Neuropsychologia 48 (2010) 394–401

Contents lists available at ScienceDirect

Neuropsychologia

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

Conflict processing of symbolic and non-symbolic numerosity

Titia Gebuis[a,*], J. Leon Kenemans[a], Edward H.F. de Haan[b], Maarten J. van der Smagt[a]

[a] Experimental Psychology, Helmholtz Institute and Utrecht University, The Netherlands
[b] Social and Behavioural Sciences, University of Amsterdam, The Netherlands

ARTICLE INFO

Article history:
Received 4 March 2009
Received in revised form
23 September 2009
Accepted 25 September 2009
Available online 3 October 2009

Keywords:
ERP
P300
Size congruency
LRP

ABSTRACT

It is commonly assumed that the processing of magnitudes occurs independent of modality or notation. Several studies have reported similar behavioural as well as neurophysiological responses to magnitudes presented in distinct modalities as well as notations, but a direct assessment of possible interactions between different modalities and notations, using measures of electro-cortical processing, is lacking. The present study investigates whether the neural activity underlying symbolic and non-symbolic numerosity processing interacts with the neural activity underlying physical size processing before, or proceeds independently until, selective activation of the motor system. We used a symbolic (Arabic numbers) and non-symbolic (arrays of dots) size congruency task and instructed subjects to judge either the numerical or the physical size of the stimuli, while event related potentials were recorded. Longer reaction times as well as a decrease in accuracy were obtained for incongruent compared to congruent trials. For the event related potential data, this congruency effect was also found with respect to the latency of the P3 component reflecting an interaction at the level of stimulus evaluation. Moreover, incongruence delayed the stimulus-locked but not the response-locked lateralized readiness potential. Together these results suggest that, irrespective of notation, the interaction between different magnitudes occurs before selective response activation.

1. Introduction

1.1. Modality and notation independent processing

Two influential models, The Triple Code (Dehaene, Piazza, Pinel, & Cohen, 2003) and the ATOM model (Walsh, 2003) propose that distinct magnitudes are encoded in an amodal format. Evidence consistent with these models comes from behavioural and imaging studies investigating the 'distance effect' and 'size effect' using distinct modalities. The distance effect refers to the faster responses obtained when discriminating between numerically far (1 and 9) compared to numerically close numbers (4 and 5) whereas the size effect refers to the longer reaction times obtained for the comparison of numerically large (8 and 9) compared to small numbers (1 and 2). Both effects have been measured on a behavioural as well as a neuronal level using modalities such as physical size (Cohen Kadosh et al., 2005; Kaufmann et al., 2005; Pinel, Piazza, Le Bihan, & Dehaene, 2004; Tang, Critchley, Glaser, Dolan, & Butterworth, 2006), luminance (Cohen Kadosh et al., 2005; Cohen Kadosh, Cohen Kadosh, & Henik, 2008; Pinel et al., 2004), line length (Fias, Lammertyn, Reynvoet, Dupont, & Orban, 2003), time (Dormal, Seron, & Pesenti, 2006) and pitch (Rusconi, Kwan, Giordano, Umilta, & Butterworth, 2006; for a review and meta-analysis on this topic see: Cohen Kadosh, Cohen Kadosh, et al., 2008 and Cohen Kadosh, Henik, & Rubinsten, 2008).

In addition to modality independence, The Triple Code model also proposes that magnitudes are processed independent of notation. In line with this hypothesis, several behavioural priming studies report within as well as cross notation priming effects (Dehaene et al., 1998; Reynvoet & Brysbaert, 2004; Reynvoet, Brysbaert, & Fias, 2002; Reynvoet, Gevers, & Caessens, 2005). These numerical priming effects are generally explained as evidence for an abstract magnitude code. However, priming effects can also arise between concepts that are processed in distinct areas (e.g. colours and objects). Such effects might arise due to strong associations between both concepts (Nijboer, van Zandvoort, & de Haan, 2006). This idea is underlined by synesthesia research revealing priming effects between colours and graphemes that cannot be explained on the basis of the convergence of both processes into a unitary code (Gebuis, Nijboer, & van der Smagt, 2009a; Gebuis, Nijboer, & van der Smagt, 2009b). More conclusive evidence for an abstract representation came from neuroimaging studies showing that number words and Arabic numbers activate similar intraparietal brain areas (Pinel, Dehaene, Riviere, & LeBihan, 2001) and lead to repetition

---

* Corresponding author at: Experimental Psychology, Utrecht University, Heidelberglaan 2, 3584 CS Utrecht, The Netherlands. Tel.: +31 30 2533409; fax: +31 30 2534511.
E-mail address: T.Gebuis@uu.nl (T. Gebuis).

0028-3932/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.neuropsychologia.2009.09.027

suppression in these areas (Naccache & Dehaene, 2001). Notation independent processing of number words and Arabic numbers was further investigated by Cohen Kadosh, Cohen Kadosh, Kaas, Henik, and Goebel (2007) and Cohen Kadosh, Cohen Kadosh, Linden, et al. (2007) using functional magnetic resonance adaptation as a measure. This method allows investigating stimulus related processes without the requirement of subjects responding to the stimuli presented. Since response selection (Gobel, Johansen-Berg, Behrens, & Rushworth, 2004; Jiang & Kanwisher, 2003) activates similar areas as those expected to subserve numerosity processes, this method overcomes a crucial problem in this domain of research. In line with the imaging results of Naccache and Dehaene (2001), Cohen Kadosh, Cohen Kadosh, Kaas, et al. (2007) and Cohen Kadosh, Cohen Kadosh, Linden, et al. (2007) revealed notation independent adaptation and recovery patterns in the left hemisphere but also areas only responsive to Arabic numbers in the right hemisphere. This only partly overlapping system for Arabic numbers and number words led them to suggest that the magnitude system might not be as notation independent as previously thought. The results of two additional behavioural studies (Cohen Kadosh, 2008; Cohen Kadosh, Henik, et al., 2008) further emphasized this notion.

In contradiction to number words, numerosity presented in non-symbolic notation (e.g. an array of dots) is often viewed as the precursor for numerosity presented in symbolic notation (e.g. Arabic number). Infants are already capable of processing numerosity presented in non-symbolic notation (Brannon, 2002; Feigenson, Dehaene, & Spelke, 2004; Jordan, Suanda, & Brannon, 2008; Xu, Spelke, & Goddard, 2005) and this capability is therefore commonly assumed to serve as the basis for our later acquired symbolic number system (Dehaene, 2001). Hence, numerosity presented in a non-symbolic notation might be subserved by the same processing mechanisms as symbolic notation. Consistent with this hypothesis, responses to numerosity presented in different notations (e.g. Arabic numbers versus dots) resulted in comparable electroencephalography (EEG) signals in adults (Libertus, Woldorff, & Brannon, 2007; Temple & Posner, 1998) as well as children (Temple & Posner, 1998). Moreover, a neural network model trained for non-symbolic notation could also process numerosities presented in a symbolic notation (Verguts & Fias, 2004). More direct evidence resulted from studies actually showing that the neural substrates subserving non-symbolic notation also become activated for symbolic notation (Diester & Nieder, 2007; Piazza, Pinel, Le Bihan, & Dehaene, 2007). Diester and Nieder (2007) revealed that monkey prefrontal cortex neurons tuned for non-symbolic numerosity fired for symbolic stimuli as well after the monkey was trained on non-symbolic to symbolic number associations. Furthermore, Piazza et al. (2007) used functional magnetic resonance adaptation and revealed adaption and recovery effects with bilateral areas responsive to arrays of dots as well as Arabic number symbols. Together, the studies presented above all suggest notation independent numerosity processing, with the exception of number *word* notation. However, interactions between different magnitudes can be manifest at multiple levels of processing; the question at what stage precisely, is still unresolved.

