Biological Psychology 149 (2020) 107782

Contents lists available at ScienceDirect

Biological Psychology

journal homepage: www.elsevier.com/locate/biopsycho

Factoring in the spatial effects of symbolic number representation

Andrey R. Nikolaev[a,1], Ann-Kathrin Beck[b,1], Steffen Theobald[a,b,1], Thomas Lachmann[a,b,c,*], Cees van Leeuwen[a,b]

[a] Laboratory for Perceptual Dynamics, Brain & Cognition Research Unit, KU Leuven - University of Leuven, Leuven, Belgium
[b] Center for Cognitive Science, Cognitive and Developmental Psychology Unit, University of Kaiserslautern, Germany
[c] Facultad de Lenguas y Educación, Universidad Nebrija, Madrid, Spain

ARTICLE INFO

Keywords:
Mental number line
EEG
Lateralization
Analytic-holistic
Healthy adults

ABSTRACT

Spatial constituents of adult symbolic number representation produce effects of size-value congruity, Spatial Numerical Association of Response Codes (SNARC), and numerical distance. According to behavioral experiments, these effects belong to distinct processing stages. Yet, these effects evoke overlapping responses in both early and late Event Related Potentials (ERPs). To probe whether these overlaps indicate sharing of resources, all relevant stimulus and response conditions were factorially combined in a numerical value comparison task. To secure ERP validity, same numbers were compared against variable reference values. This design resulted in previously unobserved interactions in behavior but inhibited late ERP effects. All effects arose early in the P1 component (around 100 ms) and most showed hemispheric specificity. Independency of congruity and SNARC effects was observed, whereas SNARC and numerical distance were closely intertwined. Differences in hemispheric specificity, rather than stage-wise separation, were key to independence.

1. Introduction

The acquisition of numerical symbols is a major feat of abstraction in human development. The ability to visually distinguish and manipulate multiple objects in space precedes the act of counting in childhood. Yet, human development does not fully detach numerical symbols from their sensorimotor origins (Fischer, 2018). Spatial associations still play a role in representing numerical magnitude in adulthood (Fias & Fischer, 2005; Moyer & Landauer, 1967; Restle, 1970). These associations are held responsible for a range of effects in symbolic number representation experiments. Among these effects, probably the three most frequently studied ones are: size-value congruity (Besner & Coltheart, 1979; Henik & Tzelgov, 1982), Spatial Numerical Association of Response Codes (SNARC) compatibility (Dehaene, Bossini, & Giraux, 1993), and numerical distance (Moyer & Landauer, 1967). All of these effects involve distinct aspects of spatial representation but potentially tap into common resources. To investigate what these effects have in common, our study combines them in a single experiment, using behavioral and Event-Related Potential (ERP) measures.

1.1. Congruity

As more objects normally occupy larger amounts of space, numerical magnitude may show associations with spatial size, leading to the *size-value congruity effect*. This effect typically occurs in speeded choice reaction time experiments when single numbers are judged against a reference number. The task may involve judging numbers explicitly according to their numerical magnitude (Schwarz & Heinze, 1998; Schwarz & Ischebeck, 2003), physical size (Besner & Coltheart, 1979; Henik & Tzelgov, 1982), or leaving these properties implicit, as in parity judgment – “odd or even” (Viarouge, Hubbard, & McCandliss, 2014). The typical result consists of faster and/or more accurate responses to small numbers presented in small font size or large ones in large font (congruent), as compared to small numbers in large font or large ones in small font (incongruent; Banks & Flora, 1977; Besner & Coltheart, 1979; Henik & Tzelgov, 1982; Paivio, 1975; Schwarz & Heinze, 1998; Schwarz & Ischebeck, 2003).

The congruity effect is often called a numerical Stroop effect (Stroop, 1935) as, akin to this effect, size-value congruity engenders interference of a task-irrelevant dimension with a task-relevant one (or between two task-irrelevant dimensions, as in the parity task). There is

* Corresponding author at: Center for Cognitive Science, Cognitive and Developmental Psychology Unit, University of Kaiserslautern, PF3049, Kaiserslautern, 67653, Germany.
E-mail address: lachmann@rhrk.uni-kl.de (T. Lachmann).
[1] Authors contributed equally (shared first authorship).

https://doi.org/10.1016/j.biopsycho.2019.107782
Received 7 February 2019; Received in revised form 8 July 2019; Accepted 8 October 2019
Available online 13 October 2019

an ongoing debate about whether the effect arises in visual encoding, where numerical magnitude and physical size might be jointly represented, and/or in the decision-making stage (Santens & Verguts, 2011). Generally speaking, decision-making has perceptual, cognitive and motor components (Jarvstad, Hahn, Rushton, & Warren, 2013). We may wonder, which of these are involved in congruity, as well as in other effects of spatial number representation. Addressing such questions experimentally benefits from high-temporal resolution measures, in particular ERPs.

ERP studies revealed that congruity affects both early and late ERP components, such as N1, P3 and the lateralized readiness potential (LRP; Kadosh, Kadosh, Kaas, Henik, & Goebel, 2007; Schwarz & Heinze, 1998; Soltész, Goswami, White, & Szűcs, 2011; Szűcs & Soltész, 2007). Size-value congruity thus seems to occur at multiple processing stages and involve a range of processes, such as contextual analysis, conflict monitoring and action selection (Szűcs & Soltész, 2007).

1.2. SNARC

A second spatial feature associated with numerical representation involves the ordering of numbers (Fischer, 2018). Small numbers are preferably associated with the left side and large numbers with the right side (at least in Western cultures). Typically, responses to small numbers using the left hand are compatible and thus facilitated, while those using the right hand are incompatible and thus inhibited, and vice versa for large numbers. This effect, known as the *Spatial Numerical Association of Response Codes (SNARC)* effect, was observed in magnitude as well as in parity judgment tasks (Dehaene et al., 1993). The implicit character of the latter suggests that the effect might reflect an intrinsic spatial property of number representation, such as a mental number line (Dehaene et al., 1993). SNARC effects have been found in children (Van Galen & Reitsma, 2008) and, with non-symbolic number presentations, even in infants (Bulf, de Hevia, & Cassia, 2016).

At the very least, the SNARC effect is well entrenched: it is independent of stimulus presentation format (symbolic vs. non-symbolic, e.g., dot patterns; Cohen Kadosh et al., 2011), and modality (e.g., auditory instead of visual; Nuerk, Wood, & Willmes, 2005; Weis, Estner, & Lachmann, 2016; Weis, Estner, van Leeuwen, & Lachmann, 2016). Moreover, SNARC-like effects are quite widespread, as they can also be found with symbolic magnitudes other than numbers, for instance with pitch (Prpic et al., 2016; Weis, Estner, Lachmann et al., 2016; Weis, Estner, van Leeuwen et al., 2016), non-symbolic and non-conceptual magnitudes such as luminance (Fumarola et al., 2014; Ren, Nicholls, Ma, & Chen, 2011), real and illusory geometrical magnitudes (Fumarola et al., 2016; Prpic et al., 2018) and conceptual magnitude (e.g., the size of animals; Shaki, Petrusic, & Leth-Steensen, 2012). The pervasive character of these effects suggests that the left-right ordering may be intrinsic to magnitude.

Then again, the SNARC effect is opposite in cultures where numbers are read from right to left (Shaki, Fischer, & Petrusic, 2009). SNARC can be overturned by reversing the order of numbers in the working memory representation (Abrahamse, van Dijck, Majerus, & Fias, 2014) or by training participants to represent numbers as time on an analog clock face (Bächtold, Baumüller, & Brugger, 1998), suggesting that the effect is mediated by the present memory representation.

The SNARC effect is commonly understood as being akin to stimulus-response (SR) compatibility, which is well known from previous studies of two-handed responses to lateralized stimuli in perceptual-motor tasks (starting from Fitts & Seeger, 1953). A difference with this classical effect is that with numbers, compatibility can be implicit as demonstrated by its occurrence in the parity task (Dehaene et al., 1993). A related issue is that in the classical SR effect, the role of the stimulus could be understood in terms of physical laterality, whereas with numbers the laterality can be a matter of their contextualized representation. For this reason, we will adopt the neutral term "response compatibility".

ERP studies initially found SNARC-related moderation of LRP and consequently associated the SNARC effect with action selection (Szűcs & Soltész, 2007) or response preparation (Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006; Keus & Schwarz, 2005) stages. However, a more recent study has revealed SNARC effects also on earlier, stimulus-related components: N1, N2, and P3 (Gut, Szumska, Wasilewska, & Jaśkowski, 2012). These effects were interpreted as evidence for digit magnitude categorization, numerical conflict detection, and monitoring. According to current *ideomotor theory* (Hommel, Müsseler, Aschersleben, & Prinz, 2001), perception shares a common code with motor behavior, such that actions are represented by their perceivable effects (Shin, Proctor, & Capaldi, 2010). Therefore, it is possible that the early effects of SNARC involve the implicit activation of anticipated motor responses, which may affect even the earliest perceptual components of ERP (Nikolaev, Ziessler, Dimova, & van Leeuwen, 2008). The SNARC effect may also be manifested in preparatory lateralization of executive attention, as suggested by the hemispheric specificity of SNARC effects on P3 (Gut et al., 2012). In sum, as with the congruity effect, numerous ERP components may be sensitive to the SNARC effect, suggesting that it involves both early and late processes.

1.3. Numerical distance

We will consider a third effect involving sensorimotor components of symbolic number representation: the *numerical distance effect* (Moyer & Landauer, 1967). Typically, this effect is observed in magnitude judgment tasks; the closer in magnitude a number is to the reference, the more time participants need to determine the correct answer. At least for symbolic numbers in their common Arabic representation, this effect cannot be ascribed to similarity in shape, as "4" and "5" differ in shape as much as "4" and "6" do.

Most explanations consider the numerical distance effect to result from noise in a continuous number representation, such that numbers are more difficult to distinguish if their difference is small. In principle, the representational continuum could be some abstract magnitude currency (Walsh, 2003). However, in the context of the other effects discussed here, a more likely substrate for this noisy continuum would be spatial representation of numerosity, such as the mental number line assumed for the SNARC effect (Dehaene et al., 1993).

