undergo abrupt changes that might have somehow resulted in stronger effects. To answer this question, we asked participants to compare flickering stimuli undergoing linear position changes that resulted in apparent motion trajectories (i.e., stimuli moved along motion trajectories, but were shown only for 32 ms at regular intervals; henceforth called the "apparent motion" condition) versus dot arrays under the flicker with random position change condition described above. Thus, the stimuli were qualitatively more similar in the two conditions, and were distinguished only by the dots jumping to new random positions or according to linear motion paths. Additionally, stimuli under the apparent motion condition were also compared to stationary flicker stimuli, in order to assess whether and to what extent motion integration might prevent distortions due to purely temporal modulation.

Flickering dots (either static or subject to position changes) were presented on the screen for 32 ms with a frequency of 6 Hz. Both flicker with random position change and apparent motion stimuli were shown in new positions at the same rate of flicker, and moving stimuli moved from one presentation to another at 6°/s. The proportions of dots changing position were 80% for the flicker with random position change and 100% for the apparent motion condition (Supplementary Movie 3). As in Experiment 1, we used a larger proportion for apparent motion stimuli to perform a more conservative comparison with randomly-changing stimuli. In Experiment 2, participants were tested only with the judge less task ($N = 33$).

Figure 2 shows the results of Experiment 2. First, flicker with random position change stimuli were significantly overestimated compared to the stationary flicker stimuli, suggesting that the overall overestimation effect could not be explained by the temporal modulation alone, and that the position changes actually contribute to the effect (5.9% ± 2.23% average change in perceived numerosity; one-sample t tests, $t[32] = 2.65$, $p = 0.012$). Second, directly comparing flicker with random position change stimuli with apparent motion stimuli confirmed that even if both stimuli involved discrete changes in position over time, a random position change resulted in a relative overestimation compared to the dots moving in implied motion trajectories (4.1% ± 2.71%; one-sample signed rank test, $Z = 2.01$, $p = 0.045$). As in Experiment 1, this result indicates that the motion processing system directly influences the spatiotemporal integration mechanism for numerosity perception. Finally, comparing stationary flicker stimuli versus apparent motion stimuli, revealed a relative overestimation of the stationary flicker dot arrays (5% ± 2.15%; one-sample t test, $t[32] = 2.36$, $p = 0.024$), demonstrating that even the purely temporal modulation of flickering dots in the absence of position changes causes a stronger effect compared to stimuli moving along implied trajectories, and showing that at least part of the effect of motion integration consists in dampening the integration of additional events created by temporal modulations.

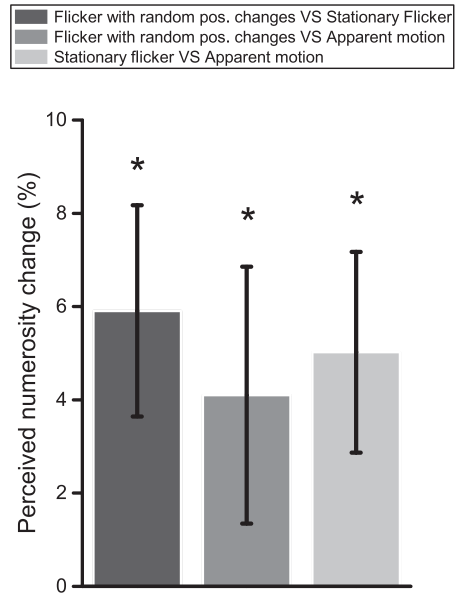

Figure 2. Results of Experiment 2. The results show that the stimuli combining flicker with random position change are overestimated compared to both the stationary flicker and apparent motion stimuli, demonstrating that position changes can be dissociated from purely temporal modulation (left bar), and that apparent motion stimuli have a weaker effect than flickering stimuli with random position change (middle bar). Also shown is that stationary flicker stimuli are overestimated when compared to apparent motion stimuli, demonstrating that motion integration can prevent overestimation driven by temporal modulation (right bar). Asterisks refer to the results of one-sample tests. *$p < 0.05$.

Experiment 3

Collectively, the results of Experiments 1 and 2 demonstrate that: (a) the spatiotemporal properties of dot array stimuli (i.e., shifts in positions) are integrated over time; (b) the effect could not be explained solely by temporal modulation, suggesting that integration occurs both in the early putative OLM stage and the later putative summation coding stage; and (c) that motion-sensitive areas are likely involved in numerosity processing, allowing to reduce the distorting effect of spatiotemporal modulations.

