Magnitude processing and integration entail perceptual processes independent from the task

Irene Togoli[1,2], Olivier Collignon[1], Domenica Bueti[2], Michele Fornaciai[1,2,*]

[1] Institut de recherche en sciences psychologiques (IPSY), Université catholique de Louvain, Louvain-la-Neuve, Belgium

[2] Cognitive Neuroscience Department, International School for Advanced Studies (SISSA), Trieste, Italy

* Correspondence:
Michele Fornaciai (michele.fornaciai@gmail.com)

ABSTRACT

The magnitude dimensions of visual stimuli, such as their numerosity, duration, and size, are intrinsically linked, leading to mutual interactions across them. Namely, the perception of one magnitude is influenced by the others, so that a large stimulus is perceived to last longer in time compared to a smaller one, and vice versa. The nature of such interaction is however still debated. In the present study we address whether magnitude integration could arise from "automatic" perceptual processes, independently from the task performed, or whether it arises from active decision making. In two separate experiments, participants watched a series of dot-array stimuli modulated in numerosity, duration, and item size. In one case (task condition), the task required them to judge the stimuli in each trial, with the specific dimension to judge indicated by a cue presented after each stimulus. In the other case (passive condition), instead, participants passively watched the stimuli. The behavioural results obtained in the task show robust magnitude integration effects across all three dimensions. Then, we identified a neural signature of magnitude integration by showing that relatively early event-related potentials can predict the effect measured behaviourally. Finally, we demonstrate an almost identical modulation of brain responses in passive viewing, occurring at the same processing stages linked to the behavioural effect. The results thus suggest that magnitude integration likely occurs via automatic perceptual processes that are engaged irrespective of the task-relevance of the stimuli, and independently from decision making.

Keywords. Magnitude perception, magnitude integration, EEG, numerosity perception, time perception, size perception.

INTRODUCTION

Magnitude dimensions such as numerosity, time, and space represent fundamental properties of the external world, as each of these dimensions provides essential information to understand and navigate the environment. Indeed, the perception of magnitudes allows us to appreciate how many objects are around us, their size and their spatial relations, and the duration and timing of the external events. While these dimensions are important in their own rights and studied in separate lines of research, a particularly interesting phenomenon is their integration and interaction. Different magnitude dimensions seem indeed linked in a way that the perception of one dimension depends on the others, usually leading to mutual biases. For instance, a large object or a numerous set of items is perceived as lasting longer in time compared to a smaller object or fewer items (e.g., Xuan et al., 2007). Vice versa, a longer stimulus may appear bigger or more numerous than a shorter one (e.g., Javadi & Aichelburg, 2012; Lambrechts et al., 2013; Togoli et al., 2021).

These mutual biases across magnitude dimensions – or "magnitude integration" effects – represent one of the core phenomena characterising magnitude perception, and have inspired several important theories in this field. One of the most influential theories in this context is the "a theory of magnitude" (ATOM) framework, proposed by Vincent Walsh in 2003 (Walsh, 2003). According to this theory, the processing of different magnitudes culminates in a generalised magnitude system encoding different dimensions with the same neural code. This in turn would allow the interaction of magnitude information in the service of perception and behaviour. This view has been however challenged by the idea that biases across magnitudes, and especially space and time, may stem from the linguistic labels assigned to them, and how we conceptualise these dimensions at a cognitive level ("metaphoric theory;" e.g., Casasanto & Boroditsky, 2008). While evidence has now been accumulated against a linguistic/conceptual view of magnitude integration (Cai & Connell, 2015; Togoli, Bueti, et al., 2022; Whitaker et al., 2022), other theories have proposed that magnitudes interact at a more cognitive rather than perceptual level, as a working memory interference (Z. G. Cai et al., 2018), or as a response bias (Yates et al., 2012). Finally, more in line with a perceptual interpretation of the integration effect, it has been recently proposed (Hendrikx et al., 2024; Tsouli et al., 2022) that the interaction could arise from the processing of different dimensions in partially overlapping cortical maps (Fortunato et al., 2023; Harvey et al., 2013, 2015; Hendrikx et al., 2022, 2024; Protopapa et al., 2019), but without involving a common neural code.

So far, the nature of the magnitude integration phenomenon thus remains a hotly debated topic. Mixed evidence indeed seems to support both the perceptual and the cognitive interpretation of this effect. For example, results from Cai et al. (2018) show that the duration judgements can be biased by the length of a stimulus only when the length information is provided before the duration judgment has started, suggesting that the bias occurs as an interference between memory traces. Furthermore, electroencephalographic (EEG) evidence from Cui et al. (2022) shows that the interference of length on duration is reflected by event-related potentials (ERPs) usually associated with working memory (i.e., the P2 and P3b component). On the other hand, we have recently shown that a congruent interference between numerosity and duration (i.e., the more numerous the stimulus is, the longer it is perceived to last) occurs only when the two dimensions are conveyed by the same stimulus (Togoli, Bueti, et al., 2022). For example, for numerosity to bias duration, the entire array of items should also convey the duration information. When duration is instead conveyed by an additional stimulus presented on top of the array, the effect reverses and becomes repulsive. We also showed that the congruent effect is independent from attention, as the bias works according to the items of the array effectively conveying duration even if the entire array is attended, and that the repulsive effect is tightly spatially localised. These results thus suggest that the interference engages a perceptual binding mechanism, possibly operating in an automatic, pre-attentive fashion.

To further address the nature of the magnitude integration phenomenon, here we assess and compare the neural (EEG) signature of magnitude integration in two distinct cases: when magnitudes are actively judged in a task, versus when they are passively watched. In two separate experiments, the participants judged either the numerosity, the duration, or the item size of dot-array stimuli against a memorised reference (task condition), or passively watched a stream of dot-array stimuli modulated in numerosity, duration, and item size (passive condition). In the task condition, to avoid participants focusing their attention on a single magnitude in each trial, we presented a cue informing them about the dimension to judge only after the offset of the stimulus. In the analysis, we first identify a neural signature of magnitude integration by assessing the brain responses evoked by the different dimensions, and by assessing the extent to which such responses could predict the integration effect measured behaviourally. Then, we assess the magnitude-sensitive brain responses during passive viewing, and compare them to the results of the task condition. If magnitude

integration occurs at the cognitive level, for instance as a working memory interference or a response bias, the task condition should show a unique signature of the effect not present in the passive condition, where the magnitudes of the stimuli are neither memorised nor judged. Instead, if the integration occurs via automatic perceptual processes, we should observe comparable signatures of magnitude processing irrespective of the task.

MATERIALS AND METHODS

Participants

A total of 51 participants were tested in the study, with 20 participants tested in the task condition (13 females; age ± SD = 24.95 ± 4.21) and 31 tested in the passive condition (19 females; age ± SD = 23.96 ± 3.73). Two participants were excluded from data analysis in the passive condition due to equipment failure (i.e., failed to record EEG data), leaving 29 participants included in the final analysis. Subjects were compensated for their participation in the study with Euro 20. All participants read and signed a written informed consent form before the start of the session. All participants had normal or corrected-to-normal vision, and reported no history of neurological, psychiatric, or developmental disorder. The study was approved by the ethics committee of the International School for Advanced Studies (Protocol 10035-III/13), and was designed to be in line with the Declaration of Helsinki. The sample size tested in the task condition was determined a priori based on the expected magnitude integration effect as observed in previous studies. Specifically, we took the average of the lowest effect sizes of the magnitude integration effects observed in Togoli et al. (2021) (i.e., average of effects at the smallest magnitude levels in the time and numerosity task of Exp. 1a; d = 0.87), the effect size observed in Togoli, Bueti, et al., (2022) (average of effects in Exp. 2; d = 0.74), and in Togoli, Fornaciai, et al. (2022) (i.e., average of effects at the intermediate magnitude levels in the time and numerosity task; d = 1.21). The resulting average effect size was d = 0.94. Considering a power of 95% and a two-tailed distribution, a power analysis indicated a total estimated sample size of 17 participants, which we rounded up to 20. The sample size of the passive-viewing condition was decided considering the more limited range of magnitudes (see below *Apparatus and stimuli*) and thus a lower expected sample size. As it was difficult to estimate a specific effect size for the magnitude effects in passive viewing, we aimed at testing a sample similar to previous passive-viewing studies in numerosity perception (Fornaciai et al., 2017; Fornaciai & Park, 2017), which included

around 30 subjects. The data from the passive condition has also been used to assess perceptual history effects across different stimuli; these results have been published in Fornaciai et al., 2023.