Krajcsi and Kojouharova (2017) provide an intriguing alternative explanation, in which the distance effect involves the categorical representations "large" and "small". The former has associations to numbers, of which the strengths are proportional to their numerical values; they are inversely proportional for the latter. Hence, the larger the numerical distance between two digits, the more distinct their categorical associations. Krajcsi and Kojouharova (2017) showed that learning of contrasting categorical associations modifies the distance effect, which puts the authors in the same camp as the opponents of intrinsic spatial representation of the SNARC effect.

In our view, Krajcsi and Kojouharova's (2017) explanation offers a new perspective on the numerical distance effect as a form of response compatibility. Both 5 and 6 are "small" compared to a reference number 7. However, 5 is more strongly associated with the "small" category than 6, which is more strongly associated with "large". Hence, for 5 there is compatibility between categorical representation and the response, but incompatibility for 6. The opposite would be the case when 5 and 6 are compared against a reference 4. Thus, response compatibility mechanisms might underlie the numerical distance effect, as they do with the SNARC effect.

In ERP studies, the numerical distance effect is most prominent over the parietal areas between 124 and 300 ms (Avancini, Galfano, & Szűcs, 2014; Dehaene, 1996; Libertus, Woldorff, & Brannon, 2007; Schwarz & Heinze, 1998; Soltész et al., 2011; Szűcs & Csépe, 2005; Szűcs & Soltész, 2007; Temple & Posner, 1998). In this interval, numerical distance moderates a number of ERP components, including N1 (Merkley, Shimi, & Scerif, 2016; Temple & Posner, 1998) as well as its frontal counterpart (Jiang et al., 2010). Curran, Tucker, Kutas, and Posner (1993) mention a second posterior positivity (210–250 ms latency), commonly denominated as P2p, that reprises the P1 at the same sites. Larger P2p amplitude was observed for close than for far numerical distances (Dehaene, 1996; Jiang et al., 2010; Libertus et al., 2007; Temple & Posner, 1998). Several studies also reported that the distance effect on ERPs may continue after 300 ms and affect the amplitude and latency of the following P3 component (Avancini et al., 2014; Dehaene, 1996; Grune, Mecklinger, & Ullsperger, 1993; Jiang et al., 2010; Libertus et al., 2007; Schwarz & Heinze, 1998). Thus, as with congruity and SNARC effects, both early and late ERP components are sensitive to numerical distance.

1.4. Three effects sharing common resources?

Even though the spatial features invoked in size-value congruity, SNARC, and numerical distance effects are distinct, they may depend on shared processing resources. In behavioral studies, this issue has traditionally been addressed with a factorial design. Any observed interactions are understood to indicate overlapping processing stages; absence of interactions is taken to indicate that factors belong to distinct processing stages (for a recent review, see Sternberg, 2018). For example, congruity and SNARC factors were combined in parity (Fitousi, Shaki, & Algom, 2009) and numerical magnitude (Weis, Theobald, Schmitt, van Leeuwen, & Lachmann, 2018) judgment tasks. Both studies showed that these effects occur independently (i.e., no interaction between congruity and SNARC factors). The authors concluded that the size-value congruity effect in behavior arises early, presumably in a perceptual decision stage (Santens & Verguts, 2011; c.f. contextual analysis; Szűcs & Soltész, 2007), while SNARC takes effect in a later stage.

On the other hand, behavioral studies have shown a correlation between SNARC and numerical distance effect sizes in individuals (Viarouge et al., 2014), and interactions between the SNARC effect and *physical* distance in the peripersonal space (Lohmann, Schroeder, Nuerk, Plewnia, & Butz, 2018). These observations would suggest that numerical distance and SNARC tap into the same resources, in accordance with Krajcsi and Kojouharova´s (2017) account of numerical distance effect in terms of response compatibility. The issue of dependency of numerical distance and SNARC, therefore, remains open. We are not aware of any behavioral studies which addressed this issue in a factorial design.

The sparseness of interactions in behavioral studies is inconsistent, at least on the face of it, with ERP studies showing that numerical concepts invoke spatial associations at multiple processing stages, affecting a range of processes. We have seen that congruity, SNARC, and numerical distance all affect both early and late ERP components. According to the common extension of Donders' subtraction method (Donders, 1868) to neuroimaging and psychophysiology, interactions in factorial designs reveal shared resources also in ERP. However, this would assume that ERP generation is a linear process, which is questionable (Alexander, Trengove, & van Leeuwen, 2015). Instead of additivity we may encounter moderation, which may give rise to false alarms and misses in linear analyses (Cohen, 2014). Thus, interactions in ERPs are neither sufficient nor necessary for observing dependencies between processes (Loizides, Achilleos, Iannetti, & Mitsis, 2015). Therefore, it is unsurprising that ERP studies have been inconclusive about interactions. For example, Soltész et al. (2011) observed an interaction between numerical distance and congruity effects in children, but failed to find one in adults.

Despite these complications, it may still be worthwhile to apply a factorial design in combination with ERP measurements. ERPs (besides their counterparts in magnetoencephalography) are the only measures having sufficient temporal resolution to study the time course of processing with high precision. False alarms and misses may be countered in ERPs on a case-by-case basis, e.g. in light of brain signals' propagation direction across the scalp and the delays in arrival this may cause in certain components across locations. None of the previous ERP studies tested the stimulus and response factors that evoke size-value congruity, SNARC compatibility, and numerical distance-related responses in a single experiment. Thus we independently varied the following factors: numerical value, font size, response laterality, and numerical distance.

We will consider sequential ERP components known to be sensitive to these factors, such as N1, P2p, and P3. In addition, we consider the P1 component, which is indicative of early perceptual processes related to the physical characteristics of stimuli, such as their size (Busch, Debener, Kranczioch, Engel, & Herrmann, 2004). Generally, since ERP components are highly sensitive to differences in physical stimulus features, we apply a fully balanced design, using comparisons between physically identical stimuli wherever possible.

Stimuli will be presented foveally, so that initial visual processing takes place in both hemispheres. Number processing may be differentiated between the two hemispheres, as has been shown for SNARC (Gut et al., 2012) and numerical distance (Avancini et al., 2014) effects. Lateralized responses will be performed with right and left hands. Hand responses are preferentially initiated in the contralateral hemisphere (Aziz-Zadeh, Iacoboni, & Zaidel, 2006). For these reasons, the left and right hemispheres are included as a separate factor in our analyses.

2. Material and method

2.1. Participants

Twenty-six healthy adults (five males, median age = 21 years, age range: 18–28 years) were recruited via the cloud-based participant management software “Sona-Systems” in the Faculty of Psychology and Educational Sciences at KU Leuven. All participants were right-handed, as indexed by a handedness inventory (Oldfield, 1971). All reported normal or corrected-to-normal vision. Six participants were excluded from further analyses: three because of failure to follow task instructions, one due to technical issues during the EEG recording, one due to more than 33% of trials removed in artifact rejection, and one due to less than 30 trials per condition remaining after artifact rejection. Thus, the data of 20 participants were included in the analysis. The study was conducted according to the guidelines of the Helsinki convention and approved by the departmental ethics committee of the KU Leuven. All participants gave their written consent after being informed about the procedure and were given the opportunity to ask questions and receive explanation about the experiment.

2.2. Stimuli

Stimuli were six digits (2, 3, 5, 6, 8, & 9) printed in grey against a black background. Stimuli were presented in the center of a 40 x 30 cm screen with a refresh rate of 75 Hz and a resolution of 1600 × 1200 pixels at a distance of 60 cm. The digits were of two font sizes: 1° and 3° of visual angle.

2.3. Procedure

The experiment was operated in Python 2.7 (Python Software Foundation) and took place in a dimly lit and shielded room. Participants were seated in front of a computer screen, placed at eye level and at arm's length on a table. Participants rested their index fingers on the Q and P keys of a QWERTY keyboard connected to the computer and responded by pressing one of the keys. The task involved numerical value comparison with a reference number. We used two sets of four numbers: 2, 3, 5, 6 in reference to 4, and 5, 6, 8, 9 in reference to 7. Both sets shared the numbers 5 and 6. In the first set, 5 and 6 were larger than reference number 4 whereas in the second set, 5 and 6 were smaller than reference number 7 (Fig. 1A). In addition, in the first set 5

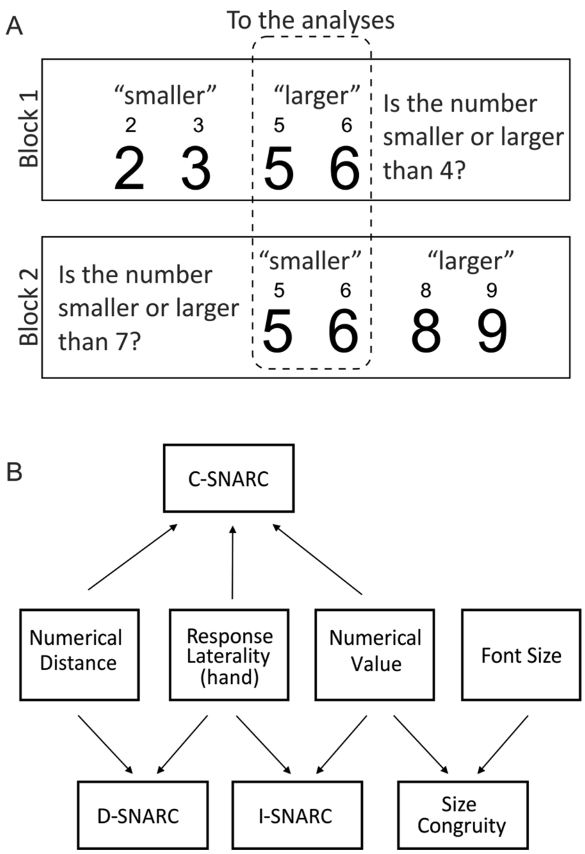

Fig. 1. A: Stimuli and task. B: Stimulus and response conditions and derived interactions.

was numerically closer than 6 to reference number 4, and in the second set, vice versa to reference 7. Thus, the numbers 5 and 6 in both sets were differently related to the reference in an entirely balanced manner.

The two sets of numbers were presented in separate blocks. Independently, in half of the blocks, numbers were responded to as small with the left hand and as large with the right hand; vice versa in the other half of the blocks. Each block was presented twice in random order. Within blocks, half of the numbers were randomly presented in small font (1°of visual angle) and in the other half in large font (3° of visual angle). Numerical value varied randomly within blocks, excluding repetitions of identical stimuli in subsequent trials. Half of the trials contained the numbers 5 or 6, both of which were repeated 50 times in each block. Eight blocks of 200 trials were presented in random order, yielding 1600 trials in total.