1.2. Interactions between distinct magnitude-processing systems

Interactions between distinct magnitude processes are frequently investigated using congruency tasks. In contrast to comparison tasks where only one dimension is manipulated, the congruency task allows to study the interaction between two magnitudes. Here, two simultaneously presented stimuli are manipulated in two magnitude dimensions (e.g. numerical and physical size) resulting in a congruent and incongruent condition. In the congruent condition the numerically larger stimulus is also physically larger whereas in the incongruent condition the numerically larger stimulus is physically smaller compared to the stimulus presented simultaneously (e.g. congruent: 2 4; incongruent: 2 4). Generally, behavioural responses are faster for congruent (facilitation) relative to incongruent stimuli (interference). There are two hypotheses that aim to explain these interaction effects (Schwarz & Heinze, 1998). First there is the *early interaction account*, which suggests that the processing streams of two distinct magnitudes merge at an "early" processing stage, before the response processes are initiated. At this level of stimulus processing the two magnitudes can interact as they are encoded in a similar format by the same system. In contrast, the *late interaction account* suggests that the two magnitudes are processed in two separate, independent processing streams up until the stage of the response. It is only at this 'response' stage that the two magnitudes can interact. The behavioural and imaging results from studies investigating the interaction between physical size and Arabic numbers or distinct levels of luminance, support the early interaction account (Cohen Kadosh, Cohen Kadosh, et al., 2008; Cohen Kadosh, Cohen Kadosh, Linden, et al., 2007). So far, no such early interaction has been obtained for a magnitude notation (e.g. number words) other than Arabic numbers (Cohen Kadosh, 2008; Cohen Kadosh, Henik, et al., 2008). The demonstration of comparable early interactions for both symbolic and non-symbolic notations is prerequisite but not conclusive evidence for the commonly held view that Arabic numbers are mapped onto a non-symbolic number system as soon as the symbolic numbers are learned. In contrast, late interactions would provide definitive evidence against an abstract code.

1.3. The current study

To investigate the timing of the interaction between two magnitudes we used a symbolic and a non-symbolic size congruency task. We specifically question whether interactions between magnitudes of distinct notation occur prior to, or only at the motor response stage. An interaction at the response stage for either task indicates that magnitude information is processed in parallel up until the response initiation or execution. Such a result would exclude the idea of a general magnitude system that is devoted to both symbolic and non-symbolic notation. Conversely, interaction effects at an "early" stage for both the symbolic and non-symbolic size congruency task would be in line with the hypothesis of a magnitude system that is notation independent, at least for Arabic numbers and dots (as opposed to number words).

Stimulus and response complexity have additive effects on reaction time. Therefore, to discriminate between the early and late interaction accounts, the electrophysiological chronometric measures: P3 and LRP are especially useful (Lansbergen & Kenemans, 2008). The latency of the P3 component reflects stimulus evaluation and categorization processes whereas processes of response selection and execution do not affect P3 latency (Donchin, 1981; Kutas, McCarthy, & Donchin, 1977; McCarthy & Donchin, 1981; Smulders, Kok, Kenemans, & Bashore, 1995). The P3 latency can thus be used to determine whether the obtained congruency effects at the behavioural level are the result of processes leading up to stimulus categorization. The LRP is a measure of selective motor activation showing larger deflections at scalp sites contralateral to the moved hand. A later onset of the stimulus-locked but not the response-locked LRP for the incongruent compared to the congruent trials indicates a stimulus conflict prior to motor preparation, whereas the opposite indicates a conflict at the response level (Smulders et al., 1995).

2. Methods

2.1. Participants

Eighteen subjects participated in the experiment of which fifteen were included in the analyses (aged between 19 and 35 years $M = 23.7$, S.D. = 1.36; 13 female, 2 male). The three subjects were discarded because more than 25% of their EEG epochs contained artefacts (for further detail see: behavioural and ERP analyses). All subjects were native Dutch speakers and had normal or corrected-to-normal vision and were paid for their participation. Written informed consent was obtained according to the Declaration of Helsinki and as approved by the Ethical Committee of the University of Utrecht.

2.2. Apparatus, stimuli and procedure

In each trial, the stimuli were displayed on a 22-in. CRT monitor using the Presentation software (Neurobehavioral Systems, Albany, CA). The paradigm was a slightly adapted version of our previous study (Gebuis, Cohen Kadosh, de Haan, & Henik, 2009).

For the symbolic comparison task, the stimuli consisted of Arabic numbers ranging from 1 to 9, which were presented in pairs with a small (1–2, 2–3, 3–4, 4–5, 5–6, 6–7, 7–8, 8–9) or large numerical distance (1–6, 2–7, 3–8, 4–9). Depending on the condition numbers were presented in a small font size (height 1.7° visual angle) or a large font size (height 2.4° visual angle). For the non-symbolic comparison task the stimuli consisted of groups of dots ranging from 8 to 16. This relatively large number of dots is used to rule out possible subitizing effects. Depending on the condition, small (0.38° in diameter) or large dots (0.53° in diameter) were presented. In each trial, groups of dots ranging from 8 to 16 were randomly distributed within a pre-specified area (3.05° square). Again, two numerical distances were used, a small numerical distance of 4 (7–11, 8–12, 9–13, 10–14, 11–15, 12–16) and a large numerical distance of 7 (7–14, 8–15, 9–16). To exclude the possibility that the participant derived the correct answer on the basis of visual sensory properties, these were controlled for: in the *congruent* condition an array of dots with more and physically larger dots (which together constitute a larger surface area), had to be compared with an array of dots with fewer and physically smaller dots (which together constitute a smaller surface area). In the *incongruent* condition an array with more but physically smaller dots (which together constitute a smaller surface area) was compared to an array with fewer but physically larger dots (which together constitute a larger surface area) (see for further details: Gebuis, Cohen Kadosh, et al., 2009). The centres of the stimuli were positioned 2.0° to the left and right side of the fixation cross on a grey background and the viewing distance was approximately 57 cm.