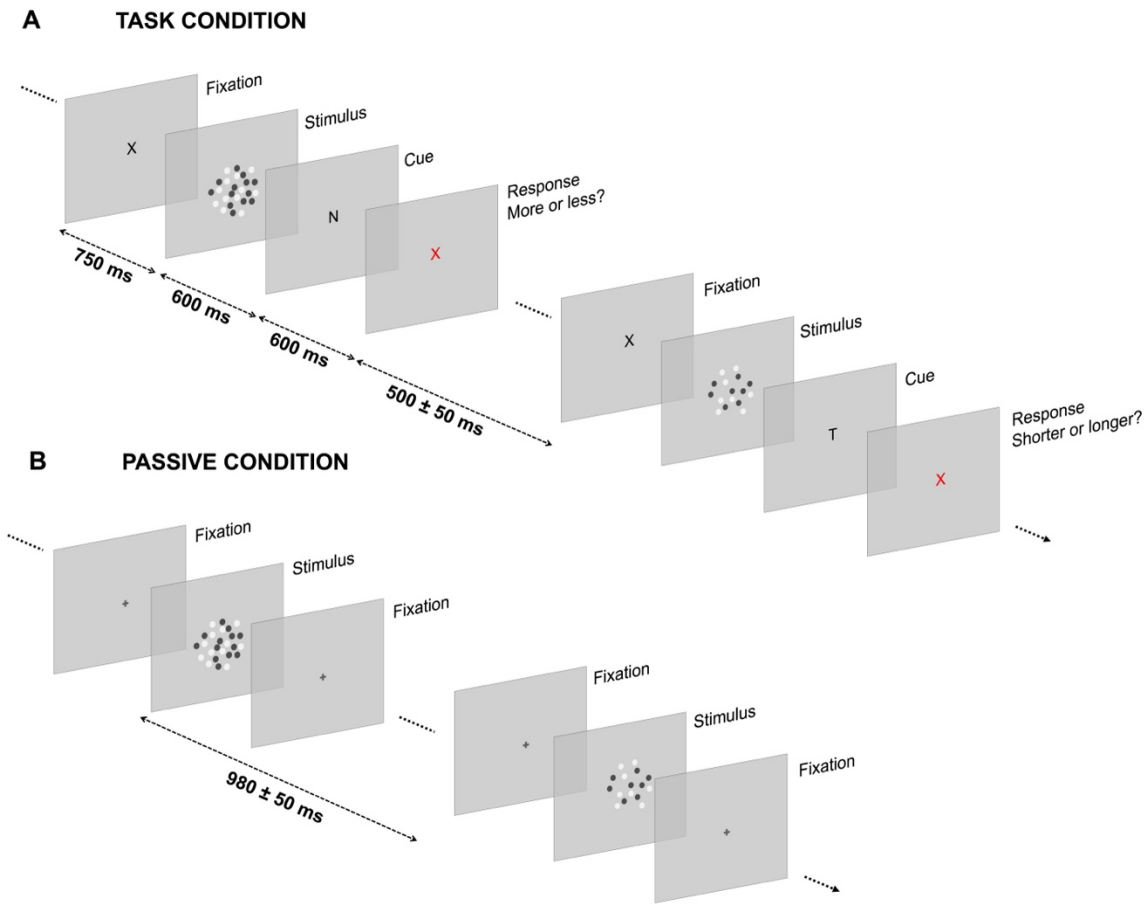

FIGURE 1. Experimental procedure. *(A) Procedure of the task condition. While participants kept their gaze at the centre of the screen (on a "X" that served as fixation cross), a stimulus was presented in each trial. The stimulus was modulated in numerosity (8-32 dots), duration (100-400 ms), and item size (i.e., the size of each item in the array; 3-10 pixels). After an interval of 600 ms from the offset of the stimulus, a cue appeared at the centre of the screen indicating which stimulus dimension the participant had to judge. Namely, the cue could be a "N" (numerosity judgment), a "T" (i.e., duration judgement), or a "S" (size judgement). The cue remained on the screen for 600 ms, after which the fixation cross turned red indicating to provide a response. The participants were then asked to report whether the magnitude indicated by the cue was bigger or smaller compared to a memorised reference corresponding to the middle of the magnitude ranges (presented before the session and before each block of trials). After providing a response, the next trial started after a variable inter-*

trial interval (500 ± 50 ms). (B) Passive viewing condition. In the passive condition, participants watched a stream of dot-array stimuli modulated in numerosity (12-24 dots), duration (140-280 ms), and item size (4-8 pixels), while keeping their gaze on a central fixation cross. In order to ensure that participants watched the stimuli, they were asked to detect an occasional oddball stimulus (i.e., a dot array with reduced contrast) presented on 3.7% of the trials. Each stimulus was separated by a variable inter-stimulus interval of 980 ± 50 ms. Stimuli are not depicted in scale.

Apparatus and stimuli

The stimuli used in both conditions were arrays of black and white dots (50/50 proportion), presented on a grey background. The stimuli were generated using the routines of the Psychophysics Toolbox (v.3; Kleiner M et al., 2007; Pelli, 1997) in Matlab (r2021b, The Mathworks, Inc.), and presented on a 1092×1080 LCD monitor running at 120 Hz, which encompassed a visual angle of 48×30 deg from a viewing distance of 57 cm. The dot-array stimuli were generated online in each trial, with the dots scattered pseudo-randomly within a circular aperture with a variable radius spanning from 200 to 400 pixels (pseudo-randomly determined in each trial). In the task condition, the dot-array stimuli could have a numerosity of 8, 12, 16, 24, or 32 dots, a duration of 100, 140, 200, 280, or 400 ms, and item size (i.e., the radius of each item in the array) of 3, 4, 6, 8, or 10 pixels, for a total of 125 unique combinations of the three magnitudes. The reference stimulus that the participants used as a comparison (presented at the beginning of the session and before each block) had the middle values of the three ranges (16 dots, 200 ms, 6 pixels). The stimuli in the passive condition had smaller magnitude ranges to make the modulation of magnitudes subtler and less obvious. Namely, the stimuli could have a numerosity of 12, 16, or 24 dots, a duration of 140, 200, or 280 ms, and item size of 4, 6, or 8 pixels. In both conditions, the stimuli were presented at the centre of the screen.

Procedure

Task condition. In the task condition, participants performed a magnitude classification task, with the dimension judged in each trial determined by a retrospective cue (i.e., presented after the offset of the stimulus). First, participants watched a reference stimulus representing the middle of the magnitude ranges used in the experiment, and were instructed to memorise it.

The reference was presented 10 times (randomising the positions of the dots) before the start of the session, and repeated 5 more time before the start of each block of trials. In the session, participants were instructed to keep their gaze on a central fixation point. Each trial started with the presentation of the fixation cross (a "X" at the centre of the screen). After 750 ms, the dot-array stimulus was presented replacing the fixation cross, and was displayed for 100-400 ms according to the duration selected in the trial. After an interval of 600 ms from the offset of the stimulus, the retrospective cue was presented at centre of the screen. The cue could be either "N," "T," or "S," respectively instructing the participant to judge the numerosity, duration (i.e., time), or item size of the stimulus. The cue remained on the screen for 600 ms. After that, the cue was replaced by a red fixation cross, instructing the participant to provide a response. According to the cue, the participant was asked to indicate whether the stimulus had a higher or lower numerosity, a longer or shorter duration, or a bigger or smaller item size compared to the memorised reference. The response was provided by pressing either the down arrow (lower/shorter/smaller) or the up arrow (higher/longer/bigger) on a standard computer keyboard. After providing a response, the next trial started after a variable inter-trial interval of 500 ± 50 ms. Each participant completed a total of 1,250 trials (10 blocks of 125 trials), corresponding to a total of 10 repetition of each unique combination of numerosity, duration, and dot size. The three tasks were randomly intermixed within the same blocks.

Passive condition. In the passive condition, the participants watched a series of stimuli modulated in numerosity, duration, and size. Each stimulus was presented centrally on the screen, and successive stimuli were separated by an inter-stimulus interval of 980 ± 50 ms. In order to make participants attend the stream of stimuli, they were asked to detect occasional oddball stimuli defined by a reduced contrast compared to the rest of the stimuli (oddball detection task). The oddball stimuli represented 3.7% of the total stimuli presented. Participants were thus instructed to press the space bar on the keyboard as fast as they could once they detected an oddball stimulus. This occasional task was designed in order to avoid drawing attention on any of the magnitude dimensions of the stimuli, while encouraging the participants to watch the stimuli. On average, the detection rates (± SD) of the oddball were 93% ± 1.3%, and the average reaction time was 313 ± 11 ms. In the passive condition, participants completed a total of 2,160 trials (8 blocks of 270 trials), for a total of 80 repetitions of each combination of stimulus magnitudes. In this condition, participants were

only instructed to watch the stream of stimuli and respond to the oddball, and the magnitudes were never mentioned in the instructions and recruiting materials.

Behavioural data analysis

In the task condition, the magnitude judgement performance and the integration effect were assessed by first computing the point of subjective equality (PSE). The PSE reflects the accuracy in the task, and the perceived magnitude of the stimuli compared to the reference magnitude. Data reflecting each specific task (i.e., all the trials in which a specific cue was presented) were used to compute the proportion of "more" (numerosity task), "longer" (duration task), or "bigger" (size task) responses as a function of both the task-relevant magnitude and the other (interfering) magnitudes. A psychometric (cumulative Gaussian) function was then fitted to the distribution of proportion of response, according to the maximum likelihood method (Watson, 1979). Specifically, within each task, the psychometric function was fitted separately for each level of each of the other interfering magnitudes. For example, when analysing the performance in the numerosity task, the fit was performed separately for each level of duration and each level of item size. To account for errors unrelated to the magnitude of the stimuli and lapses of attention, a finger-error rate correction of 2.5% (Wichmann & Hill, 2001) was applied. This correction reduces the asymptotic levels of the fit by a proportion corresponding to the rate, in order to account for the random errors preventing the proportion of responses to converge to 0% and 100% at the lower and higher end of the range, respectively. The PSE was computed as the numerosity/duration/size level corresponding to chance level (50%) responses (i.e., the median of the psychometric curve). This procedure allowed us to compute individual measures of PSE for the different combinations of task relevant and interfering dimensions (e.g., numerosity PSE when duration was 100 ms, 140 ms, 200 ms, 280 ms, and 400 ms, and the same for size, and similarly for the other tasks). When performing the fit according to a given task relevant and interfering magnitude, the other dimension was collapsed. Doing so, each data point in the fitting procedure represented the average of 25 repetitions of the specific combination of task-relevant and interfering dimension. Additionally, we assessed the precision in the task, in terms of the Weber fraction (WF). The WF was computed as the ratio between the just noticeable difference (JND; the slope of the psychometric curve) and the PSE. To assess the difference in the average WF across the three types of tasks, we used a one-way repeated measures ANOVA.

