The specious interaction of time and numerosity perception

Irene Togoli*[1], Michele Fornaciai[1], and Domenica Bueti[1]

[1] *International School for Advanced Studies (SISSA), Trieste, Italy*
[*] *Corresponding author. itogoli@sissa.it*

Article published in ***Proceedings of the Royal Society B***. Citation: Togoli I, Fornaciai M, Bueti D. 2021 The specious interaction of time and numerosity perception. *Proc. R. Soc. B* 288: 20211577. https://doi.org/10.1098/rspb.2021.1577.

Magnitude information is essential to create a representation of the external environment and successfully interact with it. Duration and numerosity, for example, can shape our predictions and bias each other (i.e., the greater the number of people queuing, the longer we expect to wait). While these biases suggest the existence of a generalized magnitude system, asymmetric effects (i.e., numerosity affecting duration but not vice versa) challenged this idea. Here we propose that such asymmetric integration depends on the stimuli used and the neural processing dynamics they entail. Across multiple behavioural experiments employing different stimulus presentation displays (static versus dynamic), and experimental manipulations known to bias numerosity and duration perceptions (i.e., connectedness and multisensory integration), we show that the integration between numerosity and time can be symmetrical if the stimuli entail a similar neural time-course and numerosity unfolds over time. Overall, these findings support the idea of a generalized magnitude system, but also highlight the role of early sensory processing in magnitude representation and integration.

1 Introduction

Magnitude dimensions like space, time, and number represent fundamental aspects of our perceptual experience and of our interaction with the external world. The importance of these dimensions for our conscious experience is also reflected by philosophical accounts conceptualizing them as a-priori forms of knowledge, through which every other experience is defined (Kant, *Critique of pure reason*; see Dehaene and Brannon, 2010). However, despite a growing amount of research in the past decades, how the brain processes, represents, and uses magnitude information is far from clear.

According to "A Theory of Magnitude" (ATOM; Walsh, 2003), space, time, and number information is processed by a generalized, supra-modal, system encoding the different magnitudes with a single metric in overlapping brain areas of the posterior parietal cortex (Bueti and Walsh, 2009). This theoretical framework is supported by empirical findings showing the existence of perceptual biases across magnitude dimensions (Burr et al., 2010; Fornaciai, Togoli, and Arrighi, 2018; Winter, Marghetis, and Matlock, 2015). For instance, perceived time has been shown to be strongly affected by numerosity, so that the larger the numerosity of a stimulus, the longer its perceived duration (Cappelletti, Freeman, and Cipolotti, 2009; Dormal and Pesenti, 2013; Oliveri et al., 2008; Roitman et al., 2007; Xuan et al., 2007). However, the idea of a generalized magnitude system has been challenged by findings showing asymmetric biases: perceived time can be biased by stimulus size and numerosity to a larger extent compared to the effect of time on these dimensions (Cappelletti, Freeman, and Cipolotti, 2009; Merritt, Casasanto, and Brannon, 2010). Such asymmetry has led, for instance, to the proposal that non-spatial magnitudes like time are simply conceptualized based on spatial knowledge (Casasanto and Boroditsky, 2008).

Here, we propose that the observation of asymmetric or symmetric biases between different magnitudes depends on the nature of the stimuli used, and particularly on the information processing time-course engaged by such stimuli. Indeed, considering time and numerosity as an example, it is clear that the time-course of their processing is likely very different. While numerical information of a set of items is immediately available after stimulus appearance and entails a relatively fast processing (i.e., 75-250 ms after stimulus onset; Park et al., 2016; Fornaciai et al., 2017; Fornaciai and Park, 2018), the duration of the same stimulus is only available after its disappearance. This difference can in turn introduce a temporal lag between the representation of duration and numerosity. Thus, while a numerical representation could interfere directly with the accumulation of temporal information, duration could affect numerosity only after its processing.

Here, in a series of experiments, we address this idea by making the processing of temporal and numerical information more similar. Our hypothesis is that while the use of static stimuli (a single array of dots) would introduce a temporal offset between the representation of time and numerosity – leading to an asymmetric bias – the use of dynamic stimuli (multiple arrays of dots) where numerosity is represented over time, would reduce this difference. Moreover, we asked when – at which processing stage – magnitude integration occurs. To answer this question, we used two experimental manipulations that are known to affect either numerosity (i.e., connectedness illusion; a systematic underestimation of pairwise-connected dots compared to unconnected dots (He et al., 2009; Fornaciai, Cicchini, and Burr, 2016) or duration perception (i.e., multisensory integration; a task irrelevant sound biases the perceived duration of a visual stimulus Heron et al., 2013) at specific information processing stages. In other words, considering the neural processing stages at which those known illusions occur allowed us to test whether the time-numerosity integration takes place before or after them.

2 Material and methods

2.1 Participants

A total of 89 participants (all naïve to the aims of the study, except two of the authors) took part in the study (58 females; mean age (± SD) = 26 ± 4.1 years). Exp. 1a, 1b, and 1c, included 15 participants each. Exp. 2a and 2b, included 24 participants each. A few participants took part in multiple experiments. Participants provided written informed consent prior to participating in the study, and received Euro 8/hour for their participation. The experimental protocol was approved by the local ethics committee, and it was in line with the declaration of Helsinki.

In Exp. 1, the sample size tested was based on a previous study that used a similar methodology (Javadi and Aichelburg, 2012). In Exp. 2 instead, we determined the sample size with a power analysis based on the results of Exp. 1a, and on previous studies measuring the effect of connectedness (Fornaciai, Cicchini, and Burr, 2016) and multisensory integration (Heron et al., 2013). Based on those previous studies, we computed a minimum expected effect size (Cohen's d) of 0.72, indicating a minimum sample size of 23 subjects. See the Supplementary Materials for more information about the power analysis.

2.2 Apparatus and stimuli

Stimuli were created using the Psychophysics Toolbox (v3) for Matlab (r2015b, The Mathworks, Inc.) and displayed on a 1092×1080 LCD monitor (120 Hz), encompassing a visual angle of 48×30 deg from a viewing distance of 57 cm.

The stimuli used in all experiments were dot-arrays presented singularly or in a sequence of multiple arrays. Each dot-array included an equal proportion of black and white dots, randomly positioned in a circular area with variable dimension (see below), presented on a grey background. In Exp. 1, dot positions were computed online on a trial-by-trial basis, and only constrained by a minimum inter-dot distance equal to the dot radius (6-10 pixels). In Exp. 2, we used connected-dot arrays including lines connecting pairs of dots. Dot positions in such stimuli were computed offline in order to avoid dots overlapping with each other (minimum inter-dot distance = 1.5 times the diameter of a single dot) or with lines (minimum dot-line distance = 0.5 deg), and intersections between lines. During the experiment, the stimuli were randomly chosen from a pool of 500 arrays created according to this procedure.

In all experiments, dot-array stimuli were constructed to vary along several non-numerical dimensions, according to a procedure used in previous studies (Park et al., 2016; DeWind et al., 2015). Note that since the main goal of the present study was to assess the integration of numerosity and duration, the different levels of the non-numerical dimensions were collapsed during data analysis. Throughout all the experiments, the radius of the dots ranged from 6 to 10 pixels. In Exp. 1, the radius of the dot-arrays ranged from 180 to 255 pixels, while in Exp. 2 it ranged from 200 to 400 pixels.

