Serial dependence generalizes across different stimulus formats, but not different sensory modalities

Michele Fornaciai[1,*], & Joonkoo Park[1,2]

[1] Department of Psychological and Brain Sciences, University of Massachusetts Amherst, 01003 Amherst, MA, USA
[2] Commonwealth Honors College, University of Massachusetts Amherst, 01003 Amherst, MA, USA

* Correspondence:
Michele Fornaciai
Department of Psychological and Brain Sciences
University of Massachusetts Amherst
mfornaciai@umass.edu

ABSTRACT

Visual perception is thought to be supported by a stabilization mechanism integrating information over time, resulting in a systematic attractive bias in experimental contexts. Previous studies show that this effect, whereby a current stimulus appears more similar to a previous one, depends on attention, suggesting an active high-level mechanism that modulates perception. Here, we test the hypothesis that such a mechanism may generalize across different stimulus formats or sensory modalities, effectively abstracting from the low-level properties of the stimuli. Participants performed a numerosity discrimination task, with task-relevant dot-array stimuli preceded by a sequence of visual (flashes) or auditory (tones) stimuli encompassing different numerosities. Our results show a clear attractive bias induced by visual sequential numerosity affecting an array of simultaneously presented dots, thus operating across different stimulus formats. Conversely, auditory sequences did not affect the judgment on visual numerosities. Overall, our results demonstrate that serial dependence in numerosity perception operates according to the abstract representation of numerical magnitude of visual stimuli irrespective of their format. These results thus support the idea that a high-level mechanism mediates visual stability and continuity, which integrates relevant information over time irrespective of the low-level sensory properties of the stimuli.

KEYWORDS

Serial dependence; numerosity perception; visual stability.

INTRODUCTION

One remarkable feature of our conscious visual perception is its stability and continuity over time: we usually experience a smooth and seamless flow despite the noisiness of brain signals and the instability of biological sensors. How does the brain keep vision stable and continuous? One possibility, raised by recent studies, is that it does so by integrating information over time to smooth out noise from neural signals and facilitate a continuous representation (e.g., Fischer & Whitney, 2014; Burr & Cicchini, 2014).

While such a process, resulting in a "continuity field," has potentially significant advantages in our everyday perception due to the stability of the external world (Burr & Cicchini, 2014; Cicchini et al., 2018), it results in systematic biases in the experimental context. Specifically, a current stimulus appears more similar to a previous one than it actually is, even if the two stimuli are completely uncorrelated. Such an attractive bias, named *serial dependence*, has been taken as a signature of a stabilization process. Furthermore, serial dependence has been observed across several very different visual dimensions, spanning from orientation (Fischer & Whitney, 2014; Cicchini et al., 2018), position (Manassi et al., 2018), and numerosity (Corbett et al., 2011; Cicchini et al., 2014; Fornaciai & Park, 2018a), to motion (Alais et al., 2017), visual variance (Suarez-Pinilla et al., 2018), face identity (Liberman et al., 2014) and attractiveness (Xia et al., 2016), which suggests a general perceptual mechanism.

Although the nature of serial dependence has been subject to debate (Fritsche et al., 2017; Bliss et al., 2018), there is increasing evidence that it directly operates on perception (Cicchini et al., 2017; Manassi et al., 2018; Fornaciai & Park, 2018a; Fornaciai & Park, 2018b; Pascucci et al., 2019). In a recent series of experiments (Fornaciai & Park, 2018a), we demonstrated that serial dependence occurs in a spatially localized fashion, which is the hallmark of a perceptual effect, but it also depends on attention. Indeed, it occurs only if stimuli (or at least their spatial positions) are attended or behaviorally relevant (Fornaciai & Park, 2018a; see also Fischer & Whitney, 2014). These results thus point to a model of serial dependence based on high-level, spatially-specific modulations affecting the low-level perceptual processing of attended/relevant stimuli, likely by means of modulatory feedback signals (e.g., see Lamme & Roelfsema, 2000).

A prediction based on the idea of a modulatory feedback mechanism is that serial dependence may operate according to a more abstract representation of the stimuli, not completely tied to their low-level sensory features. In the context of numerosity perception, a growing amount of evidence suggests that approximate numerical processing is rooted into the sensory processing stream, starting from early sensory areas (Fornaciai & Park, 2017; Fornaciai et al., 2017; Fornaciai & Park, 2018c; Cavdaroglu et al., 2015; Cavdaroglu & Knops, 2018). However, later on in the processing stream, numerosity information seems to be represented in an abstract format, relatively independent from its sensory origin (Piazza et al., 2007; Arrighi et al., 2014; Anobile et al., 2016). A serial dependence mechanism based on a high-level representation and modulatory feedback may thus be based on such an abstract representation of numerosity.

Here, we address this very hypothesis by testing whether serial dependence in numerosity perception works across different formats such as sequential (i.e., a series of events over time) and simultaneous (i.e., a set of objects in space) numerosities, and across different sensory modalities (visual and auditory). To do so, we employed a paradigm used in previous studies (Fornaciai & Park, 2018a), in which a task-irrelevant "inducer" is presented prior to task-relevant dot-array stimuli, with participants performing a numerosity discrimination task. Differently from previous studies, however, we presented the inducer as a sequence of visual (cross-format condition) or auditory (cross-modal condition) stimuli and measured its effect on subsequent task-relevant dot-arrays. If serial dependence operates at the level of abstract numerosity representation in the visual domain, an attractive bias should be observed in the visual sequential inducer condition. If serial dependence operates at an even more abstract level encompassing both visual and auditory domains, an attractive bias should be observed in both visual and auditory inducer conditions.

To preview, our results show that serial dependence generalizes across different stimulus formats, but not across different sensory modalities, thus supporting the idea of serial dependence as a high-level mechanism, operating according to an abstract representation of the stimuli irrespective of their low-level sensory properties, but limited to the modality of presentation.

METHODS

Participants

A total of 70 participants (including the author MF) took part in the study (47 females, mean age (mean ± SD) = 20.07 ± 2.11 years old). Participants were rewarded with course credit, and signed a written informed consent before participating in the study. All participants tested in the study were naïve to the purpose of the experiment (with the exception of the author MF), 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. Note that the sample size included in each experiment (see *Results*) was determined a priori based on the expected effect derived from previous studies from our group. Namely, considering an expected effect size (Cohen's *d*) of approximately 0.6 (based on previous experiments employing a similar paradigm, both published and unpublished), a power of 0.95, and a one-tailed t distribution, the total required sample size is 32 participants (for each experiment).