The two comparison tasks each consisted of two judgment conditions: (1) numerical size and (2) physical size judgment. Thus a total of four tasks were administered, each consisting of (128) congruent and (128) incongruent trials, which again consisted of (64) small and (64) large numerical distance trials. The order of the four comparison tasks was counter balanced between participants. Each trial began with a fixation cross (500 ms), followed by the stimulus (until response) and a random inter trial interval (1250–1500 ms). Participants responded by pressing the button corresponding to the side at which the target (the numerically or physically largest of the two stimuli) was presented. Half of the trials were presented with the target on the left side and half of the trials with the target on the right side. Prior to each comparison task participants received instructions and performed 20 practice trials. Between each comparison task participants could take a break.

2.3. Electrophysiological recordings

EEGs were recorded from 64 scalp electrodes according to the International 10/20 EEG system (sampling rate of 2048 Hz) using the Active Two system (BioSemi, Amsterdam, The Netherlands) relative to the common mode sense (CMS). Extremely low-noise recordings free of interference are achieved by physically integrating the first amplifier stage with a sintered Ag–AgCl electrode. This so-called Active-electrode is a sensor with very low output impedance. The ground consists of the active CMS and the passive driven right leg (DRL) electrode that form a feedback loop driving the subject's average potential as close as possible to the analog-to-digital converter (i.e. the amplifier 'zero') reference voltage in the A/D-box. Therefore, impedance measurements or gain adjustments are not required (Schutter, de Weijer, Meuwese, Morgan, & van Honk, 2008; but for a more elaborate explanation see: www.biosemi.com). The vertical electro-oculogram (VEOG) was recorded from electrodes attached above and below the left eye and the horizontal electro-oculogram (HEOG) from the outer canthi of both eyes.

2.4. Behavioural and ERP analyses

For each subject included in the analyses, median reaction times of the correct trials were calculated for each task condition.

EEG and EOG data were analyzed using Analyzer Software (1.05). Noisy electrodes were excluded from the analyses. A noisy electrode was defined as a single electrode that resulted in the rejection of more than 25% of all epochs recorded. Pz, C3, C4, the electrodes of interest, were not amongst them. EEG signals were referenced off-line to the average of all included electrodes and segmented into epochs from 200 ms prior to 1000 ms after the presentation of the stimulus. Epochs were filtered with a bandpass filter (0.05 Hz, 12 dB octave; 40 Hz, 24 dB/octave) and corrected for eye movements according to the blink algorithm of Gratton, Coles, and Donchin (1983). Trials with artefacts (difference criterion of 100 μV within an epoch; low activity criterion of 0.5 μV within a 100-ms time window) or an incorrect response were rejected from further analyses. Three subjects were discarded because more than 25% of the trials contained artefacts. The baseline was defined as the mean of the 100-ms period before stimulus onset. The baseline for the response-locked LRPs was −550 to −450 ms. For both the ERPs and LRPs grand average waveforms were created for each condition and filtered at 8 Hz, 12 dB/octave for visualization purposes only.

For correct-response trials only, P3 latency was estimated in single trials at the Pz electrode as the largest peak within the time window of 300–800 ms after stimulus presentation (Smulders, Kenemans, & Kok, 1994). The P3 amplitude was also analyzed to allow comparison with previous studies and was estimated at the Pz electrode as the largest peak within the time window of 300–800 ms after stimulus presentation. Stimulus-locked and response-locked lateralized readiness potentials (sLRP and rLRP, respectively) were computed by subtracting the electrode ipsilateral to the response hand from the electrode contralateral to the response hand. Subsequently, the obtained difference waves for both hands were averaged [mean (C3 − C4) left-hand movement + mean (C4 − C3) right-hand movement/2]. Jackknife averaging was used to accurately estimate the difference in sLRP and rLRP onset (Miller, Patterson, & Ulrich, 1998). For each condition averages were obtained by omitting one participant from the sample. Therefore, we had to correct the obtained F-values as follows: $Fc = F/(n-1)^2$ (Ulrich & Miller, 2001). As recommended by Miller et al. (1998) the relative criterion method was applied to each subsample waveform using the 50% criterion for the sLRPs and 90% for the rLRPs.

To compare the two distinct notations a repeated measures ANOVA was performed with Notation (symbols and dots) × Congruency (congruent and incongruent) × Distance (small and large) as within subject variables for both the physical size as well as the numerical size judgment task. If present, to further disentangle three-way interactions simple effects analyses were performed per notation condition.

3. Results

3.1. Reaction time results (Fig. 1a and b)

In the *numerosity comparison tasks* a distance effect was present [$F(1,14) = 226.174$, $p < 0.001$]. Subjects responded faster to large (symbolic task: 419 ms; non-symbolic task: 440 ms) compared to small (symbolic task: 475 ms; non-symbolic task: 496 ms) numerical distance trials. The main effect for congruency [$F(1,14) = 51.886$, $p < 0.001$] was also significant, reflecting the faster responses to congruent (symbolic task: 409 ms; non-symbolic task: 448 ms) compared to incongruent (symbolic task: 486 ms; non-symbolic task: 490 ms) trials. Congruency and distance interacted [$F(1,14) = 20.222$, $p = 0.001$]: the congruency effect was smaller for the large (symbolic task: 48 ms; non-symbolic task: 27 ms) compared to the small (symbolic task: 107 ms; non-symbolic task: 57 ms) numerical distance trials. Notation and congruency [$F(1,14) = 5.753$, $p = 0.031$] interacted as well, reflecting a smaller congruency effect for dot [$t(1,14) = -2.924$, $p = 0.011$] compared Arabic-number notation [$t(1,14) = -11.308$, $p < 0.001$].

In the *physical size comparison tasks* a congruency [$F(1,14) = 8.987$, $p = 0.010$] and distance [$F(1,14) = 83.367$, $p < 0.001$] effect were present. Subjects responded faster to the small compared to large (symbolic task: 417 versus 430 ms; non-symbolic task: 392 versus 395 ms) numerical distance trials. Note that this numerical distance effect is opposite to the pattern of what was found in the numerical judgment task (for an explanation see Section 4). Furthermore, faster responses were given to congruent compared to incongruent (symbolic task: 390 versus 457 ms; non-symbolic task: 372 versus 415 ms) trials. In contrast to the numerosity comparison tasks, a main effect for notation [$F(1,14) = 11.711$, $p = 0.004$] reflected faster responses in the dot notation compared to the Arabic-number notation condition. In addition, an interaction between distance and congruency was significant [$F(1,14) = 61.090$, $p < 0.001$] due to larger congruency effects for the large (symbolic task: 88 ms; non-symbolic task: 44 ms) compared to the small numerical distance trials (symbolic