The auditory stimuli used in Exp. 2 were generated using the Psychophysics Toolbox in Matlab, and presented through one of two loudspeakers located behind the screen, positioned consistently with the position of the visual stimuli. Auditory stimuli consisted of a burst of white noise (62 dB) of varying duration.

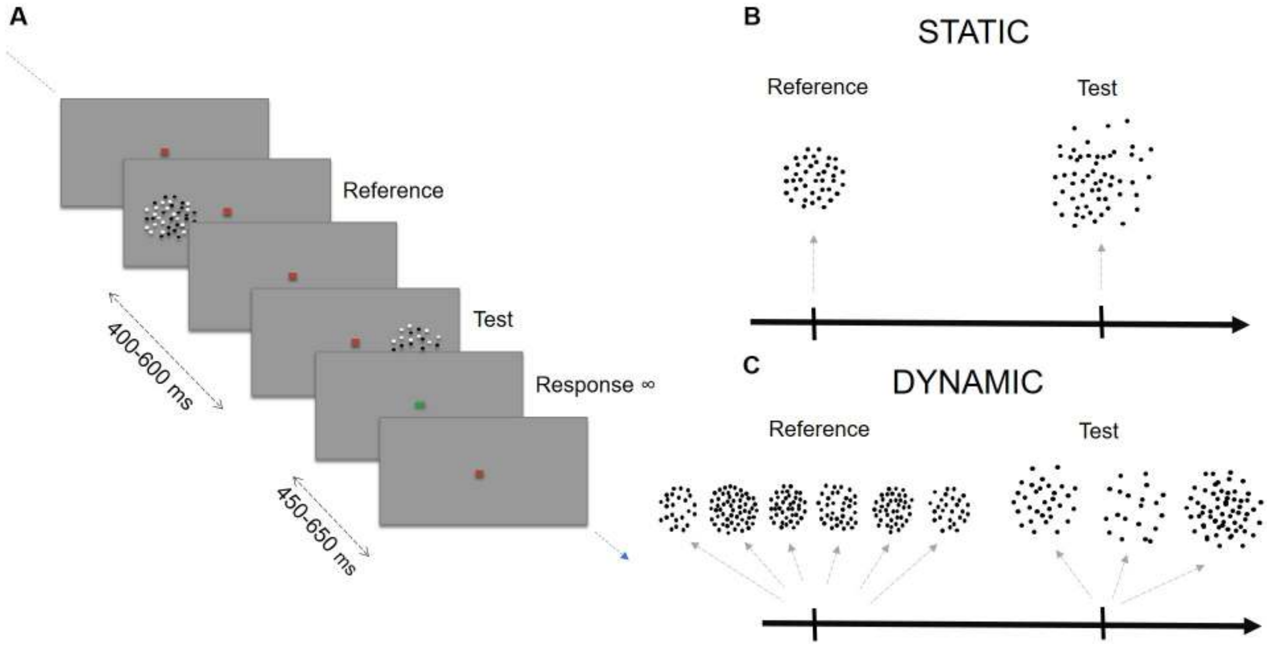

Figure 1: ***General experimental procedure and stimuli used in Exp. 1***. *(A) Example of the stimulus sequence in each trial of Exp.1a-c. In each trial, a sequence of two stimuli (reference and test) was presented on the screen, randomising both their order and position. Participants performed either a duration or a numerosity comparison task, in different sessions. The reference was always constant in the task-relevant dimension, and varied in the task-irrelevant one, while the test varied only in the task-relevant dimension. (B) The static stimuli used in Exp. 1a were defined by a single dot array with a given numerosity and duration according to the task. (C) The dynamic stimuli used in Exp. 1b-c instead included a series of briefly-flashed (50-ms) dot-arrays. This example shows two stimuli from the duration task of Exp. 1b: a 300-ms reference (six 50-ms stimuli), and a 150-ms test stimulus (three 50-ms stimuli). In Exp. 1c we used similar stimuli, but with the numerosity of each individual array computed in order to keep the total amount of the presented dots (instead of the average) equal to 37. Note that dots are depicted in black for visualization purposes, but were black and white in the actual experiment.*

2.3 Experimental design and procedure

2.3.1 General procedure

In all experiments, in separate session, participants performed either a numerosity or a duration comparison task, in a counterbalanced order. In each trial, participants fixated a central point, while a variable test dot-array and a constant reference dot-array appeared sequentially on the two sides of the screen (interstimulus interval = 400-600 ms; horizontal eccentricity from the centre of the screen = 12 deg). Both the position of the stimuli (left/right side of the screen) and their presentation order (test first/reference first) were randomized across trials. After the offset of the second stimulus, subjects were instructed to report either which stimulus contained more dots (but see below for variations in the numerosity task), or which stimulus was displayed for longer time. Responses were made by pressing either the left or the right arrow on a keyboard, to indicate the side of the screen of the more numerous/longer stimulus. After providing a response, the next trial started after 450-650 ms. No feedback about the response was provided. Fig. 1A shows an example of the general procedure used across the different experiments. In Exp. 1, participants performed 5 blocks of 196 trials for each task of each experiment. In Exp. 2, each participant performed 4 blocks of 70 trials for each task of each experiment.

2.3.2 Experiment 1

In the numerosity task of Exp. 1a (static stimuli; Fig. 1B), the test stimulus varied in numerosity from trial to trial (20, 25, 30, 37, 45, 56, or 69 dots), and had constant duration (308 ms), while the reference had always 37 dots and varied in duration (158, 200, 250, 308, 383, 483, or 600 ms). In the duration task, instead, the test stimulus varied in duration (158-600 ms) and was kept constant in numerosity (37 dots), and vice versa for the reference stimulus (20-69 dots, 308 ms).

In Exp. 1b (dynamic average numerosity), the methodology was similar to Exp. 1a, with the exception that instead of considering numerosity as the number of dots in a single dot-array, participants judged the average numerosity over multiple dot-arrays (Fig. 1C). Each stimulus was composed by a series of briefly-flashed (50-ms) dot-arrays with varying numerosity, with a variable number of arrays depending on the stimulus duration. In the numerosity task, the test stimulus varied in average numerosity (20-69), and had constant duration (300 ms; six 50-ms dot-arrays). Conversely, the reference stimulus had an average of

37 dots, but varied in duration (150, 200, 250, 300, 350, 450 or 600 ms, corresponding to a sequence of 3, 4, 5, 6, 7, 9, 12 dot-arrays). The numerosity of each dot-array was randomly chosen to span ± 50% around its base numerosity, and each sequence was computed on a trial-by-trial basis to have an average numerosity equal to the base numerosity selected. At the end of each trial, participants reported which stimulus contained, on average, more dots. In the duration task, the stimuli were constructed in a similar fashion. In this case, the reference stimulus was always 300 ms (six 50-ms dot-arrays), with sequences containing an average numerosity of 20-69 dots. The test stimulus instead varied in duration (150-600 ms; 3-12 dot-arrays), and had a constant average numerosity of 37 dots. Participants reported which of the two dot-arrays lasted longer.

