Acta Psychologica 140 (2012) 35–42

Contents lists available at SciVerse ScienceDirect

Acta Psychologica

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

What can the same–different task tell us about the development of magnitude representations?

Emmy Defever [a,b,*], Delphine Sasanguie [a,b], Mieke Vandewaetere [c,d], Bert Reynvoet [a,b]

[a] Department of Psychology, University of Leuven, 3000 Leuven, Belgium
[b] Subfaculty of Psychology and Educational sciences, University of Leuven, Kulak, 8500 Kortrijk, Belgium
[c] iTEC, Interdisciplinary Research on Technology, Communication and Education, University of Leuven, Kulak, 8500 Kortrijk, Belgium
[d] CIP&T, Centre for Instructional Psychology and Technology, University of Leuven, 3000 Leuven, Belgium

ARTICLE INFO

Article history:
Received 30 June 2011
Received in revised form 6 February 2012
Accepted 15 February 2012
Available online 14 March 2012

PsycINFO classification:
2340 (Cognitive Processes)
2800 (Developmental Psychology)
2820 (Cognitive and Perceptual Development)

Keywords:
Same–different judgment
Development
Magnitude representation
Distance effect
Mathematics achievement

ABSTRACT

We examined the development of magnitude representations in children (Exp 1: kindergartners, first-, second- and sixth graders, Exp 2: kindergartners, first-, second- and third graders) using a numerical same–different task with symbolic (i.e. digits) and non-symbolic (i.e. arrays of dots) stimuli. We investigated whether judgments in a same–different task with digits are based upon the numerical value or upon the physical similarity of the digits. In addition, we investigated whether the numerical distance effect decreases with increasing age. Finally, we examined whether the performance in this task is related to general mathematics achievement. Our results reveal that a same–different task with digits is not an appropriate task to study magnitude representations, because already late kindergarteners base their responses on the physical similarity instead of the numerical value of the digits. When decisions cannot be made on the basis of physical similarity, a similar numerical distance effect is present over all age groups. This suggests that the magnitude representation is stable from late kindergarten onwards. The size of the numerical distance effect was not related to mathematical achievement. However, children with a poorer mathematics achievement score seemed to have more difficulties to link a symbol with its corresponding magnitude.

1. Introduction

There is consistent evidence that humans have an innate capacity to represent magnitude (e.g., Feigenson, Dehaene, & Spelke, 2004; Izard, Sann, Spelke, & Streri, 2009). This ability to represent magnitude develops over time and is thought to be fundamental for acquiring the meaning of numerals. By repeatedly linking a numeral with its associated quantity, children acquire a symbolic system that is mapped onto this pre-existing representation (Barth, La Mont, Lipton, & Spelke, 2005; Mundy & Gilmore, 2009). In the last couple of years, many different studies investigated the development of magnitude representations (e.g., Holloway & Ansari, 2009; Reynvoet, De Smedt, & Van den Bussche, 2009; Soltész, Szücs, & Szücs, 2010). A task that is commonly used to study this development is the comparison task (e.g., Bugden & Ansari, 2011; Mundy & Gilmore, 2009). This task typically results in a comparison distance effect (CDE) which implies that it is easier to discriminate two magnitudes that are numerically further apart (e.g., "1" and "7") compared to numerically closer magnitudes (e.g., "5" and "7"). The distance effect is usually thought to be caused by partially overlapping representations of nearby magnitudes. A particular magnitude does not only activate its corresponding representation, but also the representations of numerically close magnitudes, according to a Gaussian distribution (Moyer & Landauer, 1967), making it more difficult to discriminate between two close magnitudes. Studies investigating the development of the CDE in children found that its size decreases with increasing age. This observation has been explained by magnitude representations that become more precise when schooling advances (Holloway & Ansari, 2009). In addition, the size of the comparison distance effect was found to be correlated with mathematics achievement, showing larger CDEs being associated with poorer mathematical ability (De Smedt, Verschaffel, & Ghesquière, 2009; Holloway & Ansari, 2009; Mussolin, Mejias, & Noël, 2010). Recent studies, however, argued that one should be careful with interpreting the CDE as reflecting the characteristics of the magnitude representation (Cohen Kadosh, Brodsky, Levin, & Henik, 2008; Holloway & Ansari, 2008; Van Opstal, Gevers, De Moor, & Verguts, 2008; Van Opstal & Verguts, 2011). For instance, Van Opstal et al. (2008) argued that the CDE can