Prior to the start of each block, participants received an instruction for minimally 10 s, e.g.: "if the number is larger than 4, press P, otherwise press Q" in a font size of 0.5° of visual angle. They then received ten practice trials, in order to avoid an increase of response times due to switching of instruction. During the practice trials, feedback about correctness and response time was given for 2000 ms.

Each trial started with a fixation cross in the center of the screen presented for a random interval ranging from 1200 to 1500 ms. The stimulus was presented until participant's response with a timeout of 2000 ms. In the inter-trial interval, a black screen was presented for a fixed 1000 ms period. With 5-minute breaks between blocks, the experiment took approximately 2.5 h to complete.

2.4. EEG recording

The EEG was recorded using a 256 channel Geodesic Sensor Net (Electrical Geodesics Inc., Eugene, OR). The net included the electrodes for recording horizontal and vertical electrooculogram. The sampling frequency was 250 Hz. The vertex electrode was used as a reference. EEG signal was filtered online with a 0.1 Hz high-pass filter and a 100 Hz low-pass filter.

2.5. Analysis of behavioral data

Error rates and reaction times (RT) were analyzed for trials containing the numbers "5" and "6" only. Responses faster than the individual average minus two standard deviations or slower than plus two standard deviations were designated as outliers. On average 4.42% trials (SD = 1.04%) per participant and per condition were outliers; these trials were excluded from analyses. Only correct responses were analyzed for RT. The data were analyzed with R (R Core Team, 2013) and Python 2.7.

2.6. EEG analysis

EEG was analyzed for trials containing the numbers "5" and "6" only, again excluding outliers. For analyzing the EEG data, we used BrainVision Analyzer (Brain Products GmbH, Gilching, Germany) and Python 2.7. Electrodes on the neck and cheeks were excluded since they are liable to neck and face muscles artifacts. Finally, of the 256 recorded channels, 161 were used in the analysis. The EEG signal was filtered using a Butterworth zero phase filter with the low cutoff at .53 Hz and the high cutoff at 30 Hz at 48 dB/oct.

We corrected for eye movement artifacts in continuous EEG recording using independent component analysis (ICA) with the Infomax Restricted algorithm (Jung et al., 1998). We selected a 400 s interval arbitrarily from the fourth block of the experiment as a training dataset for computing the unmixing matrix. ICA components were automatically identified by picking up blinks and saccades, as evidenced by their characteristic shape and maximum at frontal sites. After removing these components, the EEG was reconstructed.

Subsequently, the data were segmented into epochs of 600 ms length, including a 100 ms prestimulus interval. To remove artifacts due to bad electrode contact and large muscle movement, we applied the following thresholds on EEG amplitude: the amplitude difference between 2 sample points should not exceed 50μV; the amplitude difference should not be more than 100μV in an interval of 100 ms. The epochs containing activity below 0.5μV in an interval of 100 ms were also removed. Epochs with bad intervals were excluded from the dataset. Channels having more than 3% bad epochs were removed. We removed on average 2.5 channels (SD = 3.2 channels) per participant. After artifact removal there were on average 44.52 (SD = 3.88; range: 35–49) epochs per condition per participant.

Consistently with the previous literature on numerosity judgment, we focused our analysis on the N1, P2p, and P3, and additionally on P1, which is sensitive to stimulus size (Busch et al., 2004). Consequently, we selected the regions of interest in the areas of their typical locations (Luck, 2005): over the left and right occipital areas for P1 and N1. P2p was also analyzed over these areas, corresponding to previous reports of its location as occipito-posterior (Dehaene, 1996; Jiang et al., 2010) and its prominence in the grand-averaged curves of our analysis. P3 was analyzed over the left and right parietal areas. We averaged signals of seven electrodes, around the landmark electrodes of the International 10–20 System of Electrode Placement: O1, O2, and P3, P4. Epochs were averaged for each participant and condition separately. The averaged epochs were baseline corrected by subtracting the mean amplitude in the interval −100 ms – 0 from each point of an epoch. Using visual inspection of the grand averaged waveforms, we extracted for four components, the mean amplitude in the intervals: P1: 90–140 ms, N1: 140–200 ms, P2p: 200–250 ms, and P3: 250–350 ms.

For statistical analyses, we used repeated-measures ANOVAs with the factors: numerical value (small: 5 and large: 6), font size (small and large), numerical distance to the reference number (close and far), response laterality (left and right hand) and, in the ERP analysis, hemisphere (left and right). For post hoc analyses, we used the Fisher's Least Significant Difference (LSD) test.

2.7. Derivation of conditions

The above-mentioned factorial designs display only numerical distance as a main effect; congruity and SNARC effects would appear as interactions (Fig. 1B). Interactions involving numerical value and font size designate size-value congruity: large values in large font and small values in small font constitute congruent conditions; small values in large font and large values in small font constitute incongruent conditions.

SNARC effects are designated as interactions of numerical value and hand. Numerical value 5 responded to with the left hand, and 6 with the right hand may thus count as SNARC compatible, while 5 responded to with the right hand, and 6 with the left hand count as incompatible. On this account, numerical value 5 qualifies as intrinsically small and 6 as intrinsically large, independently of how they were responded to. Therefore this SNARC effect will be called I(ntrinsic)-SNARC effect.

Notably, because the reference numbers were varied (4 vs. 7), half of the times 5 and 6 were responded to as "small", and half of the times as "large". A second SNARC effect may occur, if compatibility depends on response categories. In this case, responses classified as "small" would preferably be given with the left-hand and responses classified as "large" preferably with the right hand. We will therefore refer to this effect as C(ategorical)-SNARC.

In our ANOVAs, C-SNARC effects are designated by three factors: numerical value, numerical distance, and hand. The number 5 qualifies as small with the far reference 7 and as large with the close reference 4, whereas 6 qualifies as large with the far reference 4 and as small with the close reference 7. These designations determine the C-SNARC compatible and incompatible hand conditions. Note that the distinction between I-SNARC and C-SNARC allows us to determine whether the left-right ordering is based on intrinsic numerical value, current categorical assignment, or a mixture of both.

Finally, we may consider a SNARC effect designated by numerical distance and hand. We will call this a D(istance)-SNARC effect. D-SNARC means that close numerical distances (i.e. between 4 and 5 or between 6 and 7) are preferably responded to with the left hand, and far distances (between 5 and 7 and between 4 and 6) preferably with the right hand. These constitute the D-SNARC compatible conditions; the D-SNARC incompatible conditions are vice versa. A D-SNARC effect would support the view elaborated in the introduction that the numerical distance effect, like the SNARC effect, is based on response compatibility, and thus that both effects are interdependent.

3. Results

3.1. Behavioral results

On average, participants gave 97.45% (*SD* = 1.6%) correct responses. The ANOVA on error rates revealed a main effect of numerical value, $F(1, 19) = 6.5$, $p = .02$, $\eta^2 = .25$, with more errors for the larger number. Additionally, there was a main effect of numerical distance, $F(1, 19) = 31.8$, $p < .001$, $\eta^2 = .63$, with more errors for the smaller distance.

Average response time was 474.5 ms (*SD* = 72.8 ms). An ANOVA on response times showed a main effect of font size, $F(1, 19) = 20.8$, $p < .001$, $\eta^2 = .52$, with fast responses to the larger stimuli. Furthermore, there was a main effect of numerical distance, $F(1, 19) = 35.6$, $p < .001$, $\eta^2 = .65$, with faster reaction times for the larger numerical distance. Additionally, there was a triple interaction between numerical value, font size, and numerical distance, $F(1, 19) = 9.0$, $p = .007$, $\eta^2 = .32$ (see Fig. 2A). Post hoc tests revealed that responses were slowest when the large numerical value was presented in small font at a small numerical distance. In contrast, participants were fastest when the large numerical value was presented in a large font at a large numerical distance. This interaction could be represented as a crossover between numerical distance and size-value congruity (Fig. 2B), with a (very small) congruity effect for the far numerical distance, and a somewhat larger, *negative* effect of congruity for the near numerical distance. This interaction will be shown in the discussion to allow diagnosing the present numerical distance effect.

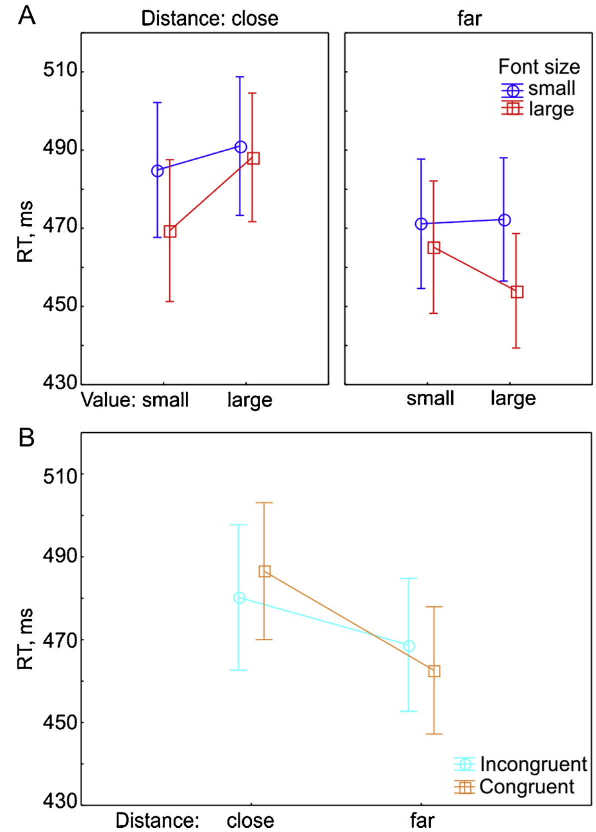

Fig. 2. RT results. A: Interaction between numerical value, font size, and numerical distance. B: Interaction between numerical distance and size-value congruity. Error bars indicate standard errors of the means across 20 participants.

I-SNARC, C-SNARC or D-SNARC effects would have been manifested in interactions involving the factor of response laterality (left-right hand). However, none of these effects were close to significance ($F < 1$).

