

[Retter Talia L.](#), [Eraßmy Lucas](#) and [Schiltz Christine](#). 2024 Identifying conceptual neural responses to symbolic numerals. *Proc. R. Soc. B.* **291**: 20240589.
<http://doi.org/10.1098/rspb.2024.0589>

Identifying conceptual neural responses to symbolic numerals

Talia L. Retter¹, Lucas Eraßmy¹ & Christine Schiltz¹

¹Department of Behavioural and Cognitive Sciences, Institute of Cognitive Science & Assessment, University of Luxembourg (Luxembourg)

Corresponding author (TLR): talia.retter@uni.lu; +352 46 66 44 6497; Université du Luxembourg \ Maison des Sciences Humaines \ 11, Porte des Sciences \ 4366 Esch-sur-Alzette \ Luxembourg

Keywords : parity; numerical cognition; EEG; frequency-tagging; human vision; Arabic numerals

Abstract

21 The goal of measuring conceptual processing in numerical cognition is distanced by the
22 possibility that neural responses to symbolic numerals are influenced by physical stimulus
23 confounds. Here, we targeted conceptual responses to parity (even vs. odd), using
24 electroencephalographic (EEG) frequency-tagging with a symmetry/asymmetry design. Arabic
25 numerals (2–9) were presented at 7.5 Hz in 50-s sequences; odd and even numbers were
26 alternated to target differential, “asymmetry” responses to parity at 3.75 Hz (7.5 Hz/2). Parity
27 responses were probed with four different stimulus sets, increasing in intra-numeral stimulus
28 variability, and with two control conditions comprised of non-conceptual numeral alternations.
29 Significant asymmetry responses were found over the occipitotemporal cortex to all conditions,
30 even for the arbitrary controls. The large physical-differences control condition elicited the
31 largest response in the stimulus set with the lowest level of variability (1 font). Only in the
32 stimulus set with the highest level of variability (20 drawn, colored exemplars/numeral) did the
33 response to parity surpass both control conditions. These findings show that physical differences
34 across small sets of Arabic numerals can strongly influence, and even account for, automatic
35 brain responses. However, carefully designed control conditions and highly variable stimulus
36 sets may be used towards identifying truly conceptual neural responses.

Introduction

There is a longstanding interest in understanding the neural bases of numerical cognition. At a fundamental level, studies have addressed basic representations of numbers, such as numerosity and relative magnitude, with electroencephalography (EEG) and functional magnetic resonance imaging (fMRI; e.g., [Dehaene, 1996](#); [Fornaciai et al., 2017](#); [Sokolowski et al., 2021](#)). At a more abstract level, studies have also addressed the neural bases of operations (e.g., solving mathematical equations; [Pauli et al., 1994](#); [Jost, Henninghausen & Rosler, 2004](#); [Park et al., 2012](#)). Investigating the neural bases of numerical cognition may contribute to reconciling different theories of number processing, understanding brain organization more generally, and ultimately informing interventions in math learning.

An important issue in such studies is whether the observed neural responses are conceptual, or rather could be accounted for by physical stimulus features, such as variations in visual properties. Visual stimulus confounds (such as dot size, density, total area, cumulative perimeter, total luminance, dot shape, array shape, etc.) have long been known to influence behavioral responses to non-symbolic representations of numbers, in particular dot arrays (see [Dehaene, Izard, & Piazza, 2005](#); [Gebuis, Kadosh & Gevers, 2016](#); [Cheng et al., 2023](#)). As in behavioral studies, neural responses intended to reflect cognitive processing for non-symbolic dot representations, such as numerosity, are typically either contrasted to responses to selected visual stimulus properties ([Soltesz & Szucs, 2014](#); [Park, 2018](#); [Van Rinsveld et al., 2020](#)) or attempted to be isolated by orthogonally varying multiple visual stimulus features ([Hyde & Spelke, 2011](#); [Libertus, Brannon & Woldorff, 2011](#); [Guillaume et al., 2018](#); [Marlair et al., 2021](#); but see [Georges, Guillaume & Schiltz, 2020](#)).

However, the issue of whether neural responses are affected by physical stimulus features is outstanding in the case of symbolic representations of numbers. One reason for this may be that it is difficult to identify physical confounds, such as line curvature, closure, or spatial frequency, that may influence responses as measured with different techniques. However, symbolic representations of numbers use exemplars that are highly distinctive in their physical features, enabling easy and rapid identification (e.g., for rapid identification of western Arabic numerals: [Starrfelt & Behrmann, 2011](#)). This may account for the possibility of above-chance decoding of the identity of Arabic numerals with fMRI ([Eger et al., 2009](#); [Bluthe, De Smet & Op de Beek, 2014](#)) and EEG ([Appelhoff, Hertwig & Spitzer, 2022](#)). Physical differences across numerals may also strongly influence behavioral responses: physical differences across Arabic numerals predict the numerical distance effect (i.e., reduced performance for close than distant numbers; since [Moyer & Landauer, 1967](#)) much better than the quantity of numerical distance ([Cohen, 2009](#)).

In addition to being highly distinctive, the limited stimulus variability of symbolic representations of numbers may contribute to the impact of their physical attributes. Arabic numerals rely on very few basic exemplars, i.e., ten numerals from 0 to 9. Additionally, numerals are typically presented in typography with approximate homogeneity in most aspects, such as size, line thickness, and color. Most studies using Arabic numerals have used a single font type and point size, in black or white (e.g., [Dehaene, 1996](#); [Fabre & Lemaire, 2005](#); [Libertus & Brannon, 2007](#); but for more variation in the context of adaptation, see [Vogel, Goffin & Ansari, 2015](#); [Finke et al., 2021](#)). Stimulus variability is important, in that it relates to the consistency, and thus impact, of physical stimulus confounds on responses: increasing stimulus variability has been identified as a way to decrease the impact of stimulus confounds (e.g., with images: [Thorpe](#)

et al., 1996; Foldiak et al., 2004; Rossion et al., 2015; with dots: Georges, Guillaume & Schiltz, 2020; with written words and/or numbers: Cohen, 2009; Lochy & Schiltz, 2019; Volfart et al., 2020; Marlair, Crollen & Lochy, 2022). Taken together, the limited variability within symbolic number exemplars, and the high variability across number exemplars, leaves responses highly susceptible to stimulus confounds.

In many previous studies using Arabic numerals, multiple number formats or modalities, or many exemplars of mathematical equations, have been used instead of controlling for numerals' visual confounds (e.g., Pauli et al., 1994; Dehaene, 1996; Liu et al., 2018; van Hoogmoed & Kroesbergen, 2018). The issue of stimulus confounds for symbolic representations of numbers is most pertinent when attempting to distinguish conceptual neural responses, such as for parity or magnitude, with restricted stimulus sets (e.g., Hsu & Szucs, 2012; Guillaume et al., 2020; Sokolowski et al., 2021). Take the example of *parity*, a conceptual distinction that divides numbers into the categories of even and odd. These categories are defined by the ending of numbers in one of five digits (0,2,4,6,8 for even; 1,3,5,7,9 for odd); given the extreme differences in shape features across Arabic numerals, the even set of numerals is more curved, closed, and symmetrical than the odd set. Therefore, neural responses distinguishing even and odd numerals (Fabre & Lemaire, 2005; Guillaume et al., 2020) might reflect a conceptual distinction of parity categories, but could also reflect the visual system's sensitivity to physical differences in numeral shape features across small groups of even and odd numbers.

Here, we targeted the conceptual numerical processing of parity, using EEG frequency-tagging to differentiate neural responses to a group of even (2,4,6,8) and a group of odd (3,5,7,9) Arabic numerals (similarly as in Guillaume et al., 2020). Importantly, we tested two control conditions of non-conceptual numeral groups, designed to differ on the extent of their physical

stimulus differences across groups (*small physical differences control*: 2,3,6,7 vs. 4,5,8,9; and *large physical differences control*: 2,4,5,7 vs. 3,6,8,9; **Fig. 1**). Moreover, we presented four different stimulus sets orthogonally, differing in their amount of variability: 1) *1 font* (lowest variability); 2) *10 fonts*; 3) *10 mixed fonts/numeral drawings*; and 4) *20 drawn* diverse, colored numerals (highest variability). We explored whether conceptual responses to parity could be distinguished from those to physical differences of non-conceptual number groups, quantitatively and qualitatively, and whether this depended on stimulus set variability. We predicted that the large physical differences control would yield larger response amplitudes than the small physical differences control, particularly with lower levels of stimulus set variability. Most importantly, we predicted that the response to parity would be larger than both control responses only with higher levels of stimulus set variability.

Methods

Participants

Participants were 15 healthy, neuro-typical human adults (age range: 18 – 27 years; age mean = 23.4 years; 2 males; 13 females). All 15 participants reported to have normal or corrected-to-normal vision, and all but two to be right-handed. The sample size approximately matched that of previous frequency-tagging studies reporting robust results, including at the individual level, e.g., using a symmetry/asymmetry paradigm: [Retter & Rossion, 2016a](#); in numerical cognition: [Park, 2018](#)). They were recruited from the University of Luxembourg community via educational and social media. Each participant was tested individually, following signed, informed consent, with procedures approved by the Ethical Review Panel of the University of Luxembourg (ERP

20-057), and consistent with the Code of Ethics of the World Medical Association (Declaration of Helsinki).

Stimuli and conditions

The stimuli were the eight Arabic numerals ranging from 2-9. The number 0 was excluded as it is not always readily identified as even (e.g., [Levenson, Tsmair & Tirosh, 2007](#)), and the number 1 was arbitrarily excluded so that two groups of four numerals could be defined. This also corresponds with the numerals that have been used in many previous EEG studies ([Fabre & Lemaire, 2005](#); as operands: [Pauli et al., 1994](#); [Niedeggen, Rosler & Jost, 1997](#); [Rosler & Jost, 1999](#); [Jost, Henninghausen & Rosler, 2004](#); [Zhou et al., 2006](#); [Domahs et al., 2007](#)). Three experimental conditions allowed the comparison of *Parity* with two controls, a *Small Physical Differences (Small Phys.-Diff.) Control* and a *Large Physical Differences (Large Phys.-Diff.) Control*. In the Parity condition, the even integers 2,4,6, and 8 defined group A, and the odd integers 3,5,7, and 9 defined group B. In the Small Phys.-Diff. control: group A consisted of 2,3,6, and 7, and group B consisted of 4,5,8 and 9. In the Large Phys.-Diff. control: group A consisted of 2,4,5 and 7, and group B consisted of 3,6,8, and 9 (**Fig. 1a**).