In Experiment 3, we further asked whether and to what extent other areas along the visual processing stream, such as color-sensitive areas along the ventral (V4) or dorsal (V3A) stream, influence the spatiotemporal integration mechanism in numerosity processing. To address this possibility, we assessed whether chromatic-contrast–defined dynamic stimuli can induce similar overestimation effects. Indeed, without any feedback from specific color-sensitive areas, the early visual cortices thought to underpin numerosity processing (i.e., V2/V3; Fornaciai et al., 2017) are generally poorly sensitive to the phase of chromatic stimuli—which make them unsuitable for segmenting chromatic objects (Castaldi, Frijia, Montanaro, Tosetti, & Morrone, 2013), and hence to discriminate changes in position across the area of the stimulus.

To achieve physical equiluminance, we measured the oscillatory signal of a photodiode driven by an alternating green and red stimulus, and chose the values of green and red minimizing the difference in the photodiode signals elicited by them. We then used such parameters for all the experimental sessions in Experiment 3. Note that physical equiluminance does not assure the elimination of all achromatic luminance-contrast–sensitive responses, as the relative luminance level required to completely eliminate luminance-sensitive responses has systematic interindividual variations. Nevertheless, physical equiluminance represents a close approximation of actual (perceived) equiluminance, likely significantly reducing the involvement of luminance-contrast–sensitive mechanisms. In this equiluminance condition, all the experimental procedures were identical to the random position change condition in Experiment 1, with the exception of the colors of background (red) and stimuli (green; Supplementary Movie 4). Participants were tested either with the judge more task ($N = 27$) or with the judge less task ($N = 34$).

Figure 3 shows the results of the equiluminance condition in Experiment 3. Results with the judge more task showed robust and significant biases in perceived numerosity with equiluminant dynamic stimuli (one-sample signed rank test, $Z = 4.51$, $Z = 4.51$, $Z = 4.54$, $Z = 4.20$, $Z = 4.46$, $Z = 4.49$, respectively for the six conditions reported in Figure 3A; all p values < 0.001). Similar significant biases were observed also with the judge less task (one-sample t tests, $t[33] = 3.5$, $t[33] = 6.3$, $t[33] = 7.6$, $t[33] = 3.9$, $t[33] = 6.9$, $t[33] = 8.8$; all p values ≤ 0.001; Figure 3B). A two-way ANOVA further showed a significant main effect of both proportion, $F(1, 198) = 6.9$, $p = 0.009$, and temporal frequency, $F(2, 198) = 23.6$, $p < 0.001$, with the effect on perceived numerosity increasing with increasing proportion of changing dots and increasing temporal frequency. Crucially, we found no difference between the effects provided by luminance- (Experiment 1) and color-defined stimuli, in both judge more (ANOVA on ranks, $H = 1.1$, $p = 0.29$) and judge less ($H = 0.21$, $p = 0.65$) tasks. Collectively, these results suggest that luminance contrast plays little role in the spatiotemporal integration mechanism of numerosity perception, as the effect in Experiment 3 is virtually identical to that in Experiment 1, and that such a mechanism could be equally driven by the color processing system in the visual cortex. However, since using physical equiluminance instead of perceived equiluminance is likely to leave some residual luminance-contrast due to interindividual differences in the equiluminance point, caution is needed when interpreting these results. Indeed, it is possible that an identical effect compared to Experiment

