EJN fens EUROPEAN JOURNAL OF NEUROSCIENCE

European Journal of Neuroscience, Vol. 30, pp. 1999–2008, 2009 doi:10.1111/j.1460-9568.2009.06994.x

COGNITIVE NEUROSCIENCE

The development of automated access to symbolic and non-symbolic number knowledge in children: an ERP study

Titia Gebuis,[1] Inkeri K. Herfs,[1] J. Leon Kenemans,[1] Edward H. F. de Haan[2] and Maarten J. van der Smagt[1]
[1]Experimental Psychology, Helmholtz Institute and Utrecht University, Utrecht, The Netherlands
[2]Social and Behavioral Sciences, University of Amsterdam, Amsterdam, The Netherlands

Keywords: children, development, ERP, SCE, Stroop

Abstract

Infants can visually detect changes in numerosity, which suggests that a (non-symbolic) numerosity system is already present early in life. This non-symbolic system is hypothesized to serve as the basis for the later acquired symbolic system. Little is known about the processes underlying the transition from the non-symbolic to symbolic code. In the current study we investigated the development of automatization of symbolic number processing in children from second (6.0 years) and fourth grade (8.0 years) and adults using a symbolic and non-symbolic size congruency task and event-related potentials (ERPs) as a measure. The comparison between symbolic and non-symbolic size congruency effects (SCEs) allowed us to disentangle processes necessary to perform the task from processes specific to numerosity notation. In contrast to previous studies, second graders already revealed a behavioral symbolic SCE similar to that of adults. In addition, the behavioral SCE increased for symbolic and decreased for non-symbolic notation with increasing age. For all age groups, the ERP data showed that the two magnitudes interfered at a level before selective activation of the response system, for both notations. However, only for the second graders distinct processes were recruited to perform the symbolic size comparison task. This shift in recruited processes for the symbolic task only might reflect the functional specialization of the parietal cortex.

Introduction

A commonly held view is that humans are endowed with 'number sense', an innate ability to work with non-symbolic numerosity. This view is based on many studies that claim that infants can detect changes in numerosity and perform simple calculations (e.g. Feigenson *et al.*, 2002; Lipton & Spelke, 2003; Wood & Spelke, 2005; Brannon *et al.*, 2007). The ability to work with non-symbolic numerosity is assumed to be a precursor for future number understanding. Children have been reported to be able to apply their non-symbolic mathematical knowledge onto symbolic math problems (Barth *et al.*, 2005), and the accuracy with which they can discriminate numerosities has been shown to correlate with performance on arithmetic at a later age (Halberda *et al.*, 2008). The idea of non-symbolic number knowledge as the precursor for symbolic number knowledge fits the hypothesis that the neural mechanisms that first subserve non-symbolic numerosities will become responsive to the later acquired number symbols. This merging of codes is the result of repeatedly associating a quantity with the number symbol it relates to. A strong association between the number symbol and its meaning results in the ability to automatically access symbolic number knowledge, which is a necessary prerequisite for complex mathematical procedures.

Correspondence: Dr T. Gebuis, as above.
E-mail: T.Gebuis@uu.nl

Received 7 July 2009, revised 24 August 2009, accepted 16 September 2009

Automatic access to symbolic number knowledge is often studied using the symbolic size congruency task (Henik & Tzelgov, 1982; Algom *et al.*, 1996). In this paradigm, two Arabic number stimuli are presented simultaneously. Each stimulus consists of two dimensions (e.g. numerical and physical size) that are manipulated independently, resulting in congruent (e.g. 3 8), incongruent (e.g. 3 8) or neutral (e.g. 3 3 or 3 8) trials. When subjects have to respond to one dimension, the other (unattended) dimension interferes only if it is automatically activated. Congruency effects obtained when subjects have to respond to the physical size of the Arabic numbers are indicative of automated access to number symbol meaning. Five-year-old children that just acquired knowledge of the Arabic numbers have automatic access to non-symbolic (e.g. arrays of dots), but not symbolic number knowledge (Gebuis *et al.*, 2009). These results implicate that the lack of automated access to number symbol meaning cannot be attributed to premature cognitive processes (e.g. inhibitory or attentional) that are necessary to perform the task. Instead, automated access to symbolic number knowledge may gradually develop with increasing age (Girelli *et al.*, 2000; Rubinsten *et al.*, 2002; Mussolin & Noel, 2007).

Passively looking at non-symbolic numerosities has been reported to result in the activation of similar parietal areas in 4-year-old children and adults (Piazza *et al.*, 2004; Cantlon *et al.*, 2006). Moreover, 5-year-old children and adults revealed similar electroencephalogram (EEG) activation patterns during the comparison of non-symbolic as well as symbolic number stimuli (Temple & Posner, 1998). However, neuroimaging studies specifically addressing symbolic number processing reported the opposite: children who were familiar with the number symbols relied more on frontal than parietal processes, compared with adults, while performing a symbolic number task (Ansari *et al.*, 2005; Rivera *et al.*, 2005; Kaufmann *et al.*, 2006). This possibly reflects the transition from a symbolic to non-symbolic numerosity system. A neural mechanism underlying this transition was suggested in a recent primate study (Diester & Nieder, 2007). In this study, Rhesus monkeys were trained to learn the association between the number symbol and the number of dots the symbol represents. Neurons in the prefrontal cortex were active for both the symbolic and non-symbolic code, whereas neurons in the parietal cortex only responded to non-symbolic notation. This result is in accordance with the idea that the frontal cortex mediates recently acquired associations; a shift towards parietal neurons can only be expected after substantial training.

To disentangle the processes subserving the acquisition of the symbolic number code in children, the current study compared children (second grade, 6.0 years) who had just become acquainted with symbolic numerosity, children (fourth grade, 8.0 years) who had worked with symbolic numerosity for about 2 years, and adults on measures of the size congruency effect (SCE) for symbolic as well as non-symbolic numerosity. We included both size congruency tasks with distinct number notations to disentangle task-related processes (e.g. working memory or inhibition) from processes specific to number notation. To investigate the neural mechanisms subserving the automatization process, we measured event-related potentials (ERPs) during both experiments. To our knowledge this is the first study about the development of automated access to number meaning using ERPs that includes children that just learned the Arabic numbers as well as a symbolic and non-symbolic size congruency task.

First, we questioned whether the SCE develops in a similar manner for symbolic and non-symbolic notation. Second, we questioned whether developmental differences in the interaction between the various numerosity dimensions reflect differences in facilitation or rather interference between dimensions. Interference effects are generally reported to arise at an earlier age than facilitation processes (Rubinsten *et al.*, 2002). The presence of both interference and facilitation would suggest full automatic number processing (Tzelgov *et al.*, 1992). Third, we explored these developmental differences with respect to more general conflict-processing mechanisms, as it may be suspected that children rely on these mechanisms in a manner that differs from adults. Specifically, we asked whether behavioral SCEs are paralleled by SCEs in the ERP signal during stimulus-evaluation time and precede selective response activation (Cohen Kadosh *et al.*, 2007; Gebuis *et al.*, in press). A differential involvement of neural processes in the younger vs. the older children or adults in the symbolic but not the non-symbolic task would reveal the neural processes involved in the initial stages of the automatization of symbolic number knowledge.