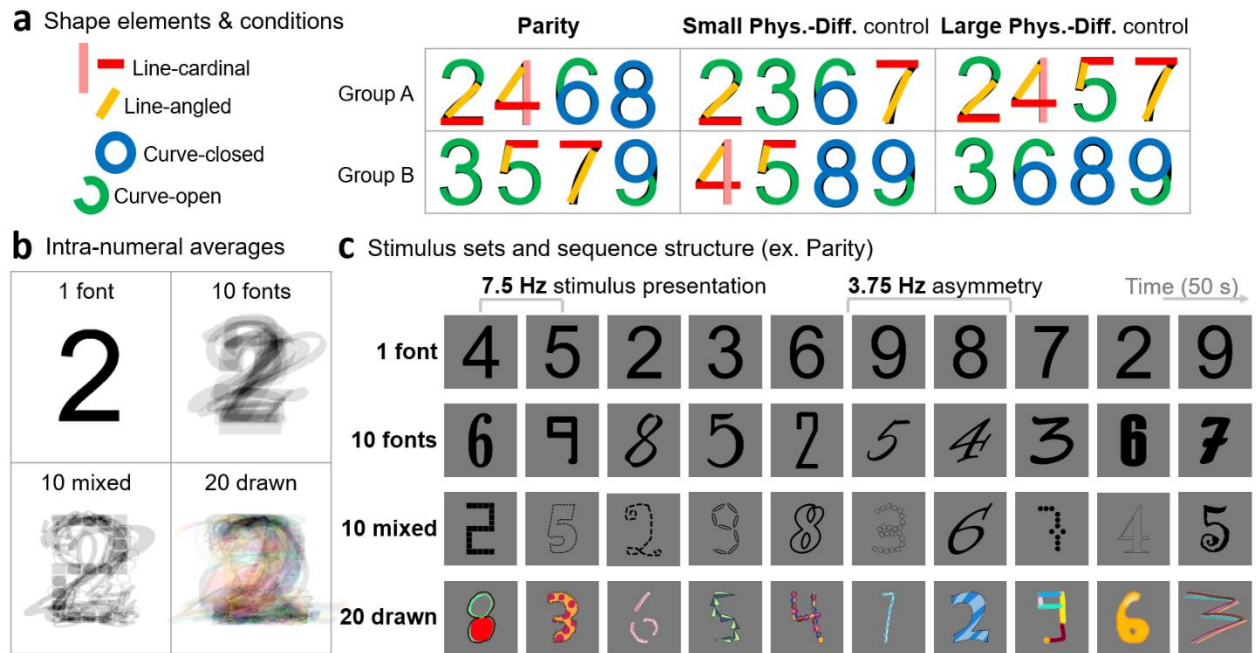


Figure 1. Experimental design. *a)* Three experimental conditions were each defined by two groups of four Arabic numerals. In the Parity condition, group A consisted of even numerals, and group B of odd numerals. In the control conditions, numerals were assigned non-conceptually to groups in consideration of the amount of shape differences across groups (small inter-group differences in the Small Phys.-Diff. control; large in the Large Phys.-Diff. control). *b)* Intra-numeral averages, with the numeral 2 as an example, demonstrating the variability of each of the stimulus sets. *c)* Stimuli were presented in 50 s sequences at 7.5 Hz (133 ms cycles), alternating between numerals of the two groups of each condition (A-B-A-B...), leading to a 3.75 Hz asymmetry frequency tag.

The control conditions were designed to provide two relative measures of non-conceptual, visual differences across numeral groups: these groups did not differ in parity, each containing 2 even and 2 odd numbers. The groups were also approximately matched for magnitude: Parity group A vs. B means: 5 vs. 6; Small Phys.-Diff. control: 4.5 vs. 6.5; Large Small Phys.-Diff. control: 4.5 vs. 6.5. In the Small Phys.-Diff. control, intended to have lower

visual differences across numeral groups, the basic shape features of straight (along the cardinal axes or angled) vs. curved contours, open vs. closed elements, and symmetry (e.g., [Triesman, 1986](#); [McCloskey & Schubert, 2014](#); for features in letter perception: [Fiest et al., 2008](#); [Grainger, Rey & Dufau, 2008](#)) of the Arial font were approximately balanced across stimulus groups by visual inspection. In the Large Phys.-Diff. control, the curviest, the most closed numeral shapes were sorted into a single group (3,6,8 and 9). None of these integers contained a straight line, and three contained at least one closed element; while all of the numerals in the other group (2,4,5 and 7) contained at least one straight line, and only one of which contained one closed element. Two numerals also had horizontal symmetry in group B of the Large Phys.-Diff. control (3 and 8), while none did in group A.

The amount of visual differences across groups for each of the conditions was also measured quantitatively, for the Arial font, in terms of non-background image space and spatial frequency (using the statistical toolbox of [Bainbridge & Oliva, 2015](#)), as well as parametric complexity ([Attneave & Arnoult, 1956](#)). The amount of physical similarity of all stimulus pairs within each group was also measured (using the method of [Cohen, 2009](#)). For all these properties, the differences across numeral groups for Parity and the Small Phys.-Diff. control were similar, tending to be somewhat smaller for Parity; the differences across numeral groups for the Large Phys.-Diff. control were considerably the largest (**Table S1**). The Large Phys.-Diff. control was thus tested as a dramatic example of the possible contribution of physical stimulus confounds, rather than a control intended to match the confounds of the Parity condition *per se*.

Stimulus sets

There were four different stimulus sets, designed to differ in their intra-numeral variability (**Fig. 1b**). In the *1 font* stimulus set, only the Arial font was used, defining one exemplar of each numeral. In the *10 fonts* set, the variable fonts Agency FB, *Segoe Print*, *Informal Român*, Poor Richard, Gloucester MT Extra Condensed, **Forte**, *Pristina*, OCR A Extended, *Vladimir Script*, and **Haettenschweiler** were used, so that there were 10 exemplars of each numeral. Note that including variability by changes in font has been applied in previous studies with Arabic numerals (e.g., Cohen, 2009; Vogel, Goffin & Ansari, 2015; with EEG: Lochy & Schiltz, 2019; Guillaume et al., 2020; Finke et al., 2021; Marlair, Crollen & Lochy, 2022). The *10 mixed* set was also comprised of 10 exemplars of each numeral, but designed to include even more variability across exemplars that included both fonts and drawings. There were five fonts used: two in plain type (*Informal Român* and *Gigli*), two in outline only, i.e., with a transparent interior (Modern No. 20 and **Forte**), and one in a dotted outline only (**SHOWCARD GOTHIC**). Additionally, there were five computer-drawn exemplars, comprised of: black squares in a grid organization, black dots in a grid with gentle edges, outlined cloud shapes, outlined ovals, and a free-form dotted line. Finally, the *20 drawn* set was comprised of 20 exemplars of each numeral, hand-drawn in color on a tablet (details below). Throughout, exemplar style (e.g., blue diagonal stripes with a thick angular shape in *20 drawn*) was applied to each of the numerals, to decrease inter-numeral variability.

In stimulus generation, the font and computer-drawn stimuli were generated in PowerPoint (Microsoft, USA). The hand-drawn stimuli were generated with a stylus (Smart Pen, Dongguan Ruibi Touch Technology Co., Ltd., China) on a tablet (iPad, 10.2 inch screen, Apple, USA). The numbers were extracted as PNGs on a white background, before being auto-cropped

to the borders of the number edges in IrfanView 64 4.53 (<https://www.irfanview.net/>). The stimuli were seized to a height of 360 pixels, based on the Arial font only for all font stimuli; variability in height across fonts of the same text size was retained. Finally, the image canvas was expanded from the center to a common size of 500x500 pixels, and the background was set to a mid-level gray (123/255). All four stimulus sets are freely available online (dataset: Retter, Eraßmy & Schiltz, 2024; including the numerals 0 and 1 for the *20 drawn* stimulus set): for some examples, see **Fig. 1**.

Paradigm and procedure

To record neural responses to parity, electroencephalogram (EEG) was recorded with a frequency-tagging approach (also referred to as “steady-state visual-evoked potentials” or “fast periodic visual stimulation”; Regan, 1989; Rossion, 2014; Norcia et al., 2015). Conceptual responses to numerical properties have been proposed in numerous EEG frequency-tagging studies (for numerosity/magnitude: Libertus, Brannon & Woldorff, 2011; Park, 2018; Guillaume et al., 2020; Lucero et al., 2020; Grasso et al., 2024; parity: Guillaume et al., 2020; cross-format integration: Marinova et al., 2021; Marlair et al., 2021; and the numerical distance effect: Marlair, Crollen & Lochy, 2022; Marlair, Lochy & Crollen, 2024). Stimuli were presented at a periodic rate, 7.5 images a second (7.5 Hz), and the recorded EEG was transformed into the frequency domain, in which responses to stimulus presentation are specifically tagged at 7.5 Hz and its harmonics (**Fig. 2**).

We applied a symmetry/asymmetry design (Tyler & Kaitz, 1977; Victor & Zemon, 1985; for more recent examples: Ales & Norcia, 2009; Ales et al., 2012; Retter & Rossion, 2016a; 2017; Retter et al., 2020; Gwinn et al., 2021). In this design, two groups of stimuli, for example leftward motion as group A, and rightward motion as group B, are presented in alternation, i.e.,

in a sequence of A-B-A-B-A-B... Responses that are “symmetrical”, i.e., general to both A and B stimulus groups, are expected at the stimulus presentation frequency, F . Responses that are “asymmetrical”, i.e., differing for A vs. B, are expected at half the of stimulus presentation frequency, $F/2$ (for a conceptual illustration, please see Fig. 2 of [Ales & Norcia, 2009](#)). Here, the two groups of images (e.g., the even numerals and odd numerals for the parity condition) within each stimulus set were alternated (even-odd-even-odd...), leading to a 3.75 Hz (i.e., 7.5 Hz/2) tag of asymmetry responses to the contrasted numeral groups. In the parity conditions, these asymmetry responses are predicted to reflect conceptual differences in the neural responses to even and odd numbers, although physical differences are also expected to contribute, as measured specifically in the control conditions.