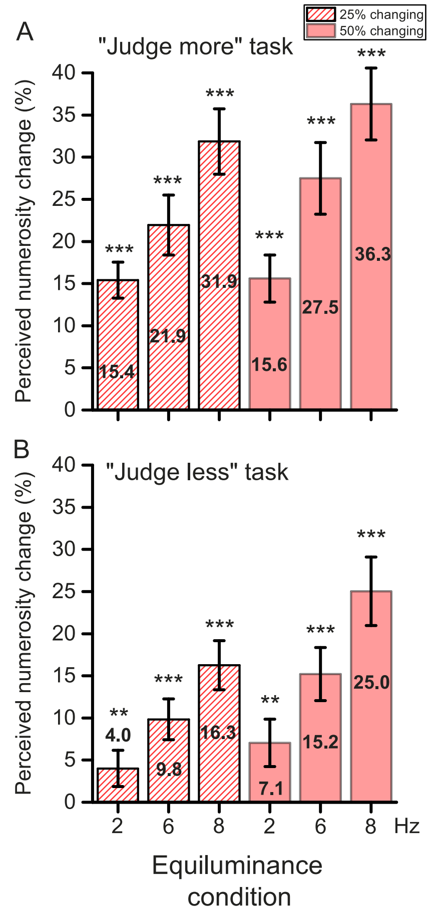

Figure 3. Results of Experiment 3. (A) Changes in perceived numerosity due to randomly changing equiluminant stimuli under the judge more task. (B) Perceived numerosity change effects under the judge less task. In all the conditions tested with equiluminant stimuli, random position changes at different temporal frequencies caused robust biases in perceived numerosity. Asterisks represent the results of one-sample tests against a null hypothesis of zero numerosity change (marked above each bar). Numbers within the bars or above them represent the numerosity change effect in the corresponding condition. Error bars represent *SEM*. $^{**}p < 0.01$, $^{***}p < 0.001$.

1 may be due to such residual luminance—a possibility that limits the conclusions that we could draw from this condition. Further studies should assess this effect using stimuli at perceived equiluminance.

Experiment 4

Finally, in Experiment 4, we investigated whether fluctuation in the numerical magnitude of stimuli gets integrated over time similarly to sequential and positional information. Indeed, in Experiments 1 through 3 we always kept the numerosity of the reference stimulus constant, while we introduced spatiotemporal manipulation aimed to reveal the integration properties of the OLM and summation coding stages. What these manipulations showed is that both temporal and positional information are integrated over time, causing overestimation of the dynamic stimuli. However, one remaining question concerning the spatiotemporal properties of the numerosity processing stream is whether fluctuations in the numerical magnitude of a stimulus (i.e., variable amount of visible dots across the presentation of the stimulus) gets similarly integrated over time.

To investigate this issue, we devised a new "dimming" condition where we presented a dot array in which a portion of dots smoothly changed luminance effectively making those dots disappear and reappear at the same locations (i.e., transitioning from black to the same gray as the background and vice versa, with 32-ms ramps at offset and onset). On the one hand, if the summation coding integrates numerosity over time, the fluctuating numerical magnitude of the dimming stimuli should be reflected by relative underestimation compared to a stable stimulus. On the other hand, however, the changes in luminance making the dots disappear and reappear provide additional temporal modulation that might actually increase the perceived magnitude of the dimming stimuli. As in the previous conditions, we tested both different proportions of dimming stimuli (25% and 50%) and different temporal frequencies (4, 6, and 10 Hz) in a 2 × 3 design. For example, a condition of 50% and 4 Hz for a dot array containing 20 dots meant 10 of the 20 dots were smoothly reduced in luminance until they disappeared for about 16 ms and then smoothly reappeared in the same positions every 250 ms. The phase of the dimming dots was synchronized (Supplementary Movie 5). Thus, by considering the magnitude and direction of perceived numerosity change, we considered three possible outcomes: (a) greater underestimation associated with higher temporal frequency and higher proportion of dimming dots, meaning that integration of numerosity over time (i.e., the fluctuating number of dots) is the leading phenomenon; (b) greater overestimation associated with higher temporal frequency and higher proportion of dimming dots, indicating that the integration of sequential information (i.e., number of events) has the strongest influence on numerical estimates; and (c) an intermediate effect, suggesting a concurrent integration of sequential and simultaneous numerical information. Again, participants were tested either with the judge more task ($N = 29$) or with the judge less task ($N = 41$).