Apparatus and stimuli

All the stimuli employed in the experiment were created using the Psychophysics Toolbox (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997) on Matlab (version r2016b; The Mathworks, Inc.). Visual stimuli were presented on a monitor screen running at 144 Hz, with a resolution of 1920 × 1080 pixel, and encompassing approximately 35×20 degrees of visual angle from a viewing distance of about 80 cm.

The visual stimuli employed in the numerosity discrimination task were arrays of black and white dots presented on a gray background. Such dot-array stimuli were systematically constructed to vary across several non-numerical dimensions, in order to span equal ranges in three orthogonal dimensions: numerosity (*N*), size (*Sz*), and spacing (*Sp*) (see Park et al., 2016; DeWind et al., 2015). Note that since the main goal of the present study was to assess serial dependence effects on approximate numerical judgments, we collapsed together the different non-numerical dimensions during data analysis. For details about this stimulus construction scheme, see Park et al. (2016) and DeWind et al. (2015). Stimulus parameters were set as follows. The reference stimulus always comprised 16 dots. Probe arrays comprised 8, 10, 13, 16, 20, 25, or 32 dots. The smaller individual area of each dot was set to 113 pixel^2 (0.038 deg^2), corresponding to a diameter of 0.11 deg (6 pixels), while the maximum individual area was 452 pixel^2 (0.15 deg^2), corresponding to a diameter of 0.22 deg (12 pixel). The minimum field area, corresponding to the circular area within which the dots were drawn, was 70,686 pixel^2 (23.9 deg^2), encompassing 5.5 degrees of visual angle in diameter (300 pixels), while the maximum field area was 282,743 pixel^2 (95.7 deg^2), encompassing 11 degrees in diameter (600 pixels). In all cases, individual dot size was kept equal within an array, and the minimum distance between any two dots was no smaller than the radius of the dots.

Inducer stimuli, on the other hand, could be either a sequence of brief sounds, or a sequence of brief visual flashes. In both cases, they included either 10 or 25 stimuli. Auditory inducer stimuli were pure tones (frequency = 700 Hz) played by means of two speakers (Logitech Multimedia Speakers 2200) located behind the screen in positions corresponding to that of the visual stimuli (sound intensity ~65 dB). Visual inducer stimuli were white circles with a constant radius of 2.5 deg, presented at 85% contrast. In Exp. 1, in both conditions (auditory and visual), 10- and 25-stimulus sequences had the same temporal frequency (10 Hz), with each stimulus (either a visual flash or an auditory tone) presented for about 20 ms with an interstimulus interval of about 80 ms. Doing so, the inducer sequence including 10 stimuli had a total duration of 1 s, while the sequence including 25 stimuli had a duration of 2.5 s. In Exp. 2, only visual stimuli were used. In this case, visual sequences could either have the same temporal frequency as in Exp. 1 (10 Hz), and thus different durations (1 s and 2.5 s respectively for 10 and 25 flashes), or the same duration (2.5 s), thus presenting different temporal frequencies (4 and 10 Hz). In this latter case, the duration of each flash was kept equal to the other conditions (20 ms), but the interstimulus interval was longer (230 ms).

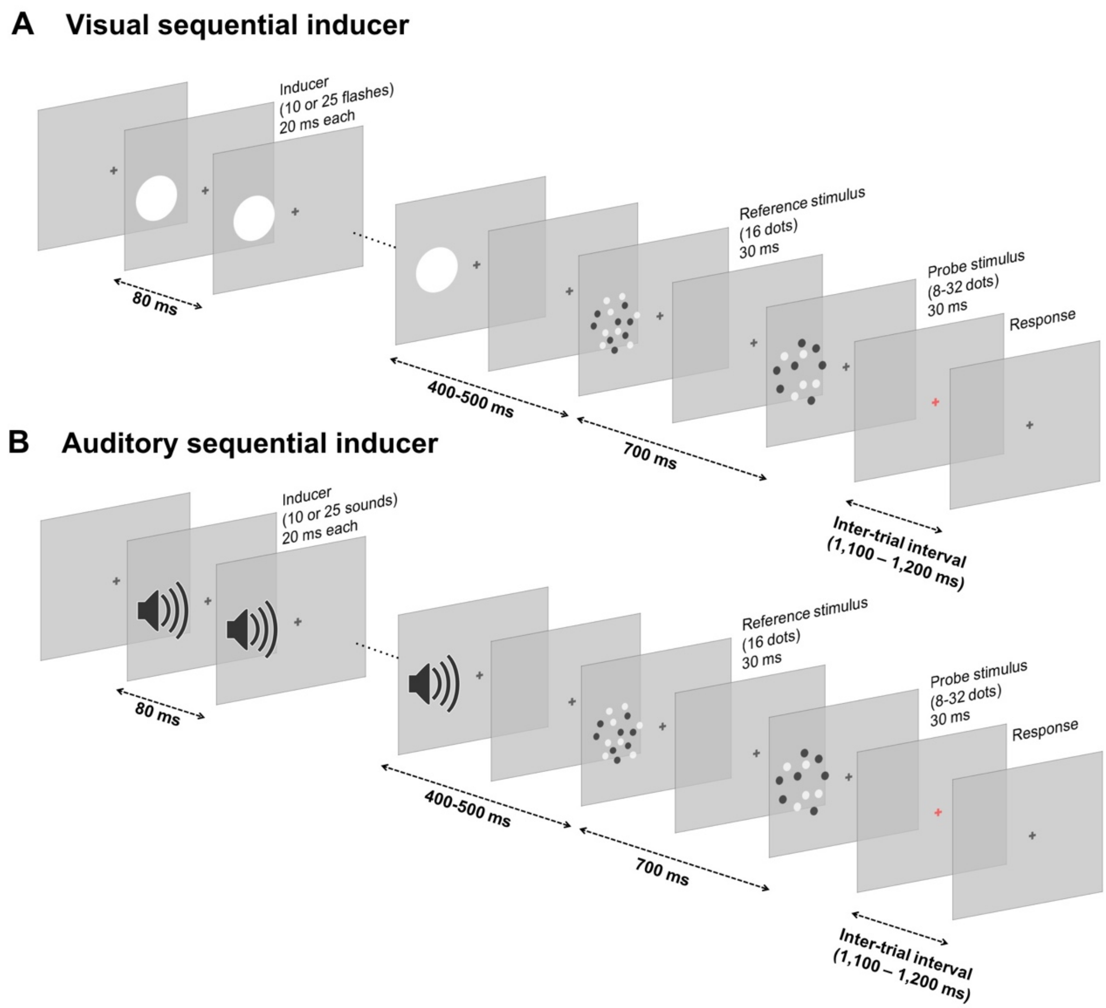

FIGURE 1. Experimental Procedures. (A) Visual sequential inducer condition in Exp. 1. The stimulation procedure in this condition involved a visual sequential inducer comprising either 10 or 25 brief flashes (each displayed for 20 ms with an interstimulus interval of 80 ms), followed by a 16-dot reference stimulus (after 400-500 ms), and finally a variable probe stimulus including 8-32 dots (after 700 ms). At