To investigate the timing of the interaction between the two manipulated magnitudes, we looked at the P3 latency, reflecting stimulus-evaluation time (Luck, 2005), and the lateralized readiness potential (LRP), reflecting the response processes involved. Together these components give a full perspective upon the conflict-processing mechanisms involved in the size congruency task. A later onset of the P3 peak as well as the stimulus-locked LRP for incongruent than congruent trials is indicative of a conflict at the semantic level that arises 'before' response selection or execution. In contrast, a later onset of the response-locked LRP for incongruent than congruent trials suggests a conflict that arises 'at' the response selection or execution stage (Kutas *et al.*, 1977; Donchin, 1981; McCarthy & Donchin, 1981; Smulders *et al.*, 1995).

Materials and methods

Participants

Three groups of subjects participated in the experiment: (i) 19 second-grade children, of which 15 were included in the analyses (M = 6.0 years, SD = 0.26, 10 males); (ii) 23 fourth-grade children, of which 19 were included in the analyses (M = 8.1 years, SD = 0.36, eight males); and (iii) 20 adults, of which 17 were included in the analyses (M = 22.5 years, SD = 1.52, six males). The data of the four second-grade children, the four fourth-grade children and the three adult subjects were discarded because more than 25% of their EEG epochs contained artifacts (for further detail, see: Behavioral and ERP analyses). The children and adults were recruited from the Utrecht area in the Netherlands. As indicated by school grades, the children included in this study revealed different levels of performance (from below average to excellent), and the adults had completed different levels of higher education. The children received presents for participation while the adults were paid. Only children that had sufficient symbolic number knowledge as indicated by their performance on a symbolic number comparison task were selected. All participants were native Dutch speakers, and had normal or corrected-to-normal vision. Written informed consent was obtained, and the study was performed according to the Declaration of Helsinki guidelines, and approved by the Ethical Committee of the Utrecht University.

Apparatus, stimuli and procedure

For the adults, the stimuli were displayed on a 22-inch CRT monitor using the Presentation software (Neurobehavioral Systems, Albany, CA, USA). The paradigm was similar to a previous study (Gebuis *et al.*, 2009). For the children, we used a 17-inch Asus laptop.

For the symbolic size congruency 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 (1.7° visual angle) or a large font size (2.4° visual angle). For the non-symbolic size congruency task, the stimuli consisted of groups of dots ranging from 7 to 16, which were randomly distributed within a pre-specified area (width and height of 3.05° visual angle). This relatively large number of dots was used to rule out possible subitizing effects. Depending on the condition, small (0.38° visual angle in diameter) or large dots (0.53° visual angle in diameter) were presented. 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). Because numerosity can be directly inferred from symbolic number notations but only for the numerosities 1–4 for the non-symbolic notation (Mandler & Shebo, 1982), using the same numerical distances for both tasks would lead to a difference in task difficulty. Therefore, larger numerical distances were used in the non-symbolic task compared with the symbolic task, which resulted in comparable error rates and response times (Fig. 1A and C). Furthermore, to exclude the possibility that the participant derived the correct answer on the basis of visual sensory properties, these were controlled for in several ways (see Gebuis *et al.*, 2009 for details). The stimuli were presented centrally on a gray background and the viewing distance was approximately 57 cm. In order to keep the children focused, we added 10 cartoon pictures, which appeared at pre-defined intervals throughout the experiment. The children were told that the task ended as soon as they 'found' the 10th picture, which was always presented at the end of each task.

For both the symbolic and non-symbolic comparison task, subjects were instructed to respond to the physical size of the stimuli by pressing the button corresponding to the side of the physically larger stimulus. The correct answer appeared on the left side in half and on the right side in the other half of the trials. Both tasks consisted of a congruent (e.g. 3 8; 96 trials), an incongruent (e.g. 3 8; 96 trials) and a neutral condition (e.g. 3 3; 48 trials). The congruent and incongruent condition consisted of (48) small and (48) large numerical distance trials. For each age group, the order of the two comparison tasks was counterbalanced between participants. Each trial began with a random inter-trial interval (1250–1500 ms), followed by a fixation cross (500 ms) and the stimulus (until response). Prior to each comparison, task participants received instructions and performed 15 practice trials. After each comparison task participants could take a break.

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; for an explanation about the system see http://www.biosemi.com or Schutter *et al.*, 2008). The vertical electro-oculogram (EOG) was recorded from electrodes attached above and below the left eye, and the horizontal EOG from the outer canthi of both eyes.

ERP preprocessing

Participants were discarded from both the ERP and the behavioral analyses when more than 25% of the trials contained artifacts or an incorrect response. This resulted in the exclusion of four second-grade and four-fourth-grade children, and three adults. EEG and EOG data were analysed using Brain Vision Analyzer software (1.05; Brian Vision, Munich). Flat as well as noisy electrodes were discarded and consequently not included in the analyses. EEG signals were off-line re-referenced to the average of all electrodes included in the analyses. The continuous EEG data were segmented into epochs from 100 ms prior until 1200 ms after target presentation. Epochs were filtered with a bandpass filter (0.3 Hz, 12 dB octave; 40 Hz, 24 dB/octave; 50 Hz Notch Filter) and corrected for eye movements according to the Gratton *et al.* (1983) algorithm. Trials with artifacts (maximum or minimum of ± 125 μV for the children and ± 80 μV for the adults) or an incorrect response were rejected from further analyses. The baseline for the stimulus-locked ERPs was defined as the mean of the 100-ms period before the onset of the target stimulus. The baseline for the response-locked LRP was −550 to −450 ms. Grand average ERPs and LRPs were created for each condition and filtered at 8 Hz, 12 dB/octave for visualization purposes only.

Behavioral and ERP analyses

For the behavioral data, accuracy and the median reaction times of the correct responses were calculated. For the children, trials with action times below 200 ms and above 1800 ms were discarded to deal with outliers due to random button presses or talking. This was not applicable to the data of the adults. For the ERP analyses, we especially looked at the P3 latency and the LRP. A later onset of the P3 peak as well as the stimulus-locked LRP is indicative of a conflict that arises before response selection or execution (Kutas *et al.*, 1977; Donchin, 1981; McCarthy & Donchin, 1981; Smulders *et al.*, 1995; Gebuis *et al.*, in press).

We estimated the P3 component at the Pz electrode, which is defined as the largest peak within the time window of 300–800 ms for adults, but 300–1000 ms for the children as they revealed a much larger spread in the timing of the P3 component. The stimulus- and response-locked LRPs (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]. [One out of the 19 fourth graders was not included in the LRP analyses due to a missing electrode (C4).] As recommended by Miller *et al.* (1998), the relative criterion method was applied to each subsample waveform using the 50% criterion to obtain the sLRP onset and 90% for the rLRP onset.