In order to better define the effect of the interfering magnitudes in each task, we then used the PSE to compute a "magnitude integration effect" index according to the following formula:

$$Magnitude\ Integration\ Effect = -1 \times ((PSE_j - PSE_{ref}) / PSE_{ref}) \times 100;$$

Where PSE_{ref} correspond to the PSE obtained when the interfering magnitude considered was the same as the reference, and PSE_j to the PSE corresponding to each other level of the interfering magnitude (either lower or higher than the reference). The change in sign (-1) was added in order to make the interpretation of the index easier. Namely, doing so a positive index means that the task-relevant magnitude is overestimated, while a negative index means that the magnitude is underestimated. To assess the significance of magnitude integration effects, we performed a series of linear mixed-effect model tests, assessing the integration biases on each type of judgement. Namely, we entered the magnitude integration effect as dependent variable, the ratio of each level of the interfering magnitude with the reference value and the magnitude itself as predictors, and the subject as the random effect (Magnitude integration effect ~ Ratio x Magnitude + (1|subj.)). Interactions found between ratio and magnitude were followed up with additional LME tests within each interfering magnitude dimension. The LME models were chosen (i.e., instead of ANOVAs) as the ratio is a continuous variable.

Finally, in order to assess the relationship between ERPs and the behavioural effect, we computed a measure of ΔPSE. This measure reflected the difference in PSE between each level of the interfering dimension and the middle magnitude level corresponding to the reference. This was done to have a similar measure that can be related to the neural effect of different magnitudes (see below for more information about the ERP analysis). All the analyses and statistical tests on behavioural data were performed in Matlab (version r2021b).

Electrophysiological recording and pre-processing

In both the task and passive condition, the EEG was recorded throughout the experimental session. EEG recording was performed by using the Biosemi ActiveTwo system (at 2048 Hz sampling rate), and a 32-channel cap based on the 10-20 system layout. To better monitor artifacts due to eye blinks and movements, we recorded the electro-oculogram (EOG) by means of a channel attached below the left eye of the subject. During the recording, we made

sure to keep the electrode offsets as low as possible. Usually, electrode offset values were kept below 20 μV, but occasionally values up to 30 μV were tolerated.

The pre-processing of EEG data was performed offline in Matlab (version R2021b), using the functions of the EEGLAB (Delorme & Makeig, 2004) and ERPlab (Lopez-Calderon & Luck, 2014) toolboxes. In both conditions, the pre-processing involved the binning and epoching of data according to each unique combination of the different magnitudes. In the task condition, the binning was also performed separately for the three cues determining the specific task in each trial. In the task conditions, epochs spanned from -300 to 1,200 ms, time-locked to the onset of each stimulus. In the passive condition, the epochs spanned from -200 to 700 ms, again time-locked to the onset of the stimuli. The rationale for using a longer epoch in the task condition was to include the presentation of the cue in the epoch. However, the data analysis was restricted to -200:700 ms as in the passive condition. In both cases, the ERPs evoked by duration were re-aligned to the offset of the stimuli. This was done as we expected an effect of duration only after the presentation of the stimuli had fully unfolded.

In both conditions, after the epoching, the EEG data was band-pass filtered with cut-offs at 0.1 and 40 Hz. Moreover, to clean up the data as much as possible from artifacts such as eye movements and blinks, we performed an independent component analysis (ICA) aimed at removing identifiable artifacts and other potential sources of systematic noise. After the ICA, we additionally applied a step-like artifact rejection procedure, with an amplitude threshold of 40 μV, a window of 400 ms, and a step size of 20 ms. This was done in order to further remove any remaining large artifact from the EEG data. On average, this led to the exclusion of 2.26% ± 1.73% of the trials in the task condition, and 0.6% ± 1.2% in the passive condition. Finally, we computed the event-related potentials (ERPs) by averaging EEG epochs within each bin, and further low-pass filtered the signal with a 30-Hz cut-off.

Event-related potentials analysis

In both conditions, the ERP analysis was performed by considering the average of the same set of three occipito-parietal target channels, selected a priori based on previous studies (Fornaciai et al., 2017; Fornaciai et al., 2023; Tonoyan et al., 2022). The chosen target channels were Oz, O1 and O2.

First, we plotted the ERPs corresponding to each magnitude and computed the linear contrast of ERPs. The weights of the linear contrast computation were [-2 -1 0 -1 -2] in the task condition, and [-1 0 1] in the passive condition, based on the number of levels of each magnitude tested in the two conditions. Moreover, we computed an additional measure of contrast by computing the difference in ERP amplitude for each level of each magnitude and the two extreme levels of the other two interfering dimensions. For example, for each level of numerosity we computed the difference in amplitude between the extreme levels of duration and size, and so on for the other dimensions. We then averaged this measure across the different levels of each magnitude. To assess the significance of the magnitude contrasts, we performed a series of one-sample t-tests against zero. To control for multiple comparisons, we applied a false discovery rate correction with q = 0.05. Finally, we computed a measure of ΔERP as the difference in amplitude between the ERPs corresponding to the middle level of the range (corresponding to the reference in the task condition) and each other level of the magnitude ranges. In the task condition, the ΔERP was computed separately according to the task. Doing so, we thus computed a measure of the effect of numerosity in the size and duration task, and so on for the other magnitudes. This measure was computed in order to have an index of the effect of magnitudes on ERPs similar to the ΔPSE computed from the behavioural results, in order to more easily relate the ERP and behavioural measures of the effect in the analysis of the task condition. To assess the significance of the ΔERP, in the task condition we performed a series of LME tests, entering ΔERP as the dependent variable, the ratio of each magnitude level and the middle level as predictor, and the subjects as the random effect (ΔERP ~ Ratio + (1|subj)). In the passive condition, we instead performed a series of paired t-tests. In both cases, the tests were performed on small 10-ms (step = 5 ms) windows in a sliding-window fashion, and the significance was corrected with FDR. Additionally, we considered significant only clusters of consecutive significant time-points larger than 10 ms.

Only in the task condition, we further assessed the relationship between ΔERP (i.e., neural effect of magnitude) and ΔPSE (i.e., behavioural effect of magnitude) via a series of LME tests. In this case, we entered the ΔPSE as the dependent variable, the ΔERP as the predictor, and the subjects as the random effect (ΔPSE ~ ΔERP + (1|subj)). This analysis was performed again across a series of 10-ms windows with step = 5 ms. Finally, the analysis was restricted to the latency windows showing a significant modulation of ΔERP by the magnitudes (see

above), and we considered statistically significant only clusters of at least three consecutive significant windows. This analysis was aimed at assessing the extent to which magnitude-sensitive brain responses can predict the effect measured behaviourally. All the analyses and statistical tests were performed in Matlab (version r2021b).

RESULTS

In this study, we measured the signatures of magnitude integration effects with subjects either engaged in actively judging the magnitude of the stimuli, or passively watching the stimuli. Doing so, we aimed at comparing such signatures to better understand the nature of the magnitude integration phenomenon. Specifically, if magnitude integration depends on decision making or other cognitive processes, we predicted to observe a unique signature of magnitude processing linked to the effect only when performing a task. Conversely, if the integration effect arises from automatic perceptual processes, then similar signatures of magnitude processing should be observed in both the task and during passive viewing.

Task condition

In the task condition, participants performed a magnitude classification task of the numerosity, duration, and item size of dot-array stimuli. Which dimension to judge was indicated to the participants via a retrospective cue (in a trial-by-trial fashion), thus forcing them to attend the stimulus as a whole rather than focusing on a single dimension. The procedure of the task condition is depicted in Fig. 1A.

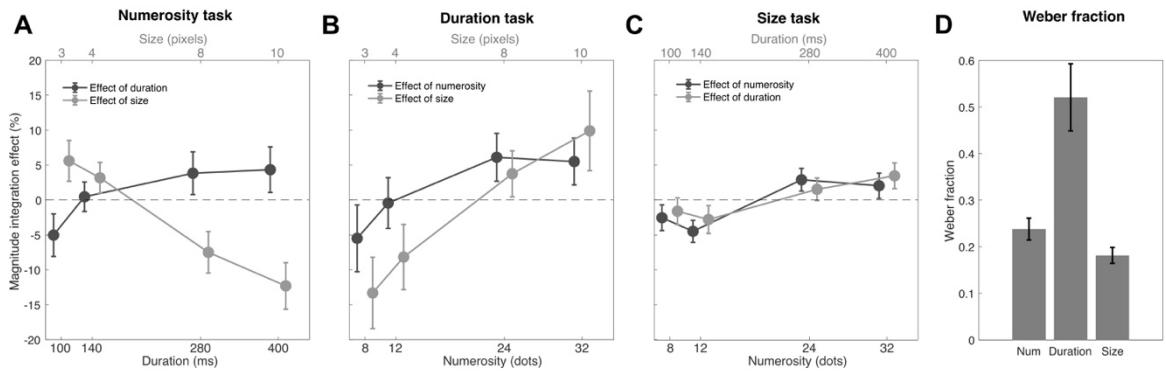

FIGURE 2. Behavioural effects of magnitude integration. *(A) Magnitude integration effects of duration and size on numerosity judgements. The values of the two interfering dimensions are reported on the upper and lower x axes. (B) Magnitude integration effects on duration*

judgements. (C) Magnitude integration effect on size judgements. The data points are jittered for the ease of visualisation. (D) Average Weber fraction in the three tasks. Error bars are SEM.

First, we assessed the behavioural effects of magnitude integration, which are shown in Fig. 2A-C. To assess the mutual biases across the different dimensions, we first computed the point of subjective equality (PSE; see *Behavioural data analysis*) as a measure of accuracy in the task. Then, we derived a magnitude integration effect index based on the difference in PSE caused by each level of the interfering magnitudes compared to the reference magnitudes (i.e., the middle levels of numerosity, duration, and size). To assess the significance of magnitude integration, we performed a series of linear mixed-effect (LME) regression models within each type of task. In the model, we entered the magnitude integration effects as the dependent variable, the ratio of each interfering magnitude level with the reference level, and the magnitude itself (e.g., "duration" and "size" in the numerosity task), as predictors, and the subjects as the random effect. The ratio (instead of the magnitude) was chosen as a predictor in order to test the effect of both interfering dimensions on each type of judgment within the same test.