Stimuli were presented as a 62.5% squarewave modulation at 7.5 Hz, i.e., appearing for 83 ms followed by a 50 ms inter-stimulus interval. Stimuli were presented through Java (Oracle, USA), with a Dell S2419HGF (USA) monitor with a screen size of about 53x30 cm and a refresh rate of 120 Hz, connected to a Dell (USA) PC with a Geforce 1050 (Nvidia, USA) graphics card. Participants had a 1 m viewing distance, such that stimuli subtended about 1.9° of vertical visual angle on average. To reduce exact image repetitions, at each presentation cycle the size of the stimulus was varied randomly between 80-120% of the original size in 10% increments ([Dzhelyova & Rossion, 2014](#)). Throughout stimulation, participants were asked to fixate on a centrally presented dark gray fixation cross, and were instructed to press on the space bar each time the cross briefly (250 ms) changed to white, which occurred 6 times per sequence at random intervals, with a minimum of 750 ms between changes (color task : [Park, 2018](#) ; [Guillaume et al., 2018](#); [Van Rinsveld et al., 2020](#); passive viewing: [Libertus, Brannon & Woldorff, 2011](#); see the Electronic supplementary material for the behavioral results).

Each stimulation sequence lasted for 50 s, throughout which stimuli from the two groups of each condition were alternated, leading to an asymmetry frequency tag at 3.75 Hz. In the full trial structure, designed to maintain fixation and avoid artifacts at the beginning and end of stimulation, there was: 1) 1-2 s of the fixation cross alone; 2) 2 s of gradual stimulus fade-in; 3) the 50 s stimulation sequence; 4) 2 s of gradual stimulus fade-out; and 5) 1-2 s of the fixation cross. For each condition there were three trial repetitions, leading to a total of 2.5 minutes of the stimulation sequence. This time encompassed 562 stimulus asymmetry presentations per condition. In total, including rest times between trials, the experimental recording session lasted about 40 minutes.

EEG acquisition

EEG was acquired with a BioSemi, ActiveTwo system, comprised of 68 active recording electrodes in the standard 10/20 locations; the standard 64, plus PO9, I1, I2, and PO10 (BioSemi B.V., Amsterdam, Netherlands). Electrode offsets were held below 40 mV, relative to the system's common-mode-sense and driven-right-leg reference loop. The EEG was digitized at a sampling rate of 512 Hz.

Analysis

Analysis was conducted with LetsWave 6 (<https://www.letswave.org/>), running over Matlab R2019b (MathWorks, USA), largely in accordance to previous symmetry/asymmetry designs (Retter & Rossion, 2016a; 2017; Retter et al., 2020). Briefly, and noting any differences: data were pre-processed, including a fourth-order Butterworth band-pass filter at 0.1-80 Hz, and coarsely segmented 1 s before and after each trial. Muscular artifacts related to eye blinks were corrected by removing a single ICA component, for participants blinking more than 0.2 times per

s during the stimulation sequences (three participants; overall $M = 0.14$ blinks/s; $SD = 0.14$ blinks/s). Noisy channels, containing multiple deflections over $\pm 100 \mu V$, were corrected for by linear interpolation with neighboring channels ($M = 1.4$ channels per participant; range = 0-3). Data were re-referenced to the average of all EEG channels, and the stimulation sequences were precisely cropped to an integer number of 3.75 Hz cycles (187 cycles per sequence = 49.87 s = 25,532 sampling bins). Sequence repetitions were averaged together by condition, before a fast Fourier transform was used to represent the data as frequency-domain amplitude and phase spectra (**Fig. 2**).

The harmonics of 3.75 Hz in the amplitude spectrum were segmented, together with a range of neighboring frequency “noise” bins. A baseline was defined by a range of about 0.4 Hz surrounding noise (20 bins; for the baseline subtraction, the minimum and maximum bin were excluded). For response amplitude measurement, a baseline subtraction was applied; for significance assessment of individual harmonics, a Z-score was computed relative to this baseline. The harmonics-of-interest were defined by those consecutively surpassing a threshold of $p < .01$, $Z > 2.32$, 1-tailed, over the average of all EEG channels, in at least one condition ([Retter, Rossion & Schiltz, 2021](#)). This led to three odd, asymmetry harmonics: 3.75, 11.25, and 18.75 Hz; and six even, symmetry harmonics: 7.5, 15, 22.5, 30, 37.5, and 45 Hz. These harmonics were then summed for further analyses, separately for the asymmetry and symmetry responses (**Fig. 2**; [Retter & Rossion, 2016b](#); [Retter, Rossion & Schiltz, 2021](#)).

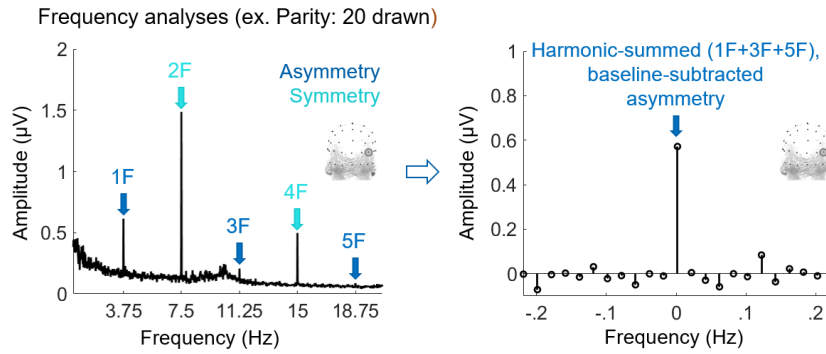


Figure 2. Harmonic-summation and baseline-subtraction. Group-level experimental data from the parity condition, with the 20 drawn stimulus set, are given as an example (electrode PO8) of an original frequency-domain amplitude spectrum, and the measurement of asymmetry (and symmetry) responses through harmonic-summation and baseline-subtraction. In the right panel, 0 Hz is the target frequency across harmonics, surrounded by a relative frequency range.

The region-of-interest (ROI) was defined as a bilateral occipitotemporal region, containing eight channels (PO8; PO10; O2; I2; and corresponding left channels; drawn in **Fig. 3**). This ROI was originally based on previous studies using numeral stimuli, in which responses were maximal over occipitotemporal regions, with responses typically peaking around PO8 (see Fig. 3 of [Libertus, Woldorff & Brannon, 2007](#); Fig. 2 of [Liu et al., 2018](#); Fig. 4 of [Akbari et al., 2022](#); [Pinel et al., 2001](#); [Plodowski et al., 2003](#); [Guillaume et al., 2020](#); [Marinova et al., 2021](#)), although sometimes labeled as an (occipito)parietal region. Post-hoc, it was verified that this ROI appropriately captured the amplitude distribution of the response over the scalp: it contained on average 76% of the top eight channels across conditions and stimulus sets for the harmonic-summed, baseline-subtracted asymmetry responses of interest. To examine response lateralization, this ROI was split into left and right sub-regions: the lateralization index was calculated as the amplitude over the (right-ROI minus the left-ROI), divided by the amplitude

over the right-ROI, at the individual participant level (overall, two individual index values were removed for exceeding two times the standard deviation of the mean).

To determine significance of asymmetry responses overall, a Z-score was assessed on the summed-harmonic responses over the occipito-temporal ROI (Retter, Rossion & Schiltz, 2021). A repeated-measured analysis-of-variance (ANOVA) was applied, with within-participants factors of *Condition* (3 levels: Parity; Small Phys.-Diff. control; Large Phys.-Diff. control), and *Stimulus Set* (4 levels: 1 font; 10 fonts; 10 mixed; 20 drawn). Follow-up one-way ANOVAs were then performed separately for each stimulus set, with the factor of *Condition*. Further, follow-up, paired-samples t-tests were performed to compare the conditions within stimulus sets in a pairwise manner (3 comparisons, Bonferroni-corrected, 2-tailed).

Results

Frequency-domain amplitude spectra were analyzed through harmonic-summation and baseline-subtraction (Fig. 2). The main responses of interest, at 3.75 Hz and its specific harmonics, tagged asymmetry responses, reflecting differences in the responses to the alternated numeral groups. The results showed asymmetry response amplitudes peaking over the occipitotemporal cortex that varied across conditions and stimulus sets (Fig. 3).

An asymmetry response to Parity (2,4,6,8 vs. 3,5,7,9) was present significantly for all stimulus sets at the occipito-temporal ROI; however, the asymmetry responses were also significant in all the control conditions, all Z 's > 2.32 , p 's $< .01$ (for exact values, see Table S2). At the individual participant level, 69% of individuals' responses were significant at $Z > 1.64$, $p < .05$, across all conditions and stimulus sets (Z 's ranging from -1.07 to 13.4).

There were significant main effects of *Condition*, $F_{2,28} = 8.16$, $p = .002$, $\eta_p^2 = 0.37$, and *Stimulus Set*, $F_{3,42} = 28.5$, $p < .001$, $\eta_p^2 = 0.67$, for the asymmetry responses. However, these

main effects were qualified by a significant interaction, $F_{6,84} = 6.04$, $p < .001$, $\eta_p^2 = 0.30$. Follow-up analyses within stimulus sets revealed a significant effect of *Condition* only for the 1 font, $F_{2,28} = 12.3$, $p < .001$, $\eta_p^2 = 0.47$, and 20 drawn, $F_{2,28} = 8.10$, $p = .002$, $\eta_p^2 = 0.37$ (10 fonts: $F_2 = 2.64$, $p = .089$, $\eta_p^2 = 0.16$; 10 mixed: $F_2 = 1.62$, $p = .22$, $\eta_p^2 = 0.10$) stimulus sets.