* Corresponding author at: Subfaculty of Psychology and Educational Sciences, Department of Psychology, University of Leuven, Kulak, Etienne Sabbelaan 53, 8500 Kortrijk, Belgium. Tel.: +32 56 24 64 43.
E-mail address: Emmy.Defever@kuleuven-kulak.be (E. Defever).

0001-6918/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.actpsy.2012.02.005

be explained on the basis of decision processes rather than from magnitude representations. These authors constructed a connectionist model that was trained to select the largest of two numbers. After training the model, it was shown that the CDE can be explained by monotonically increasing or decreasing connections weights between the magnitude layer and the response nodes and crucially, that the CDE does not require representational overlap as was assumed before (see Van Opstal et al., 2008 for a more detailed description). Also the finding of Holloway and Ansari (2008) that a developmental decrease of the CDE is common to both numerical and non-numerical comparisons, supports the idea of the CDE reflecting a decisional mechanism. More evidence came from a study of Cohen Kadosh et al. (2008), who found that comparisons of music pitches were similar to other magnitude response functions, again implying that the CDE reflects a general sensorimotor transformation rather than the mental representation of magnitudes.

Therefore, in this study we chose an alternative task that is assumed to address the magnitude representation directly, i.e. a numerical same–different task. In this task, participants have to decide whether two magnitudes are numerically the same or different. Similar to comparison tasks a numerical distance effect is typically observed which implies that it becomes easier to classify two magnitudes as numerically different when the numerical distance between the magnitudes increases. This numerical distance effect is also explained by overlapping magnitude representations. Van Opstal and Verguts (2011) simulated the numerical distance effect in a same–different task and found that, in contrast to comparison tasks, the effect is only present when overlapping representations are assumed. On the basis of these simulations, the authors argued that a same–different task is more appropriate to investigate the mental representations of magnitudes (see also Cohen Kadosh et al., 2008).

The only developmental work so far using the same–different task was carried out by Duncan and McFarland (1980). They conducted a symbolic same–different task with kindergartners, first-, third-, fifth graders and adults. These authors observed a similar symbolic distance effect over all age groups, a finding in contrast with the observation of a decreasing distance effect with increasing age observed in a comparison task. This suggests, in contrast to the conclusion based on comparison performance, that children's magnitude representation does not get more precise with increasing age from late kindergarten onwards. Recently, however, Cohen (2009) argued that the numerical distance between two numerals correlates strongly with the physical similarity between those two numerals, which may have lead researchers in previous studies to confuse the effects of physical similarity for those of numerical distance. In the study of Cohen (2009), adults were presented with digits that had to be classified as the same as or as different than five. The data revealed that the participants based their decisions entirely on the physical characteristics of the Arabic digits and not on the numerical value. It therefore remains unclear whether the previous developmental results of Duncan and McFarland (1980) tell us something about how symbolic magnitudes are represented. It can be questioned whether children use numerical value or, similar to adults, use the physical similarity to make their decisions in symbolic same–different judgments (Cohen, 2009).

In the present study, we wanted to shed light upon the contribution of numerical distance and physical similarity in a symbolic same–different task conducted in children. In addition, we examined whether the numerical distance effect in symbolic and non-symbolic same–different judgments decreases with increasing age, which would be an indication of a more precise magnitude representation with increasing age. Finally, we examined whether the performance in a numerical same–different task is related to individual differences in mathematics achievement. Therefore, in the first experiment we examined the performance of children from four different age groups (Exp 1: kindergartners, first-, second-, and sixth graders) using a symbolic and a non-symbolic same–different task. In the second experiment, in addition to a pure non-symbolic same–different task, we used a same–different task in which a digit and an array of dots had to be matched.