the end of the sequence, in most of the trials, participants had to report whether the reference or the probe stimulus contained more dots. To make subjects pay attention to the inducer stimulus, they were occasionally asked to report whether the inducer sequence contained "a few" or "a lot" of stimuli; in this case the numerosity discrimination task was skipped, and the trial was excluded from data analysis. After providing a response, the next trial started automatically after 1,100-1,200 ms. In Exp. 2, the procedure was largely identical to the visual inducer condition of Exp. 1, except that the different inducer sequences corresponding to different numerosities (10 or 25) could be presented with the same duration (2.5 s) or the same temporal frequency (10 Hz). (B) Auditory sequential inducer condition in Exp. 1. In this condition, the procedure was identical to the visual condition in Exp. 1, with the exception that instead of flashes the inducer included either 10 or 25 brief sounds (pure tones, frequency = 700 Hz). Note that the stimuli are not depicted in scale.

Procedure

All the experimental conditions of the study were performed in a quiet and dimly illuminated room. Participants sat in front of a monitor screen at a distance of about 80 cm. In all the experiments, participants fixated on a central fixation cross, while they performed a numerosity discrimination task as a main task, reporting whether a reference (16 dots) or a probe (8-32 dots) seemed to contain more dots. In order to induce serial dependencies affecting the perceived numerosity of the reference dot-array, an "inducer" stimulus was presented at the beginning of each trial. In Exp. 1, the inducer was either a visual or an auditory sequence, presented for 1 s or 2.5 s according to the number of stimuli included in the sequence (either 10 or 25). After an interval of 400-500 ms starting at the offset of the last stimulus in the inducer sequence, the reference was presented on the screen for 30 ms, followed by the probe after 700 ms (similarly presented for 30 ms). At the end of the sequence, the fixation cross turned red signaling the end of the trial, and participants were asked to report which stimulus contained the larger number of dots by pressing the appropriate key on a standard keyboard. After providing a response, the next trial started automatically after 1,100-1,200 ms. Procedure in Exp. 2 was identical to Exp. 1, with the exception that only visual sequences were employed, and the stimuli were controlled for duration and temporal frequency. Namely, in different trials randomly intermixed throughout each block, the visual inducer sequences (10 and 25 stimuli) were equalized for duration (i.e., thus spanning both 2.5 s but presenting different temporal frequencies: 4 or 10 Hz), or were equalized for temporal frequency (TF; i.e., same 10 Hz temporal frequency but different durations as in Exp. 1). In both Exp. 1 and Exp. 2, a secondary task concerning the inducer sequence was included to encourage participants to pay attention to it (i.e., similarly to Fornaciai & Park, 2018a). Namely, at the end of some trials (4 trials in each block; see below), participants were asked to report whether the inducer contained "a few" or "a lot" of stimuli.

Subjects were shown a few examples of different sequences before starting the experiment in order for them to understand how to perform the secondary task. When participants performed this secondary task, they did not have to judge numerosity. The trials that contained this secondary task were excluded from data analysis. In all the experiments, reference and probe dot-array stimuli could be presented either on the left or on the right of the central fixation cross, with a horizontal eccentricity of 11 deg. Visual inducer stimuli were presented on the screen at a position corresponding to the center of the subsequent dot-array stimuli, thus again with a 11 deg horizontal eccentricity. Auditory stimuli were played from two speakers positioned behind the screen corresponding to the possible location of the visual stimuli on the left and on the right of the central fixation cross. Participants completed 4 blocks of 60 trials in each of the two conditions of Exp. 1, and 8 blocks of 60 trials in Exp. 2. Each block in all the experiments comprised 4 "catch" trials in which participants performed the secondary task. An equal amount of trials was collected for the different combinations of inducer numerosity and probe numerosity. Participants performed a brief training session (10-15 trials) before starting the actual experiment, in order to ensure that they understand the task. An entire experimental session typically took around 50 minutes, and participants were encouraged to take frequent breaks when needed.

Behavioral data analysis

Across all the experiments, the numerosity discrimination performance was analyzed separately for each subject and condition. The serial dependence effect was assessed by analyzing trials corresponding to different inducer numerosities, which are designed to bias the perceived numerosity of the reference stimulus in opposite directions. To achieve a measure of participants' accuracy and precision in the task, the distribution of response probabilities as a function of probe numerosity was fitted with a cumulative Gaussian curve, according to the Maximum Likelihood method (Watson, 1979). The point of subjective equality (PSE), which is the probe numerosity perceptually matching the reference numerosity (thus reflecting the accuracy in the task and the reference perceived numerosity), was taken as the median of the best-fitting cumulative Gaussian curve to all the data of each participant in each condition.
Participants' precision in the task was assessed to exclude subjects showing insufficient performance. To this aim, we used the just-noticeable difference (JND), obtained as the difference in numerosity between chance level (50%) responses and 75% "probe more numerous" responses. We excluded participants showing a JND higher than 10 dots. Although this threshold is arbitrary, with the current numerosity range a JND higher than 10 dots would most likely result from a very poor discrimination performance, making the result noisy and difficult to interpret. A total of 6 subjects was excluded from data analysis based on this criterion, across all the experiments. One more participant was instead excluded because an unusually high PSE (i.e., more than three SD higher than the average of the group). Additionally, a finger

error rate correction (2%) was applied to account for lapses of attention or random response errors (Wichmann & Hill, 2001).

To assess the serial dependence effects within each condition, a paired t-test was performed comparing the distribution of PSEs corresponding to different inducer numerosities. Moreover, to obtain a direct measure of the serial dependence effect and compare it across different experiments, we calculated a serial dependence index taken as the difference in PSE in the high-numerosity inducer condition (i.e., for instance 25 stimuli) and the low-numerosity inducer condition (i.e., for instance 10 stimuli). Additionally, to achieve another index of the strength of the serial dependence effect we also computed the net difference between the two inducer conditions as follows:

$$\text{Effect} = \text{abs}(((PSE_{ind25} - PSE_{ind10})/PSE_{ind10}) \times 100);$$

Where PSE_{ind10} refers to the subject's PSE in the lower inducer condition (10 flashes/sound), and PSE_{ind25} refers to the PSE in the higher inducer condition (25 flashes/sound).