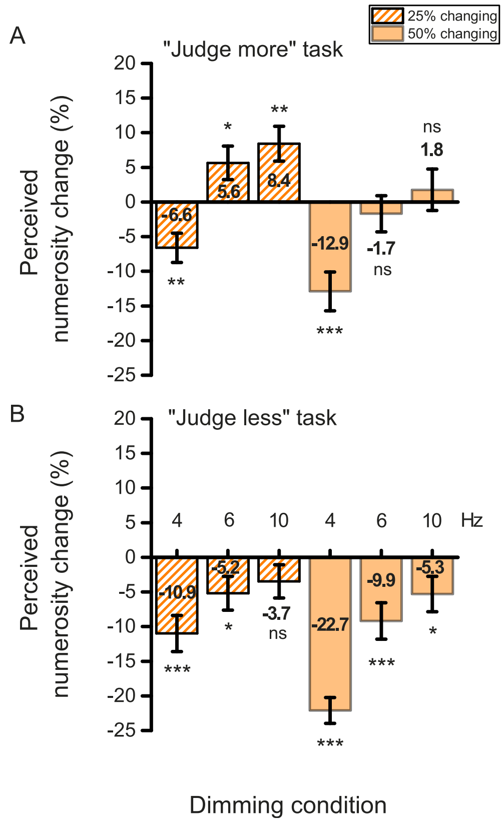

Figure 4. Results of Experiment 4. (A) Effects of dimming stimuli smoothly disappearing and reappearing, with the judge more task. (B) Effects of dimming stimuli with the judge less task. The results suggest that two processes are involved in the observed effect: decrease in perceived numerosity due to spatial integration of the number of items over time, and an increase in perceived numerosity due to integration of sequential information (i.e., the number of events, increasing with the rate of dimming). Numbers within the bars or above them represent the numerosity change effect in the corresponding condition. Asterisks represent the results of one-sample tests. Error bars represent *SEM*. ns = not significant, *$p < 0.05$, **$p < 0.01$ ***$p < 0.001$.

As shown in Figure 4A and B, we found markedly different results compared to the previous experiments. In the judge more task condition (Figure 4A), with the lower proportion of dimming stimuli (25%) we found a significant underestimation in the slower condition (4

Hz, $t[28] = -3.2$, $p = 0.003$), while the faster temporal frequencies gave rise to significant overestimation (6 Hz, $t[28] = 2.4$, $p = 0.02$; 10 Hz, $t[28] = 3.4$, $p = 0.002$). With the higher proportion (50%), we found a significant underestimation in the 4 Hz condition, $t(28) = -4.7$, $p < 0.001$, while the effects in both 6 Hz, $t(28) = -0.66$, $p = 0.51$, and 10 Hz, $t(28) = 0.59$, $p = 0.55$, conditions were not significantly different from zero. On the other hand, in the judge less condition (Figure 4B), the overall trend was toward underestimation, with almost all statistically significant effects (one-sample t tests, $t[40] = -5.3$, $p < 0.01$, $t[40] = -2.1$, $p = 0.04$, $t[40] = -1.03$, $p = 0.31$, $t[40] = -12.9$, $p < 0.001$, $t[40] = -4.8$, $p < 0.001$, $t[40] = -2.4$, $p = 0.023$). A two-way ANOVA confirmed a significant main effect of both factors proportion and temporal frequency in the judge more task, $F(1, 168) = 8.6$, $p = 0.004$, and $F(2, 168) = 15.4$, $p < 0.001$, while we found a significant interaction in the judge less task ($F[1, 2] = 6.7$, $p = 0.001$; Holm-Sidak multiple comparisons, 25% proportion: $t = 3.01$, $p = 0.009$; $t = 2.24$, $p = 0.051$; and $t = 0.77$, $p = 0.082$, respectively for 6 Hz vs. 4 Hz and 10 Hz vs. 6Hz; 50% proportion: $t = 5.7$, $p < 0.001$, $t = 3.97$, $p < 0.001$, $t = 5.17$, $p = 0.082$, respectively for 10 Hz vs. 4 Hz, 6 Hz vs. 4 Hz and 10 Hz vs. 6 Hz). No differences in WF were observed across the different conditions (ANOVAs on ranks; p values > 0.262).

Overall, these results demonstrate that the final numerical representation of a fluctuating stimulus is defined by integrating its numerosity over time—which could be described as an average of the items contained in the stimulus across the presentation—and that such integration interacts with the additional sequential information created by the fluctuations themselves. Consider for instance the first three conditions in the judge more task (Figure 4A). The slowest (4 Hz) conditions showed a significant underestimation, suggesting that the lower number of dots presented from time to time contributes to the final numerosity estimate. Increasing the flicker temporal frequency (and hence the number of events), however, caused a reversal of the effect, demonstrating that integration in the temporal domain takes a leading role with increasing frequency.

Discussion

In the present study, we aimed to characterize how the spatiotemporal features of visual stimuli affect the representation of approximate numerical magnitudes. Despite the large number of studies dedicated to uncovering the functional and neural properties of numerosity perception (e.g., Burr & Ross, 2008; Cicchini et al., 2016; Fornaciai & Park, 2017; Harvey et al., 2013; Park, 2017; Park et al., 2016; Piazza et al., 2007)—which collectively revealed the existence of a dedicated and generalized sense of number (but see Dakin, Tibber, Greenwood, Kingdom, & Morgan, 2011; Durgin, 2008; and Gebuis et al., 2016, for arguments against this view)—how the spatiotemporal properties of dynamic visual displays shape numerical processing, and at which level of the visual processing stream such properties might be integrated, is still not clear. Also, more recent reports have started to reveal a cascade of processing stages in the visual stream subserving visual numerosity perception (Fornaciai et al., 2017; Park et al., 2016), consistent with the idea put forth in computational models in which a first crucial step for numerosity processing starts from object segmentation and normalization in the so-called OLM (Dehaene & Changeux, 1993; Verguts & Fias, 2004). However, besides the putative role of the OLM in normalizing the spatial features of numerical stimuli, little is known about the other features of such processing stage. Thus, to investigate these issues, we tested whether and how concurrent spatiotemporal modulations get integrated in the process of numerosity perception, and whether the earlier levels of numerosity processing integrates information over time.