2. Experiment 1

2.1. Method

2.1.1. Participants

Participants were 140 typically developing children recruited from a primary school in Belgium. The sample consisted of 30 kindergartners (16 males, mean age = 5.067 years, $SD = .254$), 32 first graders (12 males, mean age = 6.125, $SD = .336$), 38 second graders (14 males, mean age = 7.079, $SD = .273$) and 40 sixth graders (15 males, mean age = 11.175, $SD = .385$). None was aware of the purpose of the experiment. All children received a small reward for their participation.

2.1.2. Stimuli

Number knowledge test: Before starting the same–different experiments, all kindergartners and first graders needed to conduct a simple task to measure their knowledge of the digits 1 to 9 and the magnitude they represent. First, children were asked to read aloud the digits 1 to 9 (that were printed on the left side of the sheet), to ensure that they recognized the digits. Next they had to connect the numbers to pictures on the right hand side of the sheet where collections with different numbers of clowns were presented. Children who did not score 9 out of 9 were excluded from further analyses.

Mathematics achievement: The mathematical skills of the kindergartners were assessed by the 'Toeters' (CLB Haacht, 2005). The subtest mathematics of this test covers items on counting, mathematics language and conservation. Cronbach's α for this test is .89 (CLB Haacht & Department of Psychology of the Lessius Hogeschool, 2005). Mathematical skills of the elementary school children were assessed using a curriculum-based standardized achievement test for mathematics from the Flemish Student Monitoring System (Deloof, 2005; Dudal, 2000, 2001). The scores on the test administered halfway the school year were used. The reliability index of Kuder–Richardson (KR 20) for the test is .90, .89 and .88 for first, second and sixth grades, respectively. This test includes 60 items covering number knowledge, understanding of operations, (simple) arithmetic, word problem solving, measurement and geometry. For the analyses, we transformed the raw mathematics achievement scores to z-scores per grade.

Symbolic same–different task: The task was administered using notebooks with 14-inch screens. In each trial, two Arabic digits ranging from 1 to 9 were displayed simultaneously in white on a black background using E-prime 1.0 (www.pstnet.com; Psychology Software Tools). The digits were presented in Arial font and subtended 0.57° visual angle (= .05 cm) in width and 0.69° (= .06 cm) visual angle in height from a viewing distance of about 50 cm. Trials with a numerical distance up to 5 of all possible combinations were presented. This choice was motivated based on previous research in which it was shown that the decrease in reaction times is limited for distances larger than 5 (Sasanguie, Defever, Van den Bussche, & Reynvoet, 2011). The 9 'same' trials (i.e. trials with two numerically same stimuli) were presented twice, resulting in a final list of 78 trials (60 'different' trials and 18 'same' trials).

Non-symbolic same–different task: In each trial, two white-filled circles (diameter 7 cm, 8.01° visual angle from a viewing distance of about 50 cm) containing arrays of 1 to 9 black dots were shown simultaneously. The physical properties of the stimuli (i.e. total area occupied and individual dot size) were controlled using the MatLab program as described in Dehaene, Izard, and Piazza (2005). On half of the trials dot size was held constant whereas on the other half the total occupied area was held constant. This prevented participants from reliably using these non-numerical parameters to make a correct judgment. All other details of this condition were similar to the symbolic same–different task.

2.1.3. Procedure

Participants were tested in a quiet room accompanied by two experimenters. The kindergartners were tested in groups of 4 children. The children from first-, second- and sixth grades were tested in groups of 7 to 10 children. Children were seated in such a way that they could not distract each other and the experimenters could monitor them closely. The kindergartners and first graders first completed the number knowledge test before starting the same–different experiments. In these same–different experiments, subjects were instructed to decide whether the two presented stimuli were numerically the same or different, by pressing a corresponding button on an AZERTY keyboard ('a' or 'p' which were labeled with the stickers '=' and '≠' respectively). Before each task started, 5 randomly selected practice trials were presented to familiarize the children with the task requirements. During the practice trials, feedback was provided. When the experimenters noticed that children started counting during the practice trials, they instructed the children not to count and just to estimate whether the two dot collections contained a same amount of dots or not. All participants first conducted the symbolic same–different task and then the non-symbolic same–different task. Each trial started with a fixation cross (600 ms), followed by the stimuli (i.e. two Arabic digits or arrays of dots) which were presented until a response was given and an inter-trial interval (1000 ms). Children were seated at approximately 50 cm from the screen. In total, the experiment took approximately 25 min.