In Exp. 1c (dynamic total numerosity), we used a procedure largely similar to Exp. 1b. The only difference is that the numerosity of each array in the sequences was determined in a way to have a given total number of dots presented over time, rather than having a certain average numerosity. For instance, a stimulus containing 37 dots required dividing this number of dots for the number of arrays in the sequence (i.e., 3-12 arrays, corresponding to 150-600 ms). All the other aspects of the experiment were identical to Exp. 1b. Note that to avoid the repetition of the same number over multiple presentations, the dots were not distributed evenly across the different arrays.

2.3.3 Experiment 2

In Exp. 2a, we first assessed the effects of connectedness (Fig. S1A) on numerosity and multisensory integration (Fig. S1B; see *Supplementary Materials*) on time perception. In the numerosity task, participants compared a reference stimulus (containing either 36 connected or 36 isolated dots) against a variable test stimulus (20-69 isolated dots), using a procedure identical to Exp. 1a. Stimulus duration was not modulated, and all the stimuli were presented for 300 ms. In the duration task, participants performed a duration comparison task. The reference stimulus (300 ms) was always paired with a task-irrelevant auditory stimulus (i.e., a burst of white noise lasting 200 or 500 ms, centered at the middle point of the visual duration). The visual test stimulus instead varied in duration (150-600 ms), and no sound was played during its presentation. Stimulus numerosity was not modulated, and all the stimuli contained 36 dots. The rest of the procedure was identical to Exp. 1a.

A different group of participants was tested to assess the effect of connectedness on perceived duration and multisensory integration on perceived numerosity. In the first case, participants compared the duration of a constant reference (300 ms/36 dots) and a variable test stimulus (150-600 ms/36 dots). However, the reference stimulus could either contain isolated dots, or pairwise-connected dots. In order to test the effect of multisensory integration on numerosity, participants compared the numerosity of a constant reference (36 dots/300 ms) and a variable test stimulus (20-69 dots/300 ms). To induce multisensory integration, an irrelevant burst of white noise (200 or 500 ms) was played simultaneously to the reference.

In Exp. 2b, the procedure and tasks were the same as Exp. 2a; the only difference was the use of dynamic stimuli and the judgment of average numerosity. Namely, we first assessed the effect of connectedness on the perceived average numerosity of sequences of dot arrays, and the effect of multisensory integration on the perceived duration of these stimuli. Then, we assessed the effect of connectedness on perceived duration, and the effect of multisensory integration on perceived numerosity. Except for the use of dynamic stimuli, all the other aspects of the experiment were identical to Exp. 2a.

2.3.4 Data analysis

Across all experiments, the effect of the different experimental manipulations was assessed in terms of difference in the point of subjective equality (PSE), defined as the median of a cumulative Gaussian function fitted to the distribution of "test more numerous/longer" responses at different test values (maximum likelihood method; Watson, 1979). In Exp. 1, we computed a PSE value for each reference magnitude level (i.e., numerosity in the duration task, duration in the numerosity task). The average (±SD) goodness-of-fit (R2) of the psychometric fitting procedure was 0.36 ± 0.09, 0.32 ± 0.07, and 0.26 ± 0.07, for Exp. 1a, 1b, and 1c respectively. In Exp. 2, instead, the goodness of fit was 0.40 ± 0.05 (Exp. 2a) and 0.40 ± 0.07 (Exp. 2b). To have a better index of the magnitude modulation effect, we calculated a magnitude integration (MI) index based on the difference in PSE between the cases where reference and test stimulus had the same magnitude (i.e., the middle value of the magnitude range) and the cases where the reference had either a higher or a lower magnitude level in the task-irrelevant dimension:

$$MI = \left(\frac{PSE_{ref.mag} - PSE_{baseline}}{PSE_{baseline}}\right) \times 100$$

Where $PSE_{ref.mag.}$ indicates the PSE obtained when the reference had a higher or a lower task-irrelevant magnitude level (i.e., numerosity in the duration task, duration in the numerosity task). $PSE_{baseline}$ indicates the PSE obtained when reference and test had the same level of the task-irrelevant magnitude, which was taken as a baseline measure of duration/numerosity comparison performance. This index indicates to what extent the perceived duration or numerosity of the reference changes as a function of the interfering task-irrelevant magnitude, and allows to directly compare the effect of the different magnitudes in the different tasks. A negative magnitude integration index indicates a relative underestimation compared to baseline, while a positive index indicates overestimation.

To better capture the difference in the effect yielded by static and dynamic stimuli, we also computed an "effect index" for each task in Exp.1a-c. This index is computed as the difference in the bias observed at the two extremes of the task-irrelevant dimension of the reference stimulus (i.e., 158 and 600 ms or 150 and 600 ms in the numerosity task, respectively for Exp. 1a and Exp. 1b-c, and 20 and 69 dots in the duration task). Note that for the numerosity task of Exp.1c the effect index was computed according to the average numerosity of the stimuli.

Data analysis of Exp. 2 was similarly based on PSEs and on a magnitude integration index reflecting the difference in PSE obtained with connected- versus isolated-dots, or with a short versus a long sound.

In all experiments we also assessed participants' performance in terms of precision (i.e., sensitivity), measuring and analysing Weber's fractions. More details about this analysis are included in the Supplementary Materials.

3 Results

3.1 Experiment 1

In Exp. 1, we assessed the effect of magnitude integration between time and numerosity. Participants performed either a duration or a numerosity comparison task, between a constant reference and a variable test stimulus. While participants explicitly judged the task-relevant dimension of the stimuli according to the task (either duration or numerosity), we also modulated the task-irrelevant dimension of the reference (i.e., numerosity in the duration task and duration in the numerosity task) to induce a magnitude integration bias. Importantly, while in Exp. 1a we used static stimuli (a single dot-array), in Exp. 1b-c we used dynamic stimuli (series of visual dot-arrays), in order to transform numerosity into a temporally-distributed dimension more similar to duration.

With static dot-array stimuli (Exp. 1a), we predicted an asymmetric integration effect, i.e., a greater effect of numerosity on perceived duration compared to the effect of duration on numerosity. Participants' performance in the tasks was assessed in terms of a magnitude integration index, computed as the difference in perceived magnitude (PSE) obtained by modulating the task-irrelevant dimension of the reference (i.e., either numerosity or duration in the duration and numerosity task, respectively; see *Material and Methods*). The average PSEs obtained in the different experiments and tasks are shown in Fig. S1, while measures of precision in the task (Weber's fraction) are shown in Fig. S2; see *Supplementary Materials*.

The effect of magnitude integration was assessed with a linear mixed-effect regression model, with "task" (i.e., numerosity vs. duration) and "ratio" (i.e., the ratio between each level of reference duration or numerosity and the middle value of the reference duration/numerosity range considered as baseline) as fixed effects, subjects as the random effect, and the magnitude integration index as dependent variable. This model ($R2 = 0.59$) showed a significant effect of task ($t = 4.37$, $p < 0.001$), and a significant effect of ratio ($t = 13.70$, $p < 0.001$). This means that, for instance, in the duration task, duration estimates were biased congruently by the numerosity of the reference: fewer dots than the test (20-30 vs. 37 dots) led to underestimation, while more dots (46-69 dots) led to overestimation. A similar pattern, although weaker, was also observed in the numerosity task, with an under- and over-estimation of numerosity due to the duration of the reference.