Results from Experiment 1 demonstrate that both randomly changing and smoothly moving stimuli are systematically overestimated compared to static ones, although smoothly moving stimuli result in significantly smaller effects. Regarding the effect of randomly changing stimuli, a crucial question left open by Experiment 1 is: Does the observed overestimation actually depend on the changes in position, or does the effect only reflect the temporal modulation of dots over time (i.e., the series of events resulting from the position shifts)? Results from Experiment 2 clearly show that object-position information actually plays a role in the observed effect, suggesting that the putative OLM is capable of integrating information over time (i.e., activating new positions increases the overall activation across the map). However, the relatively small effect of additional position shift in Experiment 2 (static flicker versus flicker with random position change) additionally suggest that the OLM have weaker and/or shorter integration capacities, while the overall stronger effect of temporal modulation suggests that sequential information is more likely to be integrated at a higher level stage showing stronger/longer integration properties potentially at the summation coding stage.

Regarding the smooth motion condition of Experiment 1, the relatively smaller overestimation effect in the smoothly moving dot arrays suggests the involvement of motion integration, likely by means of feedback projections from higher level motion processing areas. By analyzing moving items using spatiotemporal receptive fields—such as those hypothesized to underlie motion processing in motion-sensitive areas (e.g., Burr & Thompson, 2011), the activity generated by the same object shifting to new locations (in smooth transition) is somehow discarded as it does not represent a new object appearing in another portion of the visual field. Even if early extrastriate areas (i.e., V2/V3) have very limited motion-processing capacities, this finding seems consistent with the observation that motion areas can exert a strong influence on earlier areas. Indeed, it has been shown that motion area V5/MT sends robust feedback to early visual cortex (i.e., V1), allowing early areas to represent complex features such as apparent motion (Muckli, Kohler, Kriegeskorte, & Singer, 2005; Sterzer, Haynes, & Rees, 2006). Such reduction in the overestimation effect could result from two different processes, as motion integration might result in a reduced integration of temporal or positional features. Even if our results cannot disentangle the two possibilities, it is likely that the difference in the effect reflects a combination of the two processes—that is, both a reduced integration of new positions and a reduced accumulation of temporal events due to the smooth transition of the moving dots.

On the one hand, one may argue that the difference between the effect of random position changes and motion could be driven by purely qualitative differences in the stimuli. In the random position change condition, dots undergo sudden jumps in position, which can be described as a strong transient event. In contrast, in the smooth motion condition the dots undergo smooth changes, with stimuli appearing in adjacent positions in successive points in time. This difference (i.e., abrupt versus smooth changes) could have somehow resulted in an overall stronger effect of abruptly changing stimuli. We directly addressed this issue in Experiment 2, by using the apparent motion condition in which moving stimuli underwent similar abrupt changes in position, with the difference of following a linear apparent motion trajectory. The results showed that the apparent motion stimuli are nevertheless underestimated compared to the random position change stimuli (and even compared to static flicker stimuli), supporting the idea that motion integration is involved in the processing of such stimuli. On the other hand, one might consider stimuli with random position changes as a special case of stimuli with smooth motion, as sudden changes in position might be interpreted as motion with extremely high speed. Indeed, when the right combination of spatiotemporal parameters is used, two separate flashes can easily be perceived as continuous motion along a defined path (e.g., Grossberg & Rudd, 1992). However, the spatiotemporal parameters used in the current random position change stimuli make it extremely unlikely for such dynamics to be perceived as a continuous motion. Specifically, the spatial extent of the random position changes was quite large ($5.4° \pm 0.57°$) and position shift occurred instantaneously. Under such stimulation condition, it is very difficult—if not impossible—to induce the perception of apparent motion, ensuring that our random position change stimuli are well beyond the capacities of motion-sensitive mechanisms. However, considering that the observed overestimation effect is highly dependent on the specific parameters chosen in the two conditions, and that only a rough comparison could be made between speed and temporal frequency (i.e., see Results), caution is in order when interpreting the comparison between the effects provided by the two classes of stimuli.