In the numerosity task (Fig. 2A), we observed robust effects of both duration and size, although in opposite directions. While duration had a congruent effect (i.e., the longer the duration, the higher the perceived numerosity), size had an opposite, repulsive effect: the bigger the size, the lower the perceived numerosity. The LME test (adjusted-$R^2 = 0.22$) indeed showed a significant interaction between ratio and magnitude ($\beta = 20.96$, $t = 5.90$, $p < 0.001$). This interaction was followed up with simpler LME tests considering each interfering dimension separately. The results of these additional tests showed that both duration (adj-$R^2 = 0.56$, $\beta = 5.47$, $t = 3.23$, $p = 0.002$) and size (adj-$R^2 = 0.47$, $\beta = -15.50$, $t = -6.16$, $p < 0.001$) induced significant biases on numerosity judgements. In the duration task (Fig. 2B), we observed again a significant interaction between ratio and magnitude (adj-$R^2 = 0.37$, $\beta = -12.02$, $t = -2.48$, $p = 0.019$), this time suggesting that size had a stronger influence on duration compared to the effect of numerosity on duration. Two follow-up LME tests however showed significant congruent biases induced by both numerosity (adj-$R^2 = 0.33$, $\beta = 7.47$, $t = 2.75$, $p = 0.007$) and size (adj-$R^2 = 0.46$, $\beta = 19.30$, $t = 4.88$, $p < 0.001$). Finally, looking at Fig. 2C it

is clear that size was the magnitude most resistant to biases from other dimensions. Although weaker, the LME test (adj-R^2 = 0.27) showed a significant main effect of ratio (β = 3.99, t = 3.02, p = 0.003), no main effect of magnitude (β = -1.31, t = -0.53, p = 0.59), and no interaction (β = 0.41, t = 0.22, p = 0.82). Overall, the behavioural results of the classification task showed systematic mutual biases across all the dimensions tested, albeit with some partial asymmetries.

To assess the participants' precision in the task, we assessed the Weber fraction (WF; computed as the ratio of the just noticeable difference and the PSE), which is shown in Fig. 2D. On average, size showed the lowest WF (0.18 ± 0.08), suggesting the highest precision in the task, followed by numerosity (0.24 ± 0.10), and finally duration (0.52 ± 0.32), which was the most difficult dimension to judge. A one-way repeated measure ANOVA (with factor "task") confirmed that the WFs across the three types of task are significantly different ($F(2,38) = 18.59$, $p < 0.001$), in line with previous studies (Fornaciai et al., 2023).

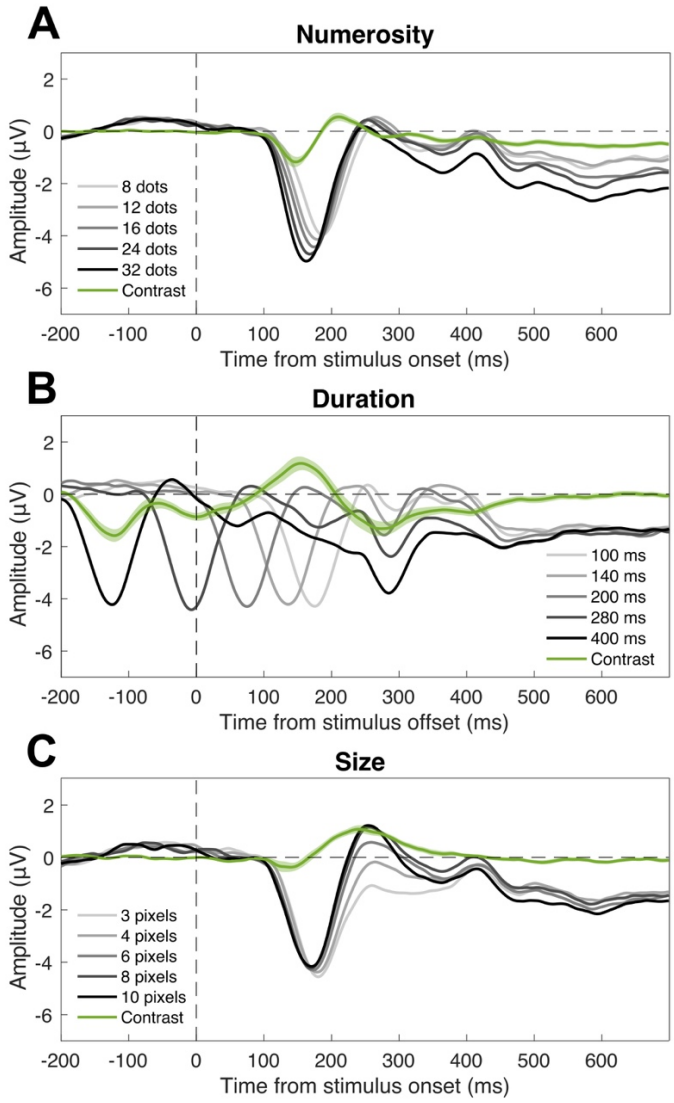

FIGURE 3. Event-related potentials (ERPs) evoked by each magnitude. *(A) ERPs evoked by the stimulus numerosity. (B) ERPs evoked by the stimulus duration. (C) ERPs evoked by item size. In all panels, the green wave indicates the linear contrast of the ERPs. Note that while numerosity and size ERPs were time-locked to the onset of the stimuli, ERPs corresponding to duration were re-aligned to the offset. The zero in panel B thus indicates the offset of the stimuli. The vertical dashed line indicates the onset or offset of the stimuli. The horizontal dashed line indicates the zero of the amplitude scale. The shaded area around the contrast wave represents the SEM.*

After assessing the behavioural effect, we went on and addressed the neural signature of magnitude integration. First, we plotted the event-related potentials (ERPs) evoked by the different levels of each magnitude, irrespective of the task performed. The ERPs are shown in

Fig. 3. In the case of numerosity (Fig. 3A), we observed a first negative peak of numerosity-sensitive responses at 150-200 ms after stimulus onset, followed by weaker but more sustained responses throughout the epoch. In the case of duration, aligning the waves to the offset of the stimuli created a misalignment of the onset responses, which introduce spurious effects. However, we also observed a large modulation at around 300 ms after stimulus offset, that seems to genuinely reflect duration. Finally, brain responses sensitive to item size showed the main peak at around 250 ms, with a large positive deflection.

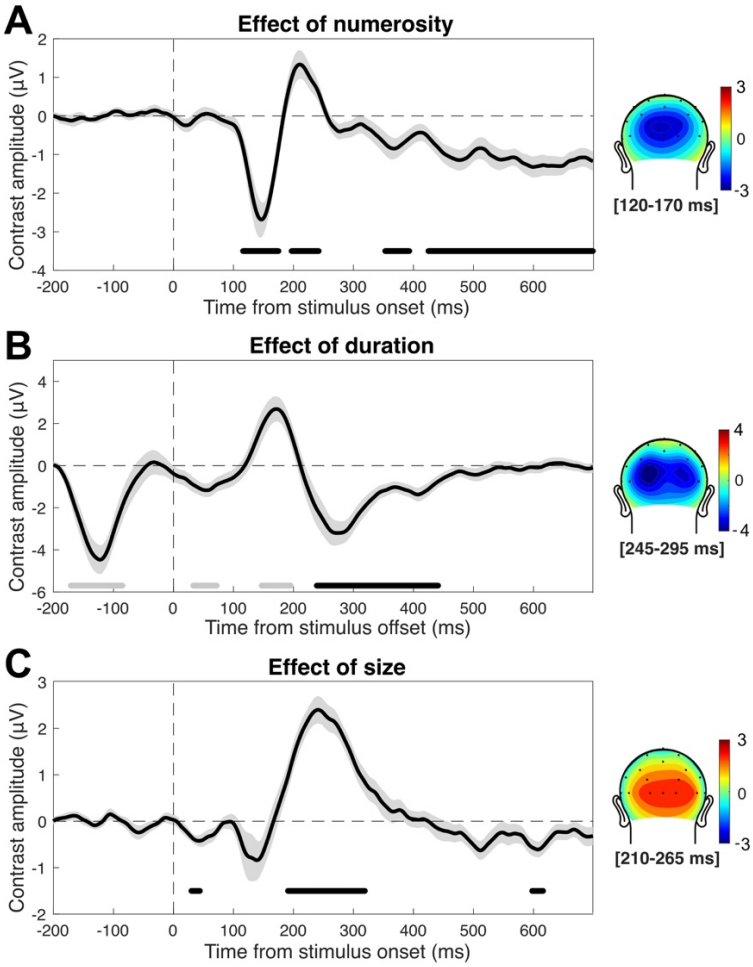

FIGURE 4. Average contrast amplitude reflecting the neural effects of the three magnitudes. *The contrast amplitude was computed as the difference between the extreme levels of each "interfering" magnitude, performed separately for each level of each magnitude. For instance, the effect of numerosity on duration was computed as the difference between ERPs corresponding to 100 ms/32 dots and 100 ms/8 dots, and so on for the other levels of duration. The same was done for the effect of numerosity on size, and the resulting contrasts were averaged together. A similar procedure was then used to compute the effect of duration and the effect of size. (A) Contrast amplitude reflecting the effect of numerosity. (B)*

Contrast amplitude reflecting the effect of duration. (C) Contrast amplitude reflecting the effect of size. The black lines at the bottom of the plots mark the significant latency windows assessed with a series of FDR-corrected one-sample t-tests. Note that, in the case of duration, significant latency windows driven by spurious differences were not considered, and are marked in grey. The vertical dashed line indicates the onset or offset of the stimuli. The horizontal dashed line indicates the zero of the amplitude scale. The shaded area around the wave represents the SEM. The topographic plots besides each panel show the distribution of scalp activity in a 50-ms window around the main peak of each wave.