As previous imaging studies revealed the reliance of children on additional frontal processes during the performance of a symbolic congruency task, a similar dissociation between children and adults was expected in our study. To identify possible congruency effects at components other than the parietal P3 component, difference waves of congruent vs. incongruent trials were inspected. A clear congruency effect appeared present in the symbolic data of the second graders at frontal electrode sites with the strongest effects being left lateralized, whereas a smaller congruency effect in the opposite direction was apparent at the right frontal electrode sites in the non-symbolic data. Due to flat or noisy electrodes in some of the subjects, only the electrodes Fz, FCz and FC3 were included in the analyses of the left frontal electrode cluster. Although this results in a decrease in statistical power, because of the slightly smaller electrode cluster taken into account, this should not yield any other problems. For the right frontal electrode cluster, the electrodes F2, FC2 and FC4 were included in the analyses. The congruency effects occurred about 200 ms at a positive component, the P2 component, which was quantified as the largest peak within the time window of 180–300 ms. For the fourth graders and adults, no congruency effects appeared present at the P2 component (or any other component besides the P3), therefore no comparable analyses were performed for these age groups.

A substantial difference can be expected in the time needed to process the information in children compared with adults. One way to account for this difference in both the behavioral and electrophysiological data for each age is by dividing the (average reaction time and average electrophysiological) result of the congruent and incongruent conditions by those of the neutral condition for each participant separately. In this manner, any interaction with the group cannot be explained by the overall difference in the timing of the response or electrophysiological process. To test whether a congruency effect was present for each group and whether this congruency effect was distinct for symbolic compared with non-symbolic notation, we performed a repeated-measures ANOVA with notation (symbolic/non-symbolic) × congruency (congruent/incongruent) × distance (small/large) as within-subject factors, and group (second grade/fourth grade/adults) as a between-subjects factor. If present, three- or four-way interactions were further investigated using a simple effects analysis for each task separately. This analysis looks at the effect of

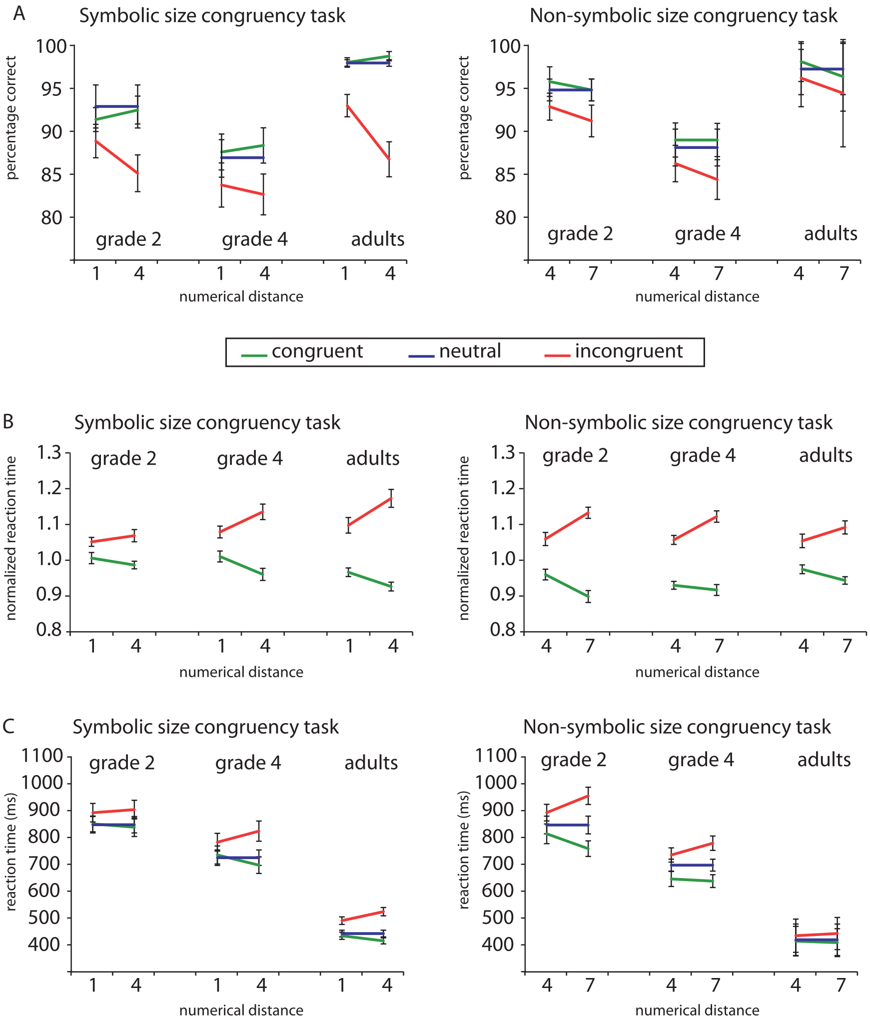

FIG. 1. The accuracy results (A), the normalized reaction time results (B) and the reaction time results (C) of the symbolic (left panels) and the non-symbolic (right panels) size congruency task. Within each panel, the results of the second- and fourth-grade children and adults are presented from left to right. A congruency effect is present for both tasks and each age group showing fewer errors (A) and faster responses (B and C) for the congruent (green lines) compared with the incongruent (red lines) trials. Of special interest is the fact that the congruency effect increased with increasing age for the symbolic size congruency task but decreased with increasing age for the non-symbolic size congruency task (B).

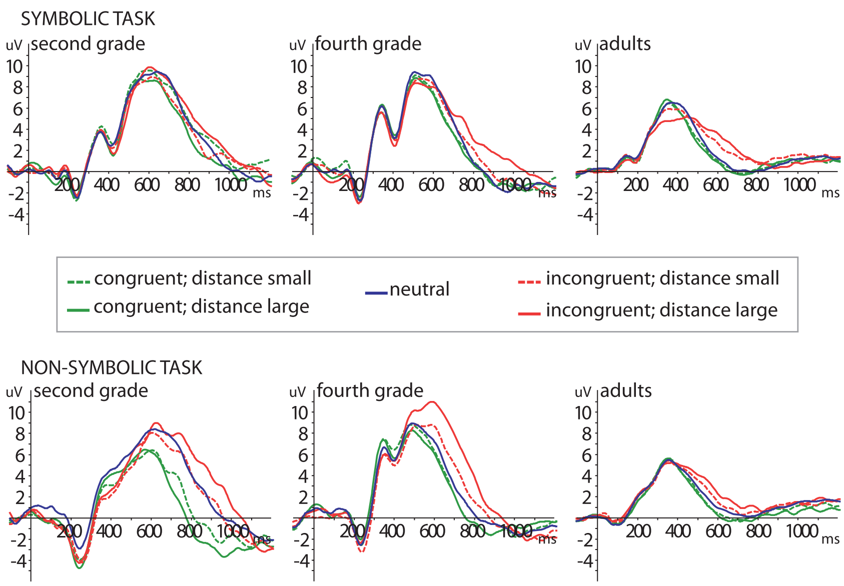

FIG. 2. The P3 component at the Pz electrode. From left to right the ERP grand average waveform at the Pz electrode is presented for second and fourth graders and adults. The upper panels reveal the results of the symbolic congruency task whereas the lower panels represent the results of the non-symbolic task. Except for second graders on the symbolic size congruency task, the P3 peaked around a later point in time in the incongruent compared with the congruent condition. Similar to the behavioral results, the congruency effect as reflected by P3 latency increased with increasing age for symbolic notation but decreased with increasing age for non-symbolic notation.

one independent variable at individual levels of the other independent variable (Keppel, 1991; Field, 2005). The presence of congruency effects led us to perform *post hoc* paired samples *t*-tests between neutral and congruent or incongruent trials to investigate whether the congruency effect is the result of a facilitation and/or interference, respectively.