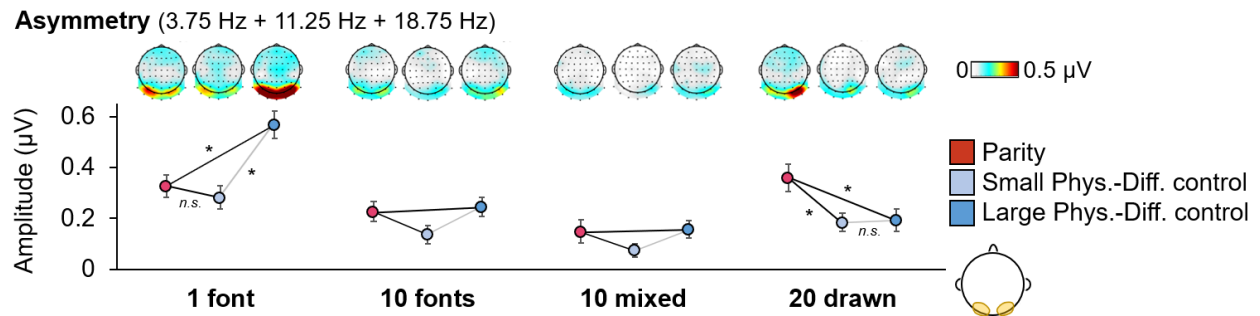


Figure 3. Asymmetry responses at 3.75 Hz and harmonics: group-level amplitudes and scalp topographies. Amplitude is quantified over the occipito-temporal ROI, indicated on the scalp plot in the bottom right, harmonic-summed and baseline-subtracted. Error bars indicate ± 1 SE of the mean. * = significant ($p < .05$, Bonferroni-corrected)

For 1 font, this was driven by the largest response for the Large Phys.-Diff. control (p 's $< .003$, d 's $> .21$); the difference between Parity and the Small Phys.-Diff. control was not significant, $t_{14} = 0.66$, $p = .16$, $d = .26$. For 20 drawn, the effect of *Condition* was driven by the largest response for Parity (p 's $< .034$, d 's $> .75$); there was not a significant difference between the two control conditions, $t_{14} = 0.21$, $p = .25$, $d = .054$. That is, while Parity responses were generally higher than those to the Small Phys.-Diff. control across stimulus sets, only the 20 drawn stimulus set produced significantly larger asymmetry responses than both the control conditions (**Fig. 4**). For this 20 drawn stimulus set, at the individual participant level, 93% of individuals' responses were significant ($Z > 1.64$, $p < .05$) for Parity, compared to 60% for each

of the control conditions; and 80% of individuals had a larger response to parity than to the average of the two controls (see Fig. 4).

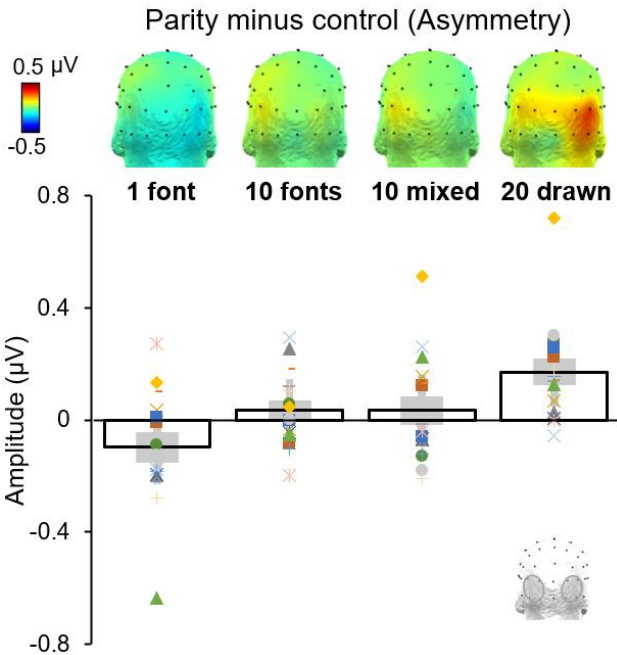


Figure 4. Asymmetry response amplitude differences for parity minus the average of the two control conditions. Amplitudes from 3.75 Hz are harmonic-summed and baseline-subtracted. Above: back-of-the-head, 3D scalp topographies. Below: the corresponding response amplitudes over the occipito-temporal ROI, as drawn in the bottom right. Mean data is plotted with a black outline bar plot, with ± 1 SE of the mean indicated by a filled gray rectangle. Individual participant data are represented with dot markers, with each individuals' marker shape and color held constant across stimulus sets.

The scalp topography of Parity asymmetry responses peaked over occipito-temporal channels, particularly over the right hemisphere: on average across all conditions and stimulus sets, 84% of individuals' responses were right-lateralized (i.e., had a right-lateralization index > 0). There were differences in lateralization apparent across stimulus sets and conditions (**Fig. 3**).

For the 20 drawn and 10 mixed stimulus sets, the right lateralization was stronger for the Parity condition than for the controls (20 drawn Parity index: 0.38; controls 0.15-0.16; see **Fig. S2a**; 10 mixed Parity index: 2.11; controls 0.08-0.13; however, this pattern was reversed for the 1 font and 10 fonts stimulus sets, with the Parity response for 1 font being the only left-lateralized response (index = -0.94; controls 0.44-0.56; 10 fonts Parity index: 0.01; controls 0.16-0.22). The mean phase of the asymmetry response at 3.75 Hz for the 20 drawn stimulus set was determined for the Parity condition, but was not reliable enough to interpret for the lower-amplitude control conditions, preventing further analyses (**Fig. S2b**).

Finally, the symmetry responses at 7.5 Hz and its harmonics, tagging symmetrical responses to visual stimulus presentation, were examined. These responses did not considerably differ across conditions for any of the stimulus sets (**Fig. S1**). There was only a significant main effect of *Stimulus Set*, $F_{3,42} = 5.46$, $p = .003$, $\eta_p^2 = 0.28$ (all other F 's < 0.55 , p 's $> .77$). This reflected small response amplitude differences overall across stimulus sets, as follows in decreasing order: 10 fonts, 1 font, 20 drawn, 10 mixed (respectively, $M = 2.14 \mu V$, $SE = 0.13 \mu V$; $M = 2.00 \mu V$, $SE = 0.15 \mu V$; $M = 1.87 \mu V$, $SE = 0.093 \mu V$; $M = 1.78 \mu V$, $SE = 0.096 \mu V$). The topography of symmetry responses appeared more consistent than the asymmetry responses across stimulus sets, also being right-lateralized, but not differing by condition (**Fig. S1**).

Discussion

To record responses to a numerical concept, parity, we used an EEG frequency-tagging, symmetry/asymmetry paradigm, in which even (2,4,6,8) vs. odd (3,5,7,9) symbolic, Arabic numerals were contrasted at 3.75 Hz. Responses at 3.75 Hz thus reflected differences in the conceptual responses to even and odd numerals – as well as to any elicited visual response

differences between the groups of numerals. To isolate responses to physical stimulus differences, we tested two control conditions in which groups of non-conceptual numerals were alternated at 3.75 Hz (Small Phys.-Diff. control, targeting small physical differences across groups: 2,3,6,7 vs. 4,5,8,9; Large Phys.-Diff. control, targeting extremely large physical differences: 2,4,5,7 vs. 3,6,8,9). To further investigate the influence of visual attributes of the EEG responses, we compared four stimulus sets, differing in their variability (from 1 font, consisting of one Arial font exemplar per numeral, up to 20 drawn, consisting of 20 hand-drawn, colored exemplars per numeral). We investigated whether Parity would give larger and/or qualitatively different 3.75 Hz asymmetry responses than the control conditions, and to what extent this was supported by increasing stimulus variability.

Identifying conceptual responses to parity with variable stimuli

The asymmetry response at 3.75 Hz to Parity over the occipitotemporal cortex was significantly larger than to the two control conditions for the most variable, 20 drawn stimulus set: this response was 2.0 times larger than to the Small Phys.-Diff. control, and 1.9 times larger than to the Large Phys.-Diff. control. The Parity asymmetry response was also larger than the Small Phys.-Diff. control for the moderately variable 10 fonts (1.7 times larger) and 10 mixed (2.0 times) stimulus sets. For the least variable, 1 font stimulus set, the Parity asymmetry response was not significantly larger than the Small Phys.-Diff. control, but was significantly smaller than the Large Phys.-Diff. control.

A higher response amplitude to Parity than the control(s), particularly as stimulus set variability increases, is in line with our prediction that conceptual responses to parity may be evidenced beyond (variable) physical stimulus influences. An automatic access to parity would be consistent with some previous studies (e.g., [Shepard et al., 1975](#); [Krueger & Hallford, 1984](#);

Reynvoet et al., 2002; with EEG: Fabre & Lemaire, 2005; and EEG frequency-tagging: Guillaume et al., 2020), and with the status of parity as a fundamental aspect of number knowledge (Ifrah, 2000). However, the interpretation of capturing automatic, conceptual responses to parity, building upon the persistent response amplitudes to physical stimulus differences as evidenced in the control condition(s), must be proposed tentatively. Physical stimulus differences can equal or surpass the amplitude of potential conceptual response differences to even and odd numerals, as demonstrated by the Large Phys.-Diff. condition with the 1 font stimulus set, and there was a lack of qualitative response specificity for Parity (as addressed in a following section). Moreover, responses to conceptual and physical stimulus response components may not be simply additive.

The higher response amplitude to Parity than the controls only for the 20 drawn stimulus set might be attributed to the increased variability of this stimulus set. It may be that the low- to mid-level visual response differences elicited by the numeral groups were more equated across conditions for this highly variable stimulus set, leading to the least ambiguous identification of conceptual Parity responses. It may be observed that the amplitude of the Parity asymmetry responses was not larger for the 20 drawn than 1 font stimulus set, and smaller responses were produced for the 10 fonts and particularly for the 10 mixed stimulus sets (**Fig. 3**): physical differences across numeral groups may not have been more reduced in the complex, colored, 20 drawn stimulus set, merely more equated across conditions. Relative to the other stimulus sets, the response to the 20 drawn Large Phys.-Diff. control appeared specifically decreased, such that the two control conditions produced similar asymmetry response amplitudes only for this stimulus set.