To better address the significance of magnitude-sensitive brain responses, we computed a measure of ERP contrast based on the difference between the extreme levels of the interfering dimensions' ranges. Namely, for each level of each magnitude, we contrasted the ERPs as a function of the extreme levels of the interfering dimensions. For example, to compute the effect of numerosity, we subtracted the ERP corresponding to the combination of 100 ms and the two extreme levels of numerosity (8 and 32 dots). The same subtraction was performed for the combination of 140, 200, 280, and 400 ms and the extreme levels of numerosity. The same was done for the combination of each level of size and the extreme levels of numerosity. The effect of duration and size were computed in the same way by switching the dimension. The average of this contrast measure, reflecting the effect of the different magnitudes in driving ERPs, is shown in Fig. 4. Additionally, Fig. 4 shows the topographic distribution of contrast amplitude in a 50-ms window around the main peak of each corresponding contrast wave. To assess the significance of the contrast amplitude, we performed a series of one-sample t-tests against zero, corrected for multiple comparisons with false discovery rate (FDR; $q = 0.05$). When reporting the results below, we indicate the range of t-values and p-values (adjusted with FDR) as [min max].

The numerosity-sensitive brain responses (Fig. 4A) showed four significant latency windows. The strongest effect was observed at a negative deflection at 120-175 ms ($t = [-6.97, -2.41]$, $p = [<0.001, 0.049]$) after stimulus onset, which was around the peak of contrast amplitude (-2.7 μV) observed at 145 ms after stimulus onset. This peak was followed by additional significant windows at 200-240 ms ($t = [2.45, 3.62]$, $p = [<0.001, 0.046]$), 355-390 ms ($t = [-3.18, -2.41]$, $p = [0.012, 0.048]$), and 425-700 ms ($t = [-5.61, -2.47]$, $p = [0.001, 0.044]$) after

stimulus onset. Regarding the effect of duration on ERPs, we excluded the latency windows driven by the spurious difference induced by the onset misalignment (see Fig. 3B). A genuine peak of activity driven by duration was observed at 270 ms after the offset (-3.2 μV), and we observed a significant latency window around this peak, spanning 240-440 ms (t = [-5.99, -2.60], p = [<0.001, 0.048]). In the case of size, the peak of activity was observed at 240 ms (2.4 μV). The largest significant latency window was observed around this peak, spanning 190-320 ms (t = [3.14, 8.44], p = [<0.001, 0.049]). Two additional, smaller significant windows were observed at very early latencies (30-45 ms; t = [-4.04, -3.23], p = [0.009, 0.042]), and at later latencies (600-615 ms; t = [-3.59, -3.16], p = [0.022, 0.048]). In all cases, the topographic distribution of scalp activity around the peaks (plotted besides each panel) showed a posterior distribution consistent with activity in occipital cortex.

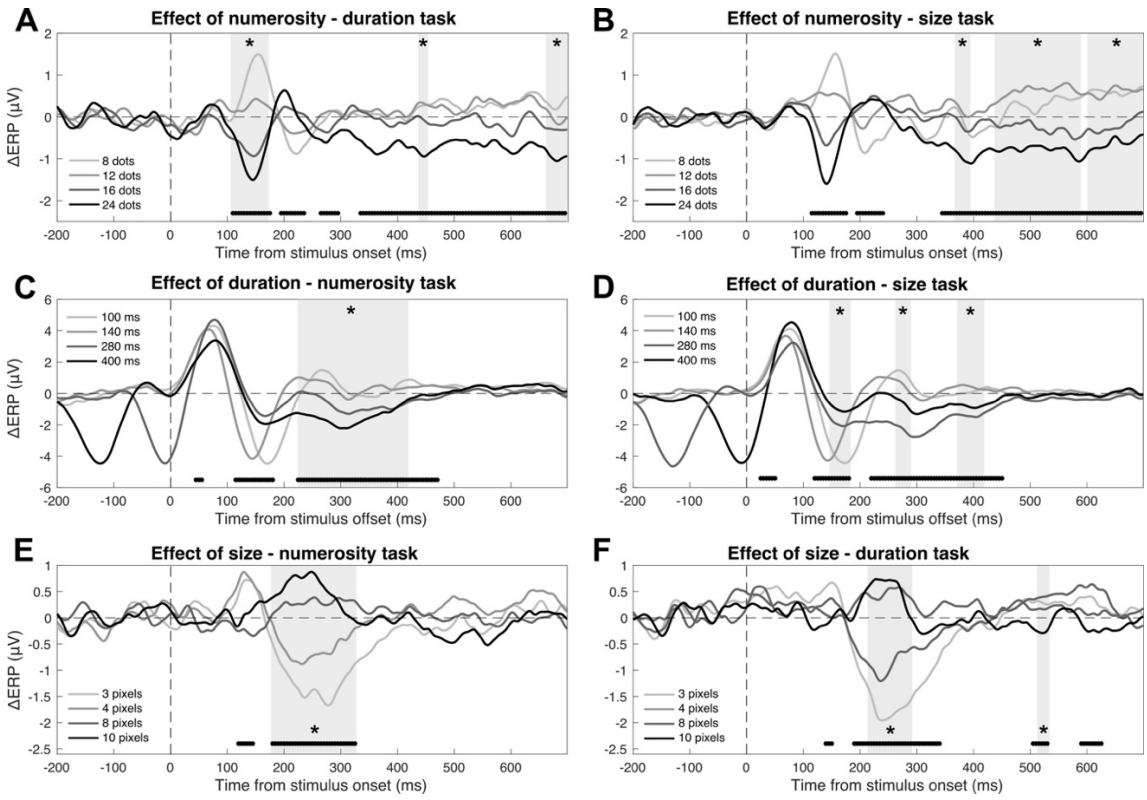

FIGURE 5. ΔERP measures and relationship with the behavioural effect. *(A) ΔERP measures reflecting the effect of numerosity in the duration task. (B) Effect of numerosity in the size task. (C) Effect of duration in the numerosity task. (D) Effect of duration in the size task. (E) Effect of size in the numerosity task. (F) Effect of size in the duration task. The black lines at the bottom of the plots mark the latency window where there is a significant difference in ΔERP as a function of the different levels of the magnitude. The vertical dashed line indicates the onset or offset of the stimuli. The horizontal dashed line indicates the zero*

of the amplitude scale. The grey shaded areas marked with stars indicate the latency windows where we observed a significant relationship between ΔERP and the behavioural effect (ΔPSE).

Our main goal in the task condition was however to identify the latency windows whereby the modulation of brain activity predicts the magnitude integration bias observed behaviourally. We then further computed two measures of the effect that could be related to each other in data analysis: ΔPSE, reflecting the behavioural effect, and ΔERP, reflecting the neural effect of magnitudes. These measures were computed by subtracting either the PSE or the ERP amplitude of each level of the interfering magnitudes from the PSE/ERP corresponding to the middle, reference level. For the ΔERP, this measure was computed at each time point throughout the epochs, and separately for the different types of tasks (see Fig. 5). Doing so, we thus computed the effect of numerosity in the duration and size task (Fig. 5A-B), the effect of duration in the numerosity and size task (Fig. 5C-D), and the effect of size in the numerosity and duration task (Fig. 5E-F).

To address the relationship between neural and behavioural measures of magnitude effects, we first looked for latency windows showing a significant modulation of ΔERP as a function of the different levels of the magnitudes. To do so, we performed a series of LME tests individually for the effect of each magnitude in each task. In the LME model, we entered the ΔERP as the dependent variable, the ratio of each magnitude level with the middle level as the predictor, and the subjects as the random effect. The LME tests were performed across a series of 10-ms windows with a 5-ms step, in a sliding-window average fashion. To control for multiple comparisons, we again applied a FDR procedure with $q = 0.05$. Clusters of less than three consecutive significant tests (after FDR) were not considered. The results of these tests are shown with black lines at the bottom of each plot in Fig. 5, marking the significant latency windows.

After identifying the latencies showing a significant modulation of ΔERP, we looked for a relationship between ΔERP and ΔPSE within these windows. We thus performed a series of LME tests (10-ms windows with 5-ms step) including ΔPSE as the dependent variable, ΔERP as the predictor, and the subjects as the random effect. The effect of numerosity on duration

(Fig. 5A) showed three windows whereby the modulation of ΔERP could predict the behavioural effect (marked with grey shaded areas in the figure), a larger early window at 110-170 ms, followed by two smaller windows at 440-450 ms and 665-695 ms (β = [0.006, 0.009], t = [2.02, 3.69], p = [<0.001, 0.047], adj-R^2 = [0.31, 0.43]). The effect of numerosity on size (Fig. 5B) showed again three significant windows, but clustered at later latencies: 370-390 ms, 440-585 ms, and 605-695 ms (β = [0.085, 0.180], t = [2.06, 4.12], p = [<0.001, 0.047], adj-R^2 = [0.25, 0.36]). The effect of duration on numerosity showed a single large significant window, spanning 225-415 ms (β = [0.233, 0.597], t = [2.20, 3.77], p = [<0.001, 0.31], adj-R^2 = [0.47, 0.54]). The effect of duration on size, showed three smaller windows, with the first at earlier latencies spanning 150-180 ms, and the following two at latencies more consistent with the effect on numerosity: 265-285 ms and 375-415 ms (β = [-0.043, 0.088], t = [-2.67, 2.78], p = [0.006, 0.047], adj-R^2 = [0.56, 0.59]). Finally, the effect of size on numerosity showed a single, large significant window at 180-325 ms (β = [0.707, 1.171], t = [2.65, 6.50], p = [<0.001, 0.009], adj-R^2 = [0.18, 0.48]), while the effect of size on duration showed two significant windows at 215-290 ms and 515-530 ms (β = [-0.012, 0.013], t = [-3.19, 2.12], p = [0.002, 0.041], adj-R^2 = [0.13, 0.26]). These results show that the behavioural effect of magnitude integration could be reliably predicted by the modulation of magnitude-sensitive responses in the different task types, providing a neural signature of the effect.