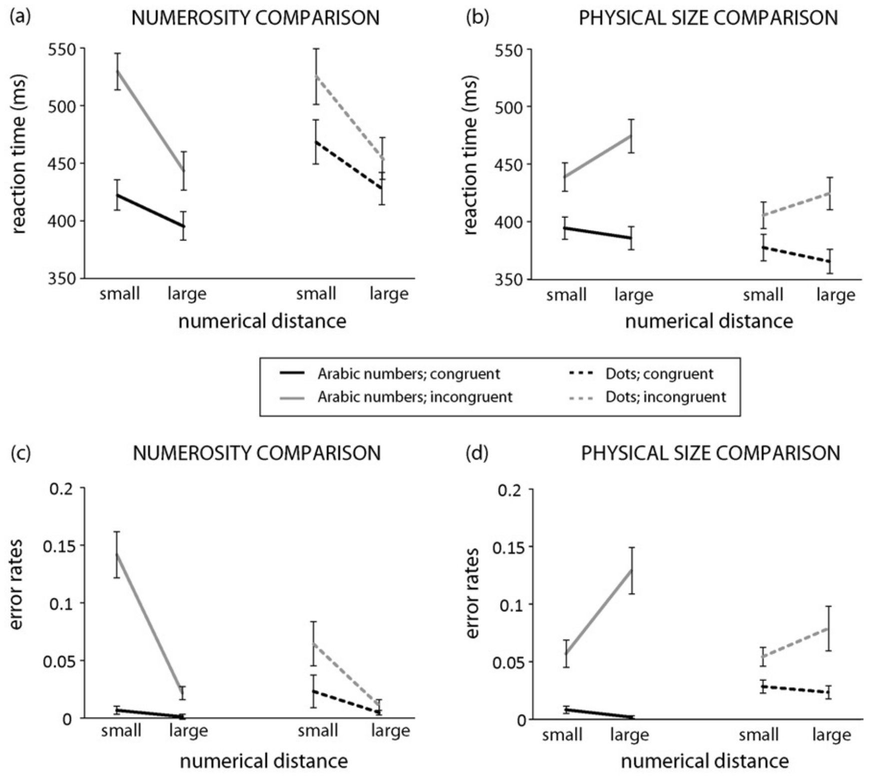

Fig. 1. Behavioural results of the four congruency tasks. The error rates (bottom panels) and the reaction times (upper panels) are shown for the numerosity (left panels) and the physical size (right panels) comparison tasks. Data for Arabic numbers are shown as solid lines and those for dots as dashed lines. The results reveal a clear effect of congruency for both notation conditions on both comparison tasks as well as measures, i.e. more accurate or faster responses are given for congruent compared to incongruent trials. In addition, the small numerical distance trials show larger congruency effects for numerosity comparison, whereas the large numerical distance trials reveal the largest congruency effects in the physical size comparison tasks.

task: 59 ms; non-symbolic task: 28 ms). The interaction between notation and congruency [$F(1,14) = 5.315$, $p = 0.037$] also reached significance reflecting a smaller congruency effect for the dot [$t(1,14) = -7.212$, $p < 0.001$] compared to Arabic-number notation [$t(1,14) = -7.181$, $p < 0.001$].

3.2. Error rates

Except for notation in the physical size comparison tasks, all main effects as well as the two and three-way interactions were significant [$F(1,14) > 4.5$, $p < 0.05$] for both comparison tasks. Fig. 1c and d reveals that the error rates of the different conditions follow the same pattern as they do for reaction time: in the incongruent compared to congruent trials, errors as well as reaction times increase. Hence, there is no sign of differential speed-accuracy trade-offs between conditions that could complicate our chronometric analysis.

3.3. ERP results

3.3.1. P3 latency (Fig. 2)

For the *numerosity comparison tasks* the results revealed a distance [$F(1,14) = 31.547$, $p < 0.001$] as well as a congruency effect [$F(2,26) = 23.976$, $p < 0.001$]. The P3 peak appeared around a later point in time for numerically small compared to numerically large distance trials as well as incongruent compared to congruent trials. Congruency and distance interacted [$F(2,26) = 23.976$, $p < 0.001$] indicating a larger congruency effect for smaller numerical distance trials. In addition, congruency interacted with notation [$F(1,14) = 12.356$, $p = 0.003$] reflecting a smaller congruency effect for the dot notation [$t(1,14) = -2.767$, $p = 0.015$] compared to Arabic-number notation [$t(1,14) = -5.692$, $p < 0.001$].

In the *physical size comparison tasks* a distance [$F(1,14) = 5.728$, $p = 0.031$] as well as a congruency [$F(1,14) = 29.456$, $p < 0.001$] effect were present. The P3 peak appeared around a later point in time for large compared to small as well as incongruent compared to congruent trials. Note that, similar to what was found in the behavioural results, the numerical distance effect for the physical size comparison is again opposite to the pattern of what was found in the numerical judgment task (for an explanation see Section 4). Moreover, distance and congruency interacted, the distance effect was larger for the large compared to the small numerical distance [$F(1,14) = 8.194$, $p = 0.013$] trials.

3.3.2. P3 amplitude (Fig. 2)

In the *numerosity comparison tasks* main effects for distance [$F(1,14) = 13.053$, $p = 0.003$], congruency [$F(1,14) = 5.806$, $p = 0.030$] and notation [$F(1,14) = 8.538$, $p = 0.011$] were present. The main effect of notation reflects the overall larger peaks obtained for the Arabic number compared to the dot notation condition. The main effects for congruency and distance refer to the larger peak amplitude for congruent compared to incongruent and large compared to small numerical distance trials.

In the *physical size comparison tasks* the results revealed a congruency effect only [$F(1,14) = 8.596$, $p = 0.011$].

3.3.3. sLRP (Fig. 3)

In the *numerosity comparison tasks* the data revealed a trend towards significance for the effect of congruency [$F(1,14) = 3.12$, $p = 0.09$] and a significant distance effect [$F(1,14) = 9.38$, $p = 0.008$]. A two-way interaction between notation and distance [$F(1,14) = 9.45$, $p = 0.008$] was present, as well as a three-way interaction between notation, congruency and distance [$F(1,14) = 14.983$, $p = 0.002$]. To disentangle the effects underlying the three-way interactions, sim-