The 20 drawn stimulus set created the highest amount of diversity within numeral exemplars (20 colored, hand-drawn per numeral), while decreasing the variability across numerals (each schema was applied to all numerals). Increasing variability intra-numerally and decreasing variability inter-numerally might decrease the relative importance of the latter, leading to a more controlled impact of visual differences across numeral sets (e.g., [Thorpe et al., 1996](#); [Foldiak et al., 2004](#); [Rossion et al., 2015](#); [Georges, Guillaume & Schiltz, 2020](#); [Marlair, Crollen & Lochy, 2022](#)). We propose that increasing variability, such as modeled most strongly by the 20 drawn stimulus set here, is therefore a promising avenue in future research targeting conceptual processing. It should be noted that there are many means to increase stimulus variability, beyond using a hand drawn stimulus set, such as using: multiple fonts (e.g., with Arabic numerals: [Cohen, 2009](#); [Vogel, Goffin & Ansari, 2015](#); [Lochy & Schiltz, 2019](#); [Guillaume et al., 2020](#); [Finke et al., 2021](#); [Marlair, Crollen & Lochy, 2022](#); larger stimulus sets, e.g., incorporating 0 and 1, and using multi-digit numbers; numerous formats, e.g., numeral, word, and dot representations of numbers¹: [Dehaene, 1996](#); [Temple & Posner, 1998](#); [Libertus, Woldorff & Brannon, 2007](#); numerous modalities across conditions: [Kiefer & Dehaene, 1997](#); [Dickson et al., 2018](#); [Finke et al., 2021](#); or numerous mixed modalities: [Marinova et al., 2021](#); [Marlair et al., 2021](#); [Marlair, Crollen & Lochy, 2022](#); and varying image parameters, such as size, at each stimulus presentation: [Dzhelyova & Rossion, 2014](#); [Vogel, Goffin & Ansari, 2015](#); [Guillaume et al., 2020](#); [Finke et al., 2021](#)).

Responses to physical stimulus differences

¹ However, while similarities in responses across formats suggest consistent conceptual processes, this approach is limited in that differences in responses across formats are difficult to attribute to stimulus properties or cognitive processing (see [Libertus, Woldorff & Brannon, 2007](#); and [Plodowski et al., 2003](#)).

It should be observed that physical stimulus differences greatly affected the EEG response amplitude. Significant asymmetry responses at 3.75 Hz were observed for all conditions of all stimulus sets, even those with high variability. These control asymmetry responses reflected contrasts of two group of four numerals, that cannot readily be explained conceptually: in both control conditions, the numeral sets were balanced for parity and approximately for magnitude.

Here, controls of non-conceptual numeral groups were used to estimate the impact of physical stimulus confounds on the parity response (as in [Guillaume et al. 2020](#), except that in that study only half of the trial repetitions were tested for the control, which could account for the lower baseline-dependent control responses reported). Two control conditions were used here to give measures of predicted low and high visual differences across numeral groups. The Large Phys.-Diff. control condition was designed intuitively to have a large amount of shape differences across numerals groups: straight and open shapes (2,4,5,7; all containing at least one straight line segment, and only one of which contains a closed element) vs. curvy and closed shapes (3,6,8,9; none of which contains a straight line segment, and three of which have at least one closure). Not unrelatedly, the Large Phys.-Diff. control also had the largest differences across numeral groups in non-background image space, perimetric complexity, spatial frequency, and physical similarity, measured with the Arial font (**Table S1**). The EEG response amplitude to the Large Phys.-Diff. control was consistently about double that to the Small Phys.-Diff. control across three stimulus sets, i.e., all but 20 drawn, and greatly exceeded the Parity response for the 1 font stimulus set (**Fig. 3**). Instead of being well-matched to the Parity condition, the Large Phys.-Diff. control served here as a demonstration of the potential impact of stimulus confounds.

This emphasizes the importance that studies address (intrinsic) physical stimulus confounds with controls, to limit the possibility that stimulus confounds, rather than conceptual representations, could account for reported effects. Importantly, this issue is not only relevant for non-symbolic numerical representations, but is demonstrated here for symbolic, Arabic numerals (see also [Cohen, 2009](#)). To be most informative, these controls should be as well-matched as possible: for example, while it is possible to use nonmeaningful characters, letters, foreign language letters, and/or symbols as a control condition (e.g., [Iguchi & Hashimoto, 2000](#); [Price & Ansari, 2011](#); [Park et al., 2014](#); [Fraga-Gonzalez et al., 2022](#)), their visual attributes might not appropriately match those of experimental interest. For another example, a control condition of non-frequency-tagged trials, leading to no frequency-tagged responses, is not informative about potential physical stimulus confounds ([Marinova et al., 2021](#)).

With the two control conditions tested here, coarsely targeting shape, we are not able to point to which stimulus attributes most affected the neural responses. By testing additional control numeral groups, it may be possible to further probe which physical differences relate to larger EEG response amplitudes (e.g., cumulative perimeter, area, spatial frequency; for other approaches to identify visual features influencing non-symbolic representations of number with fMRI: [Cohen Kadosh et al., 2007](#); [Wilkey et al., 2017](#)). We identified 15 possible control conditions that could be tested with the numerals 2-9 (each consisting of two sets of four numerals equated for parity and excluding large magnitude differences across sets). Understanding the impact of different physical features on responses with different methodologies would be highly informative for future studies' design control. It could also be argued that testing further control conditions would increase the specificity of a higher response amplitude to a targeted conceptual condition.

Beyond amplitude: phase and scalp topography

In theory, neural responses to numerals that reflect conceptual processing could be dissociated from responses merely driven by physical stimulus differences not only by amplitude, but by temporal (phase or latency) or spatial (scalp localization) response attributes. Here, little temporal information could be extracted, given the high stimulation rate of 7.5 Hz, i.e., stimulus presentation every 133 ms, and asymmetry rate of 3.75 Hz (267 ms), and low asymmetry phase reliability (**Fig. S2b**). In any case, it is unclear whether to expect a temporal dissociation of conceptual and physical stimulus responses: it has been suggested that EEG responses reflecting conceptual processing of numerosity may occur later in time than those to low-level changes in visual shape of dot arrays (Soltesz & Szucs, 2014; Akbari et al. 2022); but this is contradicted in other studies reporting early processing of numerosity with EEG (Park et al., 2016; see also: Liu et al., 2018).

An EEG scalp topography lateralized over the occipito-temporal/parietal cortex is generally interpreted as involving specialized numerical processing (e.g., Dehaene, 1996; Marinova et al., 2021; but also at medial-occipital sites: Pinel et al., 2001; Guillaume et al., 2018; George, Guillaume & Schiltz, 2020), in line with the parietal cortex being identified as highly relevant from lesion and neuroimaging studies, along with other cortical regions (reviews: Hubbard et al., 2005; Moeller, Willmes & Klein, 2015). Here, the asymmetry responses to parity at 3.75 Hz had right lateralized occipitotemporal topographies, with the exception of the 1 font stimulus set (**Fig. 3**). In previous studies on numerical cognition, response lateralization has been inconsistent with changes in numbers and stimulus properties (e.g., see Figs. 4 and 7 of Libertus, Woldorff & Brannon, 2007; Fig. 2 of Marlair, Crollen & Lochy, 2021), although a right-

lateralization for parity is in line with previous EEG reports (Fabre & Lemaire, 2005; Guillaume et al., 2020).

However, the control asymmetry topographies within each stimulus set had similar topographies as for parity (as in Guillaume et al., 2020); and the responses to visual stimulation at 7.5 Hz were also lateralized over the right occipitotemporal cortex, with some variability in the strength of lateralization across stimulus sets (**Fig. S1**). One possibility is that shape processing for Arabic numerals (or for perceptually grouped small-number dot arrays) may invoke lateralized responses through considerable visual processing beyond the early visual cortex (for less lateralized responses to larger dot numerosities: Libertus, Woldorff & Brannon, 2007; Hyde & Spelke, 2011; Guillaume et al., 2018 vs. Marinova et al., 2021; see also Han et al., 2002; Murray et al., 2002). This further suggests that in this context, scalp topography lateralization is sensitive to physical stimulus confounds, rather than being a reliable indicator of conceptual responses. Overall, a lack of qualitative distinction for parity vs. control responses here underlines the importance of controlling stimulus confounds in order to target conceptual responses to symbolic numbers.

EEG frequency-tagging

Our paradigm benefited from the advantages of EEG frequency-tagging, including automaticity, specificity, consistency, and objectivity (in numerical cognition: Libertus, Brannon & Woldorff, 2011; Park, 2018; Guillaume et al., 2020; Van Rinsveld et al., 2020; Lucero et al., 2020; Marinova et al., 2021; Marlair et al., 2021; Marlair, Crollen & Lochy, 2022; Marlair, Lochy & Crollen, 2024; Grasso et al., 2024). *Automatic* aspects of numerical processing, occurring without instruction or directed attention, can be measured because participants need not be assigned a (number-)related task. *Specificity* of numerical responses is protected from procedural

variability in instruction and task performance, i.e., participants' understanding of the language, mathematical terms, motivation, test anxiety, etc. Also, the absence of a related task enables the same paradigm to potentially be used *consistently* with adults, children, infants, patient populations, and non-human species (e.g., [Libertus, Brannon & Woldorff, 2011](#); [Park, 2018](#)). Frequency-tagging confers a high signal-to-noise ratio and *objectivity* in response identification and measurement ([Regan, 1989](#); [Rossion, 2014](#); [Norcia et al., 2015](#)). Moreover, thanks to frequency-tagging, the asymmetry responses at 3.75 Hz could also be contrasted to the general responses to visual stimulation at 7.5 Hz, which did not differ across conditions within each stimulus set.

Conclusion

In conclusion, we suggest that neural responses may reflect conceptual numerical processing of Arabic numerals, however there is an important influence of physical stimulus confounds that should be taken into account in studies targeting conceptual responses.

Acknowledgments

We thank Vera Hilger and Brenda Gilson for their valuable assistance in piloting and data collection, and the journal editor, Bruno Rossion and three anonymous reviewers for providing helpful feedback on earlier versions of this manuscript.

Funding sources

This work was supported by the SNA Math INTER project [INTER/FNRS/17/1178524 to CS] funded by the Luxembourgish Fund for Scientific Research (FNR, Luxembourg).

AI declaration

We have not used AI-assisted technologies in creating this article.

Competing interests

We declare we have no competing interests.