2.1.4. Data analyses

Children were excluded from the analyses when they scored at chance level (error rate of 45–55%) on either the same trials or the different trials (5 kindergartners, 5 first- and 2 second graders), and when they were too slow or made too many errors (3 *SD* above the grade average; 2 kindergartners, 2 first-, 4 second- and 4 sixth graders). One kindergartner was excluded because the math achievement score was missing. Finally, the data of two kindergarteners were not included because they scored not perfectly on the number knowledge test. These exclusion criteria resulted in a final sample of 20 kindergartners (10 males; mean age = 5.100 years, *SD* = .308), 25 first — (8 males; mean age = 6.160, *SD* = .374), 32 second — (12 males; mean age = 7.063, *SD* = .246) and 36 sixth graders (11 males; mean age = 11.194, *SD* = .401). Trials on which an error was made (on average 2% of the trials in the symbolic task and 12% of the trials in the non-symbolic task) were discarded from the RT analyses.

Median RTs from correct responses on the different trials and mean error rates on the different trials were separately submitted to a repeated measures analysis of variance with stimulus notation (symbolic or non-symbolic) and distance (1, 2, 3, 4 or 5) as within-subject factors, grade (kindergartners, first-, second- or sixth graders) as a between-subject factor and mathematics achievement as a covariate.

Additionally, we also examined whether the physical similarity between two digits influenced children's performance in the symbolic same–different task. To define the physical similarity between digits, we used an identical approach as in Campbell and Clarck (1988). Fifteen adults were asked to rate the physical similarity between two different Arabic numbers on a 7-point scale ranging from 1 (i.e. weak similarity) to 7 (i.e. high similarity) (see Appendix A for mean ratings). We administered the ratings in adults because we felt that young children (particularly kindergartners and first graders) would be unable to provide sufficiently fine-grained judgments on a seven-point scale. Moreover, it would be very difficult to explain the instructions of such a rating to young children. A strong correlation was found between our ratings and the ratings obtained by Campbell and Clarck (1988), $r(58) = .860$, $p < .0001$, suggesting that our ratings are reliable. For the analysis, trials were divided into trials with a low physical similarity (i.e. below the median similarity rating of 2.38, e.g., pair 6–7) and trials with a high physical similarity (i.e. above the median similarity rating 2.38, e.g., pair 8–9). Numerical distance was split up into two levels: small numerical distances (RTs on distances 1 and 2) and large numerical distances (RTs on distances 3, 4, and 5). This was done to avoid that numerical distance would have more levels than physical similarity. Median RTs of correct responses on the different trials were submitted to a repeated measures analysis of variance with physical similarity (low or high) and distance (small or large) as within-subject factors, grade (kindergartners, first-, second- or sixth graders) as a between-subject factor and mathematics achievement as a covariate.

2.2. Results

2.2.1. Reaction times

The repeated measures analysis of variance with stimulus notation and distance as within-subject factors, grade as a between-subject factor and mathematics achievement as a covariate yielded a main effect of numerical distance ($F(4,105) = 21.535$, $p < .0001$, $\eta_p^2 = .451$). This was embedded in an interaction between stimulus notation and numerical distance ($F(4,105) = 8.170$, $p < .0001$, $\eta_p^2 = .237$), showing that the median RTs decreased with increasing distance in the non-symbolic task only (see Fig. 1 and Table 1). Post hoc pairwise comparisons showed that responses were slower on trials with a smaller distance compared to trials with a larger distance (all $ps < .050$), with the exception that no difference was observed between distances 2 and 3, and between distances 4 and 5. A main effect of stimulus notation also emerged ($F(1,108) = 229.804$, $p < .0001$, $\eta_p^2 = .680$), indicating that subjects responded faster in the symbolic task (i.e. 980 ms) compared to the non-symbolic task (i.e. 1496 ms). A main effect of grade ($F(3,108) = 44.130$, $p < .0001$, $\eta_p^2 = .551$) showed that RTs decreased with increasing grade. No effect of mathematics achievement was found ($F(1,108) = 2.108$, $p = .149$), nor was there an interaction between numerical distance and mathematics achievement ($F(4,105) = 1.014$, $p = .404$). A marginal significant interaction between stimulus notation and grade was also observed ($F(3,108) = 2.457$, $p = .067$, $\eta_p^2 = .064$), suggesting that the RTs decreased with increasing age in each notation condition (all $ps < .010$), with the exception that no difference was found between the kindergartners and the first graders in the non-symbolic task ($p = .178$). The interaction between distance and grade was not significant ($F < 1$), indicating a similar distance effect over all age groups. All other effects were also not significant (all $ps > .404$).