A t-test was used to compare the distribution of serial dependence effects across different conditions (i.e., visual and auditory). Finally, we also tested for a correlation between the level of precision in the task and serial dependence. To do so, we used for each participant a measure of JND obtained by collapsing together the two inducer numerosity conditions (in Exp. 1), and also collapsing together different inducer types (same duration and same TF, in Exp. 2), in order to get a single JND value for each participant. These JND values were log-transformed to achieve normality. We then assessed the correlation of such a measure of precision with the indices of serial dependence effect calculated as reported above.

RESULTS

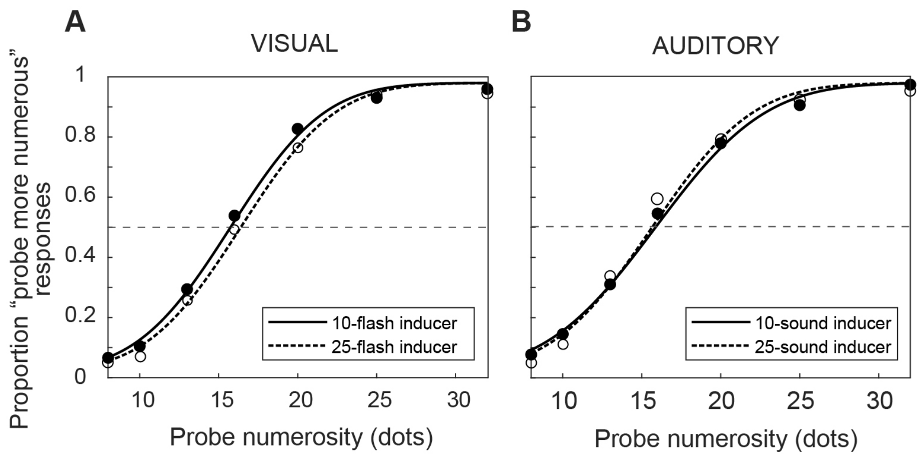

FIGURE 2. Average psychometric curves in the two conditions of Experiment 1. (A) Psychometric curves in the visual condition, representing the lower (10 flashes) and higher (25 flashes) inducer condition. (B) Psychometric curves in the auditory condition, representing the lower (10 sounds) and higher (25 sounds) inducer condition. Note that average psychometric curves were computed by pooling the data from all participants together.

Experiment 1

Participants (N = 34) compared the numerosity of the *reference* and the *probe* dot arrays, while the reference array was preceded by a sequential numerosity inducer either in visual or auditory modalities. First, by plotting the average psychometric curves of the visual condition (Fig. 2A), it is immediately clear that the two sequential inducer conditions result in a relative shift of the curves, with a more leftward-shifted curve in the lower inducer condition (10 flashes), and a relative rightward shift in the higher inducer condition (25 flashes). On the other hand, in the auditory condition (Fig. 2B), the two curves are almost perfectly superimposed, suggesting that the influence of an auditory inducer is much reduced compared to the visual one.

To achieve a more precise index of performance in the two conditions, we analyzed the data individually for each participant and condition. The results, illustrated in Figure 3, confirm that a visual sequential inducer (Fig. 3A) was highly effective in inducing attractive serial dependencies, systematically biasing the perceived numerosity of the subsequent reference dot array. In other words, when the 16-dot reference was preceded by a sequence of 10 flashes, its perceived numerosity was significantly lower compared to the same 16-dot reference preceded by 25 flashes ($t(33) = 3.094$, $p = 0.004$, $d = 0.53$), with an average difference in perceived numerosity of about 5.3% (± 1.7%, SEM). On the other hand, an auditory sequential inducer was not as effective (Fig. 3B), resulting in no significant attractive bias on the perceived numerosity of the reference dot array, and with even a non-significant trend in the opposite (repulsive) direction (t(33) = -0.895, p = 0.377, d = 0.15, average difference = 1.1% ± 1.6%). Comparing the distribution of the serial dependence effect (i.e., PSE in the higher inducer minus PSE in the lower inducer; Fig. 3C) directly between the two inducer conditions showed a statistically significant difference (t(33) = 2.77, p = 0.009, d = 0.47), again demonstrating that visual and auditory inducers yielded largely different effects.

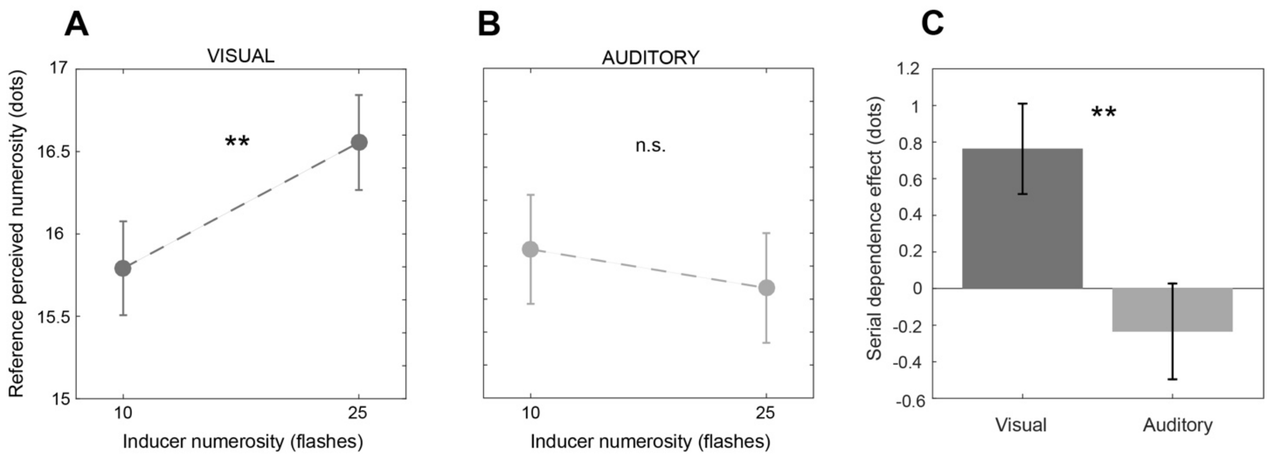

*FIGURE 3. Results of Experiment 1. (A) Visual inducer condition. In this condition, a visual sequence of flashes induced strong and systematic attractive biases affecting the perceived numerosity of the subsequent reference dot-array. (B) Auditory inducer condition. An auditory sequential inducer, however, did not affect the perceived numerosity of the dot-array stimulus. (C) Average serial dependence effect index in the visual and auditory condition. The serial dependence effect index was calculated as the PSE in the larger inducer condition (25 dots) minus PSE in the smaller inducer condition (10 dots) in each participant. Error bars are SEM. n.s. = not significant, ** p < 0.01.*

Finally, regarding the performance in the "catch" task that participants performed on some trials (i.e., discriminating whether the inducer contained a few or many flashes/sounds), the proportion of correct responses (mean ± SEM) was 79.2% ± 2.96% in the visual condition, and 90.1% ± 2.63% in the auditory condition. The proportion of correct responses was significantly higher in the auditory condition (t(33) = 4.55, p < 0.001). This may indicate that performing the attentional task in the visual modality might have been more difficult compared to performing it on auditory stimuli, possibly due to a bigger attentional load with stimuli in the same modality.