References

- Akbari, S., Soltanlou, M., Sabourimoghaddam, H., Nuerk, H.-C. & Leuthold, H. (2022). The complexity of simple counting: ERP findings reveal early perceptual and late numerical processes in different arrangements. *Scientific Reports*, 12: 6763.
<https://doi.org/10.1038/s41598-022-10206-y>
- Ales, J. M., & Norcia, A. M. (2009). Assessing direction-specific adaptation using the steady-state visual evoked potential: Results from EEG source imaging. *Journal of Vision*, 9(7):8, 1–13, <http://journalofvision.org/9/7/8/10.1167/9.7.8>

569 Appelhoff, S., Hertwig, R. & Spitzer, B. (2022). EEG-representational geometries and
 570 psychometric distortions in approximate numerical judgment. *PLOS Computational*
 571 *Biology*, 18(12): e1010747. <https://doi.org/10.1371/journal.pcbi.1010747>
 572 Attneave, F., & Arnoult, M. D. (1956). The quantitative study of shape and pattern
 573 perception. *Psychological Bulletin*, 53(6), 452–471. <https://doi.org/10.1037/h0044049>
 574 Bainbridge, W.A. & Oliva, A. (2015). A toolbox and sample object perception data for
 575 equalization of natural images. *Data in Brief* 5, 846-851.
 576 Bulthé, J., De Smedt, B., & Op de Beeck, H. P. (2014). Format-dependent representations of
 577 symbolic and non-symbolic numbers in the human cortex as revealed by multi-voxel
 578 pattern analyses. *NeuroImage*, 87, 311–322.
 579 <https://doi.org/10.1016/j.neuroimage.2013.10.049>
 580 Cheng, D., Cui, Z., Chen, C. *et al.* (2023). The database for extracting numerical and visual
 581 properties of numerosity processing in the Chinese population. *Sci Data*, 10, 28.
 582 <https://doi.org/10.1038/s41597-023-01933-6>
 583 Cohen, D. J. (2009). Integers do not automatically activate their quantity representation.
 584 *Psychonomic Bulletin & Review*, 16(2), 332-336. <https://doi.org/10.3758/PBR.16.2.332>
 585 Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-
 586 dependent and -independent representations of numbers in the parietal lobes. *Neuron*,
 587 53(2), 307–314. <https://doi.org/10.1016/j.neuron.2006.12.025>
 588 Dehaene, S., Tzourio, N., Frak, V., Raynaud, L., Cohen, L., Mehler, J. & Mazoyer, B. (1996).
 589 Cerebral activations during number multiplication and comparison: a PET study.
 590 *Neuropsychologia*, 34, 1097-1106. [https://doi.org/10.1016/0028-3932\(96\)00027-9](https://doi.org/10.1016/0028-3932(96)00027-9)

591 Dehaene, S., Izard, V., & Piazza, M. (2005). Control over non-numerical parameters in
 592 numerosity experiments. Unpublished manuscript (available on www.unicog.org).

593 Dickson, D. S., Cerda, V. R., Beavers, R. N., Ruiz, A., Casteneda, R. & Wicha, N.Y.Y. (2018).
 594 When 2×4 is meaningful: the N400 and P300 reveal operand format effects in
 595 multiplication verification. *Psychophysiology*, 55(11):e13212.
 596 <https://doi.org/10.1111/psyp.13212>

597 Domahs, F., Domahs, U., Schlesewsky, M., Ratnckx, E., Verguts, T., Willmes, K., & Nuerk, H.
 598 C. (2007). Neighborhood consistency in mental arithmetic: Behavioral and ERP
 599 evidence. *Behavioral and brain functions : BBF*, 3, 66. [https://doi.org/10.1186/1744-9081-](https://doi.org/10.1186/1744-9081-3-66)
 600 [3-66](https://doi.org/10.1186/1744-9081-3-66)

601 Dzhelyova, M. & Rossion, B. (2014). The effect of parametric stimulus size variation on
 602 individual face discrimination indexed by fast periodic visual stimulation. *BMC*
 603 *Neuroscience*, 15:87, 1-12. <https://doi.org/10.1186/1471-2202-15-87>

604 Eger, E., Michel, V., Thirion, B., Amadon, A., Dehaene, S., & Kleinschmidt, A. (2009).
 605 Deciphering cortical number coding from human brain activity patterns. *Current Biology*,
 606 19(19), 1608–1615. <https://doi.org/10.1016/j.cub.2009.08.047>

607 Fabre, L., & Lemaire, P. (2005). Age-related differences in automatic stimulus-response
 608 associations: insights from young and older adults' parity judgments. *Psychonomic bulletin*
 609 *& review*, 12(6), 1100–1105. <https://doi.org/10.3758/bf03206450>

610 Fiset, D., Blais, C., Ethier-Majcher, C., Arguin, M., Bub, D., & Gosselin, F. (2008). Features for
 611 identification of uppercase and lowercase letters. *Psychological science*, 19(11), 1161–
 612 1168. <https://doi.org/10.1111/j.1467-9280.2008.02218.x>

613 Finke, S., Kemény, F., Clayton, F.J., *et al.* (2021). Cross-format integration of auditory number
 614 words and visual-Arabic digits: An ERP study. *Front. Psychol.*, 12:765709.
 615 <https://doi.org/10.3389/fpsyg.2021.765709>

616 Földiák, P., Xiao, D., Keyser, C., Edwards, R., & Perrett, D. I. (2004). Rapid serial visual
 617 presentation for the determination of neural selectivity in area STSa. *Progress in brain*
 618 *research*, 144, 107–116. [https://doi.org/10.1016/s0079-6123\(03\)14407-x](https://doi.org/10.1016/s0079-6123(03)14407-x)

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

622 Fraga-González, G., Di Pietro, S. V., Pleisch, G., Walitza, S., Brandeis, D., Karipidis, I. I., &
 623 Brem, S. (2022). Visual Occipito-Temporal N1 Sensitivity to Digits Across Elementary
 624 School. *Frontiers in human neuroscience*, 16, 887413.
 625 <https://doi.org/10.3389/fnhum.2022.887413>

626 Gebuis, T., Cohen Kadosh, R., & Gevers, W. (2016). Sensory-integration system rather than
 627 approximate number system underlies numerosity processing: A critical review. *Acta*
 628 *psychologica*, 171, 17–35. <https://doi.org/10.1016/j.actpsy.2016.09.003>

629 Georges, C., Guillaume, M. & Schiltz, C. (2020). A robust electrophysiological marker of
 630 spontaneous numerical discrimination. *Scientific Reports*, 10, 18376.
 631 <https://doi.org/10.1038/s41598-020-75307-y>

632 Grainger, J., Rey, A., & Dufau, S. (2008). Letter perception: from pixels to
 633 pandemonium. *Trends in cognitive sciences*, 12(10), 381–387.
 634 <https://doi.org/10.1016/j.tics.2008.06.006>

635 Grasso, P. A., Petrizzo, I., Coniglio, F., & Arrighi, R. (2024). Electrophysiological correlates of
636 temporal numerosity adaptation. *Frontiers in neuroscience*, 18, 1349540.
637 <https://doi.org/10.3389/fnins.2024.1349540>

638 Guillaume, M., Mejias, S., Rossion, B., Dzhelyova, M., & Schiltz, C. (2018). A rapid, objective
639 and implicit measure of visual quantity discrimination. *Neuropsychologia*, 111, 180–189.
640 <https://doi.org/10.1016/j.neuropsychologia.2018.01.044>

641 Guillaume, M., Poncin, A., Schiltz, C., & Van Rinsveld, A. (2020). Measuring spontaneous and
642 automatic processing of magnitude and parity information of Arabic digits by frequency-
643 tagging EEG. *Scientific reports*, 10(1), 22254. [https://doi.org/10.1038/s41598-020-79404-](https://doi.org/10.1038/s41598-020-79404-w)
644 [w](https://doi.org/10.1038/s41598-020-79404-w)

645 Gwinn, O. S., Retter, T. L., O’Neil, S. F. & Webster, M. A. (2021). Contrast adaptation in face
646 perception revealed through EEG and behavior. *Frontiers in Systems Neuroscience*,
647 15:701097, 1-15. <https://doi.org/10.3389/fnsys.2021.701097>

648 Han, S., Ding, Y. & Song, Y. (2002). Neural mechanisms of perceptual grouping in humans as
649 revealed by high density event related potentials. *Neuroscience Letters*, 319, 29-32.
650 [https://doi.org/10.1016/s0304-3940\(01\)02502-2](https://doi.org/10.1016/s0304-3940(01)02502-2)

651 Hubbard, E. M., Piazza, M., Pinel, P., & Dehaene, S. (2005). Interactions between number and
652 space in parietal cortex. *Nature reviews. Neuroscience*, 6(6), 435–448.
653 <https://doi.org/10.1038/nrn1684>

654 Hyde, D. C., & Spelke, E. S. (2009). All numbers are not equal: an electrophysiological
655 investigation of small and large number representations. *Journal of cognitive*
656 *neuroscience*, 21(6), 1039–1053. <https://doi.org/10.1162/jocn.2009.21090>

Hyde, D. C., & Spelke, E. S. (2011). Neural signatures of number processing in human infants: evidence for two core systems underlying numerical cognition. *Developmental science*, 14(2), 360–371. <https://doi.org/10.1111/j.1467-7687.2010.00987.x>

Hsu, Y.-F., & Szűcs, D. (2012). The time course of symbolic number adaptation: Oscillatory EEG activity and event-related potential analysis. *NeuroImage*, 59(4), 3103–3109. <https://doi.org/10.1016/j.neuroimage.2011.11.017>

Ifrah, G. (2000). *The Universal History of Numbers*. New York: Wiley.