The asymmetry of the integration bias between duration and numerosity was revealed by a significant interaction between task and ratio ($t = -4.46$, $p < 0.001$). A series of post-hoc tests revealed that the effect of numerosity on duration perception was significantly stronger compared to the effect of duration on numerosity at intermediate (0.64 and 0.68, respectively for duration and numerosity; $t(14) = 2.07$, $p = 0.042$, Cohen's $d = 0.54$) and extreme (0.51/0.54, 1.93/1.96; $t(14) = 2.81$, $p = 0.006$, $d = 0.65$; $t(14) = 2.16$, $p = 0.035$, $d = 0.67$) ratios. At the ratios closest to 1 (0.81 and 1.23/1.22) and at the intermediate ratio above 1 (1.53/1.51) the effect of numerosity on duration perception did not significantly differ from the effect of duration on numerosity (all t-values ≤ 1.21, all p-values ≥ 0.231).

In Exp. 1b we used dynamic stimuli (see Fig. 1C) to reduce the temporal offset between the representation of time and numerosity – i.e., by making time and numerosity to unfold similarly over time. To this aim, we presented a sequence of arrays, and made participants to evaluate either their duration, or the average numerosity across them. Fig. 2B shows the results of Exp. 1b. Here we observed a much larger effect of duration on numerosity, comparable if not even slightly stronger than the effect of numerosity on duration. A linear mixed-effect model ($R2 = 0.64$) showed again an effect of ratio on the magnitude integration index ($t = 8.44$, $p < 0.001$), but no effect of task ($t = 0.39$, $p = 0.069$). Importantly, we did not observe any significant interaction between task and ratio ($t = 0.30$, $p = 0.76$), indicating that the effect in the two tasks was largely symmetrical. These results demonstrate that with dynamic stimuli where numerosity is defined over time, magnitude integration operates in a similar fashion across different magnitudes.

However, in Exp. 1b, it remains unclear whether the effect in the numerosity task is driven by duration, or by the total number of items displayed over time, which increased with duration. To dissociate the ef-

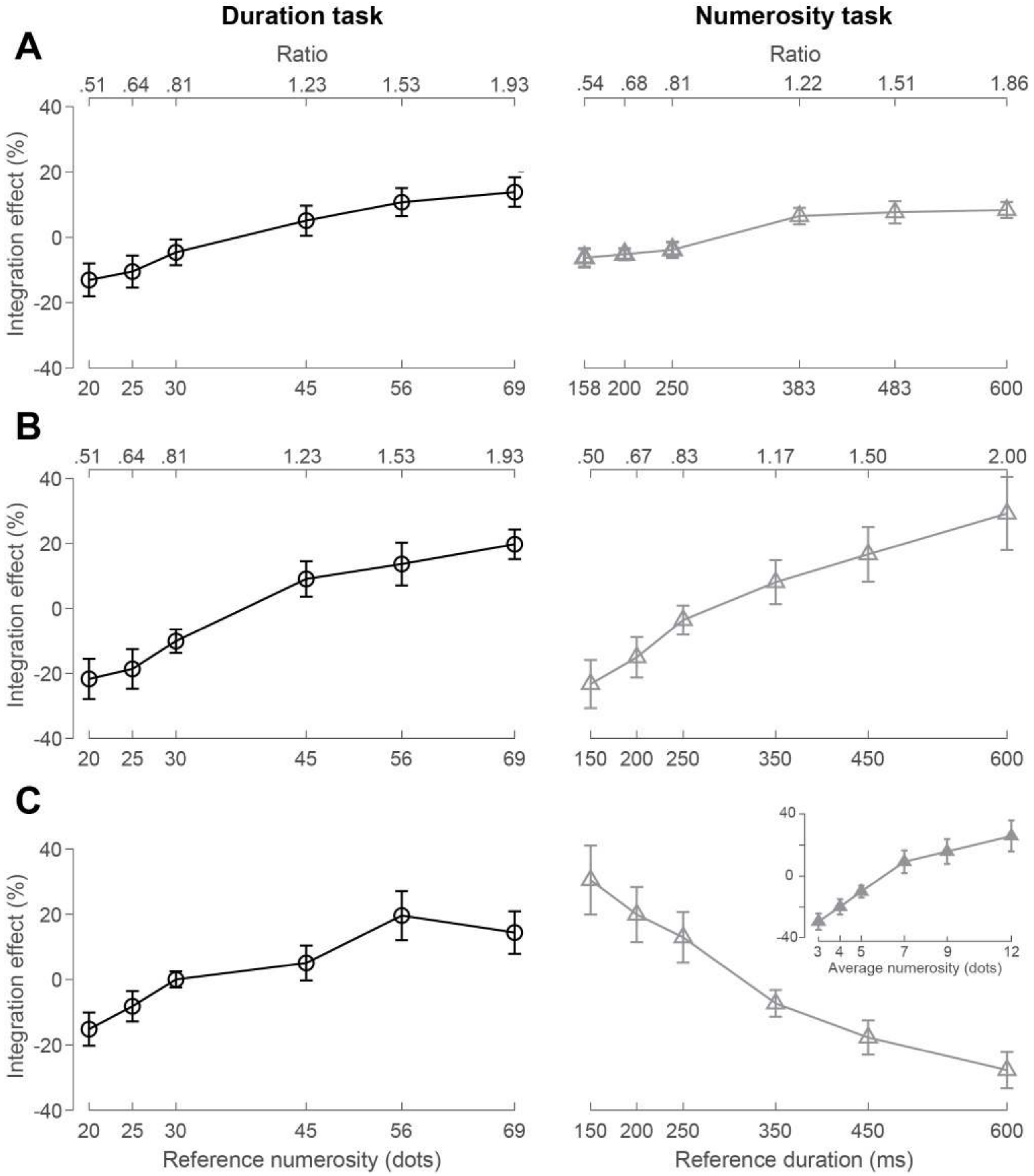

Figure 2: ***Results of Exp. 1***. *All the plots show the magnitude integration index (i.e., reflecting the perceptual bias) as a function of the different levels of reference magnitude. (A) Results of Exp. 1a where we used static stimuli. (B) Results of Exp. 1b where we used dynamic stimuli based on average numerosity. (C) Results of Exp. 1c where we used dynamic stimuli based on total numerosity. The inset in the bottom-right panel shows the integration effect re-coded as a function of average numerosity, rather than duration (i.e., corresponding to 12, 9, 7, 5, 4, and 3 dots for durations ranging 150-650 ms). Error bars represent 95% confidence intervals (CI). The upper x-axes report the ratio of the different levels of the reference magnitude with the middle value of the range. The lower x-axes report the actual magnitude of the reference stimulus in each task.*

fect of the average numerosity from the total numerosity over time, we performed an additional experiment (Exp. 1c) in which we kept fixed the total number of dots over time, introducing an opposite relationship between duration and average numerosity (the shorter the duration, the higher the average numerosity).