Experiment 2

Results from Experiment 1 show that serial dependence operates across stimulus formats (although only in the same modality), suggesting that the effect arises at a relatively high – more abstract – representational level, beyond low-level sensory properties of the dot array stimuli. Nevertheless, because the two inducer stimuli were kept in a constant temporal frequency and thus had different durations, the results could be interpreted as having the duration information rather than numerosity driving the attractive biases in numerosity judgment. Indeed, interactions between magnitude dimensions such as time and numerosity are well known (Dormal et al., 2006; Xuan et al., 2007; Lambrechts et al., 2013; Tsouli et al., 2018; Fornaciai et al., 2018). Thus, in Exp. 2, we tested to what extent the attractive serial dependence found in Experiment 1 can be generalized in cases where duration is not always correlated with numerosity.

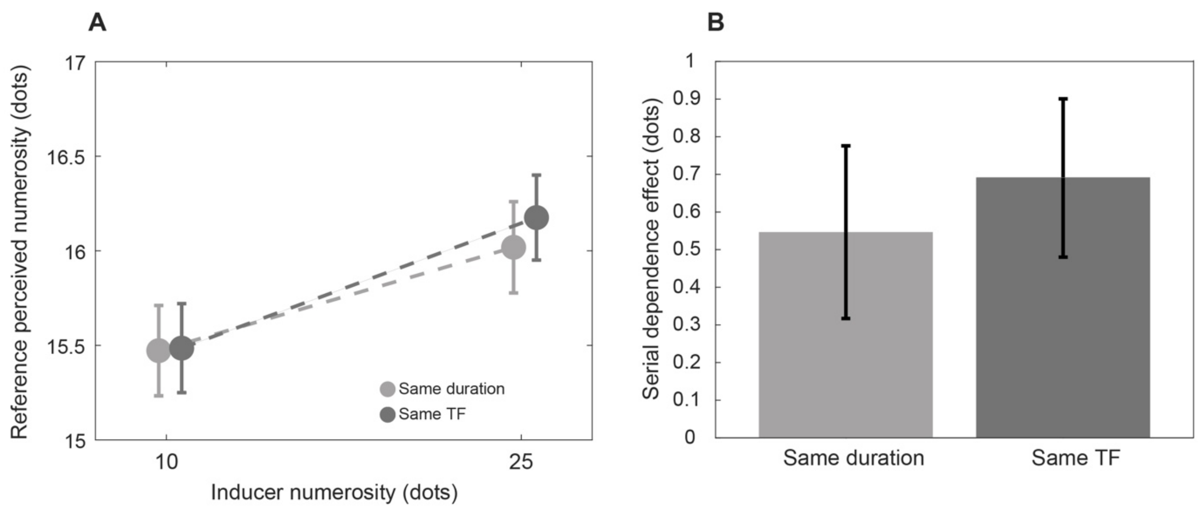

FIGURE 4. Results of Experiment 2. (A) Average PSE as a function of inducer numerosity in the same duration (light gray) and same temporal frequency (dark gray) condition. (B) Average serial dependence effect indexes in the two conditions of Exp. 2. Error bars are SEM.

Experiment 2 (N = 30) was identical to the visual inducer condition in Exp. 1, except the two sequential inducers were on half of the trials in same duration or on the other half in same temporal frequency, but importantly always in different numerosities (10 dots versus 25 dots). The results, as illustrated in Figure 4, indicate that visual sequential inducers again significantly biased the representation of the subsequent reference dot-array stimulus in the attractive direction regardless of whether the two inducers were in same duration ($t(29) = 2.3812$, $p = 0.0240$, $d = 0.435$) or in same temporal frequency ($t(29) = 3.2814$, $p = 0.0027$, $d = 0.60$). The average difference between the two inducer conditions was 3.9% ± 1.5% and 4.8% ± 1.4%, respectively for the same duration and same TF condition. While the effect seems somehow stronger in the same TF condition (which is basically the same as in Exp. 1), comparing the distribution of serial dependence effects across the two conditions did not reveal any statistically significant difference (Fig. 4B; $t(29) = 0.4844$, $p = 0.632$, $d = 0.09$). Such a comparable effect across the two control conditions suggests that the effect is most likely determined by the numerosity of the sequence, with smaller if any influence of other attributes such as duration or temporal frequency.

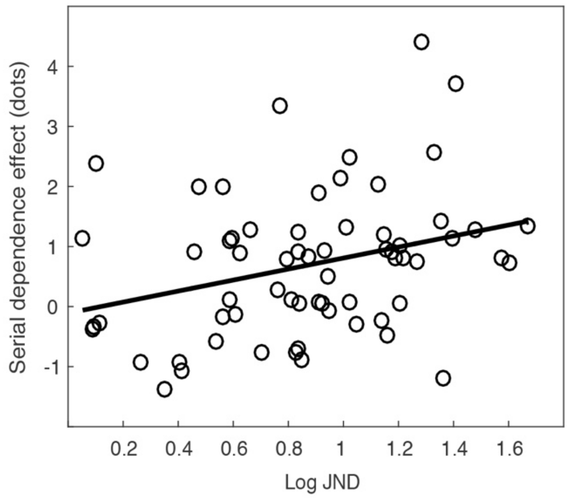

FIGURE 5. Correlation between serial dependence and precision in the task. Scatterplot illustrating the correlation between the log-transformed JND and the serial dependence effect, including data from both Exp. 1 (only the visual condition) and Exp. 2.