Besides the feedback from motion-sensitive areas, we also investigated whether color-sensitive areas are involved in the numerosity processing stream, by testing the effect of random position changes with physically equiluminant stimuli. The results of this condition revealed an effect very similar to what we observed with luminance-defined dot arrays. A first interpretation of this result concerns the possibility that the observed effect is relatively independent from luminance contrast. According to this view, the virtually identical effects found in Experiment 1 (random position change condition) and Experiment 3 suggest that a combination of temporal as well as positional modulation is involved in the effects in both experiments, which in turn implies that the OLM processes object-position information even in the case of equiluminant stimuli. This is surprising, since a relatively low-level (V2/V3: Fornaciai et al., 2017) OLM by itself should be poorly sensitive to position changes in chromatic-contrast–defined dot arrays. Namely, while areas like V2 or V3 can respond to purely chromatic information, their responses are not selective for phase-congruency of the visual images (Castaldi et al., 2013), suggesting a little involvement in contours detection and segmentation. Phase sensitivity for chromatic stimuli emerges only at later stages, at the level of V3A in the dorsal stream and V4 in the ventral stream (Castaldi et al., 2013). Thus, an OLM located at the level of V2/V3 should be poorly suited to segment chromatic objects, and changes in position would be indistinguishable at this level. So overall, according to this interpretation, the results of Experiment 3 suggest that color-sensitive areas may contribute to numerosity processing by means of feedback to the OLM, in line with previous findings showing robust feedback signals from V4 to earlier visual areas (Zeki & Shipp, 1989). Also, these results are in line with previous studies showing that numerosity perception does not depend on the first-order luminance-contrast properties of visual stimuli, but works equally well with second-order contrast-defined stimuli (Kramer, Di Bono, & Zorzi, 2011). However, our behavioral results do not allow us to pinpoint the target and the specificity of the color-sensitive signals needed to segment equiluminat stimuli. While a color-sensitive mechanism specifically subserving numerosity perception—hence enabling object segmentation at the level of the putative object location map—is possibile, it is equally possible that such signals serve a more general purpose. According to this latter idea, color-sensitive signals might be directed to other lower level areas like V1, with the purpose of equiluminant edge detection. After that, numerosity-specific mechanisms could use the newly segmented output from low-level areas to detect changes in position.

A second interpretation concerns the possibility that the effect in Experiment 3 might be driven by residual luminance contrast. Indeed, since we used physical, instead of perceived, equiluminance, it is likely that residual luminance-contrast might have triggered the response of luminance-contrast-sensitive mechanisms. In this case, the identical results in Experiments 1 and 3 would be due to the same luminance-contrast–sensitive mechanisms, which could provide similar responses even with different ranges of luminance. As we could not disentangle these two interpretations, limiting the possibility to draw firm conclusions from physically equiluminant stimuli, further studies controlling for residual luminance contrast are needed in order to clarify the involvement of color-sensitive mechanisms in numerosity processing.

Regarding the integration of dot arrays showing a fluctuating numerical magnitude over time, our results in Experiment 4 suggest that the final numerical representation is defined by taking into account the variable number of dots contained in the stimulus. However, this biased representation seems to be counteracted by the presence of temporal modulations, which additionally demonstrate that different integration processes operates in parallel to define numerical representations.

One remaining question from an anatomical point of view is whether the putative OLM resides in one early sensory region receiving feedback from other higher level areas or whether multiple brain regions function as the putative OLM. Our interpretation of the current results is more aligned with the former view where the OLM resides in V2/V3 receiving feedback from color- and motion-sensitive regions, and integrating positional information over time. However, it is also possible that those higher level cortical areas themselves may have the capacity to process object segmentation and normalization, thus representing additional stations in the numerosity processing stream. Future studies with a focus on anatomical origins of the current effects should be able to distinguish these hypotheses. Regardless of these two hypotheses, we argue that the current effects of spatiotemporal modulations at least partially occur prior to the summation coding stage in the parietal cortex: while the effect of temporal modulation and fluctuating numerosity are well explained by integration at the summation coding, the effect of positional information in Experiments 1, 2, and 3 is better explained by integration at the level of the OLM. This interpretation is based on the rationale that summation coding should be blind to object locations in the putative OLM and should only receive information about the overall activation across the OLM, an idea that is consistent with the influential framework by Lennie (1998).