TABLE 1. The congruency effects, averaged over distance, for the different behavioral and electrophysiological measures

	Size congruency effects (incongruent – congruent)		Facilitation (neutral – congruent)		Interference (incongruent – neutral)	
	Symbolic	Non-symbolic	Symbolic	Non-symbolic	Symbolic	Non-symbolic
Accuracy (percentage correct)						
2nd grade	−5*	−3*	1	−1	−6*	−3*
4th grade	−5	−4*	−1	−1	−4*	−3*
Adults	−9*	−3*	−1	0	−8*	−2
Reaction time (ms)†						
2nd grade	54* (6)	138* (17)	3	60*	51*	77*
4th grade	87* (12)	115* (17)	9	56*	78*	60*
Adults	83* (19)	47* (11)	24*	17*	59*	29*
P3 latency (ms)						
2nd grade	4	85*	15	53*	−10	32
4th grade	31*	54*	10	5	21*	49*
Adults	49*	42*	18	30*	30*	12*
sLRP onset (ms)						
2nd grade	68*	112*	45	16	23	96
4th grade	63*	114*	−3	−12*	66	126*
Adults	68*	69*	5*	1*	63	68*

*Results which are significant or reveal a trend towards significance ($P < 0.06$). †The congruency effects of the normalized data are shown in parentheses. sLRP, stimulus-locked lateralized readiness potential.

Results

An overview of the SCEs, as well as facilitation and interference effects for the behavioral and ERP measures can be found in Table 1.

Results number comparison task second graders

The number comparison task was included as a prerequisite for inclusion in the analyses, as it shows whether the children have sufficient knowledge of the number symbols. One child responded at chance level and was discarded from further analyses. The remaining children ($N = 15$) showed an accuracy score of 92.5%. Moreover, they were faster to respond ($t_{1,14} = 6.397$, $P < 0.001$) and were more accurate ($t_{1,14} = -3.833$, $P = 0.002$) in large compared with small numerical distance trials (1347 ms/98% vs. 1791 ms/87%, respectively). These results are indicative of the presence of symbolic number knowledge. Consequently, any absence of a congruency effect in the behavioral or ERP data on the symbolic comparison task can not be due to insufficient knowledge of the number symbols. Instead, a not yet fully automated link between the number symbols and their meaning would be a more plausible explanation.

Accuracy results (Fig. 1A)

Congruency effects

A significant main effect was obtained for notation, distance as well as congruency ($F_{1,48} > 14.460$, $P < 0.001$ for all effects). In addition, except for the two-way interaction between notation and distance ($F_{1,48} = 0.011$, $P = 0.916$), all remaining two-way interactions (notation × congruency; notation × group; congruency × distance; congruency × group; distance × congruency) and all three-way interactions (notation × congruency × distance; notation × congruency × group; notation × distance × group; congruency × distance × group) were significant ($F_{2,48} > 3.243$, $P < 0.048$ for all mentioned interactions). To disentangle the mechanisms underlying these two- and three-way interactions, simple effects analyses were performed. The obtained results showed a congruency effect for all age groups for the symbolic task ($F > 4.12$, $P < 0.046$). In contrast, a congruency effect for the non-symbolic task was only significantly present for the adults ($F_{1,16} = 4.36$, $P = 0.041$), and showed a trend towards significance for the second graders ($F_{1,14} = 3.91$, $P = 0.053$) but not for the fourth graders ($F_{1,18} = 3.05$, $P = 0.085$). The SCE in the symbolic and non-symbolic tasks imply that both adults and children made more errors in the incongruent compared with the congruent condition. Besides congruency effects, the results also revealed a symbolic distance effect for the adults ($F_{1,16} = 44.28$, $P < 0.001$) as well as an interaction between distance and congruency ($F_{1,16} = 7.69$, $P = 0.007$), suggesting larger congruency effects with increasing numerical distance.

Facilitation and interference effects

For the symbolic task, an interference ($t > 3.202$, $P < 0.005$), but no facilitation ($t > -1.387$, $P > 0.182$), effect was present for all age groups and both numerical distances. For the non-symbolic task, the results revealed a significant interference effect for the second and fourth graders ($t > 2.489$, $P < 0.023$), but not the adults ($t = 1.546$, $P = 0.142$). Similar as in the symbolic task, facilitation effects were absent for all age groups ($t > -0.871$, $P > 0.395$ for all comparisons).

Reaction time results (Fig. 1B and C)

Congruency effects

The main effects for notation, congruency and distance were all significant ($F_{1,48} > 4.383$, $P < 0.042$). In addition, the results revealed a two-way interaction between distance and congruency ($F_{1,48} = 125.792$, $P < 0.001$), a triple interaction between notation, congruency and group ($F_{2,48} = 11.489$, $P < 0.001$), and a four-way interaction between congruency, notation, distance and group