Passive condition

In the passive-viewing condition, participants watched a stream of dot-array stimuli modulated in numerosity, duration, and item size, and responded to occasional oddball stimuli defined by a lower contrast. No instruction suggested the participants to explicitly attend the magnitudes of the stimuli. This passive-viewing protocol thus provides a cleaner index of the responses to the different magnitudes, not confounded by decision making or other task-related processes. If magnitude integration arises from automatic perceptual processes, then we expected to observe a similar modulation of brain responses consistent with the timing observed in the task condition (see Fig. 5). Instead, if decision making is necessary for magnitude integration to occur, the modulation of brain responses linked to the behavioural effect should not occur during passive viewing.

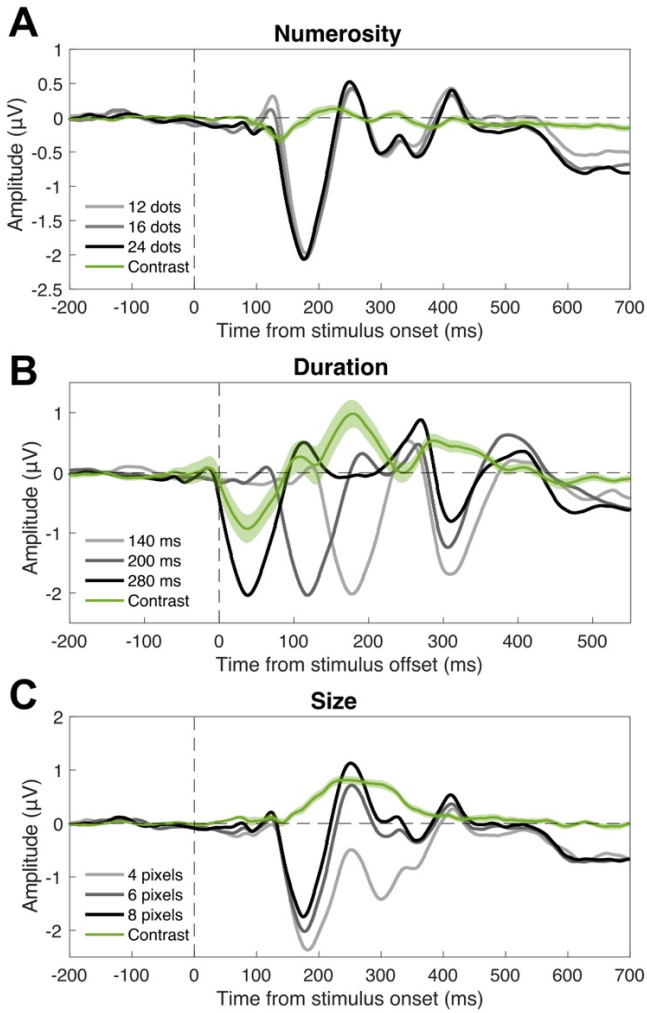

FIGURE 6. ERPs in the passive condition. *(A) ERPs relative to the different levels of numerosity. (B) ERPs corresponding to duration. (C) ERPs corresponding to item size. The green wave represents the linear contrast of ERPs. The vertical dashed line indicates the onset or offset of the stimuli. The horizontal dashed line indicates the zero of the amplitude scale. The shaded area around the contrast wave represents the SEM. Note that while in the case of numerosity and size the ERPs are time-locked to the onset of the stimuli, duration ERPs are time-locked to the offset of the stimuli.*

First, we assessed the ERPs corresponding to the different levels of the three magnitudes (Fig. 6). The overall pattern was largely consistent with what observed in the task condition (see Fig. 3), with however some differences. Numerosity (Fig. 6A) showed an early positive deflection that we did not observe in the task condition, with the magnitude of the stimuli however modulating the stimuli negatively (i.e., the larger the numerosity the more negative, or less positive, the amplitude, similarly to the task condition). More importantly, the ERPs

seemed to lack the later sustained modulation that was evident in the task condition. Duration (Fig. 6B) showed a similar deflection compared to the task condition, but the modulation of amplitude was in the opposite direction. Finally, size (Fig. 6C) showed instead ERPs consistent with the task condition.

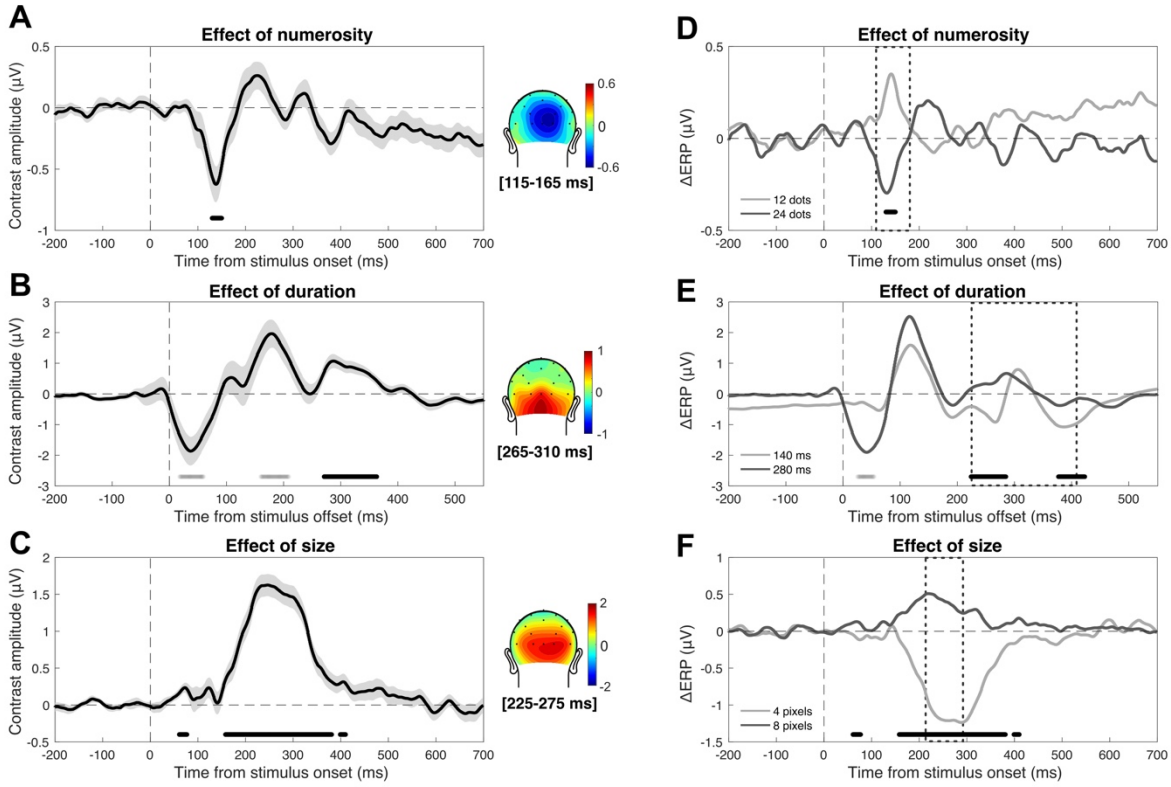

FIGURE 7. Magnitude contrasts and ΔERPs in the passive condition. *(A) Magnitude contrast reflecting the effect of numerosity. (B) Magnitude contrast reflecting the effect of duration. (C) Magnitude contrast reflecting the effect of Size. The topographic plots besides each panel show the distribution of contrast amplitude in a 50-ms windows around the peak. The shaded area around the wave represents the SEM. (D) ΔERP measures reflecting the effect of numerosity. (E) ΔERP measures reflecting the effect of duration. (F) ΔERP measures reflecting the effect of size. The dotted boxes show the latency windows whereby the brain responses in the task condition predicted the behavioural effect. The black lines at the bottom of the plot indicate the significant latency windows. In the case of duration, the earliest window was not considered because it was clearly driven by the misalignment of the onset responses, and has been marked in grey. The vertical dashed line indicates the onset or offset of the stimuli. The horizontal dashed line indicates the zero of the amplitude scale.*

To better assess the significance of the magnitude-sensitive brain responses, we computed again a measure of contrast based on the difference between the extreme levels of each magnitude (Fig. 7A-C). As in the task condition, the difference was computed separately for each combination of the different levels of the three dimensions. The contrast amplitude was then tested with a series of one-sample t-tests against zero, corrected with FDR (q = 0.05). In the case of numerosity (Fig. 7A), we observed a significant early window (130-150 ms; t = [-4.34, -3.64], p = [0.026, 0.047]), showing a negative deflection consistent with the task condition (see Fig. 4A). Differently from the task condition, we did not observe significant latency windows later on in the epoch. The effect of duration showed a few significant windows, although we excluded the ones clearly driven by the misalignment of the onset responses (see Fig. 6B). The significant latency window showing genuine duration-sensitive responses spanned 270-360 ms (t = [2.71, 4.76], p = [0.002, 0.043]), consistently with the task condition. Differently from the task condition, however, the contrast amplitude here showed a positive, rather than negative, deflection. Finally, size showed a large main window at 155-380 ms (t = [2.62, 10.88], p = [<0.001, 0.048]), consistent with the effect of size in the task condition. In addition, we observed two additional, smaller windows at 60-75 ms and 400-410 ms (t = [2.62, 3.63], p = [0.005, 0.049]). Similarly to the task condition, the topography of peak contrast amplitude over the scalp showed a posterior distribution consistent with occipital cortex (see plots beside panel A-C).