To investigate whether the absence of a numerical distance effect in the symbolic task is caused by a response strategy based on the physical similarity, we performed a second repeated measures analysis of variance with physical similarity (low-high) and numerical distance (small-large) as within-subject factors, grade as a between subject factor and mathematics achievement as a covariate. This yielded a main effect of physical similarity ($F(1,108) = 14.111$, $p < .0001$, $\eta_p^2 = .116$), showing that responses were faster when the physical similarity between two digits was low (on average 975 ms) compared to when the physical similarity was high (on average 1028 ms). The main effect of grade indicated that RTs decreased with increasing age ($F(3,108) = 51.003$, $p < .0001$, $\eta_p^2 = .586$). No main effect of mathematics achievement was found ($F(1,108) = 3.089$, $p = .082$) nor were there any other significant effects (all $ps > .103$).

2.2.2. Error rates

The repeated measures analysis of variance with stimulus notation and distance as within-subject factors, grade as a between-subject factor and mathematics achievement as a covariate revealed a main

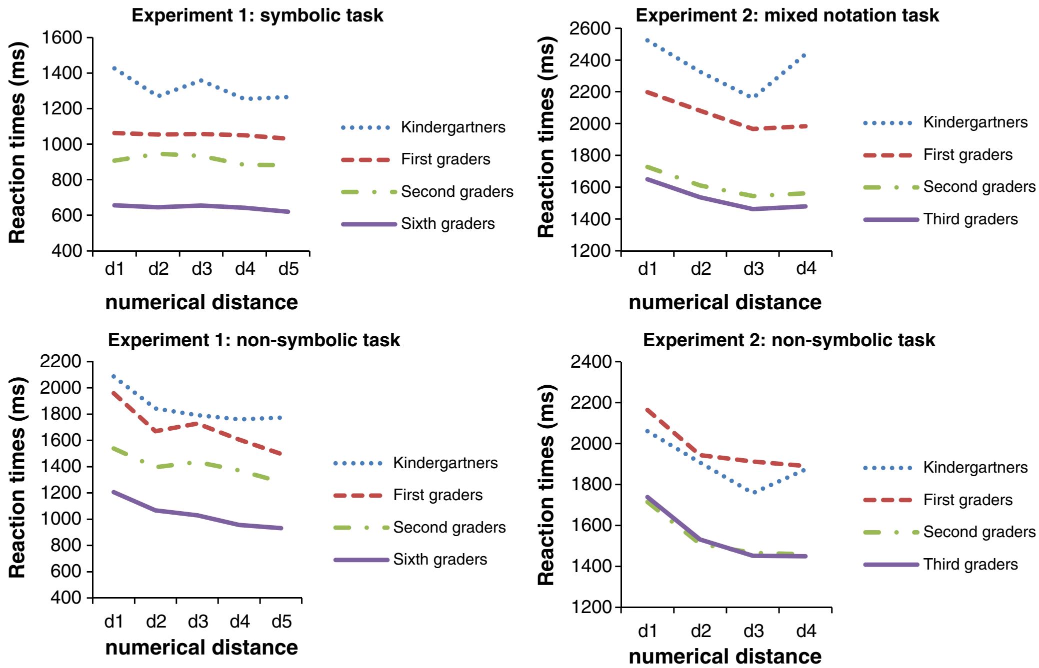

Fig. 1. The reaction times (ms) in function of numerical distance and grade for each task in Experiments 1 and 2 (for the standard deviations see Tables 1 and 2).