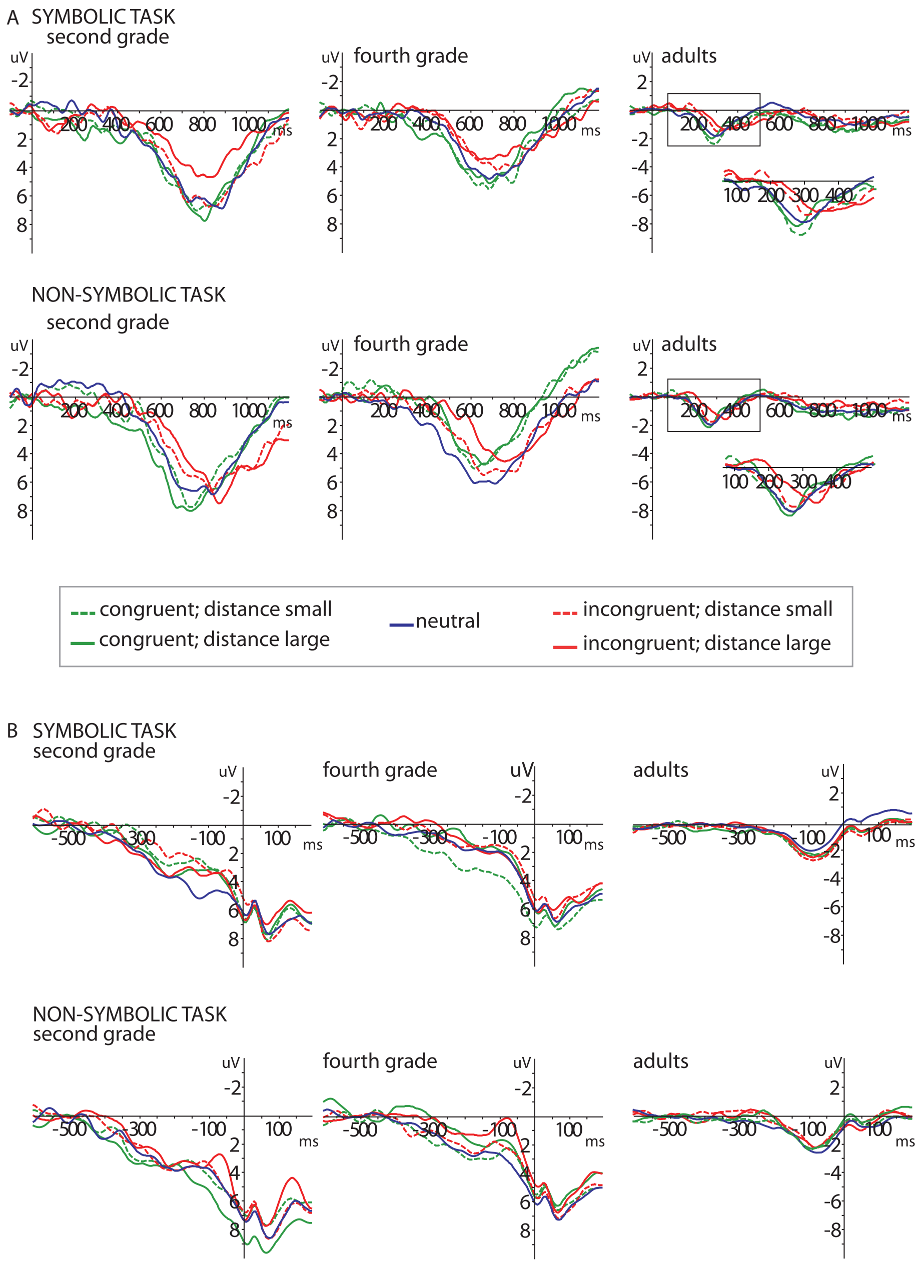

FIG. 3. The sLRPs (A) and rLRPs (B). The sLRPs showed an onset latency effect in both the symbolic and non-symbolic task for all age groups. For the adults a close-up of the onset latency effect of the sLRP is given as well (A). The onset of the incongruent condition started around a later time point compared with the congruent condition. The rLRPs do not reveal an onset latency effect in the symbolic and non-symbolic task for any of the age groups (B). It should be noted that in contrast to the adults, the axes of the data of the children have negative up and positive down. Such positive instead of negative going LRPs have been shown before (Szucs *et al.*, 2007). The cause of this effect is beyond the scope of this article and will be addressed elsewhere.

($F_{2,48} = 6.145$, $P = 0.004$). Subsequent simple effect analyses were performed to disentangle the mechanisms subserving these interactions. The results revealed a significant SCE for all age groups in the symbolic ($F > 20.26$, $P < 0.001$) as well as non-symbolic size congruency task ($F > 51.68$, $P < 0.001$), implicating faster responses to congruent compared with incongruent trials. Moreover, an interaction between congruency and distance was shown for the symbolic size congruency task for the fourth graders and adults only ($F > 7.89$, $P < 0.006$), and for the non-symbolic size congruency task for all age groups ($F > 5.36$, $P < 0.024$). The presence of an interaction between congruency and distance suggests that the congruency effect increases with increasing distance.

Facilitation and interference

For both the symbolic and non-symbolic tasks, interference effects were present in all age groups ($t < -4.412$, $P < 0.001$). In contrast to interference, the facilitation effect of the symbolic task was significant for adults only ($t_{16} = 4.5325$, $P < 0.001$), whereas the facilitation effect of the non-symbolic task appeared present for all age groups ($t > 3.534$, $P < 0.003$). The results of the symbolic task reveal the emergence of interference before facilitation, whereas the results of the more familiar non-symbolic notation show interference and facilitation effects already present in second graders. Surprisingly, the normalized results reveal that the SCE increased for symbolic notation but decreased for non-symbolic notation with increasing age (Fig. 1B), as reflected by a trend towards significance for the triple interaction between congruency, task and group ($F_{2,48} = 3.105$, $P = 0.056$).

P3 latency analyses of the congruency tasks (Fig. 2)

Congruency effects

A significant main effect for congruency ($F_{1,48} = 69.117$, $P < 0.001$), an interaction between congruency and notation ($F_{1,48} = 10.921$, $P = 0.002$), and a triple interaction between notation, congruency and group ($F_{2,48} = 6.337$, $P = 0.004$) were present. Simple effects analyses were performed to disentangle the two- and three-way interactions. For the symbolic task, a congruency effect was present for the fourth graders and adults ($F > 5.16$, $P < 0.026$), whereas results for the non-symbolic task showed congruency effects for all age groups ($F > 8.88$, $P < 0.04$). These results indicate that stimulus-evaluation time as reflected by the P3 latency is affected by the congruency effect. The P3 peak amplitudes appeared at a later point in time for incongruent compared with congruent trials, irrespective of notation for the fourth graders and adults. The lack of a congruency effect for the second graders on the symbolic task only ($F_{1,14} = 0.18$, $P = 0.675$) is surprising considering the presence of a congruency effect for symbolic notation at the behavioral level.

Facilitation and interference

For the second graders, not surprisingly, no facilitation ($t_{14} = 0.937$, $P = 0.365$) or interference ($t_{14} = 0.670$, $P = 0.514$) was found for the symbolic notation (note that no SCE was present for P3 latency of second graders on the symbolic task). In contrast, a significant facilitation ($t_{14} = 2.366$, $P = 0.033$) but no interference effect ($t_{14} = -1.807$, $P = 0.092$) was obtained for the non-symbolic notation. For the fourth graders no facilitation ($t_{18} < 2.366$, $P > 0.384$) but only interference effects were present for the symbolic ($t_{18} = -2.160$, $P = 0.044$) and the non-symbolic task ($t_{18} = -3597$, $P = 0.002$). For the adults a significant facilitation effect was present for the non-symbolic ($t_{16} = 2.165$, $P = 0.046$) but not the symbolic notation ($t_{16} = 1.592$, $P = 0.131$), whereas a (trend towards a) significant interference effect was present for both the non-symbolic ($t_{16} = -3.746$, $P = 0.002$) and symbolic notation ($t_{16} = -2.108$, $P = 0.051$). Together, interference effects appear to be the dominant cause of the P3 latency effects on both the symbolic and non-symbolic task for fourth graders and adults.

sLRP and rLRP onset latency analyses (Fig. 3A and B, respectively)

Congruency effects

For the sLRP the SCE was significant ($F_{1,47} = 81.257$, $P < 0.001$). In addition, a significant interaction between group and congruency ($F_{2,47} = 3.924$, $P = 0.027$), notation and congruency ($F_{1,47} = 6.003$, $P = 0.018$), as well as distance and congruency ($F_{1,47} = 15.726$, $P < 0.001$) were obtained. To indicate whether the congruency effects were present for the distinct groups, notation and distances, respectively, Bonferroni-corrected *post hoc* pairwise comparisons were performed. These results revealed a significant congruency effect for each age group, each notation condition and each numerical distance ($P < 0.001$ for all of the comparisons). Together, it can be concluded that the onset of the sLRP was delayed for incongruent compared with congruent trials, for non-symbolic compared with symbolic notation, and for large- compared with small-distance trials. The congruency effect for non-symbolic notation was larger for the children compared with that of the adults.