In order to better compare the modulation of brain responses in the task and in the passive condition, we also computed the ΔERP measure reflecting the neural effects of the three magnitudes (Fig. 7D-F). The ΔERPs were assessed with a series of paired t-tests, performed considering 10-ms windows (step = 5 ms) as in the task condition, corrected with FDR (q = 0.05). The overall timing of the significant windows was consistent with the contrast measure (Fig. 7A-C). ΔERPs reflecting the effect of numerosity showed a significant modulation at 130-150 ms (t = [3.64, 4.33], p = [0.025, 0.047]). The effect of duration showed a modulation at 225-285 ms (t = [-6.91, -2.78], p = [<0.001, 0.049]), and at 380-420 ms (t = [-3.68, -2.79], p = [0.016, 0.049]) after stimulus offset. Finally, in the case of size, we observed the main modulation in a large window spanning 160-380 ms (t = [-10.87, -2.62], p = [<0.001, 0.049]). We again observed two smaller significant windows at 60-75 ms (t = [-3.63, -2.65], p = [0.005, 0.046]) and 400-410 ms (t = [-2.85, -2.62], p = [0.032, 0.049]). As a comparison, Fig.

7D-F shows with dotted boxes the latency windows where we observed a significant relationship between neural and behavioural measures of magnitude integration in the task condition. In all cases, we observed an overlap between the significant windows in the passive and task condition.

DISCUSSION

In the present study, we assessed and compared the signatures of magnitude integration in two different conditions: when participants are engaged in actively judging the magnitude of the stimuli, or when they passively watched the stimuli. The phenomenon of magnitude integration – i.e., the mutual biases usually observed across different dimensions – is a hallmark of magnitude perception. Indeed, stimulus dimensions such as numerosity, duration, and size systematically interact with each other, leading to biases when judging them. Such mutual interactions have played a pivotal role in the development of influential theories like "a theory of magnitude" (ATOM; Walsh, 2003) and the "metaphor" theory (Casasanto & Boroditsky, 2008). However, the nature of this bias and its underlying mechanisms remain unclear.

Different mechanisms have been proposed to explain the interaction of magnitude dimensions. On the one hand, according to ATOM, the interaction would occur in perceptual processing due to the encoding of different dimensions with a common neural code (Walsh, 2003). Recent neuroimaging studies, however, albeit showing common neural substrates failed to provide evidence for a shared neural code (Borghesani et al., 2019; Hendrikx et al., 2024; Tsouli et al., 2022). According to the metaphor theory, on the other hand, the effect would instead arise at the conceptual or linguistic level, due to the use of "spatial" concepts to describe time (e.g., a "long" time; Bottini & Casasanto, 2013; Casasanto & Boroditsky, 2008; but see Whitaker et al., 2022). This theory however relied on asymmetric effects across temporal and non-temporal dimensions, which depend on the type of stimuli used (Javadi & Aichelburg, 2012; Lambrechts et al., 2013; Togoli et al., 2021). The original effect (i.e., space biasing time but not vice versa; (Casasanto & Boroditsky, 2008)) has also proven to be difficult to replicate (Whitaker et al., 2022). Moreover, other authors proposed that magnitudes could interact during working memory maintenance, nudging each other while stored in memory (Cai et al., 2018; Cui et al., 2022), or bias the response selection in comparisons tasks (Yates et al., 2012). Finally, in support of a perceptual account of

magnitude integration, we have recently shown that the effect relies on perceptual binding, show spatial selectivity, and does not depend on attention (Togoli, Bueti, et al., 2022). Considering the results from these studies, the nature of the magnitude integration phenomenon remains a debated topic.

In the present study, we further addressed the nature of the magnitude integration effect by assessing a new prediction. Namely, a cognitive effect hinging upon memory or decision should show a unique neural signature not present when the magnitudes are neither explicitly attended nor judged. Conversely, a perceptual effect is expected to occur in a more automatic fashion, independently from the relevance or judgment of magnitude. Thus, similar signatures should be observable with or without a magnitude judgment task.

Our behavioural results show systematic biases across the three magnitudes. First, numerosity was biased by both duration and item size. However, while duration show a congruent effect (the longer the duration, the higher the perceived numerosity) as in previous studies (Javadi & Aichelburg, 2012; Togoli et al., 2021), size induce an opposite bias (the larger the size, the lower the perceived numerosity). While perhaps unusual in the context of magnitude integration, this is in line with previous studies showing indeed a negative relationship between numerosity and items size (DeWind et al., 2015; Fornaciai et al., 2019). Duration is instead similarly affected by both numerosity and size in a congruent fashion, in line with previous studies (e.g., Lambrechts et al., 2013; Xuan et al., 2007), although the latter exerts a stronger influence. Finally, size seems the dimension most resistant to integration biases, and shows only modest, albeit significant, influences from the other magnitudes. Size is also the dimension that is easiest to judge (Fig. 2D; see also Fornaciai et al., 2023), and the generally low variability of responses might explain its robustness to biases. Considering this, the results thus show some (partial) asymmetries in the effects as some dimensions are more vulnerable to biases than others, in line with previous studies using similar stimuli (Togoli, Bueti, et al., 2022; Togoli et al., 2021).

In terms of event-related potentials, we found robust brain responses to the different magnitudes in the task condition. Overall, our analyses identified a set of latency windows that shows the stronger peaks of activity driven by the different dimensions. Namely, around 150 ms and 250 ms after stimulus onset in the case of numerosity and size, and around 300 ms after stimulus offset for duration. Brain activity in these latency windows appears to be

modulated by the different dimensions in a parametric fashion, according to the magnitude of the stimuli. Crucially, with just one exception (i.e., the effect of numerosity in the size task), brain activity at or around such peaks can significantly predict the bias observed behaviourally: the larger the brain responses, the stronger the magnitude integration bias. In the case of numerosity, the timing observed here is consistent with numerosity-sensitive responses measured in previous studies (Libertus et al., 2007; Park et al., 2016; Temple & Posner, 1998). However, recent studies also show that numerosity processing starts very early in visual cortex, around 75-100 ms after stimulus onset (Fornaciai et al., 2017; Park et al., 2016). The numerosity processing stage linked to the magnitude integration effect observed here thus likely represents a later stage compared to the early processing of numerosity information. In the case of duration, previous studies highlight a variety of possible correlates of duration processing, like the contingent negative variation (CNV; Damsma et al., 2021; but see Kononowicz & Penney, 2016), the N2 (Tonoyan et al., 2022), the P2 (Li et al., 2017), and the P3 (Ernst et al., 2017) ERP components. The timing shown in our results appears to be consistent with the results of Benau et al. (2018), showing duration sensitivity at around 350 ms after stimulus offset. Finally, in terms of size, the timing of responses sensitive to size of the items appears to be roughly consistent with previous results (Park et al., 2016) showing a peak at around 200 ms.

Comparing the timing of magnitude-sensitive brain response in passive viewing revealed similar evoked activity in most of the cases, closely mirroring the responses observed in the task condition. Especially in the case of numerosity and size, the peaks of magnitude-sensitive activity (i.e., ΔERP; compare Fig. 5 and Fig. 7D-F) show indeed a one-to-one correspondence, with similar timing and polarity. In duration perception, however, although the timing is very similar, brain responses showed an opposite polarity. This may additionally suggest that while the processing of numerosity and size is largely invariant to task relevance, the brain responses to duration may at least partially depend on the task relevance of this dimension. Namely, while the same duration processing stage seems to get engaged (i.e., as suggested by brain responses at the same latency and with the same scalp topography), actively attending the magnitudes of the stimuli may modulate how duration information is processed. This is not completely surprising, as duration shows different properties compared to the other dimensions (i.e., duration information needs to be accumulated, while the other dimensions can be processed from the onset), and the encoding of duration information is

notoriously poorer in vision compared to other senses (e.g., Alais & Burr, 2004; Cai & Connell, 2015). Thus, the processing of duration may be more sensitive to the modulatory effects of attention or task-relevance.

The lack of decision-making in the passive viewing parading represents its major strength, as it allows to exclude the involvement of decision making, task-related brain processes (e.g., working memory encoding and maintenance; Cai et al., 2018; Cui et al., 2022), and response biases (Yates et al., 2012). However, it also has the obvious weakness that magnitude integration could not be directly measured to confirm the effect. The striking similarity in the brain responses to the magnitudes, peaking at the same latencies where we demonstrated a relationship with the behavioural effect and showing the same scalp topography, nevertheless provides evidence that magnitude integration likely occurs even in the absence of a task.

Differently from the present study, previous EEG results concerning magnitude integration (involving duration and length) suggested the involvement of working memory interference (Cui et al., 2022). Cui et al. (2022) indeed observed effects of duration and length after the offset of the intervals, at ERP components usually associated with working memory maintenance, such as the P2 and P3b. While the effect of duration in Cui et al.'s work shows a timing consistent with the present results (~250-300 ms after stimulus offset), length has an effect at a much different timing (~300 ms after stimulus *offset*) compared to our earlier peak of the size dimension (~250 ms after stimulus *onset*). Additionally, the scalp topography of the magnitude effects had a much more anterior distribution, peaking at parieto-frontal locations, as opposed to our results showing a posterior, occipital distribution. However, Cui et al. also employed much different stimuli: longer intervals (800-1,200 ms) and quite large lengths up to 15 degrees of visual angle. Considering the relatively long durations and the fact that the stimuli were marked only at the beginning (onset) and end (offset) point, it is not surprising that they engaged memory processes (e.g., Rammsayer & Lima, 1991). Both magnitudes in such a task indeed rely on the memory trace of the first marker presented rather than on sustained sensory stimulation, making it more likely that any interference would involve higher-level cognitive processes. Our results however do not conflict with such an interpretation. Perceptual and mnemonic interferences are indeed not mutually exclusive processes, and can both occur depending on the nature of the stimuli and paradigm used. Our results however show that magnitude interactions can be perceptual in nature when based on stimuli relying on sensory/perceptual processing rather than memory.