Iguchi, Y., & Hashimoto, I. (2000). Sequential information processing during a mental arithmetic is reflected in the time course of event-related brain potentials. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, 111(2), 204–213. [https://doi.org/10.1016/s1388-2457\(99\)00244-8](https://doi.org/10.1016/s1388-2457(99)00244-8)

Jost, K., Hennighausen, E., & Rösler, F. (2004). Comparing arithmetic and semantic fact retrieval: effects of problem size and sentence constraint on event-related brain potentials. *Psychophysiology*, 41(1), 46–59. <https://doi.org/10.1111/1469-8986.00119>

Kiefer, M. & Dehaene, S. (1997). The time course of parietal activation in single-digit multiplication: Evidence from event-related potentials *Mathematical Cognition*, 3(1), 1-30. <https://doi.org/10.1080/135467997387461>

Krueger, L. E., & Hallford, E. W. (1984). Why $2 + 2 = 5$ looks so wrong: on the odd-even rule in sum verification. *Memory & cognition*, 12(2), 171–180. <https://doi.org/10.3758/bf03198431>

Levenson, E., Tsamir, P. & Tirosh, D. (2007). Neither even nor odd: Sixth grade students' dilemmas regarding the parity of zero. *The journal of mathematical behavior*, 26(2), 83-95. <https://doi.org/10.1016/j.jmathb.2007.05.004>

- Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and brain functions : BBF*, 3, 1. <https://doi.org/10.1186/1744-9081-3-1>
- Libertus, M. E., Brannon, E. M., & Woldorff, M. G. (2011). Parallels in stimulus-driven oscillatory brain responses to numerosity changes in adults and seven-month-old infants. *Developmental neuropsychology*, 36(6), 651–667. <https://doi.org/10.1080/87565641.2010.549883>
- Liu, R., Schunn, C.D., Fiez, J.A. & Libertus, M.A. (2018). The integration between nonsymbolic and symbolic numbers: Evidence from an EEG study. *Brain and Behavior*, 8(4):e00938. <https://doi.org/10.1002/brb3.938>
- Lochy, A., & Schiltz, C. (2019). Lateralized Neural Responses to Letters and Digits in First Graders. *Child development*, 90(6), 1866–1874. <https://doi.org/10.1111/cdev.13337>
- Lucero, C., Brookshire, G., Sava-Segal, C., Bottini, R., Goldin-Meadow, S., Vogel, E. K., & Casasanto, D. (2020). Unconscious number discrimination in the human visual system. *Cerebral cortex*, 30(11), 5821–5829. <https://doi.org/10.1093/cercor/bhaa155>
- Marinova, M., Georges, C., Guillaume, M., Reynvoet, B., Schiltz, C., & Van Rinsveld, A. (2021). Automatic integration of numerical formats examined with frequency-tagged EEG. *Scientific reports*, 11(1), 21405. <https://doi.org/10.1038/s41598-021-00738-0>
- Marlair, C., Lochy, A., Buyle, M., Schiltz, C., & Crollen, V. (2021). Canonical representations of fingers and dots trigger an automatic activation of number semantics: an EEG study on 10-year-old children. *Neuropsychologia*, 157, 107874. <https://doi.org/10.1016/j.neuropsychologia.2021.107874>

702 Marlair, C., Crollen, V. & Lochy, A. (2022). A shared numerical magnitude representation
 703 evidenced by the distance effect in frequency-tagging EEG. *Sci Rep*, 12, 14559.
 704 <https://doi.org/10.1038/s41598-022-18811-7>

705 Marlair, C., Lochy, A. & Crollen, V. (2024). Frequency-tagging EEG reveals the effect of
 706 attentional focus on abstract magnitude processing. *Psychon Bull Rev*.
 707 <https://doi.org/10.3758/s13423-024-02480-w>

708 McCloskey, M. & Schubert, T. (2014). Shared versus separate processes for letter and digit
 709 identification. *Cognitive Neuropsychology*, 31(5-6), 437-60.
 710 <https://doi.org/10.1080/02643294.2013.869202>

711 Moeller, K., Willmes, K., & Klein, E. (2015). A review on functional and structural brain
 712 connectivity in numerical cognition. *Frontiers in human neuroscience*, 9, 227.
 713 <https://doi.org/10.3389/fnhum.2015.00227>

714 Moyer, R. S. & Landauer, T. K. (1967). Time required for judgements of numerical inequality.
 715 *Nature*, 215, 1519-1520. <https://doi.org/10.1038/2151519a0>

716 Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape
 717 perception reduces activity in human primary visual cortex. *Proceedings of the National*
 718 *Academy of Sciences of the United States of America*, 99(23), 15164–15169.
 719 <https://doi.org/10.1073/pnas.192579399>

720 Niedeggen, M., Rösler, F., & Jost, K. (1999). Processing of incongruous mental calculation
 721 problems: evidence for an arithmetic N400 effect. *Psychophysiology*, 36(3), 307–324.
 722 <https://doi.org/10.1017/s0048577299980149>

723 Norcia, A. M., Appelbaum, L. G., Ales, J. M., Cottareau, B. R., & Rossion, B. (2015). The
724 steady-state visual evoked potential in vision research: A review. *Journal of vision*, 15(6),
725 4. <https://doi.org/10.1167/15.6.4>

726 Park, J., Chiang, C., Brannon, E. M., & Woldorff, M. G. (2014). Experience-dependent
727 hemispheric specialization of letters and numbers is revealed in early visual
728 processing. *Journal of cognitive neuroscience*, 26(10), 2239–2249.
729 https://doi.org/10.1162/jocn_a_00621

730 Park, J., DeWind, N. K., Woldorff, M. G., & Brannon, E. M. (2016). Rapid and Direct Encoding
731 of Numerosity in the Visual Stream. *Cerebral cortex (New York, N.Y. : 1991)*, 26(2), 748–
732 763. <https://doi.org/10.1093/cercor/bhv017>

733 Park J. (2018). A neural basis for the visual sense of number and its development: A steady-state
734 visual evoked potential study in children and adults. *Developmental cognitive*
735 *neuroscience*, 30, 333–343. <https://doi.org/10.1016/j.dcn.2017.02.011>

736 Park, J., Hebrank, A., Polk, T. A., & Park, D. C. (2012). Neural dissociation of number from
737 letter recognition and its relationship to parietal numerical processing. *Journal of Cognitive*
738 *Neuroscience*, 24(1), 39–50. https://doi.org/10.1162/jocn_a_00085

739 Pauli, P., Lutzenberger, W., Rau, H., Birbaumer, N., Rickard, T. C., Yaroush, R. A., & Bourne,
740 L. E., Jr (1994). Brain potentials during mental arithmetic: effects of extensive practice and
741 problem difficulty. *Brain research. Cognitive brain research*, 2(1), 21–29.
742 [https://doi.org/10.1016/0926-6410\(94\)90017-5](https://doi.org/10.1016/0926-6410(94)90017-5)

743 Pinel, P., Dehaene, S., Rivière, D., & LeBihan, D. (2001). Modulation of parietal activation by
744 semantic distance in a number comparison task. *NeuroImage*, 14(5), 1013–1026.
745 <https://doi.org/10.1006/nimg.2001.0913>

746 Plodowski, A., Swainson, R., Jackson, G. M., Rorden, C., & Jackson, S. R. (2003). Mental
 747 representation of number in different numerical forms. *Current biology*, 13(23), 2045–
 748 2050. <https://doi.org/10.1016/j.cub.2003.11.023>

749 Price, G. R., & Ansari, D. (2011). Symbol processing in the left angular gyrus: evidence from
 750 passive perception of digits. *NeuroImage*, 57(3), 1205–1211.
 751 <https://doi.org/10.1016/j.neuroimage.2011.05.035>

752 Regan, D. (1989). *Human Brain Electrophysiology*. Elsevier: New York.

753 Retter, T. L., & Rossion, B. (2016a). Visual adaptation provides objective electrophysiological
 754 evidence of facial identity discrimination. *Cortex*, 80, 35–50.
 755 <https://doi.org/10.1016/j.cortex.2015.11.025>

756 Retter, T. L., & Rossion, B. (2016b). Uncovering the neural magnitude and spatio-temporal
 757 dynamics of natural image categorization in a fast visual stream. *Neuropsychologia*, 91, 9–
 758 28. <https://doi.org/10.1016/j.neuropsychologia.2016.07.028>

759 Retter, T. L., & Rossion, B. (2017). Visual adaptation reveals an objective electrophysiological
 760 measure of high-level individual face discrimination. *Scientific Reports*, 7(1), 3269.
 761 <https://doi.org/10.1038/s41598-017-03348-x>

762 Retter, T. L., Gwinn, O. S., O'Neil, S. F., Jiang, F., & Webster, M. A. (2020). Neural correlates
 763 of perceptual color inferences as revealed by #thedress. *Journal of vision*, 20(3), 7.
 764 <https://doi.org/10.1167/jov.20.3.7>

765 Retter, T. L., Rossion, B., & Schiltz, C., (2021). Harmonic amplitude summation for frequency-
 766 tagging analysis. *J Cog Neurosci*, 16, 1-22. https://doi.org/10.1162/jocn_a_01763

767 Retter, T. L., Eraßmy, L. & Schiltz, C. (2024). Identifying conceptual neural responses to
 768 symbolic numerals [Dataset]. Dryad. <https://doi.org/10.5061/dryad.1zcrjdg13>

769 Reynvoet, B., Caessens, B., & Brysbaert, M. (2002). Automatic stimulus-response associations
 770 may be semantically mediated. *Psychonomic Bulletin & Review*, 9(1), 107–112.
 771 <https://doi.org/10.3758/BF03196263>

772 Rossion, B. (2014). Understanding face perception by means of human electrophysiology.
 773 *Trends in Cognitive Sciences*, 18, 310-318. <https://doi.org/10.1016/j.tics.2014.02.013>

774 Rossion, B., Torfs, K., Jacques, C., & Liu-Shuang, J. (2015). Fast periodic presentation of
 775 natural images reveals a robust face-selective electrophysiological response in the human
 776 brain. *Journal of Vision*, 15(1), 15.1.18. <https://doi.org/10.1167/15.1.18>

777 Shepard, R.N., Kilpatrick, D.W., & Cunningham, J.P. (1975). The internal representation of
 778 numbers. *Cognitive Psychology*, 7, 82-138. [https://doi.org/10.1016/0010-0285\(75\)90006-7](https://doi.org/10.1016/0010-0285(75)90006-7)

779 Sokolowski, H. M., Hawes, Z., Peters, L., & Ansari, D. (2021). Symbols Are Special: An fMRI
 780 Adaptation Study of Symbolic, Nonsymbolic, and Non-Numerical Magnitude Processing
 781 in the Human Brain. *Cerebral Cortex Communications*, 2(3), tgab048.
 782 <https://doi.org/10.1093/texcom/tgab048>