Moreover, we tested for a correlation between participants' precision in the task (JND) and the magnitude of the serial dependence effect across both Exp. 1 and Exp. 2 (Fig. 5). Doing so, we observed a statistically significant correlation ($r = 0.304$, $p = 0.0146$) between JND and the index of serial dependence effect. This shows that the strength of serial dependence varies as a function of the precision of perceptual estimates. Although the correlational nature of this result does not allow us to pinpoint the exact direction of the effect, a possibility is that in conditions of higher uncertainty or noisier representation (i.e., high JND), the visual system may give a stronger weight to past information, resulting in stronger serial dependencies (see *Discussion*).

Finally, the proportion of correct responses across the two conditions of Exp. 2 was 80.8% ± 2.75%.

DISCUSSION

In the present study, we addressed the hypothesis that serial dependence in numerosity perception may operate according to an abstract representation of approximate numerical magnitude. This prediction comes from the idea that serial dependence results from a high-level modulatory process affecting perception, as suggested by recent studies (Fornaciai & Park, 2018a; Pascucci et al., 2019). Indeed, previous studies show that serial dependence is both spatially localized and dependent on attention (Fornaciai & Park, 2018a) suggesting the involvement of a spatially-localized modulation of early visual activity mediated by attention (e.g., Somers et al., 1999). The current results support this idea by showing that the effect generalizes across different visual presentation formats. Namely, the numerosity of a sequential inducer affects the perceived numerosity of a subsequently presented dot-array stimulus, with a clear attractive trend. A control experiment (Exp. 2) showed that other features of the sequential inducer like duration or temporal frequency have little influence on serial dependence in this context, similarly to what has been previously demonstrated with cross-format numerosity adaptation (Arrighi et al., 2014). In contrast to the case of a visual sequential inducer, an auditory sequential inducer did not systematically affect the perceived numerosity of visual dot-array stimuli, showing that serial dependence operates in a within-modality fashion. This additionally suggests that serial dependence is a distinct mechanism compared to perceptual adaptation, which has been shown to generalize also across different sensory modalities (Arrighi et al., 2014). Finally, the specificity for sensory modality also suggests that the effect is perceptual in nature, and that it is distinct from priming, which has been demonstrated to work across modalities (e.g., McKone & Dennis, 2000; Greene et al., 2001; Buchner et al., 2003; Vallet et al., 2010). Note that in this context a *perceptual* effect does not necessarily involve a bias in the early *sensory* processing. While the effect may arise at higher-level processing stages after sensory encoding (see below for further discussion about possible neural underpinnings), the crucial point is that such an effect would operate directly on perception, distorting the subjective experience (i.e., appearance) of a stimulus.

How do the present results compare with previous findings? On the one hand, such a cross-format effect is surprising, as earlier studies on serial dependence show that the effect depends on the similarity between two stimuli. For instance, several studies since Fischer & Whitney (2014), using a continuous series of stimuli and a reproduction task, show that the effect of past stimuli declines as the difference compared to the current one becomes too large (Fischer & Whitney, 2014; Fritsche et al., 2017; Manassi et al., 2018). This led some authors to conceptualize serial dependence as a sort of weighted average akin to a Kalman filter (Burr & Cicchini, 2014). In our paradigm, stimuli in different formats had completely different low-level features. Yet, we observed a systematic attractive effect very similar to previous results (Fornaciai & Park, 2018a). This pattern is not irrelevant to our previous finding in which using 8 and 32-dot inducers was equally effective as using 12 and 24-dot inducers in eliciting serial dependence for a 16-dot reference (Fornaciai & Park, 2018a), suggesting a relatively low specificity of the effect.

These results indicate some qualitative differences between our results using numerosity information in a discrimination task and some other previous studies using orientation information in a reproduction task.

One possibility is that such a selectivity in previous studies may be a feature of serial dependence in circular dimensions (i.e., orientation as in Fischer & Whitney, 2014 and Fritsche et al., 2017; Pascucci et al., 2019; or the circular position space in Manassi et al., 2018), while it may not emerge in magnitude dimensions such as numerosity. Another possibility is instead related to the difference in the paradigm used. While a reproduction task requires participants to reproduce a stimulus with a high degree of precision – thus requiring paying more attention to its specific value – a two-alternative forced-choice paradigm like the one used in the present study may shift the focus to broader categories such as "more" versus "less," irrespective of the specific magnitude of the stimuli. While our data cannot disentangle these possibilities, this remains an interesting question for future studies. Irrespective from this, our data show that what actually matters is not the similarity in the superficial sensory feature of the stimuli, but the information that can be extracted from them.

Our results are in line with earlier studies suggesting a generalized and abstract representation of numerosity (Piazza et al., 2007; Arrighi et al., 2014; Anobile et al., 2016). Indeed, while numerosity processing seems to be deeply rooted into modality-specific sensory pathways (Cavdaroglu et al., 2015; Cavdaroglu & Knops, 2018; Fornaciai et al., 2017; Fornaciai & Park, 2018c) and interacting with several sensory systems like color and motion perception (Fornaciai & Park, 2017), both psychophysical (e.g., Arrighi et al., 2014; Anobile et al., 2016) and neuroimaging (e.g., Piazza et al., 2007) data show that numerical magnitude is represented in a more abstract fashion at a relatively high-level in the cortical processing hierarchy (e.g., in parietal cortex). According to this idea, then, it is not surprising that serial dependence operates according to the numerical magnitude of the stimuli irrespective of their format. In this scenario, serial dependence would operate according to a relatively high-level representation of past stimuli (i.e., in this case the inducer), to affect the processing of the current stimulus starting from the earliest levels of visual processing (Fornaciai & Park, 2018b).

It is interesting to note that the paradigm used in the present study closely resembles the one used by Arrighi et al. (2014) to assess cross-format numerosity adaptation effects. Strikingly, while Arrighi et al. (2014) observed a strong repulsive effect consistent with perceptual adaptation, here we observed a systematic attractive bias. The different outcomes of a similar paradigm, however, may depend on the specific parameters of inducer/adaptor stimuli used in different context. The most important one, in our interpretation, is the duration of the stimuli: while the long adaptation procedure used by Arrighi et al.

(2014) – more focused on the temporal frequency of the stimuli rather than on their number – would more easily induce a repulsive effect, the shorter stimuli used in the present study would provide an attractive bias. Another example of this difference in the effect due to the duration of the stimuli is also evident in previous studies. For instance, in a previous study from our group (Fornaciai & Park, 2018), we used a paradigm (i.e., see Exp. 3 in Fornaciai & Park, 2018) very similar to the classic numerosity adaptation procedure (e.g., Burr & Ross, 2008), except for the duration of the inducer/adaptor stimuli. Similarly, while using very long sustained stimuli yields a repulsive effect, our results employing very brief stimuli showed systematic attractive effects.