How would this perceptual interaction occur in the brain? Interestingly, our results show that the effects across the different magnitudes do not occur at a unique, generalised stage, but show different timings consistent with different brain processing stages. The interference between magnitudes thus depends on the specific processing dynamics of the different dimensions (see also Togoli et al., 2021), rather than a common processing stage. Our results are thus not fully consistent with the idea of a generalised magnitude processing system, as proposed by the ATOM framework (e.g., Walsh, 2003). Instead, the results seem more in line with recent findings of topographical cortical maps of different magnitudes, overlapping with each other (Y. Cai et al., 2021; Fortunato et al., 2023; Harvey et al., 2013, 2015; Harvey & Dumoulin, 2017; Hendrikx et al., 2024; Protopapa et al., 2019). Recently, it has indeed been proposed that the interaction between different magnitudes could arise from the overlap of neural populations sensitive to different dimensions (Hendrikx et al., 2024; Tsouli et al., 2022), without the need for a centralised mechanism or a common magnitude neural code (Walsh, 2003).

To conclude, our results show that the neural signatures of magnitude processing and integration are very similar whether participants explicitly attend and judge magnitude information or passively watch the stimuli. This in turn suggests that similar brain processing stages are engaged irrespective of the task, and thus that magnitude integration likely occur even in the absence of decision making. Overall, our results thus provide new evidence supporting the idea that magnitude integration is an automatic perceptual phenomenon independent from the task performed, affecting the phenomenological appearance of the stimuli rather than their memory or the way they are judged.

Acknowledgements. This project has received funding from the European Union's Horizon Europe research and innovation programme under the Marie Sklodowska-Curie grant agreement No. 101103020 "PreVis" to MF, 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 to DB, and from the Italian Ministry of University and Research under the call FARE (project ID: R16X32NALR) and under the call PRIN2017 (project ID: XBJN4F) to DB.

REFERENCES

Alais, D., & Burr, D. (2004). The Ventriloquist Effect Results from Near-Optimal Bimodal Integration. *Current Biology*, *14*(3), 257–262. https://doi.org/10.1016/j.cub.2004.01.029

Benau, E. M., DeLoretta, L. C., & Moelter, S. T. (2018). The time is "right:" Electrophysiology reveals right parietal electrode dominance in time perception. *Brain and Cognition*, *123*, 92–102. https://doi.org/10.1016/j.bandc.2018.03.008

Borghesani, V., de Hevia, M. D., Viarouge, A., Pinheiro-Chagas, P., Eger, E., & Piazza, M. (2019). Processing number and length in the parietal cortex: Sharing resources, not a common code. *Cortex*, *114*, 17–27. https://doi.org/10.1016/j.cortex.2018.07.017

Bottini, R., & Casasanto, D. (2013). Space and time in the child's mind: metaphoric or ATOMic? *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00803

Cai, Y., Hofstetter, S., van der Zwaag, W., Zuiderbaan, W., & Dumoulin, S. O. (2021). Individualized cognitive neuroscience needs 7T: Comparing numerosity maps at 3T and 7T MRI. *NeuroImage*, *237*, 118184. https://doi.org/10.1016/j.neuroimage.2021.118184

Cai, Z. G., & Connell, L. (2015). Space–time interdependence: Evidence against asymmetric mapping between time and space. *Cognition*, *136*, 268–281. https://doi.org/10.1016/j.cognition.2014.11.039

Cai, Z. G., Wang, R., Shen, M., & Speekenbrink, M. (2018). Cross-dimensional magnitude interactions arise from memory interference. *Cognitive Psychology*, *106*, 21–42. https://doi.org/10.1016/j.cogpsych.2018.08.001

Casasanto, D., & Boroditsky, L. (2008). Time in the mind: Using space to think about time. *Cognition*, *106*(2), 579–593. https://doi.org/10.1016/j.cognition.2007.03.004

Cui, M., Peng, C., Huang, M., & Chen, Y. (2022). Electrophysiological Evidence for a Common Magnitude Representation of Spatiotemporal Information in Working Memory. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhab466

Damsma, A., Schlichting, N., & van Rijn, H. (2021). Temporal Context Actively Shapes EEG Signatures of Time Perception. *The Journal of Neuroscience*, *41*(20), 4514–4523. https://doi.org/10.1523/JNEUROSCI.0628-20.2021

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

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

Ernst, B., Reichard, S. M., Riepl, R. F., Steinhauser, R., Zimmermann, S. F., & Steinhauser, M. (2017). The P3 and the subjective experience of time. *Neuropsychologia*, *103*, 12–19. https://doi.org/10.1016/j.neuropsychologia.2017.06.033

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., Farrell, A., & Park, J. (2019). Looking for more food or more people? Task context influences basic numerosity perception. *Cortex*, *114*. https://doi.org/10.1016/j.cortex.2018.05.021

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

Fornaciai, M., Togoli, I., & Bueti, D. (2023). Perceptual History Biases Are Predicted by Early Visual-Evoked Activity. *The Journal of Neuroscience*, *43*(21), 3860–3875. https://doi.org/10.1523/JNEUROSCI.1451-22.2023

Fortunato, G., Togoli, I., & Bueti, D. (2023). The more numerous the longer: how the integration between numerosity and time leads to a common neural response. *Proceedings of the Royal Society B: Biological Sciences*, *290*(1998). https://doi.org/10.1098/rspb.2023.0260

Harvey, B. M., & Dumoulin, S. O. (2017). A network of topographic numerosity maps in human association cortex. *Nature Human Behaviour*, *1*(2), 0036. https://doi.org/10.1038/s41562-016-0036

Harvey, B. M., Fracasso, A., Petridou, N., & Dumoulin, S. O. (2015). Topographic representations of object size and relationships with numerosity reveal generalized quantity processing in human parietal cortex. *Proceedings of the National Academy of Sciences*, *112*(44), 13525–13530. https://doi.org/10.1073/pnas.1515414112

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic Representation of Numerosity in the Human Parietal Cortex. *Science*, *341*(6150), 1123–1126. https://doi.org/10.1126/science.1239052

Hendrikx, E., Paul, J. M., van Ackooij, M., van der Stoep, N., & Harvey, B. M. (2022). Visual timing-tuned responses in human association cortices and response dynamics in early visual cortex. *Nature Communications*, *13*(1), 3952. https://doi.org/10.1038/s41467-022-31675-9

Hendrikx, E., Paul, J. M., van Ackooij, M., van der Stoep, N., & Harvey, B. M. (2024). Cortical quantity representations of visual numerosity and timing overlap increasingly into superior cortices but remain distinct. *NeuroImage*, *286*, 120515. https://doi.org/10.1016/j.neuroimage.2024.120515

Javadi, A. H., & Aichelburg, C. (2012). When Time and Numerosity Interfere: The Longer the More, and the More the Longer. *PLoS ONE*, *7*(7), e41496. https://doi.org/10.1371/journal.pone.0041496

Kleiner M, Brainard D, & Pelli D. (2007). What's new in Psychtoolbox-3? *Perception 36 ECVP Abstract Supplement*.

Kononowicz, T. W., & Penney, T. B. (2016). The contingent negative variation (CNV): timing isn't everything. *Current Opinion in Behavioral Sciences*, *8*, 231–237. https://doi.org/10.1016/j.cobeha.2016.02.022

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

Li, B., Chen, Y., Xiao, L., Liu, P., & Huang, X. (2017). Duration adaptation modulates EEG correlates of subsequent temporal encoding. *NeuroImage*, *147*, 143–151. https://doi.org/10.1016/j.neuroimage.2016.12.015

Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and Brain Functions*, *3*(1), 1. https://doi.org/10.1186/1744-9081-3-1

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

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

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

Protopapa, F., Hayashi, M. J., Kulashekhar, S., van der Zwaag, W., Battistella, G., Murray, M. M., Kanai, R., & Bueti, D. (2019). Chronotopic maps in human supplementary motor area. *PLOS Biology*, *17*(3), e3000026. https://doi.org/10.1371/journal.pbio.3000026

Rammsayer, T. H., & Lima, S. D. (1991). Duration discrimination of filled and empty auditory intervals: Cognitive and perceptual factors. *Perception & Psychophysics*, *50*(6), 565–574. https://doi.org/10.3758/BF03207541

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences*, *95*(13), 7836–7841. https://doi.org/10.1073/pnas.95.13.7836

Togoli, I., Bueti, D., & Fornaciai, M. (2022). The nature of magnitude integration: Contextual interference versus active magnitude binding. *Journal of Vision*, *22*(11), 11. https://doi.org/10.1167/jov.22.11.11

Togoli, I., Fornaciai, M., & Bueti, D. (2021). The specious interaction of time and numerosity perception. *Proceedings of the Royal Society B: Biological Sciences*, *288*(1959), 20211577. https://doi.org/10.1098/rspb.2021.1577

Togoli, I., Fornaciai, M., Visibelli, E., Piazza, M., & Bueti, D. (2022). The neural signature of magnitude integration between time and numerosity. *BioRxiv*, 2022.08.29.505731. https://doi.org/10.1101/2022.08.29.505731

Tonoyan, Y., Fornaciai, M., Parsons, B., & Bueti, D. (2022). Subjective time is predicted by local and early visual processing. *NeuroImage*, *264*, 119707. https://doi.org/10.1016/j.neuroimage.2022.119707

Tsouli, A., Harvey, B. M., Hofstetter, S., Cai, Y., van der Smagt, M. J., te Pas, S. F., & Dumoulin, S. O. (2022). The role of neural tuning in quantity perception. *Trends in Cognitive Sciences*, *26*(1), 11–24. https://doi.org/10.1016/j.tics.2021.10.004

Walsh, V. (2003). A theory of magnitude: common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences*, *7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002

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

Whitaker, M. M., Hansen, R. C., Creem-Regehr, S. H., & Stefanucci, J. K. (2022). The relationship between space and time perception: A registered replication of Casasanto and Boroditsky (2008). *Attention, Perception, & Psychophysics*, *84*(2), 347–351. https://doi.org/10.3758/s13414-021-02420-z

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. https://doi.org/10.1167/7.10.2

Yates, M. J., Loetscher, T., & Nicholls, M. E. R. (2012). A generalized magnitude system for space, time, and quantity? A cautionary note. *Journal of Vision*, *12*(7), 9–9. https://doi.org/10.1167/12.7.9