783 Soltész, F., & Szűcs, D. (2014). Neural adaptation to non-symbolic number and visual shape: an
 784 electrophysiological study. *Biological psychology*, 103, 203–211.
 785 <https://doi.org/10.1016/j.biopsycho.2014.09.006>

786 Starrfelt, R., & Behrmann, M. (2011). Number reading in pure alexia--a review.
 787 *Neuropsychologia*, 49(9), 2283–2298.
 788 <https://doi.org/10.1016/j.neuropsychologia.2011.04.028>

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

792 Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual
793 system. *Nature*, 381(6582), 520–522. <https://doi.org/10.1038/381520a0>

794 Triesman, A. (1986). Features and objects in visual processing. *Scientific American*, 255, 114-
795 125.

796 Tyler, C. W., & Kaitz, M. (1977). Movement adaptation in the visual evoked response.
797 *Experimental Brain Research*, 27(2), 203–209. <https://doi.org/10.1007/BF00237698>

798 van Hoogmoed, A. H., & Kroesbergen, E. H. (2018). On the Difference Between Numerosity
799 Processing and Number Processing. *Frontiers in psychology*, 9, 1650.
800 <https://doi.org/10.3389/fpsyg.2018.01650>

801 Van Rinsveld, A., Guillaume, M., Kohler, P. J., Schiltz, C., Gevers, W., & Content, A. (2020).
802 The neural signature of numerosity by separating numerical and continuous magnitude
803 extraction in visual cortex with frequency-tagged EEG. *Proceedings of the National*
804 *Academy of Sciences of the United States of America*, 117(11), 5726–5732.
805 <https://doi.org/10.1073/pnas.1917849117>

806 Victor, J. D., & Zemon, V. (1985). The human visual evoked potential: analysis of components
807 due to elementary and complex aspects of form. *Vision research*, 25(12), 1829–1842.
808 [https://doi.org/10.1016/0042-6989\(85\)90006-9](https://doi.org/10.1016/0042-6989(85)90006-9)

809 Vogel, S. E., Goffin, C. & Ansari, D. (2015). Developmental specialization of the left parietal
810 cortex for the semantic representation of Arabic numerals: an fMR-adaptation
811 study. *Developmental Cognitive Neuroscience*, 12, 61–73.
812 <https://doi.org/10.1016/j.dcn.2014.12.001>

813 Volfart, A., Jonas, J., Maillard, L., Colnat-Coulbois, S. & Rossion, B. (2020).
814 Neurophysiological evidence for crossmodal (face-name) person-identity representation in

the human left ventral temporal cortex. *PLoS Biology*, 18(4): e3000659.

<https://doi.org/10.1371/journal.pbio.3000659>

Wilkey, E. D., Barone, J. C., Mazzocco, M. M. M., Vogel, S. E., & Price, G. R. (2017). The effect of visual parameters on neural activation during nonsymbolic number comparison and its relation to math competency. *NeuroImage*, 159 (August), 430–442. <http://doi.org/10.1016/j.neuroimage.2017.08.023>

Zhou, X., Chen, C., Dong, Q., Zhang, H., Zhou, R., Zhao, H., Chen, C., Qiao, S., Jiang, T., & Guo, Y. (2006). Event-related potentials of single-digit addition, subtraction, and multiplication. *Neuropsychologia*, 44(12), 2500–2507.

<https://doi.org/10.1016/j.neuropsychologia.2006.04.003>

826

Electronic supplementary material

827

For the article: *Identifying conceptual neural responses to symbolic numerals*

828

Authors: Talia L. Retter, Lucas Eraßmy & Christine Schiltz

829

Journal: Proceedings of the Royal Society B: Biological Sciences

830

Article DOI: <https://doi.org/10.1098/rspb>

831

832 Behavioral results

833 Performance at the fixation cross luminance change task (brief changes from dark gray to white)
834 was performed with high accuracy ($M = 96.4\%$; $SE = 1.14\%$; range = 84.7% to 100%) and fast
835 response times (RT; $M = 476$ ms; $SE = 13.9$ ms; range = 408 to 594 ms) overall, suggesting that
836 participants paid attention to the screen and understood the task instructions. For both accuracy
837 and RT, there was a main effect of Stimulus Set (accuracy: $F_{3,42} = 5.07$, $p = .004$, $\eta_p^2 = 0.27$; RT:
838 $F_{1.8,25.4} = 44.1$, $p < .001$, $\eta_p^2 = 0.76$). There was not a main effect of Condition or an interaction
839 (accuracy: F 's < 2.2 , p 's $> .14$, η_p^2 's < 0.14 ; RT: F 's < 0.97 , p 's $> .41$, η_p^2 's < 0.07). Performance
840 was lowest for the 20 drawn stimulus set (accuracy: $M = 90.0\%$; $SE = 3.8\%$; RT: $M = 552$ ms; SE
841 $= 21.4$ ms), with all other means over 98.0% for accuracy and ranging from 453 to 457 ms for RT.
842 Worse performance with the 20 drawn stimulus set is likely due to the distracting complexity and
843 color, including white, of this stimulus set.

844

845

	Parity	Small Phys.- Diff. control	Large Phys.- Diff. control
Non-background image space	1.35	1.67	2.20
Perimetric complexity	2.90	3.66	37.7
Image energy	0	0	1.5
High spatial frequency	0.024	0.25	0.65
Physical similarity	0.11	1.04	1.39

846 *Table S1. Image statistics for the 1 font stimulus set: values represent the differences between*
847 *numeral groups. Therefore, the higher the values here, the more that visual property might be*
848 *expected to contribute to higher EEG asymmetry responses at 3.75 Hz. Non-background image*
849 *space represents a measure of “retinal size”. Image energy was defined as the lowest frequency*
850 *(cycles/image) that contains 80% of image energy. High spatial frequency was defined as the*
851 *percent of image energy above 10 cycles/image. Physical similarity is the ratio of overlapping to*
852 *different line segments (as in [Cohen, 2009](#)).*

853

	Parity	Small Phys.- Diff. control	Large Phys.- Diff. control
1 font	14.4	12.8	32.1
10 fonts	8.70	4.94	13.3
10 mixed	7.94	2.74	8.12
20 drawn	21.2	8.51	9.11

854 *Table S2. Z-scores for asymmetry responses, harmonic-summed at the occipito-temporal ROI (Z*
855 *> 1.64 corresponds to $p < .05$, 1-tailed; $Z > 2.32$ corresponds to $p < .01$, 1-tailed).*

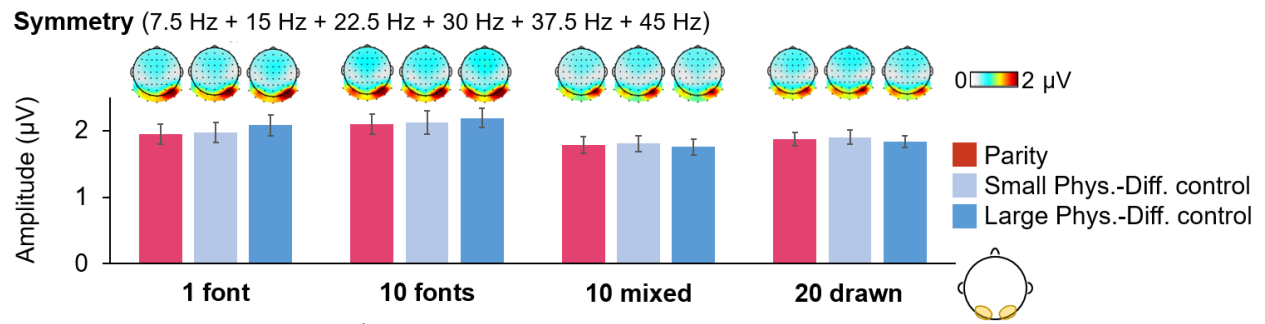
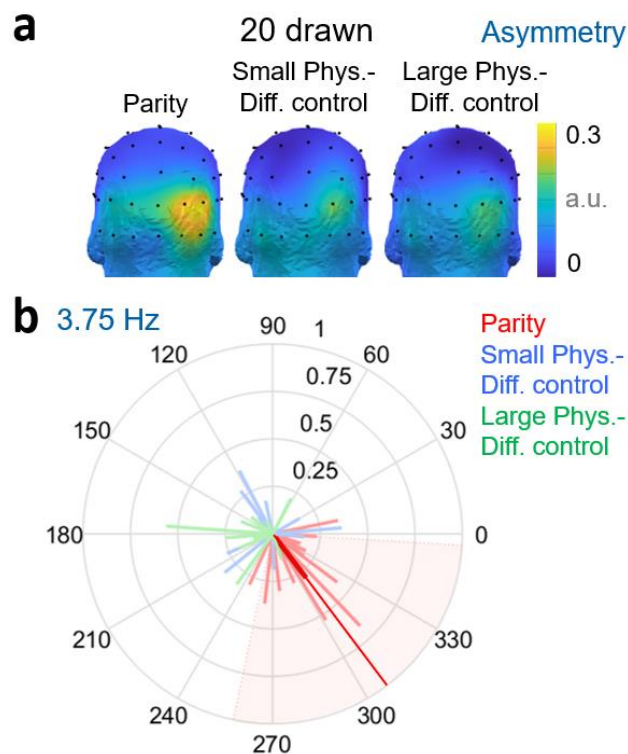


Figure S1. Symmetry responses at 7.5 Hz and harmonics (notations as in Fig. 3).



860

861 **Figure S2. Individually normalized topographies and phase for the 20 drawn stimulus set. a)**

862 *Individuals' topographies were normalized before being averaged together: these head plots thus*

863 *indicate the reliability of spatial response distributions across participants, rather than absolute*

864 *amplitude differences across conditions. b) Vectors represent individual participant data, with the*

865 *amplitude defining the length of the vector, and phase at 3.75 Hz its angle. Only the Parity*

866 *condition was reliable enough across participants to calculate a mean value (thick dark line, with*

867 *extension to the edge of the graph for simpler value estimation) with a 95% confidence interval*

868 *(indicated with light red shading).*