The results (Fig. 2C) show again a robust effect of numerosity on perceived duration, but an opposite effect of duration on numerosity: shorter durations led to a relative numerosity overestimation compared to longer durations. A linear mixed-effect model (R2 = 0.63) showed a significant effect of task (t = 15.01, p $<$ 0.001), and ratio (t = 8.44, p $<$ 0.001), and a significant interaction between them (t = -16.69, p $<$ 0.001). Post-hoc tests showed that the effect of numerosity on duration was significantly different from the effect of duration on numerosity, at all the ratio levels (t-values $\geq$ 2.825, p-values $\leq$ 0.007).

These results suggest that the numerical estimates in this context are critically driven by the average numerosity of the sequences, rather than by the total number of dots displayed. By re-coding the different levels of reference magnitude according to average numerosity (inset in the rightmost panel of Fig. 2C) in the numerosity task, we observed a pattern consistent with the results of Exp.1b. A linear mixed-effect model similarly showed an effect of both task (t = -3.98, p < 0.001) and ratio (t = 8.33, p $<$ 0.001), and a significant interaction between them (t = 4.10, p < 0.001). Differently from Exp. 1a, this interaction shows that the effect of duration on numerosity is actually stronger that the effect of numerosity on duration. Post-hoc tests showed a significant difference between the effects on numerosity and duration at the extremes of the range (ratio 0.51/0.50: t(14) = 2.539, p = 0.014, d = 0.80; 1.93/2: t(14) = 3.236, p = 0.002, d = 0.64; no significant differences were observed elsewhere, t-values $\leq$ 1.927, p-values $\geq$ 0.059).

Finally, we computed an "effect index" in each task and compared it across the experiments (see Material and Methods). The average effect index across Exp. 1a-c is shown in the Supplementary Fig. S3. A two-way ANOVA with factors "experiment" and "task" showed a main effect of both factors (F(2,28) = 16.42, p < 0.001, η_p^s = 0.28, and F(1,14) = 5.03, p = 0.028, η_p^s = 0.06, respectively), and a significant interaction (F(2,28) = 8.43, p $<$ 0.001, η_p^s = 0.17). A series of Bonferroni-corrected post-hoc tests (alpha = 0.008) showed no statistically significant differences across the different duration tasks of Exp. 1 (t $\leq$ -2.64, p $\geq$ 0.013). Conversely, across the numerosity tasks, the effect obtained with the average and total numerosity stimuli was significantly different from the effect obtained with static stimuli (t(28) = 4.89, p < 0.001, d = 1.78; t(28) = 7.24, p $<$ 0.001, d = 2.64). No significant difference was observed between the average and the total numerosity stimuli (t(28) = -0.61, p = 0.55). Finally, we compared the effect obtained in the duration vs. the numerosity task separately for each experiment (alpha = 0.017). The results show that while there is a significant difference between the effect of numerosity on duration and duration on numerosity with static stimuli (Exp. 1a; t(28) = 2.92, p = 0.007, d = 1.06) – highlighting indeed an asymmetric pattern of magnitude integration – there is no difference in Exp. 1b (t(28) = 1.30, p = 0.20, d = 0.48). Additionally, we also observed a significant difference in Exp. 1c, reflecting the stronger bias found in the numerosity compared to the duration task (t(28) = 3.71, p $<$ 0.001, d = 1.35). However, since the bias in the numerosity task is driven by the average numerosity, this result cannot be interpreted as reflecting an asymmetry between duration and numerosity.

3.2 Experiment 2

To further asses the nature of magnitude integration and shed light into its time-course, we investigated the effects of illusory changes in the perception of the task-irrelevant magnitude dimension of the reference stimulus on the perception of the task-relevant dimension of the test. To this aim, we exploited the connectedness illusion [17] to distort perceived numerosity, and multisensory integration to change perceived duration [19]. We chose these manipulations because each of them is known to happen at a specific processing stage: a relatively "early" processing stage in the case of connectedness (i.e., 150 ms from stimulus onset [16]), or a later stage in case of multisensory integration (i.e., since the integration between the two durations should occur after their offset).

In Exp. 2a we used static stimuli. First, we assessed these effects within the single appropriate dimension (i.e., the effect of connectedness on numerosity, Fig. 3A, and of multisensory integration on duration, Fig. 3B). Connectedness should indeed reduce the perceived numerosity of an array when compared to the same number of isolated dots, while the presence of an irrelevant sound should bias the perceived duration of a visual stimulus via multisensory integration. Then, we tested the effect of connectedness on duration perception (Fig. 3C) and the effect of multisensory integration on numerosity perception (Fig. 3D).

As shown in Fig. 3A-B, in both conditions we observed strong and systematic perceptual biases. The numerosity of a connected dot-array was strongly underestimated compared to the same array of isolated dots, by on average 34% (paired sample t-test, t(23) = 11.07, p < 0.001, d = 2.86). Similarly, the presence of a task-irrelevant sound paired with a visual stimulus strongly affected its perceived duration, leading to under- or over-estimation according to the sound duration (average effect of ±29% around the physical reference duration of 300 ms; short vs. long sound, t(23) = 9.19, p $<$ 0.001, d = 2.98).

Next, we tested how such illusory modulations of perceived numerosity and duration impact magnitude integration. Specifically, in the duration task, the reference stimulus (300 ms/36 dots) could either contain isolated or pairwise connected dots, whereas in the numerosity task the reference stimulus (36 dots/300 ms) was paired with a task-irrelevant sound (200 or 500 ms). By using static stimuli (Fig. 3C-D), however, we did not observe any significant modulation of magnitude integration. Specifically, we did not observe any effect of connectedness (average effect of $\sim$3%; t(23)