Finally, as we did not include any control for nonnumerical visual attributes (i.e., density, area), one may argue that the observed effects might concern other continuous magnitude properties of the visual images instead of numerosity. However, much evidence from previous studies suggests that using dot array stimuli numerosity represents the most relevant information even in passive-viewing paradigms (Fornaciai et al., 2017; Fornaciai & Park, 2017; Park et al., 2016), and likely even more so when it comes to explicit numerical judgments (e.g., Anobile et al., 2016a; Cicchini et al., 2016; DeWind, Adams, Platt, & Brannon, 2015; but see Durgin, 2008, and Dakin et al., 2011, for positions opposing to this ideas). Also, while there is a compelling rationale to hypothesize an effect of spatiotemporal modulations on numerical estimates (i.e., increased number of events, modulation of positional information), effect at the level of other dimensions (i.e., texture density) would be much harder to explain. For these reasons, we believe that numerical discrimination performances measured under such circumstance actually reflect processes related to numerosity perception, rather than other visual properties.

To conclude, our results show that the spatiotemporal properties of visual dot array stimuli strongly affect numerical discrimination performances, via integration of additional temporal and positional information likely occurring at both the early and higher level stages of numerosity processing. These integration properties involving signals from multiple midlevel visual areas are in line with the idea of an abstract sense of number, where the numerosity readout at higher level stages benefits from processing across several lower level visual areas specifically tuned to different attributes. Our findings demonstrate that the putative object location map stage is much more complex than previously thought, integrating positional information over time and, at least indirectly, exploiting different feedback inputs from higher level visual areas. These findings illustrate that early sensory areas carry out a large amount of work for the processing of numerical information, integrating several spatiotemporal features of the stimuli, and that numerosity processing is not a linear feed-forward mechanism, but involves interactions and feedback across a network of visual areas.

Keywords: numerosity perception, visual cortex, spatiotemporal integration, motion perception, color perception

Acknowledgments

This study was supported by the National Science Foundation Career Award (BCS1654089) to JP. We thank Abigail Farrell and Brianna Southwick for their assistance in data collection. MF and JP devised the experiments. MF collected and analyzed the data. MF and JP interpreted the results and wrote the manuscript.

Commercial relationships: none.
Correspondence: Michele Fornaciai.
Email: mfornaciai@umass.edu.
Address: Department of Psychological & Brain Sciences and Commonwealth Honors College, University of Massachusetts, Amherst, MA, USA.

References

Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Animal Cognition*, *11*(3), 495–503, doi:10.1007/s10071-008-0140-9.

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

Anobile, G., Cicchini, G. M., & Burr, D. C. (2016). Number as a primary perceptual attribute: A review. *Perception*, *45*(1–2), 5–31, doi:10.1177/0301006615602599.

Arrighi, R., Togoli, I., & Burr, D. C. (2014). A generalized sense of number. *Proceedings of the Royal Society B: Biological Sciences*, *281*(1797), 20141791, doi:10.1098/rspb.2014.1791.

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436, doi:10.1163/156856897X00357.

Burr, D., & Ross, J. (2008). A visual sense of number. *Current Biology*, *18*, 425–428, doi:10.1016/j.cub.2008.02.052.

Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985–2010. *Vision Research*, *51*(13), 1431–1456. https://doi.org/10.1016/j.visres.2011.02.008

Castaldi, E., Aagten-Murphy, D., Tosetti, M., Burr, D., & Morrone, M. C. (2016). Effects of adaptation on numerosity decoding in the human brain. *NeuroImage*, *143*, 364–377. https://doi.org/10.1016/j.neuroimage.2016.09.020

Castaldi, E., Frijia, F., Montanaro, D., Tosetti, M., & Morrone, M. C. (2013). BOLD human responses to chromatic spatial features. *European Journal of Neuroscience*, *38*(2), 2290–2299, doi:10.1111/ejn.12223.

Chen, Q., & Verguts, T. (2013). Spontaneous summation or numerosity-selective coding? *Frontiers in Human Neuroscience*, *7*, 886, doi:10.3389/fnhum.2013.00886.

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

Dakin, S. C., Tibber, M. S., Greenwood, J. A., Kingdom, F. A. A., & Morgan, M. J. (2011). A common visual metric for approximate number and density. *Proceedings of the National Academy of Sciences*, *108*(49), 19552–19557, doi:10.1073/pnas.1113195108.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience*, *5*, 390–407, doi:10.1162/jocn.1993.5.4.390.