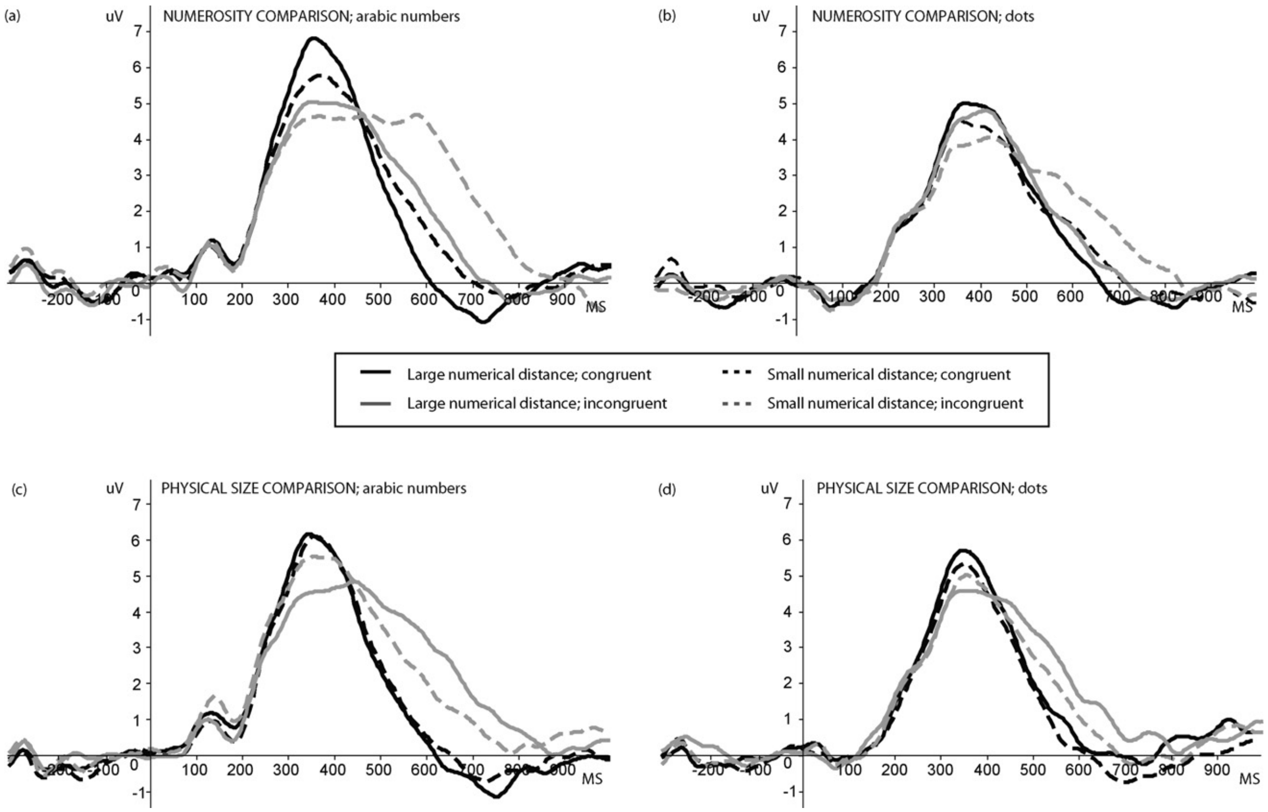

Fig. 2. Grand average ERPs at the Pz electrode. The results of the numerosity comparison (upper panels) and the physical size comparison (lower panels) of both the Arabic-number notation (left panels) and the dot notation (right panels) conditions are presented. Black lines represent congruent, grey lines incongruent trials; Solid lines represent large, dashed lines represent small numerical distance conditions. In both tasks and for both notations, a congruency effect is present in the latency (peaks appear around a later time point for incongruent trials compared to congruent trials) and peak amplitude data (peaks are larger or congruent compared to incongruent trials). Note, that the numerical distance effects for the incongruent trials are opposite for both comparison conditions. A larger delay and smaller peak amplitude are present in the physical size compared to numerosity comparison tasks for the large numerical distance trials and vice versa for small numerical distance trials.

ple effects analyses were performed. In the Arabic-number notation condition the results revealed a trend towards significance for congruency [$F(1,14) = 3.53$, $p = 0.07$] as well as a significant distance effect [$F(1,14) = 14.054$, $p = 0.001$] and interaction [$F(1,14) = 4.45$, $p = 0.045$]. In the dot notation condition an interaction between distance and congruency was present [$F(1,14) = 9.12$, $p = 0.006$]. In the *physical size comparison tasks* there was a significant congruency [$F(1,14) = 18.14$, $p = 0.001$] as well as a distance [$F(1,14) = 13.44$ $p = 0.003$] effect.

3.3.4. rLRP (Fig. 4)

For both the *numerosity comparison* as well as *the physical size comparison tasks,* no significant main effect or interaction was present.

4. Discussion

In the present study we investigated at which stage in the transition from perception to action, different magnitude systems interact. Using a conflict paradigm and electro-cortical readout measures for stimulus evaluation and selective motor activation, the answer is clear cut: interactions for both symbolic and non-symbolic notation take place before the end of stimulus evaluation (P3 peak latency), and before the start of selective motor preparation (LRP). These results support the early interaction account and are in line with the notion of a magnitude system that subserves numerosity presented both in symbolic and in non-symbolic notation.

The behavioural data revealed similar interaction patterns for the two notation conditions, Arabic numbers and dots, which is in agreement with the results of our previous study (Gebuis, Cohen Kadosh, et al., 2009). We extended the results of our previous study by revealing numerical distance effects as well as the interaction of numerical distance and congruency at the behavioural and ERP level. When numerosity interacted with physical size as well as the reverse, response time and the number of errors increased for the incongruent compared to the congruent trials. In addition, the congruency effect was affected by the numerical distance, though in opposite direction for the two comparison tasks. In the numerosity comparison task the effect of congruency increased with decreasing numerical distance while the reverse occurred in the physical size comparison task, a result, also obtained in previous studies (Cohen Kadosh, Cohen Kadosh, Linden, et al., 2007; Schwarz & Heinze, 1998). A tentative explanation for these results is as follows: When subjects have to judge numerical size, numbers with a large numerical distance are easier to compare than numbers with a small numerical distance. Consequently, when number has to be judged, the relative fast responses to the large compared to the small numerical distance trials will receive *less* interference from the unattended physical size dimension. However, when physical size has to be judged, the subject has to suppress the response to the incorrect numerical dimension. Since the large compared to the small numerical distance trials are still easier to discriminate, they are also less easy to suppress. Therefore, large compared to small numerical distance trials induce *larger* interference effects when physical size is compared. These behavioural results cannot be explained on the basis of the relative speed of processing account (Schwarz & Ischebeck, 2003). In the ERP data, an onset latency effect for the sLRP (which is explained below) was present. This finding indicates equal processing time for both magnitude dimen-