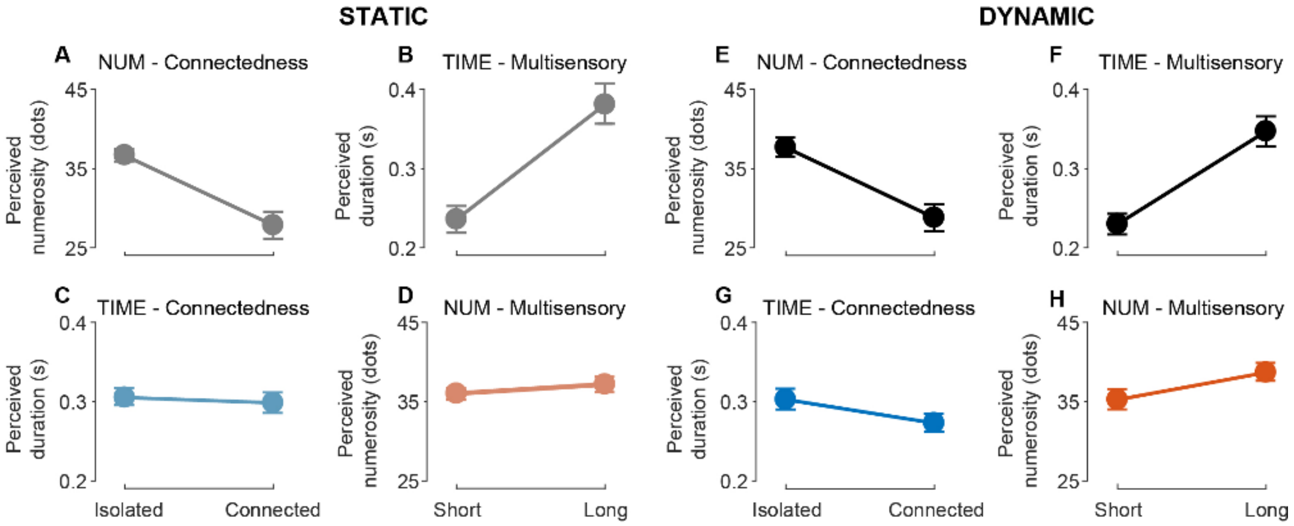

Figure 3: ***Results of Exp. 2****. (A-D) Results of Exp. 2a using static stimuli, in terms of perceived numerosity or duration (average PSE). (A) Effect of connectedness on numerosity perception. (B) Effect of multisensory integration on perceived duration. (C) Effect of connectedness on perceived duration. (D) Effect of multisensory integration on numerosity. (E-H) Results of Exp. 2b using dynamic stimuli. (E) Effect of connectedness on the perceived average numerosity of sequences of dot-array stimuli. (F) Effect of multisensory integration on perceived duration. (G) Effect of connectedness on duration. (H) Effect of multisensory integration on average numerosity. Error bars represent 95% CI.*

= 0.97, p = 0.342, d = 0.24; Fig. 3C) on duration comparison, suggesting that the bias of numerosity on time perception is based on the veridical number of dots, irrespective of connectedness. Similarly (Fig. 3D), we did not observe any significant effect of the task-irrelevant sound on perceived numerosity (average effect = ∼3.5%; t(23) = 1.97, p = 0.061, d = 0.58).

In Exp. 2b we used again dynamic stimuli (based on average numerosity as in Exp. 1b). As shown in Fig. 3E-F, both illusions appear to be robust also in this context. Namely, connectedness strongly reduced the perceived average numerosity of a sequence of pairwise connected dot-arrays (average effect ∼34%; t(23) = 8.71, $p < 0.001$, d = 2.62). Also multisensory integration biased the perceived duration of dynamic stimuli, with under- or over-estimation according to the sound duration (average effect ±24% around the physical duration of 300 ms; t(23) = 11.46, $p < 0.001$, d = 3.09). Differently from Exp. 2a, when testing how these two illusions impact the numerosity-time integration, we found robust effects. Namely, the average numerosity perceptually reduced by connectedness caused a significant underestimation of perceived duration (average effect ∼12%; t(23) = 4.34, $p < 0.001$, d = 1.10). Similarly, the distorted perceived duration of the visual stimulus caused by the sounds affected perceived numerosity (average effect ∼10%; t(23) = 4.73, $p < 0.001$, d = 1.33), resulting in a relative under- or over-estimation according to the sound duration. Precision measures obtained in the different conditions of Exp. 2 are shown in Fig. S5 (*Supplementary Materials*).

To better assess the effects of connectedness and multisensory integration on time and numerosity perception – and the role of dynamic stimuli – we directly compared them (Fig. 4). Namely, we performed a two-way (independent-sample) ANOVA, with factors "stimulus type" (static vs. dynamic) and "task" (time vs. numerosity). The results showed a significant effect of stimulus type (F(1,92) = 11.88, $p < 0.001$, η_p^s = 0.114), and no effect of task (F(1,92) = 0.15, p =

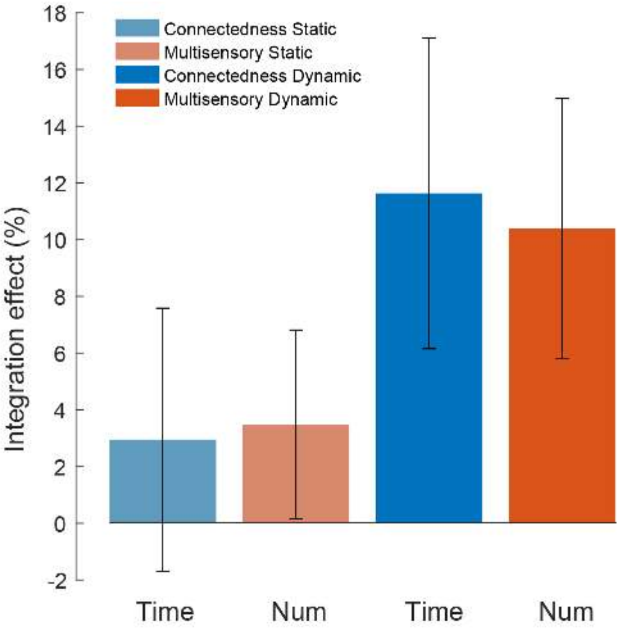

Figure 4: ***Comparison between cross-magnitude effects in Exp. 2****. (Average magnitude integration indexes measured when testing the effect of connectedness on duration (blue bars) and the effect of multisensory integration on numerosity (red bars). The effect was calculated from the PSE data shown in Fig. 3 (panels C, D, G, and H). Error bars represent 95% CI.*

0.70) or interaction between the two factors (F(1,92) = 0.03, p = 0.87). This suggests that the effect of illusory perception of time and numerosity on magnitude integration strongly depends on the stimulus presentation format (i.e., yielding stronger effects with dynamic stimuli) and also in this case dynamic stimuli give rise to a symmetrical interaction between the different magnitudes.

4 Discussion

The present work addresses two important questions about magnitude perception and integration: what causes asymmetries in magnitude integration and when does magnitude integration occur? An important feature of magnitude perception is the integration of information across different dimensions and the perceptual biases linked to it. The presence of biases across magnitude dimensions has suggested the existence of a generalized system representing magnitude information with the same neural code Walsh, 2003), putatively located in high-level associative cortices like for instance the posterior parietal cortex (Walsh, 2003; Harvey et al., 2013; Harvey et al., 2015). However, results showing asymmetric integration effects across different dimensions challenged this idea (Casasanto and Boroditsky, 2008). The inconsistency of previous results has left many doubts about the mechanism underlying magnitude integration. Here we showed that a symmetrical interaction between time and numerosity can be observed if the two dimensions become temporally equivalent and numerosity unfolds over time.