DeWind, N. K., Adams, G. K., Platt, M. L., & Brannon, E. M. (2015). Modeling the approximate number system to quantify the contribution of visual stimulus features. *Cognition*, *142*, 247–265, doi:10.1016/j.cognition.2015.05.016.

Durgin, F. H. (2008). Texture density adaptation and visual number revisited. *Current Biology*, *18*(18), R855–R856, doi: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, doi:10.1016/j.neuroimage.2017.05.069.

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

Fornaciai, M., & Park, J. (2017). Distinct neural signatures for very small and very large numerosities. *Frontiers in Human Neuroscience*, *11*, 21, doi:10.3389/fnhum.2017.00021.

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, doi:10.1016/j.actpsy.2016.09.003.

Gepshtein, S., Tyukin, I., & Kubovy, M. (2007). The economics of motion perception and invariants of visual sensitivity. *Journal of Vision*, *7*(8):8, 1–18, doi:10.1167/7.8.8. [PubMed] [Article]

Grossberg, S., & Rudd, M. E. (1992). Cortical dynamics of visual motion perception: Short-range and long-range apparent motion. *Psychological Review*, *99*(1), 78–121, doi:10.1037/0033-295X.99.1.78.

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

Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences, USA*, *106*(25), 10382–10385, doi:10.1073/pnas.0812142106.

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

Kramer, P., Di Bono, M. G., & Zorzi, M. (2011). Numerosity estimation in visual stimuli in the absence of luminance-based cues. *PLoS ONE*, *6*(2), e17378, doi:10.1371/journal.pone.0017378.

Lennie, P. (1998). Single units and visual cortical organization. *Perception*, *27*(8), 889–935, doi:10.1068/p270889.

Muckli, L., Kohler, A., Kriegeskorte, N., & Singer, W. (2005). Primary visual cortex activity along the apparent-motion trace reflects illusory perception. *PLoS Biology*, *3*(8), e30265, doi:10.1371/journal.pbio.0030265.

Nieder, A. (2016). The neuronal code for number. *Nature Reviews Neuroscience*, *17*(6), 366–382. https://doi.org/10.1038/nrn.2016.40

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

Park, J., Dewind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and direct encoding of numerosity in the visual stream. *Cerebral Cortex*, *26*(2), 748–763, doi:10.1093/cercor/bhv017.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442, doi:10.1163/156856897X00366.

Pepperberg, I. M. (2006). Cognitive and communicative abilities of Grey parrots. *Applied Animal Behaviour Science*, *100*(1–2), 77–86, doi:10.1016/j.applanim.2006.04.005.

Piantadosi, S. T., & Cantlon, J. F. (2017). True numerical cognition in the wild. *Psychological Science*, *28*(4), 462–469, doi:10.1177/0956797616686862.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293–305, doi:10.1016/j.neuron.2006.11.022.

Rugani, R., Vallortigara, G., Priftis, K., & Regolin, L. (2015). Number-space mapping in the newborn chick resembles humans' mental number line. *Science*, *347*(6221), 534–536, doi:10.1126/science.aaa1379.

Sengupta, R., Surampudi, B. R., & Melcher, D. (2014). A visual sense of number emerges from the dynamics of a recurrent on-center off-surround neural network. *Brain Research*, *1582*, 114–124, doi:10.1016/j.brainres.2014.03.014.

Sterzer, P., Haynes, J. D., & Rees, G. (2006). Primary visual cortex activation on the path of apparent motion is mediated by feedback from hMT+/V5. *NeuroImage*, *32*(3), 1308–1316, doi:10.1016/j.neuroimage.2006.05.029.

Stoianov, I., & Zorzi, M. (2012). Emergence of a "visual number sense" in hierarchical generative models. *Nature Neuroscience*, *15*(2), 194–196, doi: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, doi:10.1162/0898929042568497.

Watson, A. B. (1979). Probability summation over time. *Vision Research*, *19*(5), 515–522, doi:10.1016/0042-6989(79)90136-6.

Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, *89*(1), B15–B25, doi:10.1016/S0010-0277(03)00050-7.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*(1), B1–B11, doi:10.1016/S0010-0277(99)00066-9.

Zeki, S., & Shipp, S. (1989). Modular connections between areas V2 and V4 of macaque monkey visual cortex. *European Journal of Neuroscience*, *1*(5), 494–506, doi:10.1111/j.1460-9568.1989.tb00356.x.