effect of numerical distance ($F(4,105) = 35.568$, $p < .0001$, $\eta_p^2 = .575$). This effect was imbedded in a two-way interaction between notation and numerical distance ($F(4,105) = 22.870$, $p < .0001$, $\eta_p^2 = .466$), indicating that a distance effect only emerged in the non-symbolic task (see Table 1). The error rates in the non-symbolic task decreased with increasing distance (all $ps < .050$), with the exception that no difference was found between distances 3 and 5 and between distances 4 and 5. A main effect of stimulus notation was also present ($F(1,108) = 134.401$, $p < .0001$, $\eta_p^2 = .554$), indicating that more errors were made in the non-symbolic task (on average 9.60%) compared to the symbolic task (on average 2.00%). The effect of mathematics achievement was also significant ($F(1,108) = 4.435$, $p < .050$, $\eta_p^2 = .039$). Correlations controlling for grade were calculated to investigate the association between the error rates in each task and mathematics achievement. A marginal significant correlation was found for the non-symbolic task ($r_{partial}(110) = -.179$, $p = .058$), but not for the symbolic task ($r_{partial}(110) = -.127$, $p = .180$). A higher mathematics achievement tended to be associated with a lower error rate in the non-symbolic task. The significant main effect of grade ($F(3,108) = 4.660$, $p < .010$, $\eta_p^2 = .115$) was embedded in a two-way interaction with stimulus notation and grade ($F(3,108) = 4.778$, $p < .010$, $\eta_p^2 = .117$). Post hoc pairwise comparisons showed no significant difference in error rate between the groups in the symbolic task (all $ps > .291$). In the non-symbolic task, no difference between the two youngest groups nor

Table 1
Mean RTs (SD) and mean error rates (SD) as a function of numerical distance and grade in the symbolic and the non-symbolic task of Experiment 1.

	Numerical distance				
	Distance 1	Distance 2	Distance 3	Distance 4	Distance 5
Reaction times (ms)					
Symbolic task					
Kindergartners	1426.33 (369.17)	1268.83 (234.65)	1358.33 (280.87)	1254.18 (277.75)	1264.88 (246.90)
First graders	1062.68 (213.50)	1054.02 (265.39)	1056.84 (149.70)	1050.34 (340.74)	1031.32 (252.11)
Second graders	907.23 (231.94)	945.95 (253.80)	933.45 (277.55)	884.00 (226.62)	881.75 (208.02)
Sixth graders	656.33 (132.70)	645.54 (124.40)	655.17 (128.92)	641.67 (124.89)	619.53 (108.60)
Non-symbolic task					
Kindergartners	2086.00 (790.86)	1842.13 (549.37)	1792.52 (619.69)	1760.85 (470.40)	1772.50 (552.90)
First graders	1959.14 (586.56)	1669.08 (632.35)	1727.52 (648.97)	1605.84 (359.87)	1496.32 (532.39)
Second graders	1538. 53 (529.41)	1395.92 (430.44)	1434.66 (439.89)	1371.39 (462.40)	1283.20 (383.66)
Sixth graders	1206.75 (311.54)	1066.63 (220.58)	1029.54 (288.76)	956.88 (231.33)	932.76 (207.13)
Error rates (%)					
Symbolic task					
Kindergartners	1.20 (.03)	1.75 (.04)	2.00 (.04)	3.50 (.05)	1.20 (.04)
First graders	1.20 (.03)	1.68 (.04)	2.24 (.04)	1.60 (.05)	1.92 (.04)
Second graders	2.47 (.05)	1.75 (.05)	3.03 (.05)	1.25 (.05)	3.41 (.06)
Sixth graders	3.53 (.05)	0.78 (.05)	1.81 (.04)	1.94 (.04)	1.69 (.05)
Non-symbolic task					
Kindergartners	10.80 (.11)	3.85 (.06)	6.15 (.06)	4.00 (.08)	6.80 (.10)
First graders	15.04 (.12)	8.52 (.09)	7.00 (.09)	4.00 (.08)	3.44 (.08)
Second graders	21.97 (.12)	13.03 (.14)	11.34 (.11)	6.88 (.09)	9.56 (.11)
Sixth graders	26.78 (.13)	14.94 (.10)	6.86 (.07)	4.17 (.07)	6.78 (.10)