A crucial question, then, is: what are the possible neural substrates of this cross-format serial dependence bias? While an account based on a bias at the early levels of sensory processing (e.g., Fischer & Whitney, 2014) does not fit the observed effect due to wide differences in the stimuli used, the present results seem more consistent with an account of serial dependence based on the persistence of perceptual decision templates at a read-out/decision stage (Pascucci et al., 2019). In a recent study, Pascucci et al. (2019) conceptualized attractive serial dependence as a bias occurring at a relatively high-level processing stage, affecting the decoding of early sensory activity performed by read-out "decision" units. Namely, traces of past perceptual decisions would linger at such high-level stage, effectively modulating the read-out weights applied to interpret low-level activity. Such a modulation, in turn, would bias the resulting perceptual representation of the current stimulus, directly affecting the appearance of the stimulus. A similar account of serial dependence based on a high-level neural mechanism was also provided by Fritsche et al. (2017). However, such an account was mostly based on a working memory bias, which is not consistent with results showing that serial dependence affects perception directly (e.g., Cicchini et al., 2018, Fornaciai & Park, 2018; Manassi et al., 2018; Pascucci et al., 2019).

Taking into account previous research on numerosity perception, our results thus fit with this model (Pascucci et al., 2019) in pinpointing a relatively high level locus for serial dependence. At the neural level, parietal areas were often associated with an abstract representation of numerosity independent from the stimulation format. In this scenario, therefore, parietal neurons would thus represent the read-out units biased by attractive serial dependence, while numerosity-related activity in early visual cortex (e.g., Fornaciai & Park, 2018; DeWind et al., 2018) would not be affected directly. The fact that auditory inducers do not affect the perceived numerosity of visual stimuli, however, limits this interpretation by suggesting that such read-out units only receive visual signals, while previous studies observed other types of perceptual effects (such as adaptation) generalizing also across sensory modalities (Arrighi et al., 2014).

Additionally, it is also interesting to note that the present findings are consistent with recent results concerning a different process: trans-saccadic integration. Indeed, it has been shown that integration of numerosity information across eye movements is insensitive to changes in the low-level features of the stimuli, such as the luminance/color of the items (Hübner & Schütz, 2017). Although the manipulations performed by Hübner & Schütz (2017) are not as extensive as presenting different stimulus formats, this finding also suggests that integration of numerosity information occurs at a relatively high level in the visual hierarchy. Moreover, it has been also shown that trans-saccadic integration depends on attention (Stewart & Schütz, 2018), again similarly to what has been previously demonstrated for serial dependence (Fischer & Whitney, 2014; Fornaciai & Park, 2018a), suggesting that the two integrative processes may even involve partially overlapping or similar neural mechanisms.

Another interesting point concerns the correlation observed between serial dependence and JND (Fig. 5). This correlation may suggest that the magnitude of the effect varies as a function of the precision of perceptual estimates – that is, in conditions of higher uncertainty (i.e., as indexed by lower precision), the visual system may rely to a greater extent on past inputs to disambiguate or improve the representation of current sensory stimulation. However, although this explanation is in line with previous studies (Cicchini et al., 2018b), the correlational nature of this result does not allow us to draw a strong conclusion about this point. In fact, across several previous experiments from our group (Fornaciai & Park, 2018a; Fornaciai & Park, in press), we rarely observed a correlation between JND and serial dependence (i.e., only in 3 out of 10 independent experiments). Although this is a potentially interesting point, thus, the fact that this correlation is very difficult to replicate makes it difficult to draw any conclusion from it. Looking at Fig. 5, however, it is also evident that there is a relatively large variability in the magnitude of the serial dependence effect across participants, with a few data points actually showing negative (repulsive) effects. As serial dependence is highly dependent on attention, a possibility is that the degree to which participants paid attention to the inducer may have determined the strength of the attractive effect. Unfortunately, too few trials were collected in the catch task to achieve a realistic index of participants' attention. An intriguing possibility for future studies would thus be to modulate attention more extensively, for instance by using a double-task design in every trial, and/or secondary tasks loading attention to different extents.

Finally, another potentially important point to consider when interpreting the results, concerns the specific paradigm employed in this study. Indeed, the fully sequential stimulation procedure used here represents a limitation: when measuring the effect of the inducer on the reference stimulus, some effect might have extended to the subsequent probe stimulus, effectively reducing the magnitude of serial dependence. In fact, in a recent study from our group (Fornaciai & Park, in press), we observed a stronger effect when presenting reference and probe simultaneously in two different portions of the visual field. However, although a sequential presentation is not the optimal procedure in this context, this only makes our test more conservative. Another possibility related to this point, on the other hand, is that the absence of an effect in the auditory condition might reflect a particularly long effect, extending to the probe and thus compensating the change in perceived numerosity of the reference. While we cannot conclusively rule out this possibility, previous research show that serial dependence in numerosity perception sharply decreases after one stimulus (Cicchini et al., 2014), making more plausible that the null effect in the auditory condition is genuine. Additionally, also having a constant reference stimulus may not be optimal in this context. Indeed, the reduced uncertainty of such a constant stimulus may in turn reduce the magnitude of the serial dependence effect. Thus, another interesting possibility for future studies is to use more variable stimuli to increase uncertainty.

To conclude, our results show that, at least in the context of a numerosity discrimination task, serial dependence affects stimuli irrespective of their presentation format, and hence irrespective from their sensory properties. This finding advances our understanding of serial dependence by showing that at least in some contexts it operates on an abstract representation of the stimuli, affecting similar representations extracted from widely different stimuli. This in turn converges with recent evidence ascribing the phenomenon of serial dependence to a relatively high-level processing stage, possibly at the level of read-out decision units interpreting low-level sensory activity to form a perceptual representation.

AUTHOR CONTRIBUTIONS

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

CONFLICT OF INTEREST

The authors declare no competing financial interests.

ACKNOWLEDGMENTS

This study was supported by the National Science Foundation CAREER Award BCS1654089 to J. P.