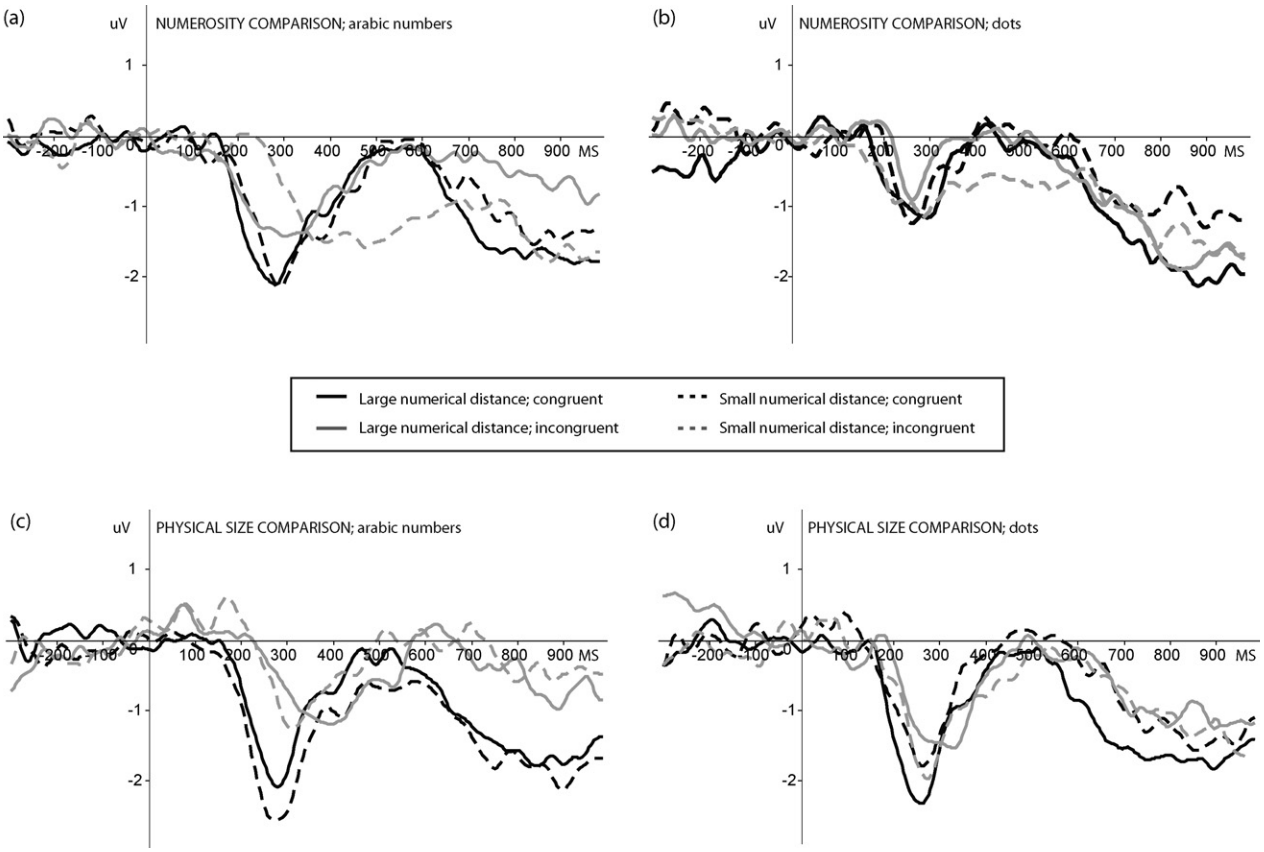

Fig. 3. Grand average stimulus-locked LRPs. The sLRPs elicited by the congruent (black lines) and incongruent (grey lines) trials of the numerosity comparison (upper panel) and the physical size comparison (lower panel) tasks and the Arabic-number notation (left panels) and the dot notation (right panels) conditions. Note, that the numerical distance effects (numerically small condition is dashed line; numerically large condition is solid line) for the incongruent trials are again opposite for both comparison conditions.

sions before the initiation or execution of the response (Lansbergen & Kenemans, 2008). Therefore it is more likely that the relative strength (rather than speed) of the two competing processes determines the degree of interference (Cohen, Dunbar, & McClelland, 1990).

The comparable behavioural patterns for Arabic numbers as well as dots hint at a similar timing of interaction, however the behavioural results cannot differentiate between the early and late interaction account. To be able to distinguish between both hypotheses we looked at the P3 component at a parietal electrode site. The results of the P3 latency nicely mimicked the behavioural congruency effects with its peaks appearing around a later point in time for incongruent compared to congruent trials. Such a congruency effect for P3 latency is indicative of conflict processing at the stimulus evaluation level (De Houwer, 2003; Schmidt & Cheesman, 2005; van Veen & Carter, 2005) and consequently is in agreement with the early interaction account. Interestingly, similar to the behavioural results, the numerical distance effect influenced the P3 latency differentially for both comparison tasks. Larger delays for small numerical distances in the numerosity comparison task and the reverse for the physical size comparison task were obtained.

To allow comparison with previous studies, we also investigated the P3 amplitude, which is indicative of cognitive load (for a review see: Kok, 2001). Larger P3 amplitudes were found for congruent compared to incongruent trials, suggesting an increase in cognitive load for incongruent compared to congruent trials. The P3 amplitude (intensity of processing) and the latency (timing of mental processes) are always discussed separately as if only affected by distinct cognitive processes. Therefore it is not expected that either one of the findings are confounded by the presence of the other (Kok, 2001; Luck, 2005). The result of a conflict at the level of stimulus evaluation as determined by the P3 latency (Schwarz & Heinze, 1998) as well as the effect of cognitive load as reflected by the P3 amplitude are in line with previous studies investigating the symbolic size congruency conflict (Cohen Kadosh, Cohen Kadosh, Linden, et al., 2007; Szucs & Soltesz, 2007, 2008).

As hypothesized, the presence of P3 latency effects should coincide with a delayed onset of the sLRP for incongruent trials and the absence of such a delay for the rLRP. This was indeed the pattern of results obtained. Delayed onset latencies were found for incongruent compared to congruent trials for the sLRP only. These effects were present for both notation conditions when physical size was compared and only for symbolic notation when numerical size was compared. Possibly, the less accurate representation of dots in comparison to Arabic numbers could account for the absence of the sLRP onset latency effect in the dot notation condition. This idea is in agreement with Piazza et al. (2007) showing more accurate neuronal tuning for Arabic numbers compared to dots. The sLRP onset latency effects suggest that numerosity and physical size were processed in equal speed and competed before selective motor preparation started, therefore similarly to the P3 latency effect, the sLRP onset latency effect underscores the notion of an early interaction account.

Whether a general magnitude system that processes distinct magnitudes in a common code exists is a matter of debate (Cohen Kadosh, 2008; Cohen Kadosh & Walsh, 2009). Our results demonstrate that for both symbolic and non-symbolic notation the