In contrast to the sLRP, no significant results were obtained for the rLRP onset latency data ($F < 2.353$, $P > 0.10$ for all of the comparisons).

Congruency effects at left frontal electrode sites (second-grade children; Fig. 4)

Only for the second graders, the P3 latency effect was absent for the symbolic notation. sLRP onset latency effects, however, suggest that the conflict (as manifested in the behavioral data) did arise at a premotor level. Moreover, previous studies have reported the activation of frontal instead of, or in addition to, parietal processes in children during a symbolic comparison task (Ansari *et al.*, 2005; Rivera *et al.*, 2005; Kaufmann *et al.*, 2006). Together, it appears likely that in our study the children recruited distinct processes in the symbolic size congruency task as well. Difference waves of the incongruent – the congruent condition revealed SCEs at about 200 ms mainly at the left frontal electrode sites for the symbolic task and at right frontal electrode sites for the non-symbolic task (P2 component), but only for second graders. For the ‘symbolic task’, the SCE was found to be significant at both the right ($t_{14} = 7.290$, $P = 0.017$) and left ($t_{14} = 5.089$, $P = 0.041$) frontal electrode cluster. In contrast, for the ‘non-symbolic task’, the results did not reveal a significant congruency effect at neither the left ($t_{14} = 0.775$, $P = 0.394$) nor the right electrode cluster ($t_{14} = 1.212$, $P = 0.298$). Thus, SCEs at components other than the P3 component were present for the symbolic task only. It should be noted that the difference waves for the fourth graders and adults did not reveal any SCE at components other than the parietal P3 component.

Discussion

In this study we sought to identify the neural mechanisms underlying the development of automated symbolic number knowledge. We first asked whether symbolic and non-symbolic numerosity processes develop in a similar manner. A second question was whether the developmental differences in the interaction between the various numerosity dimensions reflect differences in facilitation or rather interference between dimensions. Third, we questioned whether the tested SCEs reflect similar neural interactions in children as in adults.

To summarize the results: on the behavioral level, (1a) for all age groups and both tasks a congruency effect was present. (1b) In the symbolic task, the congruency effects consisted mainly of an interfering effect; the facilitatory component became more apparent with increasing age. (1c) In the non-symbolic task, the congruency effect consisted of both a facilitatory and interfering component. (1d) The congruency effect increased with increasing age for the symbolic task but decreased with increasing age for the non-symbolic task. On the electrophysiological level, (2a) except for the second graders on the symbolic task, all age groups revealed a P3 latency effect on both tasks. (2b) Similar to the reaction time data, the congruency effect consisted mainly of an interference component. (2c) SCEs were only present for sLRP, not for rLRP onset latencies. (3) Of special interest was the recruitment of frontal processes (P2 component) by second graders in the symbolic task only.

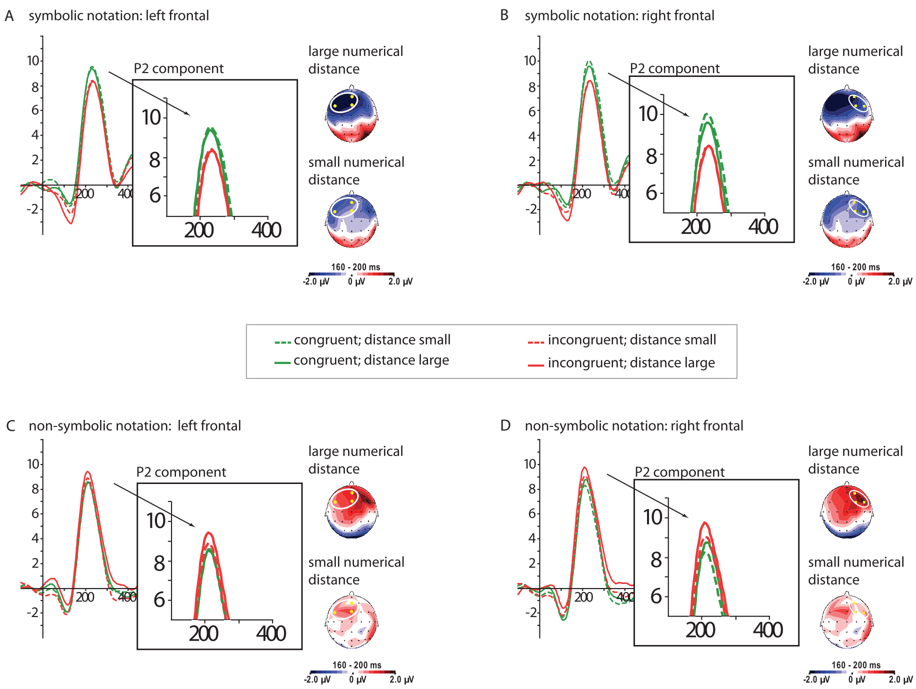

FIG. 4. The scalp maps and ERP data of frontal electrode sites of second-grade children. Difference waves (incongruent – congruent) revealed a large frontal, mainly left lateralized SCE in the symbolic (upper panels), but a smaller and right lateralized SCE in the non-symbolic task (lower panels). Only for the symbolic task, the congruency effect was present at the P2 component for both electrode clusters. No significant SCEs were present for the non-symbolic task.

The development of the SCE at the behavioral level

The behavioral congruency effect in both the symbolic and non-symbolic size congruency task for all age groups suggests the presence of automated number knowledge already for the second graders (6.0 years). Such an early SCE for both notations has been reported for Chinese children (Zhou *et al.*, 2007; 5.8 years), but not for European children (Girelli *et al.*, 2000; 8.3 years; Rubinsten *et al.*, 2002; 7.3 years). The large age range reported at which automatic access to number symbol meaning emerges has led to suggestions about the cause of this diversity. Zhou *et al.* (2007) proposed that Chinese children have several cultural advantages: Chinese numbers need less pronunciation time and are more incorporated in daily life (e.g. days of the week). The second graders in our current and in our previous study were recruited from the same city. In contrast to our previous study (Gebuis *et al.*, 2009), the children of the current study revealed congruency effects, a discrepancy in results, which clearly cannot be explained by cultural differences. We propose instead that the differences in onset of automatic access to number symbol meaning are highly related to informal or parental schooling in addition to formal schooling. Another relevant factor may be that the present second graders were on average 5 months older compared with our previous study, yet the discrepancy with the above-mentioned Girelli *et al.* (2000) and Rubinsten *et al.* (2002) studies remains.