REFERENCES

Alais, D., Leung, J., & Van der Burg, E. (2017). Linear Summation of Repulsive and Attractive Serial Dependencies: Orientation and Motion Dependencies Sum in Motion Perception. *The Journal of Neuroscience*, *37*(16), 4381–4390. https://doi.org/10.1523/JNEUROSCI.4601-15.2017

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

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

Bliss, D. P., Sun, J. J., & D'Esposito, M. (2017). Serial dependence is absent at the time of perception but increases in visual working memory. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-15199-7

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

Buchner, A., Zabal, A., & Mayr, S. (2003). Auditory, visual, and cross-modal negative priming. *Psychonomic Bulletin and Review*. https://doi.org/10.3758/BF03196552

Burr, D., & Cicchini, G. M. (2014). Vision: Efficient adaptive coding. *Current Biology*, *24*(22), R1096–R1098. https://doi.org/10.1016/j.cub.2014.10.002

Cavdaroglu, S., Katz, C., & Knops, A. (2015). Dissociating estimation from comparison and response eliminates parietal involvement in sequential numerosity perception. *NeuroImage*, *116*, 135–148. https://doi.org/10.1016/j.neuroimage.2015.04.019

Cavdaroglu, S., & Knops, A. (2018). Evidence for a Posterior Parietal Cortex Contribution to Spatial but not Temporal Numerosity Perception. *Cerebral Cortex*, (1993), 1–13. https://doi.org/10.1093/cercor/bhy163

Cicchini, G. M., Anobile, G., & Burr, D. C. (2014). Compressive mapping of number to space reflects dynamic encoding mechanisms, not static logarithmic transform. *Proceedings of the National Academy of Sciences*, *111*(21), 7867–7872. https://doi.org/10.1073/pnas.1402785111

Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, *17*(14), 6. https://doi.org/10.1167/17.14.6

Cicchini, G. M., Mikellidou, K., & Burr, D. C. (2018). The functional role of serial dependence. *Proceedings of the Royal Society B: Biological Sciences*, *285*(1890), 20181722. https://doi.org/10.1098/rspb.2018.1722

Corbett, J. E., Fischer, J., & Whitney, D. (2011). Facilitating stable representations: Serial dependence in vision. *PLoS ONE*, *6*(1). https://doi.org/10.1371/journal.pone.0016701

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*, 247–65. https://doi.org/10.1016/j.cognition.2015.05.016

Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity-duration interference: A Stroop experiment. *Acta Psychologica*, *121*(2), 109–124. https://doi.org/10.1016/j.actpsy.2005.06.003

Fischer, J., & Whitney, D. (2014). Serial dependence in visual perception. *Nature Neuroscience*, *17*(5), 738–743. https://doi.org/10.1038/nn.3689

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

Fornaciai, M., & Park, J. (2018b). Attractive Serial Dependence in the Absence of an Explicit Task. *Psychological Science*, *29*(3), 437–446. https://doi.org/10.1177/0956797617737385

Fornaciai, M., & Park, J. (2018a). Serial dependence in numerosity perception. *Journal of Vision*, *18*(9), 15. https://doi.org/10.1167/18.9.15

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

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

Fornaciai, M., & Park, J. (In press). Spontaneous repulsive adaptation in the absence of attractive serial dependence. Journal of Vision.

Fornaciai, M., Togoli, I., & Arrighi, R. (2018). Motion-induced compression of perceived numerosity. *Scientific Reports*, *8*(1), 6966. https://doi.org/10.1038/s41598-018-25244-8

Fritsche, M., Mostert, P., & de Lange, F. P. (2017). Opposite Effects of Recent History on Perception and Decision. *Current Biology*, *27*(4), 590–595. https://doi.org/10.1016/j.cub.2017.01.006

Greene, A. J., Easton, R. D., & LaShell, L. S. R. (2001). Visual–Auditory Events: Cross-Modal Perceptual Priming and Recognition Memory. *Consciousness and Cognition*, *10*(3), 425–435. https://doi.org/10.1006/ccog.2001.0502

Hübner, C., & Schütz, A. C. (2017). Numerosity estimation benefits from transsaccadic information integration. *Journal of Vision*, *17*(13), 12. https://doi.org/10.1167/17.13.12

Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3? *Perception ECVP 2007 Abstract Supplement*, *36*(14), 1–16. https://doi.org/10.1068/v070821

Lambrechts, A., Walsh, V., & van Wassenhove, V. (2013). Evidence Accumulation in the Magnitude System. *PLoS ONE*, *8*(12), e82122. https://doi.org/10.1371/journal.pone.0082122

Lamme, V. A., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences*, *23*(11), 571–9.

Liberman, A., Fischer, J., & Whitney, D. (2014). Serial dependence in the perception of faces. *Current Biology : CB*, *24*(21), 2569–74. https://doi.org/10.1016/j.cub.2014.09.025

Manassi, M., Liberman, A., Kosovicheva, A., Zhang, K., & Whitney, D. (2018). Serial dependence in position occurs at the time of perception. *Psychonomic Bulletin & Review*. https://doi.org/10.3758/s13423-018-1454-5

McKone, E., & Dennis, C. (2000). Short-term implicit memory: Visual, auditory, and cross-modality priming. *Psychonomic Bulletin & Review*, *7*(2), 341–346. https://doi.org/10.3758/BF03212991

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. https://doi.org/10.1093/cercor/bhv017

Pascucci, D., Mancuso, G., Santandrea, E., Della Libera, C., Plomp, G., & Chelazzi, L. (2019). Laws of concatenated perception: Vision goes for novelty, decisions for perseverance. *PLOS Biology*, *17*(3), e3000144. https://doi.org/10.1371/journal.pbio.3000144

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

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. https://doi.org/10.1016/j.neuron.2006.11.022

Somers, D. C., Dale, a M., Seiffert, a E., & Tootell, R. B. (1999). Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(4), 1663–8. https://doi.org/10.1073/pnas.96.4.1663

Stewart, E. E. M., & Schütz, A. C. (2018). Attention modulates trans-saccadic integration. *Vision Research*, *142*, 1–10. https://doi.org/10.1016/j.visres.2017.11.006

Suárez-Pinilla, M., Seth, A. K., & Roseboom, W. (2018). Serial dependence in the perception of visual variance. *Journal of Vision*, *18*(7), 4. https://doi.org/10.1167/18.7.4

Tsouli, A., Dumoulin, S. O., te Pas, S. F., & van der Smagt, M. J. (2018). Adaptation reveals unbalanced interaction between numerosity and time. *Cortex*, (April). https://doi.org/10.1016/j.cortex.2018.02.013

Vallet, G., Brunel, L., & Versace, R. (2010). The Perceptual Nature of the Cross-Modal Priming Effect. *Experimental Psychology*, *57*(5), 376–382. https://doi.org/10.1027/1618-3169/a000045

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

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception & Psychophysics*, *63*(8), 1293–1313. https://doi.org/10.3758/BF03194544

Xuan, B., Zhang, D., He, S., & Chen, X. (2007). Larger stimuli are judged to last longer. *Journal of Vision*, *7*(10), 2.1-5. https://doi.org/10.1167/7.10.2