3.2. EEG results

3.2.1. P1

An ANOVA on P1 amplitude revealed no main effects; numerical value still provided a tendency, $F(1, 19) = 4.1$, $p = .056$, $\eta^2 = .18$. An interaction between numerical value and hemisphere, $F(1, 19) = 12.6$, $p = .002$, $\eta^2 = .40$, revealed that higher amplitude for the large (6) rather than small (5) numerical value was more prominent over the right hemisphere ($p < .001$) than over the left ($p = .03$).

Note also that large and small font sizes differed in latency on all ERP components. This effect may easily be understood, e.g. in terms of stimulus discriminability. This illustrates the principled issue that comparing across non-identical stimuli is a potential source of confounding. Size-value congruity effects compare between different numbers (5 and 6) in different font sizes. Given that numerical value and font size are balanced over congruent and incongruent conditions, this does not threaten the validity of the congruity effect.

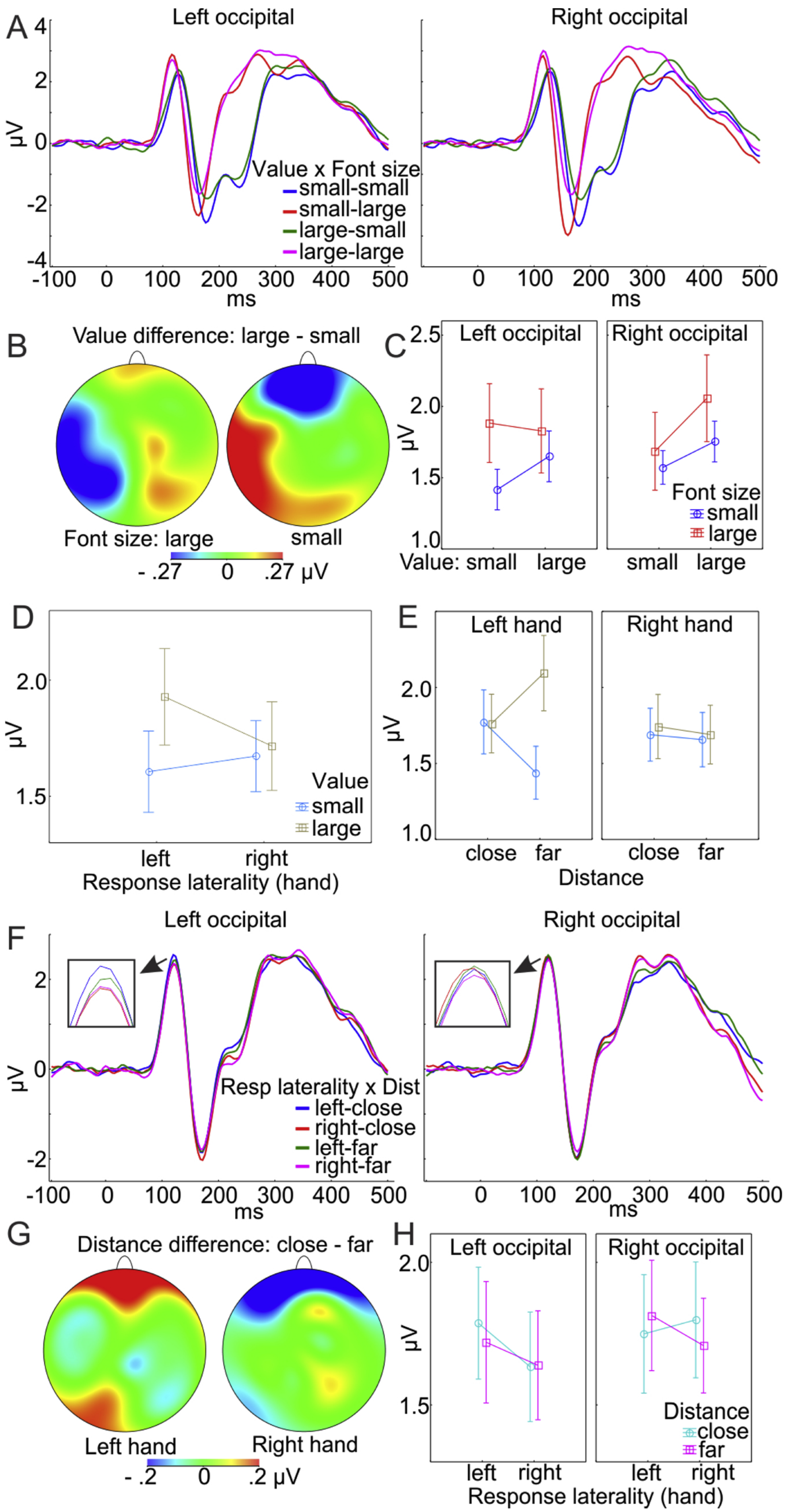

Fig. 3. P1 results. A-C: Interaction between hemisphere, numerical value, and font size. A: grand-averaged potentials for left and right occipital ROIs. B: difference maps (large minus small numerical value) for large and small font size at the P1 peak latency. C: mean amplitude in the interval 90–140 ms after stimulus onset. D: mean amplitude for the interaction between numerical value and response laterality (I-SNARC). E: mean amplitude for the interaction between numerical distance, numerical value and response laterality (C-SNARC). F-H: Interaction between hemisphere, numerical distance, and response laterality (D-SNARC). D: grand-averaged potentials for left and right occipital ROIs. E: difference maps (close minus far distance) for the left and right hands at P1 peak latency. F: mean amplitude in the 90–140 ms interval after stimulus onset. Error bars indicate standard errors of the means across 20 participants.

Fig. 3 shows the interactions obtained for P1 between the stimulus and response conditions of our design; Fig. 4 shows the same effects regraphed to visualize the effects involving congruity, SNARC, and distance effects. There was an interaction between hemisphere, numerical value, and font size, $F(1, 19) = 17.3$, $p < .001$, $\eta^2 = .48$. Fig. 3A–C shows that the previously observed effect of numerical value is moderated by both hemisphere and font size: In the left hemisphere, a larger amplitude was observed for large numbers in small font while the right

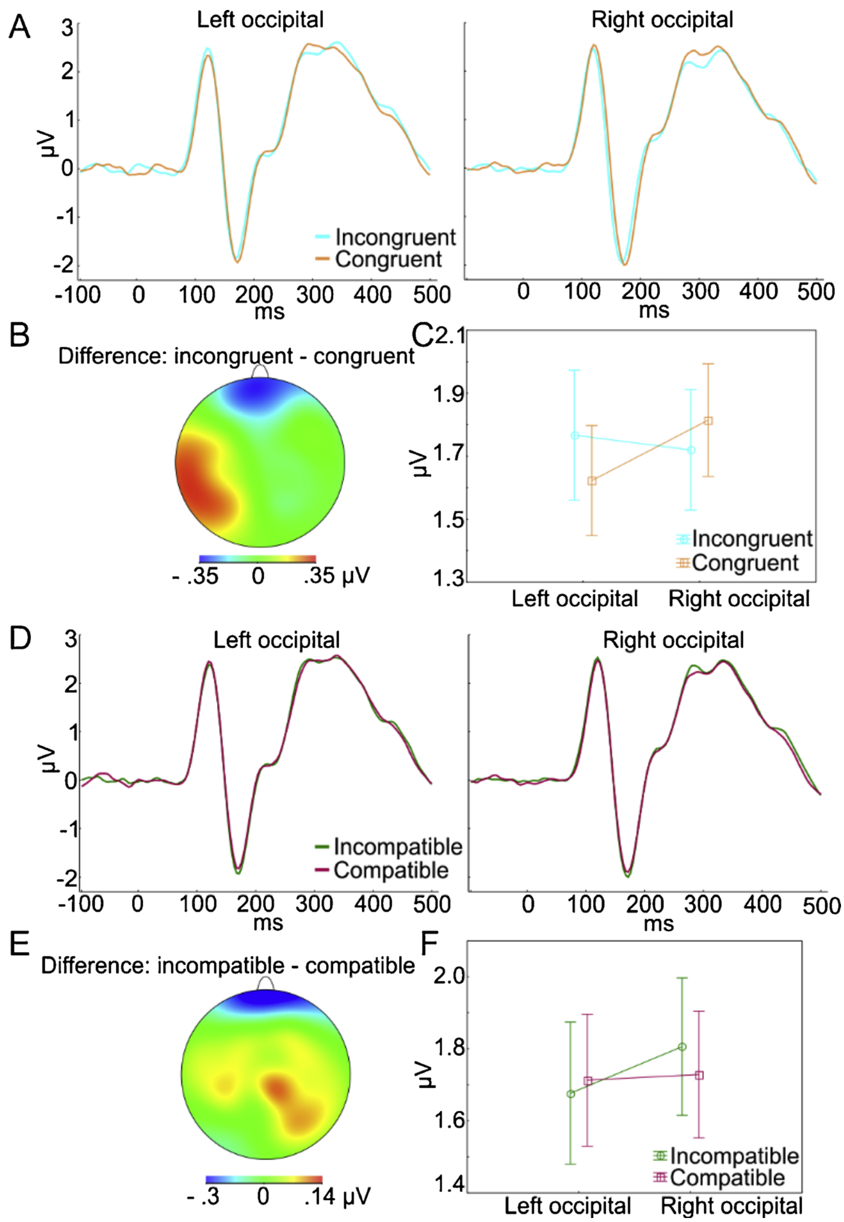

Fig. 4. P1 results. A-C: Interaction between hemisphere and congruity. A: the grand-averaged potentials for the left and right occipital ROIs. B: the difference map (incongruent minus congruent) at the P1 peak latency. C: the mean amplitude in the interval 90–140 ms after stimulus onset. D-F: Interaction between hemisphere and D-SNARC effect. D: grand-averaged potentials for left and right occipital ROIs. E: difference map (D-SNARC incompatible minus compatible) at P1 peak latency. F: mean amplitude in the interval 90–140 ms after stimulus onset. Error bars indicate standard errors of the means across 20 participants.

hemisphere showed a larger amplitude for large values in large font. As a size by value interaction designates size-value congruity, the result suggests opposing congruity effects between the hemispheres. However, post hoc analyses showed that the interaction was not significant over the right hemisphere ($p = .3$), but significant over the left hemisphere ($p = .002$; Fig. 4A–C). According to the interaction, incongruent stimuli are reflected in higher amplitude than congruent ones. We conclude that the effect in the left hemisphere is a negative congruity effect.