In line with previous findings (Cappelletti, Freeman, and Cipolotti, 2009; Oliveri et al., 2008; Roitman et al., 2007; Dormal, Seron, and Pesenti, 2006), our results with static stimuli (Exp. 1a) showed an asymmetric relationship between time and numerosity. In the presence of static stimuli this is not surprising, since the processing of numerosity and duration likely entails very different time-courses. Namely, while temporal information necessarily unfolds over time, and representing duration is possible only at the offset of the stimulus itself, numerosity processing has been shown to be relatively fast, starting very early after stimulus onset (i.e., 75-100 ms; Fornaciai, Cicchini, and Burr, 2016; Fornaciai and Park, 2018) and likely ending 250 ms after. Numerosity could then interfere with temporal processing even before duration representation, while duration could interfere with numerosity only after numerosity has already been represented. According to this, using static stimuli the interaction effects are likely of very different nature: while numerosity would affect how time is processed (i.e., during a hypothetical accumulation of temporal information), duration could affect numerosity only at a post-perceptual processing stage, where numerosity representations are presumably less prone to biases. Thus, according to our hypothesis, if there is a temporal offset between the representation of two dimensions, the faster one would have an advantage over the slower one. However, when numerosity is defined over time by using dynamic stimuli (Exp. 1b), its representation becomes much more vulnerable to duration. In this scenario, numerical and temporal information would unfold in parallel and the lack of temporal offset between them leads to symmetrical biases. The asymmetry between dimensions would thus be related to differences in the processing upstream to a hypothetical high-level magnitude system (Walsh, 2003). An alternative interpretation, however, is that with dynamic stimuli duration could be exploited by the visual system as additional sensory evidence to compute numerosity. While this may seem intuitively plausible – especially in the case of total numerosity, where the longer the sequence, the more items could be displayed within it – it is less plausible for the average numerosity. Indeed, although the average numerosity of a sequence requires integrating information over time, it does not require any temporal information per se, and in Exp. 1b was completely uncorrelated to duration.

Results from Exp. 2 provide further evidence concerning the nature of magnitude integration. First, the fact that multisensory integration does not affect the perceived numerosity of static dot-arrays suggests again that once a numerosity representation is formed, it is not prone to biases. Surprisingly though, with static stimuli, changes in perceived numerosity induced by connectedness do not affect perceived duration. Since the connectedness illusion is based on segmentation processes occurring relatively early in the visual processing stream (i.e., 150 ms, in early visual areas like V2-V3; Fornaciai et al., 2017), this result suggests that the effect of numerosity on time may occur even before perceptual segmentation. In other words, this effect is based on an early representation of numerosity which is not affected by connectedness (i.e., arising around 100 ms; Fornaciai et al., 2017). This finding supports the idea that – with static stimuli – the effect of numerosity on time arises very early, possibly affecting directly the encoding of duration. In the presence of dynamic stimuli, instead, changes in perceived numerosity due to connectedness affected time perception, and changes in perceived duration due to multisensory integration affected numerosity perception. Both these findings further support the idea that when the neural time-course of the different dimensions is "equated," magnitude integration becomes symmetrical.

Importantly, our results also suggest that the magnitude integration bias observed in the present study is perceptual in nature, rather than based on decision, memory, or a linguistic/conceptual representation (Casasanto and Boroditsky, 2008; Cai et al., 2018). Indeed, for instance, a memory or decision bias should always be linked to perceived magnitude irrespective of the physical properties of the stimuli (i.e., static versus dynamic) and of the manipulation used to bias their perception (i.e., connectedness and multisensory integration), which is not the case considering the present results. However, it is also important to note that perceptual biases and other types of interactions between magnitudes are not mutually exclusive, and may be elicited in different contexts or even co-exist.

How can this framework explain previous reports of symmetric biases across magnitudes? For instance, Javadi and Aichelburg (Javadi and Aichelburg, 2012), by modulating the duration and numerosity of static dot-array stimuli, observed similar biases across numerosity and time. This symmetrical effect could be explained by the short duration range used (53-106 ms), likely yielding a minimal offset between numerosity and duration representation, in line with the idea that the time-course of information processing determines the interaction between magnitudes. Furthermore, it has been hypothesized [27] that the asymmetry between space and time, usually observed with visual stimuli, might be due to the different sensitivity of the visual system to spatial and temporal information. In support to this idea, Cai and Connell (Cai and Connell, 2015) using tactile stimuli show a reversed asymmetry, with duration biasing spatial judgments but not the opposite. In this case, the poor spatial resolution of the tactile modality, together with the fact that spatial estimates required the integration of haptic signals from the two hands (i.e., participants touched the two ends of a stick with the two index fingers), might have contributed to a sluggish encoding of spatial information. And this, in turn, might have reduced the temporal offset between the processing of spatial and temporal information.

Finally, another aspect of the time-numerosity integration is the directionality of the perceptual bias, determined by the contribution of different stimulus dimensions to the overall energy of the stimuli. In general, increasing or decreasing either the duration or the numerosity of the reference stimulus (compared to the test) either increases or decreases the relative energy of the stimulus, determining the direction of the bias. Two exceptions to this trend have been observed. The first is in the numerosity task of Exp.1a, where changes in the duration of the reference did not proportionally affect its perceived numerosity (i.e., asymmetric interaction between time and numerosity). This suggests that, with static stimuli, the contribution of duration to the energy of the stimuli is limited. The second exception is in the numerosity task of Exp. 1c, where keeping the total numerosity of the reference constant introduced a negative correlation between numerosity and time: the shortest stimulus was paired with the highest numerosity, and vice-versa. This resulted in an incongruent bias: the shorter the duration the higher the perceived numerosity. This finding first shows that the stimulus energy, in terms of numerosity, is determined by the number of items presented in a single array, rather than by the total amount of items presented across multiple arrays. Second, it suggests that to observe a symmetric bias it is important not only to make the stimuli temporally equivalent, but also to make sure that the different dimensions provide congruent contributions to the overall stimulus energy.

To conclude, our results shed new light on the nature of integration between different magnitude dimensions and its neural mechanisms. Overall, our findings suggest that the dynamics of low-level magnitude processing play an important role in determining the symmetry of magnitude integration: a stimulus dimension rapidly processed would more easily affect a slowly processed one than vice versa. When the stimulus dynamics are equated by introducing a temporal component into a spatial dimension like numerosity, the resulting integration can be perfectly symmetrical. We thus conclude that the operation of a high-level generalized magnitude system could be symmetrical in nature, and that any asymmetry in integration effects may arise with differences in the nature of the stimuli used and in the underlying sensory processing.

Supplementary materials. Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5604593.

Data availability. All the data generated during the experiments described in this manuscript are available on Open Science Framework at this link: https://osf.io/93uwb/.

Acknowledgements. We thank Matteo Fantuz for his assistance during the data collection. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme grant agreement No. 682117 BIT-ERC-2015-CoG, and from the Italian Ministry of University and Research under the call FARE (project ID: R16X32NALR) to DB, and from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 838823 "NeSt" to MF. This open access manuscript has been formatted in LaTeX using the Wenneker article template (https://www.latextemplates.com/template/wenneker-article).

Conflict of interest. The Authors declare no competing financial interest.

References

Bueti, Domenica and Vincent Walsh (2009). "The parietal cortex and the representation of time, space, number and other magnitudes." In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364.1525, pp. 1831–40. ISSN:

1471-2970. DOI: 10.1098/rstb.2009.0028. URL: http://www.ncbi.nlm.nih.gov/pubmed/19487186http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC2685826.

Burr, David C. et al. (2010). "Saccades Compress Space, Time, and Number". In: *Trends in cognitive sciences* 14.12, pp. 528–533. ISSN: 13646613. DOI: 10.1016/B978-0-12-385948-8.00012-8. URL: http://dx.doi.org/10.1016/j.tics.2010.09.005.

Cai, Zhenguang G and Louise Connell (2015). "Space–time interdependence: Evidence against asymmetric mapping between time and space". In: *Cognition* 136, pp. 268–281. ISSN: 00100277. DOI: 10.1016/j.cognition.2014.11.039. URL: https://linkinghub.elsevier.com/retrieve/pii/S0010027714002650.