When number knowledge is fully automated, a congruency effect consists of both a facilitatory and an inhibitory component (Henik & Tzelgov, 1982). In the symbolic task, the interference component was present, but the facilitatory component gradually developed with age. In contrast, in the non-symbolic task, both facilitatory and inhibitory effects were present from second grade onwards. The later emergence of facilitatory effects in the symbolic task indicates that the link between the number symbol and its meaning is not yet fully automated in children that just gained understanding of the number symbols. This notion is further emphasized by the fact the youngest children in our study did not show an interaction between congruency and distance in the symbolic task, while this interaction was apparent in the non-symbolic task. Together, even though all age groups revealed a congruency effect for both the symbolic and non-symbolic task, the lack of a facilitatory effect for second graders hints at a symbolic automatization process that is not fully developed.

The neurophysiological correlates underlying the SCE at distinct ages

While the children and adults revealed comparable behavioral results on the symbolic and non-symbolic size congruency task, the electrophysiological correlates revealed a clear distinction for the second graders only. For both fourth graders and adults, a SCE was present for P3 latency irrespective of notation, whereas second graders revealed a congruency effect for non-symbolic notation only. For all age groups, however, a sLRP and no rLRP onset latency effect was apparent for both notations, which implies that the conflict between the two magnitudes (size and numerosity) arose before motor processes were activated, and therefore at an 'early stimulus level'. Thus, for second graders, the SCE present in the reaction time data of the symbolic task was not paralleled by a P3 latency effect even though the conflict between both magnitudes arose at a pre-response level. Instead of a P3 latency effect in the symbolic task, a congruency effect was present at the amplitude of the frontal P2 component for these children. The frontal P2 component is suggested to relate to the evaluation of task-relevant stimuli but not to orienting of attention or preparation and execution of motor responses (Potts, 2004), and is as such suggested to be comparable to the P3 wave (Luck, 2005). This frontal congruency effect could only have arisen when number was accessed automatically (as the task was to judge physical size), and therefore can explain the congruency effects obtained at the behavioral level.

The increase in symbolic SCE but the decrease in non-symbolic SCE

Even though the behavioral and electrophysiological congruency effects were present for all age groups as well as the symbolic and non-symbolic task, the effects appeared to increase with increasing age for the symbolic notation while the opposite pattern appeared for non-symbolic notation. This is a remarkable effect as it is argued that the symbolic congruency effect decreases with age as a result of developing inhibitory mechanisms (Szucs *et al.*, 2007). These contradictory results can easily be explained. In contrast to the Szucs *et al.* (2007) study, we normalized our reaction times to account for the overall difference in response times between children and adults. Normalization of the data of the Szucs *et al.* (2007) study results in a comparable increase in interference effect as obtained in our study.

However, the question remains what caused, in our study, the opposite patterns in the development of the symbolic and non-symbolic SCE. One perspective is that until the age that children are taught about the symbolic numbers they are mainly confronted with non-symbolic notation, but as soon as the symbolic notation is mastered this notation is predominantly used, as it allows more precise calculations and is necessary for everyday activities. The lack of continued training and thus the decreased utilization of the structures subserving non-symbolic numerosity processing could well account for the reduction of non-symbolic SCE with increasing age. This hypothesis is in agreement with a study reporting reduced adults' sensitivity to numerosity changes, compared with what is known from infants (T. Gebuis, J.L. Kenemans, E.H.F. de Haan & M.J. van der Smagt, unpublished data). Furthermore, as stated above, it can be suggested that the decrease in non-symbolic congruency is the result of the development of inhibitory mechanisms. Indeed, interference effects are often reported to be stronger in children than in adults, although the extent of this interaction is dependent on the type and complexity of the task (Gerstadt *et al.*, 1994; Wright *et al.*, 2003; West *et al.*, 2004; Hanauer & Brooks, 2005). Children's inhibitory mechanisms develop most rapidly between 3.6 and 6.0 years (Gerstadt *et al.*, 1994; Wright *et al.*, 2003), and are expected to be on the same level as adults at about the age of 12 years (Durston & Casey, 2006). For symbolic notation then, the congruency effect would reflect the combined contributions of the increase in automatization of symbolic number knowledge and the increase in the efficacy of inhibitory mechanisms with advancing age.

The mechanisms underlying the recruitment of additional processes

The non-symbolic size congruency task was included in the experiment to allow a direct comparison with the symbolic task, i.e. to disentangle general task- and numerosity-related processes. On a behavioral level, the second graders revealed both facilitation and interference for non-symbolic notation, but only interference for symbolic notation. On the neuronal level, a P3 latency effect was present in the non-symbolic task but not in the symbolic task. Instead, in the symbolic task, the congruency effect modulated the anterior P2 amplitude, suggesting the recruitment of distinct processes.

Our results are in line with previous neuroimaging studies reporting a shift from frontal to parietal areas of processes related to automated symbolic number processing (Ansari *et al.*, 2005; Rivera *et al.*, 2005; Kaufmann *et al.*, 2006). Two hypotheses have been raised in previous studies to explain this ontogenetic shift (Ansari *et al.*, 2005; Kaufmann *et al.*, 2006). First, increased automatization of Arabic symbol knowledge could result in less reliance on attentional and working memory processes. Our results contradict this first hypothesis. If the second graders in our study recruited distinct resources because of the increased attentional or working memory load for the number symbols compared with the dots, it should have been reflected in prolonged reaction times or less accurate responses in the symbolic but not the non-symbolic task. Even though this was not the case [*post hoc* analyses did not reveal a significant difference between the behavioral or accuracy results of the symbolic and non-symbolic task in the second graders (symbolic 853 ms/94% correct; non-symbolic 860 ms/90% correct)], it could still be argued that similar behavioral responses do not exclude the recruitment of different resources for the distinct cognitive processes involved (attention or working memory; Cohen Kadosh *et al.*, 2008). The ERP results revealed no differences in the timing of the P3 peak (related to stimulus-evaluation time) or its amplitude (related to attention or working memory processes), excluding the first hypothesis as a potential explanation of the results. [*Post hoc* analyses did not reveal a significant difference between the P3 latency or amplitude data of the symbolic and non-symbolic task in the second graders (symbolic 853 ms/12 μV; non-symbolic 860 ms/10.5 μV).] The second hypothesis holds that the parietal cortex might not be functionally specialized for the processing of numerical magnitude in second graders yet (Ansari *et al.*, 2005). Our results together with an earlier study revealing intra parietal sulcus involvement in non-symbolic numerosity processing in 4-year olds (Cantlon *et al.*, 2006) suggest that the parietal cortex is already specialized for non-symbolic but not symbolic magnitude. This parietal functional specialization is probably directly related to the level of automatization of symbolic number processes.