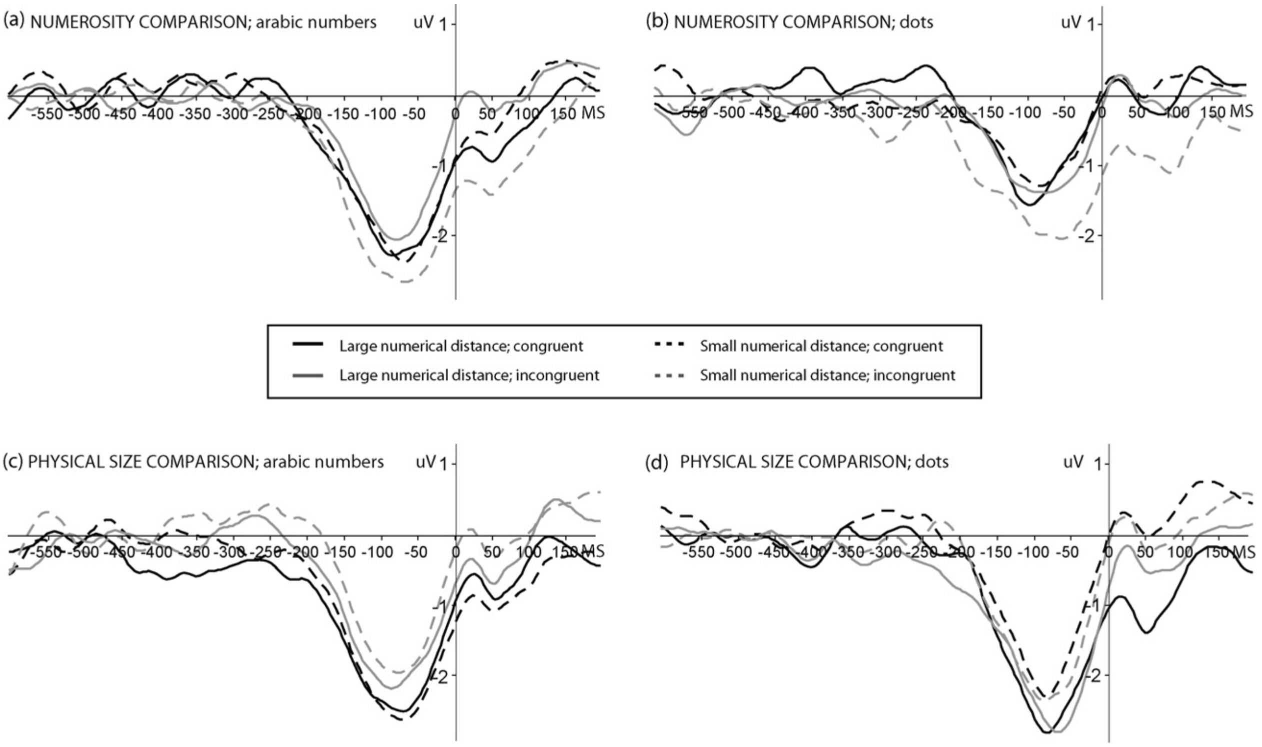

Fig. 4. Grand average response-locked LRPs. The rLRPs elicited by the congruent (black lines) and incongruent (grey lines) stimuli of the numerosity comparison (upper panel) and the physical size comparison (lower panel) tasks with Arabic-number notation (left panels) and the dot notation (right panels). In both tasks and for both notations, no congruency or distance effect was present in the onset latency.

interaction with physical size occurs *before* selective response activation, which is consistent with the early interaction account and thus the notion of a general magnitude system. Although it has been proposed that an early interaction between magnitudes provides direct evidence of abstract coding (Schwarz & Heinze, 1998), colour-word Stroop literature suggests otherwise. Early interactions between colours and words have been obtained in colour-word Stroop paradigms (De Houwer, 2003; Lansbergen & Kenemans, 2008; van Veen & Carter, 2005). Since colours and words are processed in different areas of the brain, these interactions cannot arise due to the convergence onto an abstract neural code in a similar way as is proposed for magnitudes. Instead, it is suggested that these interactions arise at the semantic level between the two semantic codes, and are the result of long-term semantic associations, or of short-term associations induced by task instructions (Cohen et al., 1990; Zhang, Zhang, & Kornblum, 1999). The latter explanation of a conflict induced by response codes was also proposed by Cohen Kadosh, Cohen Kadosh, et al. (2008) and Cohen Kadosh, Henik, et al. (2008), but they explained it in terms of conflict at the response level only (for a similar view see: Gevers, Verguts, Reynvoet, Caessens, & Fias, 2006; Notebaert, Gevers, Verguts, & Fias, 2006).

In conclusion, in our study we elucidate the timing of the interaction between physical size and symbolically and non-symbolically presented numerosity. The presence of P3 peak and sLRP onset latency effects, together with the absence of rLRP onset latency effects suggests that these magnitudes interacted before the preparation or initiation of a response, excluding the hypothesis of a late interaction account. The fact that the early interaction account holds for both the dot and Arabic number notation adds evidence to the hypothesis of a general magnitude system that subserves both symbolic and non-symbolic magnitude, yet alternative explanations that can account for the early interaction effect remain possible.

References

Brannon, E. M. (2002). The development of ordinal numerical knowledge in infancy. *Cognition*, *83*(3), 223–240.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, *97*(3), 332–361.

Cohen Kadosh, R. (2008). Numerical representation: Abstract or nonabstract? *The Quarterly Journal of Experimental Psychology (Colchester)*, *61*(8), 1160–1168.

Cohen Kadosh, R., Cohen Kadosh, K., & Henik, A. (2008). When brightness counts: The neuronal correlate of numerical–luminance interference. *Cerebral Cortex*, *18*(2), 337–343.

Cohen Kadosh, R., Cohen Kadosh, K., Kaas, A., Henik, A., & Goebel, R. (2007). Notation-dependent and -independent representations of numbers in the parietal lobes. *Neuron*, *53*(2), 307–314.

Cohen Kadosh, R., Cohen Kadosh, K., Linden, D. E., Gevers, W., Berger, A., & Henik, A. (2007). The brain locus of interaction between number and size: A combined functional magnetic resonance imaging and event-related potential study. *Journal of Cognitive Neuroscience*, *19*(6), 957–970.

Cohen Kadosh, R., Henik, A., & Rubinsten, O. (2008). Are Arabic and verbal numbers processed in different ways? *Journal of Experimental Psychology. Learning, Memory and Cognition*, *34*(6), 1377–1391.

Cohen Kadosh, R., Henik, A., Rubinsten, O., Mohr, H., Dori, H., van de Ven, V., et al. (2005). Are numbers special? The comparison systems of the human brain investigated by fMRI. *Neuropsychologia*, *43*(9), 1238–1248.

Cohen Kadosh, R., & Walsh, V. (2009). Numerical representation in the parietal lobes: abstract or not abstract? *Behav Brain Sci*, *32*(3–4), 313–328.

De Houwer, J. (2003). On the role of stimulus-response and stimulus-stimulus compatibility in the Stroop effect. *Memory & Cognition*, *31*(3), 353–359.

Dehaene, S. (2001). Is the number sense a patchwork? *Memory & Language*, *16*, 89–100.

Dehaene, S., Naccache, L., Le Clec, H. G., Koechlin, E., Mueller, M., Dehaene-Lambertz, G., et al. (1998). Imaging unconscious semantic priming. *Nature*, *395*(6702), 597–600.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology*, *20*, 487–506.

Diester, I., & Nieder, A. (2007). Semantic associations between signs and numerical categories in the prefrontal cortex. *PLoS Biology*, *5*(11), e294.

Donchin, E. (1981). Presidential address, 1980. Surprise!. . .Surprise? *Psychophysiology*, *18*(5), 493–513.

Dormal, V., Seron, X., & Pesenti, M. (2006). Numerosity–duration interference: A Stroop experiment. *Acta Psychologia*, *121*(2), 109–124.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science*, *8*(7), 307–314.

Fias, W., Lammertyn, J., Reynvoet, B., Dupont, P., & Orban, G. A. (2003). Parietal representation of symbolic and nonsymbolic magnitude. *Journal of Cognitive Neuroscience*, *15*(1), 47–56.