between the two oldest groups was observed (all $ps > .541$). However, surprisingly, the two youngest groups made significantly less errors than the two oldest groups (all $ps < .050$). We believe this is due to the exclusion of the worst performing kindergartners and first graders on the basis of the present exclusion criteria. There was also a significant interaction between numerical distance and grade ($F(12,278) = 4.368$, $p < .0001$, $\eta_p^2 = .141$), which was imbedded in a three-way interaction between stimulus notation, numerical distance and grade ($F(12,278) = 1.968$, $p < .050$, $\eta_p^2 = .069$). To disentangle this interaction, separate analyses per notation condition were performed. A significant interaction between numerical distance and grade was only present in the non-symbolic task ($F(12,278) = 2.947$, $p < .010$, $\eta_p^2 = .100$), showing that the difference in error rate between the youngest and the oldest children were the largest at the smallest distances (see Table 1).

2.3. Discussion

In this first experiment we investigated whether the distance effect observed in a symbolic (i.e. digits) and non-symbolic (i.e. arrays of dots) same–different task decreases with increasing age. Moreover, we examined whether the performance on this task was related to mathematics achievement. Finally, we also investigated whether children use magnitude information in a symbolic same–different task or whether they base their responses on the physical similarity of the digits.

The results of our *non-symbolic* task showed that all age groups had a similar numerical distance effect. Children responded faster when the numerical distance between the two presented stimuli increased. In addition, no relationship between the distance effect and mathematics achievement was observed, suggesting a similar magnitude representation in children irrespective of their mathematics achievement. However, it should be mentioned that the manner we used to control for the visual cues of our non-symbolic stimuli is debated. Although the participants could not rely on a single visual property to conduct the non-symbolic task, it cannot be excluded that the visual cues could have covaried with numerical distance (Gebuis & Reynvoet, 2011). Hence, the non-symbolic numerical distance effect might have been influenced by the visual features of the stimuli.

In contrast to the non-symbolic task, a numerical distance effect was absent in the *symbolic* same–different task. This finding contradicts several previous studies in which a distance effect in a same–different task with digits was found (Dehaene & Akhavein, 1995; Duncan & McFarland, 1980; Ganor-Stern & Tzelgov, 2008; Verguts & van Opstal, 2005). However, to date, studies using this task are limited and the analyses and designs that were applied in these studies differ from our approach. More specifically, the previous studies grouped the several distances (i.e. 'small distance trials' and 'large distance trials') to perform analyses on the distance effect. Goldfarb, Henik, Rubinsten, Bloch-David, and Gertner (2011) examined the emergence of a distance effect in a single-digit same–different task when analyzed in different ways. When they analyzed the data as this was done in the previous studies using grouped distances, a significant distance effect emerged. However, when they conducted similar analyses as the ones we used, they did not find a distance effect anymore. This latter finding is in line with our results and thus suggests that the obtained distance effect in the previous studies is limited to the specific analyses that were performed. Our results also demonstrated that when the physical similarity between pairs of digits was taken into account, children responded slower when the physical similarity between the digits increased. This is in line with the observations by Cohen (2009) and means that the performance in a numerical same–different task, in which two digits have to be judged as same or different, does not allow to draw clear conclusions about the development of symbolic magnitude representations.