We obtained an interaction tendency of numerical value and response laterality, $F(1, 19) = 3.78$; $p = .067$; $\eta^2 = .17$. As shown in Fig. 3D, the opposing tendencies are in the direction expected for the I-SNARC effect. However, this effect only reached significance for the non-dominant hand ($p = .005$).

There was a marginally significant triple interaction of numerical value, distance, and response laterality, $F(1, 19) = 4.3$, $p = .051$, $\eta^2 = .19$ (Fig. 3E). The opposing tendencies in Fig. 3E are in the direction expected for the C-SNARC effect, but significance was restricted to the non-dominant hand, where higher amplitude was observed for large than for small values for the far numerical distance ($p < .001$).

As shown in Fig. 3F–H, there was an interaction between hemisphere, numerical distance and response laterality, $F(1, 19) = 8.4$, $p = .009$, $\eta^2 = .31$. The numerical distance effect is moderated by both hemisphere and response laterality. Whereas the left hemisphere shows a larger amplitude for the ipsilateral hand independently of numerical distance ($p = .2$), the right hemisphere shows a cross-over interaction of response laterality and numerical distance, $F(1, 19) = 8.4$, $p = .009$, $\eta^2 = .31$, in which the larger amplitude is reserved for the far-left and close-right conditions, as opposed to the close-left and far-right conditions. This effect could be understood as a right-hemispheric D-SNARC effect: close numerical distances are preferably associated with a left-hand response; far distances preferably with a right-hand response (Fig. 4D–F).

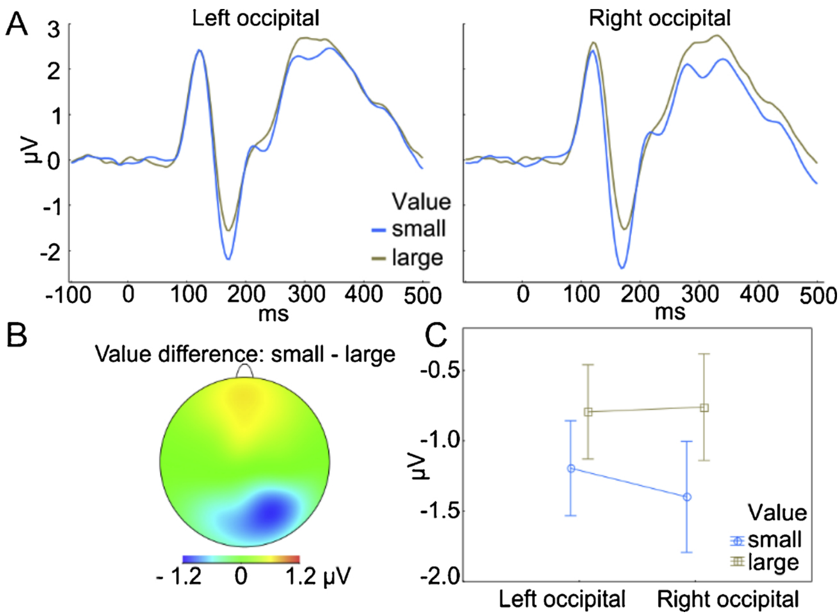

Fig. 5. N1 results. Interaction between hemisphere and absolute value. A: the grand-averaged potentials for the left and right occipital ROIs. B: the difference map (small minus large numerical value) at the N1 peak latency. C: mean amplitude in the interval 140–200 ms after stimulus onset. Error bars indicate standard errors of the means across 20 participants.

3.2.2. N1

An ANOVA revealed higher N1 amplitude for the small rather than large numerical value, $F(1, 19) = 38.1, p < .001, \eta^2 = .67$. This effect was more prominent over the right than left occipital areas, as was evident from the interaction between hemisphere and numerical value, $F(1, 19) = 8.2, p = .01, \eta^2 = .30$ (Fig. 5). The post hoc test showed that the difference between large and small numerical values was significant in both hemispheres (both $p < .001$). A marginally significant interaction of hemisphere and response laterality, $F(1,19) = 4.28, p = .053, \eta^2 = .18$, showed a slightly larger amplitude for the left hand over the left hemisphere only.

3.2.3. P2p

An ANOVA revealed a much higher P2p amplitude for the large than for the small font size, $F(1, 19) = 113, p < .001, \eta^2 = .86$, which was partially explained by the latency shift between these conditions (Fig. 6A). The interaction between hemisphere and response laterality, $F(1, 19) = 6.1, p = .02, \eta^2 = .24$, showed that the amplitude was higher for the left than for the right hand over the left hemisphere ($p = .003$), but there was no difference over the right hemisphere ($p = .9$). Furthermore, there was an interaction between hemisphere, distance, and font size, $F(1, 19) = 5.1, p = .04, \eta^2 = .21$ (Fig. 6). The post hoc test revealed a higher amplitude for the close than for the far distance, but only for large font size over the right hemisphere ($p < .001$).

3.2.4. P3

An ANOVA on P3 amplitude revealed no main effects (a marginally higher amplitude was observed over the right than left hemisphere, $F(1, 19) = 4.1, p = .058, \eta^2 = .18$, but two interactions (both of which also occurred for P1). The interaction between hemisphere and numerical value, $F(1, 19) = 16.5, p < .001, \eta^2 = .46$, showed that the amplitude was higher for large than small numerical value over the right ($p < .001$) but not over the left hemisphere ($p = .5$).

There was an interaction between hemisphere, numerical value, and font size, $F(1, 19) = 6.0, p = .02, \eta^2 = .24$. Fig. 7A–C shows higher amplitude for the large numerical value, independently of font size ($p = .3$), for the right hemisphere. For the left hemisphere, the graph shows a crossover interaction ($p = .02$), with higher amplitudes for

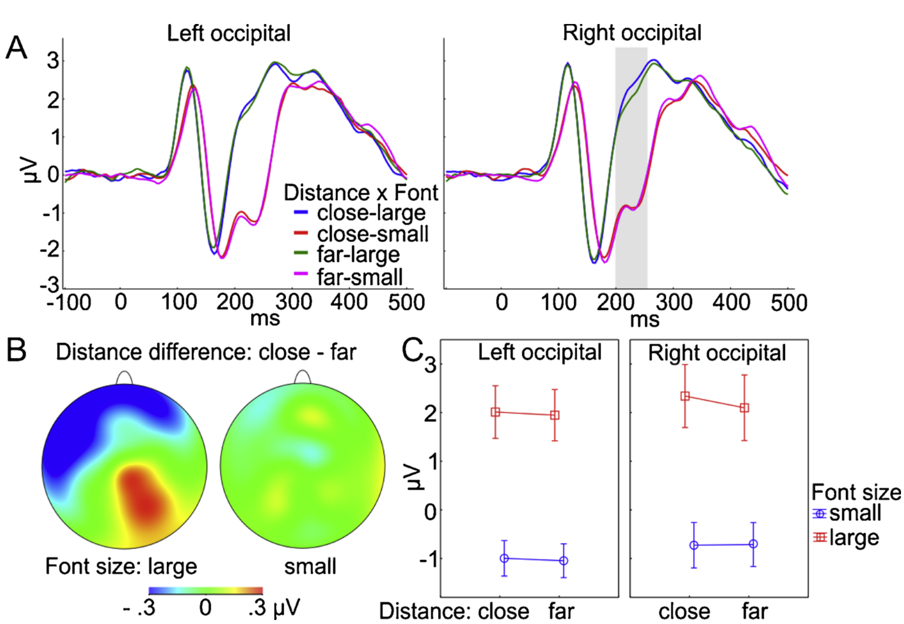

Fig. 6. P2p results. Interaction between hemisphere, distance, and font size. A: grand-averaged potentials for the left and right occipital ROIs. B: difference maps (close minus far numerical distance). C: mean amplitude in the interval 200–250 ms after stimulus onset marked by a grey bar in A. Error bars indicate standard errors of the means across 20 participants.