In conclusion, the present study reveals the development of automated symbolic number knowledge, which is contrasted with automated non-symbolic number knowledge. First, we demonstrate that automatic access to number symbol meaning can already be present in second graders (6.0 years). We propose that instead of cultural influences the diversity in informal as well as formal teaching determines the onset time of the congruency effects and thus explains the large range in onset times reported. Second, for all age groups and both tasks the magnitudes interacted at the stimulus-evaluation level, before motor preparation or execution. Third, second graders recruited distinct processes in the symbolic task only. This suggests that the ontogenetic shift from frontal to parietal areas is due to increased automatization of Arabic number knowledge, which underlies the functional specialization of the parietal areas. Fourth, the demonstrated congruency effects increased in size for the symbolic but decreased for the non-symbolic task. However, further research is needed to disentangle the origins and mechanisms underlying these diverging developmental processes.

Acknowledgement

This research was supported by NWO Grant 051.04.050 of the Cognition Program of the Netherlands Organization for Scientific Research (NWO).

Abbreviations

EEG, electroencephalogram; EOG, electrooculogram; ERP, event-related potential; LRP, lateralized readiness potential; rLRP, response-locked LRP; SCE, size congruency effects; sLRP, stimulus-locked LRP.

References

Algom, D., Dekel, A. & Pansky, A. (1996) The perception of number from the separability of the stimulus: the Stroop effect revisited. *Mem. Cognit.*, **24**, 557–572.

Ansari, D., Garcia, N., Lucas, E., Hamon, K. & Dhital, B. (2005) Neural correlates of symbolic number processing in children and adults. *Neuroreport*, **16**, 1769–1773.

Barth, H., La Mont, K., Lipton, J. & Spelke, E.S. (2005) Abstract number and arithmetic in preschool children. *Proc. Natl. Acad. Sci. U S A*, **102**, 14116–14121.

Brannon, E.M., Suanda, S. & Libertus, K. (2007) Temporal discrimination increases in precision over development and parallels the development of numerosity discrimination. *Dev. Sci.*, **10**, 770–777.

Cantlon, J.F., Brannon, E.M., Carter, E.J. & Pelphrey, K.A. (2006) Functional imaging of numerical processing in adults and 4-y-old children. *PLoS Biol.*, **4**, e125.

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. *J. Cogn. Neurosci.*, **19**, 957–970.

Cohen Kadosh, R., Lammertyn, J. & Izard, V. (2008) Are numbers special? An overview of chronometric, neuroimaging, developmental and comparative studies of magnitude representation. *Prog. Neurobiol.*, **84**, 132–147.

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

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

Durston, S. & Casey, B.J. (2006) What have we learned about cognitive development from neuroimaging? *Neuropsychologia*, **44**, 2149–2157.

Feigenson, L., Carey, S. & Spelke, E. (2002) Infants' discrimination of number vs. continuous extent. *Cogn. Psychol.*, **44**, 33–66.

Field, A.P. (2005) *Discovering Statistics Using SPSS: And Sex and Drugs and Rock 'n' Roll*, 3rd Edn. Sage, London.

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

Gebuis, T., Kenemans, J.L., de Haan, E.H.F. & van der Smagt, M.J. (in press). Conflict processing of symbolic and non-symbolic numerosity. *Neuropsychologia*. DOI:10.1016/j.neuropsychologia.2009.09.027.

Gerstadt, C.L., Hong, Y.J. & Diamond, A. (1994) The relationship between cognition and action: performance of children 3 1/2-7 years old on a Stroop-like day-night test. *Cognition*, **53**, 129–153.

Girelli, L., Lucangeli, D. & Butterworth, B. (2000) The development of automaticity in accessing number magnitude. *J. Exp. Child Psychol.*, **76**, 104–122.

Gratton, G., Coles, M.G. & Donchin, E. (1983) A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.*, **55**, 468–484.

Halberda, J., Mazzocco, M.M. & Feigenson, L. (2008) Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, **455**, 665–668.

Hanauer, J.B. & Brooks, P.J. (2005) Contributions of response set and semantic relatedness to cross-modal Stroop-like picture-word interference in children and adults. *J. Exp. Child Psychol.*, **90**, 21–47.

Henik, A. & Tzelgov, J. (1982) Is three greater than five: the relation between physical and semantic size in comparison tasks. *Mem. Cognit.*, **10**, 389–395.

Kaufmann, L., Koppelstaetter, F., Siedentopf, C., Haala, I., Haberlandt, E., Zimmerhackl, L.B., Felber, S. & Ischebeck, A. (2006) Neural correlates of the number-size interference task in children. *Neuroreport*, **17**, 587–591.

Keppel, G. (1991) *Design and Analysis: A Researcher's Handbook*. Prentice Hall, Upper Saddle River, NJ.

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

Lipton, J.S. & Spelke, E.S. (2003) Origins of number sense. Large-number discrimination in human infants. *Psychol. Sci.*, **14**, 396–401.

Luck, S.J. 2005. *An Introduction to the Event-Related Potential Technique*. MIT Press, Cambridge, Mass.

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

Mussolin, C. & Noel, M.P. (2007) The nonintentional processing of Arabic numbers in children. *J. Clin. Exp. Neuropsychol.*, **29**, 225–234.

Piazza, M., Izard, V., Pinel, P., Le Bihan, D. & Dehaene, S. (2004) Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron*, **44**, 547–555.

Potts, G.F. (2004) An ERP index of task relevance evaluation of visual stimuli. *Brain Cogn.*, **56**, 5–13.

Rivera, S.M., Reiss, A.L., Eckert, M.A. & Menon, V. (2005) Developmental changes in mental arithmetic: evidence for increased functional specialization in the left inferior parietal cortex. *Cereb. Cortex*, **15**, 1779–1790.

Rubinsten, O., Henik, A., Berger, A. & Shahar-Shalev, S. (2002) The development of internal representations of magnitude and their association with Arabic numerals. *J. Exp. Child Psychol.*, **81**, 74–92.

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. *Hum. Brain. Mapp.*, **29**, 574–580.

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 Psychol. (Amst)*, **90**, 97–109.

Szucs, D., Soltesz, F., Jarmi, E. & Csepe, V. (2007) The speed of magnitude processing and executive functions in controlled and automatic number comparison in children: an electro-encephalography study. *Behav. Brain Funct.*, **3**, 23.

Temple, E. & Posner, M.I. (1998) Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proc. Natl. Acad. Sci. U S A*, **95**, 7836–7841.

Tzelgov, J., Henik, A. & Berger, J. (1992) Controlling Stroop effects by manipulating expectations for color words. *Mem. Cognit.*, **20**, 727–735.

West, R., Bowry, R. & McConville, C. (2004) Sensitivity of medial frontal cortex to response and nonresponse conflict. *Psychophysiology*, **41**, 739–748.

Wood, J.N. & Spelke, E.S. (2005) Infants' enumeration of actions: numerical discrimination and its signature limits. *Dev. Sci.*, **8**, 173–181.

Wright, I., Waterman, M., Prescott, H. & Murdoch-Eaton, D. (2003) A new Stroop-like measure of inhibitory function development: typical developmental trends. *J. Child Psychol. Psychiatry*, **44**, 561–575.

Zhou, X., Chen, Y., Chen, C., Jiang, T., Zhang, H. & Dong, Q. (2007) Chinese kindergartners' automatic processing of numerical magnitude in stroop-like tasks. *Mem. Cognit.*, **35**, 464–470.