Cai, Zhenguang G. et al. (2018). "Cross-dimensional magnitude interactions arise from memory interference". In: *Cognitive Psychology* 106, pp. 21–42. ISSN: 00100285. DOI: 10.1016/j.cogpsych.2018.08.001. URL: https://linkinghub.elsevier.com/retrieve/pii/S0010028518300100.

Cappelletti, Marinella, Elliot D. Freeman, and Lisa Cipolotti (2009). "Dissociations and interactions between time, numerosity and space processing". In: *Neuropsychologia* 47.13, pp. 2732–2748. ISSN: 00283932. DOI: 10.1016/j.neuropsychologia.2009.05.024.

Casasanto, Daniel and Lera Boroditsky (2008). "Time in the mind: Using space to think about time". In: *Cognition* 106.2, pp. 579–593. ISSN: 0010-0277.

Dehaene, Stanislas and Elizabeth M. Brannon (2010). "Space, time, and number: a Kantian research program." In: *Trends in cognitive sciences* 14.12, pp. 517–9. ISSN: 1879-307X. DOI: 10.1016/j.tics.2010.09.009. URL: http://www.ncbi.nlm.nih.gov/pubmed/20980194.

DeWind, Nicholas K et al. (2015). "Modeling the approximate number system to quantify the contribution of visual stimulus features." In: *Cognition* 142, pp. 247–65. ISSN: 1873-7838. DOI: 10.1016/j.cognition.2015.05.016. URL: http://www.ncbi.nlm.nih.gov/pubmed/26056747.

Dormal, Valérie and Mauro Pesenti (2013). "Processing numerosity, length and duration in a three-dimensional Stroop-like task: towards a gradient of processing automaticity?" In: *Psychological Research* 77.2, pp. 116–127. ISSN: 0340-0727. DOI: 10.1007/s00426-012-0414-3. URL: http://link.springer.com/10.1007/s00426-012-0414-3.

Dormal, Valérie, Xavier Seron, and Mauro Pesenti (2006). "Numerosity-duration interference: A Stroop experiment". In: *Acta Psychologica* 121.2, pp. 109–124. ISSN: 00016918. DOI: 10.1016/j.actpsy.2005.06.003.

Fornaciai, Michele, Guido Marco. Cicchini, and David C. Burr (2016). "Adaptation to number operates on perceived rather than physical numerosity". In: *Cognition* 151, pp. 63–67. ISSN: 00100277. DOI: 10.1016/j.cognition.2016.03.006. URL: https://linkinghub.elsevier.com/retrieve/pii/S0010027716300610.

Fornaciai, Michele and Joonkoo Park (2018). "Early Numerosity Encoding in Visual Cortex Is Not Sufficient for the Representation of Numerical Magnitude". In: *Journal of Cognitive Neuroscience* 30.12, pp. 1788–1802. ISSN: 0898-929X. DOI: 10.1162/jocn_a_01320. URL: https://www.mitpressjournals.org/doi/abs/10.1162/jocn_a_01320.

Fornaciai, Michele, Irene Togoli, and Roberto Arrighi (2018). "Motion-induced compression of perceived numerosity". In: *Scientific Reports* 8.1, p. 6966. ISSN: 2045-2322. DOI: 10.1038/s41598-018-25244-8. URL: http://www.nature.com/articles/s41598-018-25244-8.

Fornaciai, Michele et al. (2017). "Numerosity processing in early visual cortex". In: *NeuroImage* 157. ISSN: 10959572. DOI: 10.1016/j.neuroimage.2017.05.069.

Harvey, Ben M. et al. (2013). "Topographic representation of numerosity in the human parietal cortex". In: *Science* 341.September, pp. 1123–1126. ISSN: 1095-9203. DOI: 10.1126/science.1239052.

Harvey, Ben M. et al. (2015). "Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex". In: *Proceedings of the National Academy of Sciences* 112.44, pp. 13525–13530. ISSN: 0027-8424. DOI: 10.1073/pnas.1515414112. URL: http://www.pnas.org/lookup/doi/10.1073/pnas.1515414112.

He, Lixia et al. (2009). "Connectedness affects dot numerosity judgment: implications for configural processing." In: *Psychonomic bulletin & review* 16, pp. 509–517. ISSN: 1069-9384. DOI: 10.3758/PBR.16.3.509.

Heron, James et al. (2013). "A neural hierarchy for illusions of time : Duration adaptation precedes multisensory integration". In: *Journal of Vision* 13.14, pp. 1–12. ISSN: 1534-7362. DOI: 10.1167/13.14.4.doi.

Javadi, Amir Homayoun and Clarisse Aichelburg (2012). "When Time and Numerosity Interfere: The Longer the More, and the More the Longer". In: *PLoS ONE* 7.7. Ed. by Trevor Bruce Penney, e41496. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0041496. URL: https://dx.plos.org/10.1371/journal.pone.0041496.

Merritt, Dustin J., Daniel Casasanto, and Elizabeth M. Brannon (2010). "Do monkeys think in metaphors? Representations of space and time in monkeys and humans". In: *Cognition* 117.2, pp. 191–202. ISSN: 00100277. DOI: 10.1016/j.cognition.2010.08.011. URL: http://dx.doi.org/10.1016/j.cognition.2010.08.011.

Oliveri, Massimiliano et al. (2008). "Perceiving numbers alters time perception". In: *Neuroscience Letters*

438.3, pp. 308–311. ISSN: 03043940. DOI: 10.1016/j.neulet.2008.04.051.

Park, Joonkoo et al. (2016). “Rapid and Direct Encoding of Numerosity in the Visual Stream”. In: *Cerebral Cortex* 26.2, pp. 748–763. ISSN: 14602199. DOI: 10.1093/cercor/bhv017.

Roitman, Jamie D. et al. (2007). “Nonverbal representation of time and number in adults”. In: *Acta Psychologica* 124.3, pp. 296–318. ISSN: 00016918. DOI: 10.1016/j.actpsy.2006.03.008.

Walsh, Vincent (2003). “A theory of magnitude: common cortical metrics of time, space and quantity”. In: *Trends in Cognitive Sciences* 7.11, pp. 483–488. ISSN: 13646613. DOI: 10.1016/j.tics.2003.09.002. URL: https://linkinghub.elsevier.com/retrieve/pii/S1364661303002304.

Watson, Andrew B. (1979). “Probability summation over time”. In: *Vision Research* 19.5, pp. 515–522. ISSN: 00426989. DOI: 10.1016/0042-6989(79)90136-6.

Winter, Bodo, Tyler Marghetis, and Teenie Matlock (2015). “Of magnitudes and metaphors: Explaining cognitive interactions between space, time, and number”. In: *Cortex* 64, pp. 209–224. ISSN: 00109452. DOI: 10.1016/j.cortex.2014.10.015.

Xuan, Bin et al. (2007). “Larger stimuli are judged to last longer.” In: *Journal of vision* 7.10, pp. 2.1–5. ISSN: 1534-7362. DOI: 10.1167/7.10.2. URL: http://www.ncbi.nlm.nih.gov/pubmed/17997671.