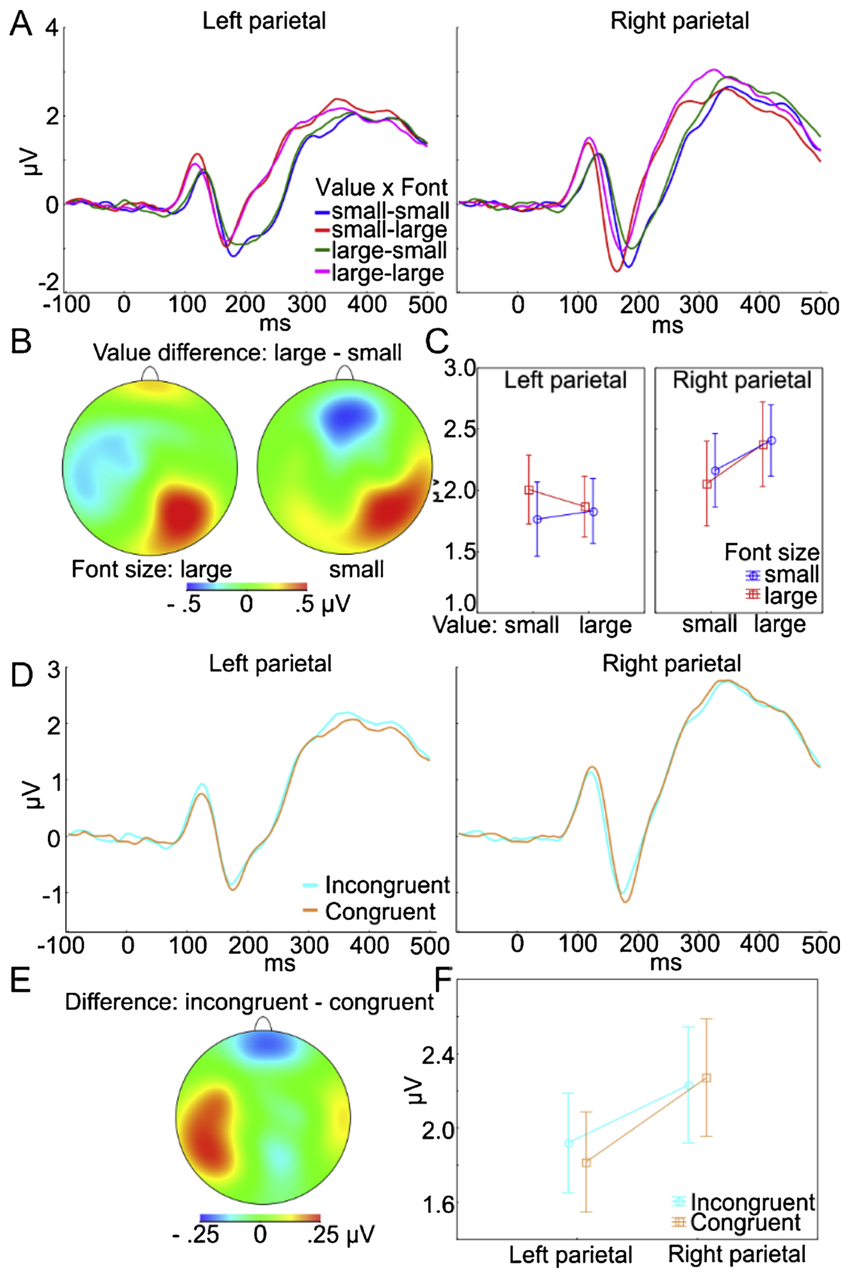

Fig. 7. P3 results. A–C: interaction between hemisphere, absolute value, and font. A: the grand-averaged potentials for the left and right occipital ROIs. B: difference maps (large minus small values) for large and small font in the time window of P3. C: mean amplitude in the interval 300–500 ms after stimulus onset. D–F: interaction between hemisphere and congruity. D: grand-averaged potentials for the left and right occipital ROIs. E: difference map (incongruent minus congruent) in the time window of P3. F: mean amplitude in the interval 300–500 ms after stimulus onset. Error bars indicate standard errors of the means across 20 participants.

incongruent conditions, in other words, a negative congruity effect. The same effect as over the left hemisphere in P1, therefore, returns in P3 (see Fig. 7D–F).

Table 1 shows a schematic overview of results (above-mentioned marginally significant ones are shown in parentheses). Behavioral results (error rates and RT) show robust effects of numerical distance. ERP results show a great deal of hemispheric specificity for all components. ERP effects and interactions occurred most frequently with numerical value and, to some extent, font size. A noticeable feature is the predominant involvement of P1. Effects and interactions of P3 to some degree mirror those of P1. Effects and interactions for N1 and P2p deviate from this pattern.

4. Discussion

We used a task involving numerical comparison to a variable reference, in order to study the combined effects of numerical distance, size-value congruity, and SNARC compatibility. Numerical distance was varied independently of numerical value, font size, and response laterality. Congruity effects were designated as interactions involving numerical value and font size. Varying response laterality (left or right hand) enabled us to study SNARC compatibility effects. We distinguished intrinsic, categorical and distance-SNARC (I-, C-, and D-SNARC, respectively). I-SNARC effects were considered as hand compatibility with numerical value (5 = small or 6 = large) irrespective of the reference (4 or 7); C-SNARC as hand compatibility with value ("large" or "small") relative to the reference; D-SNARC as hand compatibility with numerical distance from the reference (close = small or

Table 1
Schematic overview of all effects and interactions for error rates (Err), response time (RT), and for the amplitude of ERP components P1, N1, P2p, and P3.

Err	RT	P1	N1	P2p	P3
	S			S	
V		(V)	V		
					(H_r)
		H_r x V	H_r x V		H_r x V
			(H_l x L)	H_l x L	
		H_r x V x S = H_r x Cong			H_r x V x S = H_r x Cong
		(V x L = I-SNARC)			
D	D				
		(V x D x L = C-SNARC)			
		H x L x D = H_r x D-SNARC			
	V x D x S				
				H_r x D x S	

Note. Main effects and interactions observed in our experiment. H = hemisphere (left/right), H_r indicates an effect over the right hemisphere; H_l an effect over the left hemisphere; V = numeral value (5/6), S = font size, L = response laterality: (left/right hand), D = Numerical Distance (large/small), Cong = Size-value congruity (congruent/incongruent), I(ntrinsic)-SNARC, C(ategorial)-SNARC and D(istance)-SNARC (compatible/incompatible) effects are as differentiated in the main text. Marginally significant results are shown in parentheses.

far = large).

In the behavioral results, we observed main effects of numerical value and distance in the error rates, and of font size and numerical distance in the reaction times. The triple interaction in Fig. 2 between numerical value, font size, and numerical distance may appear surprising. Note, however, that Krajcsi and Kojouharova (2017) ascribed the numerical distance effect to differences in association strength between numerical values and the categories "large" and "small". For the close numerical distances, these associations may conflict with the response categories. Presumably, the number 5 is associated more strongly to the "small" category and 6 more to the "large" category. But 5 is responded to as "large" compared to its close reference 4, and 6 as "small" to its close reference 7. When numbers are far from the reference (6 to 4 and 5 to 7), the categorical association and the response are not in conflict. On the right in Fig. 2A, this condition shows a size-value congruity effect, with faster responses for congruent than for incongruent stimuli – discounting the overall preference for larger fonts. By contrast, the close numerical distance conditions in the panel on the left show a *negative* congruity effect. Here, congruity may have strengthened the conflict between the categorical associations of the numbers and their response categories.

No effects of SNARC were obtained in the behavioral data. Thus, previous observations of independency between SNARC and congruity (Fitousi et al., 2009; Weis et al., 2018) could not be substantiated. SNARC effects are typically predominant at the extremes of a numerical range rather than at the intermediate values (e.g., Figs. 2 and 3 in Gevers et al., 2006). Whereas the range extremes in our experiment varied between blocks, across the entire experiment they were 2 and 9, which qualifies our target numbers, 5 and 6, as intermediate. This may not have been optimally conducive to SNARC effects but was necessary for being able to study all effects in a single experiment.

Dependencies between congruity, numerical distance, and SNARC compatibility could, however, be substantiated by our ERP findings (Table 1). Most of these effects are concentrated on P1. P1 is generally associated with early visual perception and, specifically, with physical properties of visual stimuli (Luck, 2005). In particular, P1 amplitude increases with stimulus size (Busch et al., 2004). Here, P1 showed an earlier onset for digits in larger font (visible in Fig. 3A). Only in P2p did the difference in latency lead to a main effect of font size on amplitude. There were interactions involving font size for all components, except for N1.

Independently of stimulus size, we observed an effect of numerical value (5 vs. 6) on P1 in the right hemisphere. P1 (together with N1) has been found sensitive to spatial attention (Hillyard & Anllo-Vento, 1998). For physically balanced stimuli, this sensitivity could reflect their spatial representation. Thus, P1 is not unlikely to reflect spatial representation of numerical value. Effects of numerical value occur in all ERP components except P2p and are predominantly right-hemispheric, but lead to a main effect only in N1.

There was a hemisphere-specific interaction of font size and numerical value in P1, with a congruity tendency over the right hemisphere but inversely, a *negative* congruity effect over the left hemisphere. A later negative congruity effect over the left hemisphere on P3 "echoes" that of P1. The persistence of this effect across earlier and later ERP components, and its presence in the dominant left hemisphere, may explain why congruity, in contrast to SNARC, has a moderating effect in the behavioral data.

The opposing congruity effects between the hemispheres may be understood in relation to hemispheric specialization for, respectively, more holistic processing of visual stimuli in the right hemisphere versus more analytic processing in the left hemisphere (Van Kleeck, 1989). Hemisphere specific, Stroop-like interference of holistic and analytic features has been reported in the literature (Alivisatos & Wilding, 1982).

A right-hemispheric congruity effect suggests that, similar to the Stroop effect, early processing integrates holistically the relevant stimulus dimension (numerical magnitude) and the irrelevant one (physical size). The left-hemispheric negative congruity effect may then indicate a process Szűcs and Soltész (2007) identified as contextual analysis. Analysis of the target dimension may require suppressing the irrelevant one. This is harder to achieve if both stimuli are congruent, and thus a negative congruity effect results.

In behavioral studies, reversals of Stroop-like effects have generally been found as a function of stimulus type, rather than hemisphere, e.g. in the perception of shapes of different complexity (Bavelier, Deruelle, & Proksch, 2000; Briand, 1994; Van Leeuwen & Bakker, 1995) or between letters and shapes (Van Leeuwen & Lachmann, 2004).

We observed SNARC effects involving greater amplitude for incompatible conditions in P1. This may be a matter of early conflict resolution, for which the response-incompatible condition receives extra spatial attention. Sensitivity to a response conflict in such an early stage should not be surprising. According to *ideomotor theory* (Hommel et al., 2001), early activation of a percept may facilitate later motor responses through their common code.

Of the SNARC effects observed, two were only marginally significant. Firstly, the I-SNARC effect: 5 is compatible with the left hand and 6 with the right hand. This effect arises when spatial ordering is an intrinsic property of number representation (Dehaene et al., 1993). Secondly, the C-SNARC effect takes into account the response categories "large" or "small", which depend on the reference value. Between blocks, the reference varied, such that in one case, both 5 and 6 qualified as "small" and thus as left-compatible, and in the other as "large", hence right-compatible. The C-SNARC effect implies that spatial ordering of numbers is context-dependent (Abrahamse et al., 2014; Bächtold et al., 1998). Assuming both I- and C-SNARC effects were obtained, a mixture of intrinsic and context-specific preferences may characterize the spatial ordering of numbers. This would explain why the proponents of intrinsic and context-specific SNARC effects both claim positive results.

Both effects were weak and restricted to the non-dominant hand. This might suggest that both effects are in competition. Alternatively, the reason for the weakness of these effects may be the same as for their absence in behavior, i.e., the target numbers 5 and 6 were intermediate in the range of numbers used in the experiment.

By contrast, a strong D-SNARC-compatibility was found on P1. Large numerical distances showed greater amplitude than small ones for the left hand, and the opposite for the right hand. We may think of numerical distance as represented in the same way as numerical value: ordered from small to large as from left to right. Thus, like the other SNARC effects before, the D-SNARC effect on P1 suggests early conflict resolution. This was restricted to the right hemisphere. Thus, it seems that conflict resolution took place by allocating processing resources to the holistic characteristics of number representation, such as its position on the mental number line. This account would be consistent with an interpretation of the distance effect involving the mental number line. Earlier, in our account of the behavioral effects, we had invoked Krajcsi and Kojouharova's (2017) concept of categorical representation for the distance effect. However, spatial and categorical representations are not mutually exclusive; we may encounter the former, in particular, in the early processing stages around P1.