Gebuis, T., Cohen Kadosh, R., de Haan, E., & Henik, A. (2009). Automatic quantity processing in 5-year olds and adults. *Cognitive Processing*, *10*(2), 133–142.

Gebuis, T., Nijboer, T. C., & van der Smagt, M. J. (2009b). Multiple dimensions in bi-directional synesthesia. *European Journal of Neuroscience*, *29*(8), 1703–1710.

Gebuis, T., Nijboer, T. C., & van der Smagt, M. J. (2009a). Of colored numbers and numbered colors. *Experimental Psychology*, *56*(3), 180–187.

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.

Gobel, S. M., Johansen-Berg, H., Behrens, T., & Rushworth, M. F. (2004). Response-selection-related parietal activation during number comparison. *Journal of Cognitive Neuroscience*, *16*(9), 1536–1551.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology*, *55*(4), 468–484.

Jiang, Y., & Kanwisher, N. (2003). Common neural substrates for response selection across modalities and mapping paradigms. *Journal of Cognitive Neuroscience*, *15*(8), 1080–1094.

Jordan, K. E., Suanda, S. H., & Brannon, E. M. (2008). Intersensory redundancy accelerates preverbal numerical competence. *Cognition*, *108*(1), 210–221.

Kaufmann, L., Koppelstaetter, F., Delazer, M., Siedentopf, C., Rhomberg, P., Golaszewski, S., et al. (2005). Neural correlates of distance and congruity effects in a numerical Stroop task: An event-related fMRI study. *Neuroimage*, *25*(3), 888–898.

Kok, A. (2001). On the utility of P3 amplitude as a measure of processing capacity. *Psychophysiology*, *38*(3), 557–577.

Kutas, M., McCarthy, G., & Donchin, E. (1977). Augmenting mental chronometry: The P300 as a measure of stimulus evaluation time. *Science*, *197*(4305), 792–795.

Lansbergen, M. M., & Kenemans, J. L. (2008). Stroop interference and the timing of selective response activation. *Clinical Neurophysiology*, *119*(10), 2247–2254.

Libertus, M. E., Woldorff, M. G., & Brannon, E. M. (2007). Electrophysiological evidence for notation independence in numerical processing. *Behavioral and Brain Functions*, *3*, 1.

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

McCarthy, G., & Donchin, E. (1981). A metric for thought: A comparison of P300 latency and reaction time. *Science*, *211*(4477), 77–80.

Miller, J., Patterson, T., & Ulrich, R. (1998). Jackknife-based method for measuring LRP onset latency differences. *Psychophysiology*, *35*(1), 99–115.

Naccache, L., & Dehaene, S. (2001). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11*(10), 966–974.

Nijboer, T. C., van Zandvoort, M. J., & de Haan, E. H. (2006). Seeing red primes tomato: Evidence for comparable priming from colour and colour name primes to semantically related word targets. *Cognitive Processing*, *7*(4), 269–274.

Notebaert, W., Gevers, W., Verguts, T., & Fias, W. (2006). Shared spatial representations for numbers and space: The reversal of the SNARC and the Simon effects. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(5), 1197–1207.

Piazza, M., Pinel, P., Le Bihan, D., & Dehaene, S. (2007). A magnitude code common to numerosities and number symbols in human intraparietal cortex. *Neuron*, *53*(2), 293–305.

Pinel, P., Dehaene, S., Riviere, D., & LeBihan, D. (2001). Modulation of parietal activation by semantic distance in a number comparison task. *Neuroimage*, *14*(5), 1013–1026.

Pinel, P., Piazza, M., Le Bihan, D., & Dehaene, S. (2004). Distributed and overlapping cerebral representations of number, size, and luminance during comparative judgments. *Neuron*, *41*(6), 983–993.

Reynvoet, B., & Brysbaert, M. (2004). Cross-notation number priming investigated at different stimulus onset asynchronies in parity and naming tasks. *Experimental Psychology*, *51*(2), 81–90.

Reynvoet, B., Brysbaert, M., & Fias, W. (2002). Semantic priming in number naming. *The Quarterly Journal of Experimental Psychology A*, *55*(4), 1127–1139.

Reynvoet, B., Gevers, W., & Caessens, B. (2005). Unconscious primes activate motor codes through semantics. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *31*(5), 991–1000.

Rusconi, E., Kwan, B., Giordano, B. L., Umilta, C., & Butterworth, B. (2006). Spatial representation of pitch height: The SMARC effect. *Cognition*, *99*(2), 113–129.

Schmidt, J. R., & Cheesman, J. (2005). Dissociating stimulus-stimulus and response-response effects in the Stroop task. *Canadian Journal of Experimental Psychology*, *59*(2), 132–138.

Schutter, D. J., de Weijer, A. D., Meuwese, J. D., Morgan, B., & van Honk, J. (2008). Interrelations between motivational stance, cortical excitability, and the frontal electroencephalogram asymmetry of emotion: A transcranial magnetic stimulation study. *Human Brain Mapping*, *29*(5), 574–580.

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.

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*(3), 507–522.

Smulders, F. T., Kenemans, J. L., & Kok, A. (1994). A comparison of different methods for estimating single-trial P300 latencies. *Electroencephalography and Clinical Neurophysiology*, *92*(2), 107–114.

Smulders, F. T., Kok, A., Kenemans, J. L., & Bashore, T. R. (1995). The temporal selectivity of additive factor effects on the reaction process revealed in ERP component latencies. *Acta Psychology (Amsterdam)*, *90*(1–3), 97–109.

Szucs, D., & Soltesz, F. (2007). Event-related potentials dissociate facilitation and interference effects in the numerical Stroop paradigm. *Neuropsychologia*, *45*(14), 3190–3202.

Szucs, D., & Soltesz, F. (2008). The interaction of task-relevant and task-irrelevant stimulus features in the number/size congruency paradigm: An ERP study. *Brain Research*, *1190*, 143–158.

Tang, J., Critchley, H. D., Glaser, D. E., Dolan, R. J., & Butterworth, B. (2006). Imaging informational conflict: A functional magnetic resonance imaging study of numerical stroop. *Journal of Cognitive Neuroscience*, *18*(12), 2049–2062.

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 Science of the United States of America*, *95*(13), 7836–7841.

Ulrich, R., & Miller, J. (2001). Using the jackknife-based scoring method for measuring LRP onset effects in factorial designs. *Psychophysiology*, *38*(5), 816–827.

van Veen, V., & Carter, C. S. (2005). Separating semantic conflict and response conflict in the Stroop task: A functional MRI study. *Neuroimage*, *27*(3), 497–504.

Verguts, T., & Fias, W. (2004). Representation of number in animals and humans: A neural model. *Journal of Cognitive Neuroscience*, *16*(9), 1493–1504.

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Science*, *7*(11), 483–488.

Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*(1), 88–101.

Zhang, H. H., Zhang, J., & Kornblum, S. (1999). A parallel distributed processing model of stimulus–stimulus and stimulus–response compatibility. *Cognitive Psychology*, *38*(3), 386–432.