Previous studies reported SNARC effects at various processing stages, particularly in the P3 component, (Avancini et al., 2014; Dehaene, 1996; Grune et al., 1993; Jiang et al., 2010; Libertus et al., 2007; Schwarz & Heinze, 1998). In our experiment by contrast, these effects were restricted to P1. The reason for their absence in later components may again be the same as for their absence in behavior: the intermediate character of the numbers 5 and 6 within the entire range of numbers in the experiment. This might play a smaller role in the earlier, perceptual stages, whereas the later ERP components may be more sensitive to the broader context, i.e., the entire range. Thus, the difference in sensitivity between early and late ERP components is consistent with the notion that the former operate in a narrower context than the latter (Boenke, Ohl, Nikolaev, Lachmann, & van Leeuwen, 2009).

In contrast with behavior, no main effect of numerical distance was observed in the ERP data. However, the D-SNARC effect may be regarded as a mixture of SNARC and numerical distance effects. Numerical distance did, moreover, modulate the P2p component. Generally, the P2p reprises the P1 effect (Curran, Tucker, Kutas, & Posner, 1993). In line with previous observations, P2p amplitude over the right hemisphere was higher for the close than for the far distance (Dehaene, 1996; Jiang et al., 2010; Libertus et al., 2007; Temple & Posner, 1998). The right-hemisphere specificity of this effect accords to that of P1. In contrast with P1, this effect was restricted to large font sizes.

Size-value congruity and numerical distance effects showed an interaction in behavior that is not reflected in any of the ERPs analyzed. We explained this effect, based on the categorical associations of the numerical values. Failure to observe corresponding interactions in ERP may be related to the non-identical size of the stimuli compared across congruity conditions. In Fig. 3, this is shown to give rise to differences in ERP latencies that may obscure effects on amplitude. Consequently, P2p is the only ERP component that showed a main effect of size on amplitude. The response laterality effects on N1 and P2p were restricted to the left hemisphere. The left hemisphere is also the locus of the congruity effect on P1. Therefore, it is possible that our analyses missed an interaction of congruity and numerical distance, presumably over the left hemisphere in P2p. The latency shift, as a function of font size, might have obscured this effect.

Whereas effects of congruity on P1 were concentrated mainly on the left-hemisphere, those of SNARC compatibility were centered on the right-hemisphere. Consequently, they do not interact in this stage of processing and should be considered independent. This view is in accordance with behavioral data (Fitousi et al., 2009; Weis et al., 2018). As both effects coincide in time, their independency cannot be a matter of different processing stages.

Rather, their *lateralization* may provide the key to understanding their independence. The size-value congruity effect involves analytically discriminating the relevant from the irrelevant stimulus dimension, while the SNARC effect involves evaluating the stimulus holistically in the context of a set of numbers. The distinction between these processes corresponds to the differentiation of function between the hemispheres.

Following this line of reasoning, the numerical distance and SNARC effects should be considered interdependent: all SNARC effects occur in the right hemisphere. In particular, the D-SNARC effect is a compatibility effect of numerical *distance*. Thus, distance processing taps into the same resource as SNARC, namely holistic spatial representation of numerical value.

In sum, spatial association effects of symbolic numbers all appear around the same time. All these effects occur early in ERP: P1 is sensitive to congruity, SNARC and numerical distance. However, this does not imply that they all tap into the same resources. Dependency appears to be a matter of hemispheric specialization. Effects of size-value congruity are predominantly left-hemispheric, while D-SNARC shows the combined effect of compatibility and numerical distance in the right hemisphere. The independence of the first and the latter two contrasts with the interaction of congruity and distance effects in behavior. However, the interaction appeared to be the result of a categorical representation of numerical distance, while the early effects observed in EEG were based on spatial representation.

Declaration of Competing Interest

None.

Acknowledgments

A.R.N., S.T., and C.v.L. were supported by an Odysseus grant (G.0003.12) from the Flemish Organization for Science (FWO) to Cees van Leeuwen. A.B. and S.T. were supported by the Rhineland-Palatinate Research Initiative (Potentialbereich Cognitive Science) of the Federal Ministry of Science, Further Education and Culture (MWWK).

References

Abrahamse, E., van Dijck, J. P., Majerus, S., & Fias, W. (2014). Finding the answer in space: The mental whiteboard hypothesis on serial order in working memory. *Frontiers in Human Neuroscience, 8*, 932. https://doi.org/10.3389/fnhum.2014.00932.

Alexander, D. M., Trengove, C., & van Leeuwen, C. (2015). Donders is dead: Cortical traveling waves and the limits of mental chronometry in cognitive neuroscience. *Cognitive Processing, 16*(4), 365–375. https://doi.org/10.1007/s10339-015-0662-4.

Alivisatos, B., & Wilding, J. (1982). Hemispheric differences in matching Stroop-type letter stimuli. *Cortex, 18*(4), 619–620. https://doi.org/10.1016/S0010-9452(82)80015-4.

Avancini, C., Galfano, G., & Szűcs, D. (2014). Dissociation between arithmetic relatedness and distance effects is modulated by task properties: An ERP study comparing explicit vs. implicit arithmetic processing. *Biological Psychology, 103*, 305–316. https://doi.org/10.1016/j.biopsycho.2014.10.003.

Aziz-Zadeh, L., Iacoboni, M., & Zaidel, E. (2006). Hemispheric sensitivity to body stimuli in simple reaction time. *Experimental Brain Research, 170*(1), 116–121. https://doi.org/10.1007/s00221-005-0194-8.

Bächtold, D., Baumüller, M., & Brugger, P. (1998). Stimulus-response compatibility in representational space. *Neuropsychologia, 36*(8), 731–735. https://doi.org/10.1016/S0028-3932(98)00002-5.

Banks, W. P., & Flora, J. (1977). Semantic and perceptual processes in symbolic comparisons. *Journal of Experimental Psychology Human Perception and Performance, 3*(2), 278–290. https://doi.org/10.1037/0096-1523.3.2.278.

Bavelier, D., Deruelle, C., & Proksch, J. (2000). Positive and negative compatibility effects. *Perception & Psychophysics, 62*(1), 100–112. https://doi.org/10.3758/BF03212064.

Besner, D., & Coltheart, M. (1979). Ideographic and alphabetic processing in skilled reading of English. *Neuropsychologia, 17*(5), 467–472. https://doi.org/10.1016/0028-3932(79)90053-8.

Boenke, L. T., Ohl, F., Nikolaev, A. R., Lachmann, T., & van Leeuwen, C. (2009). Different time courses of Stroop and Garner effects in perception — An event-related potentials study. *NeuroImage, 45*(4), 1272–1288. https://doi.org/10.1016/j.neuroimage.2009.01.019.

Briand, K. A. (1994). Selective attention to global or local-structure of objects –alternative measures of nontarget processing. *Perception & Psychophysics, 55*(5), 562–575. https://doi.org/10.3758/BF03205313.

Bulf, H., de Hevia, M. D., & Cassia, V. M. (2016). Small on the left, large on the right: Numbers orient visual attention onto space in preverbal infants. *Developmental Science, 19*(3), 394–401. https://doi.org/10.1111/desc.12315.

Busch, N. A., Debener, S., Kranczioch, C., Engel, A. K., & Herrmann, C. S. (2004). Size matters: Effects of stimulus size, duration and eccentricity on the visual gamma-band response. *Clinical Neurophysiology, 115*(8), 1810–1820. https://doi.org/10.1016/j.clinph.2004.03.015.

Cohen, M. X. (2014). *Analyzing neural time series data. Theory and practice.* Boston, MA: MIT Press.

Cohen Kadosh, R., Bahrami, B., Walsh, V., Butterworth, B., Popescu, T., & Price, C. J. (2011). Specialization in the human brain: The case of numbers. *Frontiers in Human Neuroscience, 5*, 62. https://doi.org/10.3389/fnhum.2011.00062.

Curran, T., Tucker, D. M., Kutas, M., & Posner, M. (1993). Topography of the N400: Brain electrical activity reflecting semantic expectancy. *Electroencephalography and Clinical Neurophysiology, 88*(3), 188–209. https://doi.org/10.1016/0168-5597(93)90004-9.

Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and number magnitude. *Journal of Experimental Psychology General, 122*(2), 371–396. https://doi.org/10.1037/0096-3445.122.3.371.

Dehaene, S. (1996). The organization of brain activations in number comparison: Event-related potentials and the additive-factors method. *Journal of Cognitive Neuroscience, 8*(1), 47–68. https://doi.org/10.1162/jocn.1996.8.1.47.

Donders, F. C. (1868). Die schnelligkeit psychischer processe: Erster artikel. *Archiv für Anatomie, Physiologie und wissenschaftliche Medizin,* 657–681.

Fias, W., & Fischer, M. H. (2005). Spatial representation of number. In J. I. D. Campbell (Ed.). *Handbook of mathematical cognition* (pp. 43–66). New York, NY: Psychology Press.

Fischer, M. H. (2018). Why numbers are embodied concepts. *Frontiers in Psychology, 8*, 2347. https://doi.org/10.3389/fpsyg.2017.02347.

Fitousi, D., Shaki, S., & Algom, D. (2009). The role of parity, physical size, and magnitude in numerical cognition: The SNARC effect revisited. *Perception & Psychophysics, 71*(1), 143–155. https://doi.org/10.3758/APP.71.1.143.

Fitts, P. M., & Seeger, C. M. (1953). SR compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46*(3), 199–210. https://doi.org/10.1037/h0062827.

Fumarola, A., Prpic, V., Da Pos, O., Murgia, M., Umiltà, C., & Agostini, T. (2014). Automatic spatial association for luminance. *Attention, Perception & Psychophysics, 76*(3), 759–765. https://doi.org/10.3758/s13414-013-0614-y.

Fumarola, A., Prpic, V., Fornasier, D., Sartoretto, F., Agostini, T., & Umiltà, C. (2016). The spatial representation of angles. *Perception, 45*(11), 1320–1330. https://doi.org/10.1177/0301006616661915.

Gevers, W., Verguts, T., Reynvoet, B., Caessens, B., & Fias, W. (2006). Numbers and space: A computational model of the SNARC effect. *Journal of Experimental Psychology Human Perception and Performance, 32*(1), 32–44. https://doi.org/10.1037/0096-1523.32.1.32.

Grune, K., Mecklinger, A., & Ullsperger, P. (1993). Mental comparison: P3 component of the ERP reflects the symbolic distance effect. *Neuroreport, 4*(11), 1272–1274.

Gut, M., Szumska, I., Wasilewska, M., & Jaśkowski, P. (2012). Are low and high number magnitudes processed differently while resolving the conflict evoked by the SNARC effect? *International Journal of Psychophysiology, 85*(1), 7–16. https://doi.org/10.1016/j.ijpsycho.2012.02.007.

Henik, A., & Tzelgov, J. (1982). Is three greater than five: The relation between physical and semantic size in comparison tasks. *Memory & Cognition, 10*(4), 389–395. https://doi.org/10.3758/BF03202431.

Hillyard, S. A., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings of the National Academy of Sciences, 95*(3), 781–787. https://doi.org/10.1073/pnas.95.3.781.

Hommel, B., Müsseler, J., Aschersleben, G., & Prinz, W. (2001). The theory of event coding (TEC): A framework for perception and action planning. *The Behavioral and Brain Sciences, 24*(5), 849–878. https://doi.org/10.1017/S0140525X01000103.

Jarvstad, A., Hahn, U., Rushton, S. K., & Warren, P. A. (2013). Perceptuo-motor, cognitive, and description-based decision-making seem equally good. *Proceedings of the National Academy of Sciences of the United States of America, 110*(40), 16271–16276. https://doi.org/10.1073/pnas.1300239110.

Jung, T. P., Humphries, C., Lee, T. W., Makeig, S., McKeown, M. J., Iragui, V., et al. (1998). Extended ICA removes artifacts from electroencephalographic recordings. *Advances in Neural Information Processing Systems, 10*, 894–900.

Jiang, T., Qiao, S., Li, J., Cao, Z., Gao, X., Song, Y., et al. (2010). Effects of symbol type and numerical distance on the human event-related potential. *Neuropsychologia, 48*(1), 201–210. https://doi.org/10.1016/j.neuropsychologia.2009.09.005.

Kadosh, R. C., Kadosh, K. C., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and-independent representations of numbers in the parietal lobes. *Neuron, 53*(2), 307–314. https://doi.org/10.1016/j.neuron.2006.12.025.

Keus, I. M., & Schwarz, W. (2005). Searching for the functional locus of the SNARC effect: Evidence for a response-related origin. *Memory & Cognition, 33*(4), 681–695. https://doi.org/10.3758/BF03195335.

Krajcsi, A., & Kojouharova, P. (2017). Symbolic numerical distance effect does not reflect the difference between numbers. *Frontiers in Psychology, 8*, 2013. https://doi.org/10.3389/fpsyg.2017.02013.

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

Lohmann, J., Schroeder, P. A., Nuerk, H.-C., Plewnia, C.h., & Butz, M. V. (2018). How deep is your SNARC? Interactions between numerical magnitude, response hands, and reachability in peripersonal space. *Frontiers in Psychology, 9*, 622. https://doi.org/10.3389/fpsyg.2018.00622.

Loizides, C., Achilleos, A., Iannetti, G. D., & Mitsis, G. D. (2015). Assessment of nonlinear interactions in event-related potentials elicited by stimuli presented at short inter-stimulus intervals using single-trial data. *Journal of Neurophysiology, 113*(10), 3623–3633. https://doi.org/10.1152/jn.00523.2014.

Luck, S. J. (2005). *An introduction to the event-related potential technique.* Cambridge, MA: MIT press.

Merkley, R., Shimi, A., & Scerif, G. (2016). Electrophysiological markers of newly acquired symbolic numerical representations: The role of magnitude and ordinal information. *ZDM Mathematics Education, 48*(3), 279–289. https://doi.org/10.1007/s11858-015-0751-y.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature, 215*, 1519–1520. https://doi.org/10.1038/2151519a0.

Nikolaev, A. R., Ziessler, M., Dimova, K., & van Leeuwen, C. (2008). Anticipated action consequences as a nexus between action and perception: Evidence from event-related potentials. *Biological Psychology, 78*(1), 53–65. https://doi.org/10.1016/j.biopsycho.2007.12.010.

Nuerk, H.-C., Wood, G., & Willmes, K. (2005). The universal SNARC effect: The association between number magnitude and space is amodal. *Experimental Psychology, 52*(3), 187–194. https://doi.org/10.1027/1618-3169.52.3.187.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia, 9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4.

Paivio, A. (1975). Perceptual comparisons through the mind's eye. *Memory & Cognition, 3*(6), 635–647. https://doi.org/10.3758/BF03198229.

Prpic, V., Fumarola, A., De Tommaso, M., Luccio, R., Murgia, M., & Agostini, T. (2016). Separate mechanisms for magnitude and order processing in the spatial-numerical association of response codes (SNARC) effect: The strange case of musical note values. *Journal of Experimental Psychology Human Perception and Performance, 42*(8), 1241–1251. https://doi.org/10.1037/xhp0000217.

Prpic, V., Soranzo, A., Santoro, I., Fantoni, C., Galmonte, A., Agostini, T., et al. (2018). SNARC-like compatibility effects for physical and phenomenal magnitudes: A study on visual illusions. *Psychological Research.* https://doi.org/10.1007/s00426-018-1125-1 online first.

R Core Team (2013). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org.

Ren, P., Nicholls, M. E. R., Ma, Y., & Chen, L. (2011). Size matters: Non-numerical magnitude affects the spatial coding of response. *PloS One, 6*(8), e23553. https://doi.org/10.1371/journal.pone.0023553.

Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 83*(2), 274–278. https://doi.org/10.1037/h0028573.

Santens, S., & Verguts, T. (2011). The size congruity effect: is bigger always more? *Cognition, 118*(1), 94–110. https://doi.org/10.1016/j.cognition.2010.10.014.

Schwarz, W., & Heinze, H.-J. (1998). On the interaction of numerical and size information in digit comparison: A behavioral and event-related potential study. *Neuropsychologia, 36*(11), 1167–1179. https://doi.org/10.1016/S0028-3932(98)00001-3.

Schwarz, W., & Ischebeck, A. (2003). On the relative speed account of number-size interference in comparative judgments of numerals. *Journal of Experimental Psychology Human Perception and Performance, 29*, 507–522. https://doi.org/10.1037/0096-1523.29.3.507.

Shaki, S., Petrusic, W. M., & Leth-Steensen, C. (2012). SNARC effects with numerical and non-numerical symbolic comparative judgments: Instructional and cultural dependencies. *Journal of Experimental Psychology Human Perception and Performance, 38*(2), 515–530. https://doi.org/10.1037/a0026729.

Shaki, S., Fischer, M. H., & Petrusic, W. M. (2009). Reading habits for both words and numbers contribute to the SNARC effect. *Psychonomic Bulletin & Review, 16*, 328–331. https://doi.org/10.3758/PBR.16.2.328.

Shin, Y. K., Proctor, R. W., & Capaldi, E. J. (2010). A review of contemporary ideomotor theory. *Psychological Bulletin, 136*(6), 943–974. https://doi.org/10.1037/a0020541.

Soltész, F., Goswami, U., White, S., & Szűcs, D. (2011). Executive function effects and numerical development in children: Behavioural and ERP evidence from a numerical Stroop paradigm. *Learning and Individual Differences, 21*(6), 662–671. https://doi.org/10.1016/j.lindif.2010.10.004.

Sternberg, S. (2018). Some constraints on reaction-time distributions for sequential processes. In T. Lachmann, & T. Weis (Eds.). *Invariances in human information processing* (pp. 101–116). Abingdon, UK: Routledge.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643–662. https://doi.org/10.1037/h0054651.

Szűcs, D., & Csépe, V. (2005). The effect of numerical distance and stimulus probability on ERP components elicited by numerical incongruencies in mental addition. *Cognitive Brain Research, 22*(2), 289–300. https://doi.org/10.1016/j.cogbrainres.2004.04.010.

Szűcs, D., & Soltész, F. (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia, 45*(14), 3190–3202. https://doi.org/10.1016/j.neuropsychologia.2007.06.013.

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.

Van Galen, M. S., & Reitsma, P. (2008). Developing access to number magnitude: A study of the SNARC effect in 7- to 9-year-olds. *Journal of Experimental Child Psychology, 101*(2), 99–113. https://doi.org/10.1016/j.jecp.2008.05.001.

Van Kleeck, M. H. (1989). Hemispheric differences in global versus local processing of hierarchical visual stimuli by normal subjects: New data and a meta-analysis of previous studies. *Neuropsychologia, 27*(9), 1165–1178. https://doi.org/10.1016/0028-3932(89)90099-7.

Van Leeuwen, C., & Bakker, L. (1995). Stroop can occur without Garner interference: Strategic and mandatory influences in multidimensional stimuli. *Perception & Psychophysics, 57*(3), 379–392. https://doi.org/10.3758/BF03213062.

Van Leeuwen, C., & Lachmann, T. (2004). Negative and positive congruence effects in letters and shapes. *Perception & Psychophysics, 66*(6), 908–925. https://doi.org/10.

3758/BF03194984.
Viarouge, A., Hubbard, E. M., & McCandliss, B. D. (2014). The cognitive mechanisms of the SNARC effect: An individual differences approach. *PloS One, 9*(4), e95756. https://doi.org/10.1371/journal.pone.0095756.
Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Science, 7*(11), 483–488. https://doi.org/10.1016/j.tics.2003.09.002.
Weis, T., Estner, B., & Lachmann, T. (2016). When speech enhances SMARC: Joint spatial associations of pitch and timbre in nonmusicians. *The Quarterly Journal of Experimental Psychology, 69*, 1687–1700. https://doi.org/10.1080/17470218.2015.1091850.
Weis, T., Estner, B., van Leeuwen, C., & Lachmann, T. (2016). SNARC meets SPARC: Automaticity and interdependency in compatibility effects. *The Quarterly Journal of Experimental Psychology, 69*, 1366–1383. https://doi.org/10.1080/17470218.2015.1082142.
Weis, T., Theobald, S., Schmitt, A., van Leeuwen, C., & Lachmann, T. (2018). There's a SNARC in the size congruity task. *Frontiers in Psychology, 9*, 1978. https://doi.org/10.3389/fpsyg.2018.01978.