Therefore, in a second experiment, in addition to a pure non-symbolic same–different task, we conducted a mixed notation same–different task in which a digit had to be matched with a non-symbolic magnitude. This way, children were prevented from using a physical similarity strategy and the capacity to associate a symbol with its quantity could be investigated. Furthermore, we made some additional modifications to our design. First, the amount of same and different trials was equated because it could be argued that the unequal balance in our first experiment might have evoked a response anticipation towards different trials, hereby possibly affecting the distance effect. Second, we replaced the group of sixth graders by a group of third graders because we wanted to focus on age groups in which developmental changes in the distance effect are more likely to occur. Finally, only magnitudes up to five were presented for two reasons. The first reason concerns the fact that the youngest children in this study recently started learning to associate the symbols from 1 to 10 with their corresponding magnitude. Therefore, for the mixed notation task we expected that they would be more inclined to start counting the larger dot collections to verify whether a symbol would match with an amount of dots. Hence, we decided to use magnitudes that are in the subitizing range (Trick & Pylyshyn, 1993). Secondly, to keep the duration of the experiments reasonable while presenting an equal number of same and different trials, we had to reduce the range of presented magnitudes.

3. Experiment 2

3.1. Method

3.1.1. Participants

A total of 172 typically developing children took part in the experiment. Participants were 33 kindergartners (17 males; mean age = 5.424 years, $SD = .561$), 44 first graders (25 males; mean age = 6.114 years, SD = .321), 43 second graders (17 males; mean age = 7.512 years, SD = .593) and 52 third graders (23 males; mean age = 8.173 years, SD = .430). The participants were recruited from three different primary schools in Belgium. None was aware of the purpose of the experiment. All children received a small reward for their participation.

3.1.2. Stimuli and procedure

The procedure was identical to Experiment 1.

Mixed notation same–different task: In each trial, two white-filled circles (diameter 7 cm, 8.01° visual angle) were displayed simultaneously: one containing a digit between 1 and 5 presented in black, and one containing an array of 1 to 5 black dots. The digits were presented in Arial font and subtended a visual angle of approximately 1.72° (= 1.5 cm) in width and 2.52° (= 2.2 cm) in height from a viewing distance of about 50 cm. All displays contained dots of the same size but the configuration of each display was different. The presentation order of the digit and dot collection was counterbalanced such that in half of the trials the digit was on the left hand side and in the other half the digit was on the right hand side of the screen. There were 25 possible combinations (20 'different' trials, 5 'same' trials). The different trials were presented twice and the same trials were presented 8 times, resulting in a final list of 80 trials (40 same trials and 40 different trials).

Non-symbolic same–different task: In each trial, two white-filled circles (diameter 7 cm, 8.01° visual angle) containing arrays of 1 to 5 black dots were shown simultaneously. The physical properties of the stimuli (i.e. total area occupied and individual dot size) were controlled using the MatLab program as described by Dehaene et al. (2005). Participants were presented with the same trial list of numerical values as in the mixed notation same–different task.

The number knowledge and mathematics achievement test were identical to Experiment 1. In contrast to Experiment 1, third graders were included instead of sixth graders. The reliability index of Kuder–Richardson (KR 20) for the mathematics test is .90 for the third graders (Dudal, 2002).

3.1.3. Data analyses

As in Experiment 1, participants were excluded from the analyses when they scored at chance level for the same or different trials (3 kindergartners) or when they were too slow or made too many errors (3 *SD* above the grade average; 1 kindergartner, 6 first-, 5 second- and 6 third graders). In addition, 3 kindergartners who did not succeed in the number knowledge test were excluded. This resulted in a final sample of 26 kindergartners (13 males; mean age = 5.423 years, *SD* = .504 years), 38 first- (21 males; mean age = 6.105 years, *SD* = .311), 38 second graders (15 males; mean age = 7.553 years, *SD* = .602) and 46 third graders (20 males; mean age = 8.196 years, *SD* = .453). Trials on which an error was made (on average 6.70% in the mixed notation task and 6.30% in the non-symbolic task) were excluded from the RT analyses.

Median RTs from correct responses and mean error rates on trials with two numerically different stimuli were separately submitted to a repeated measures analysis of variance with stimulus notation (mixed or non-symbolic) and distance (1, 2, 3 or 4) as within-subject factors, grade (kindergartners, first-, second- or third graders) as a between-subject factor and mathematical ability as a covariate.

3.2. Results

3.2.1. Reaction times

A main effect of numerical distance was observed ($F(3,141) = 41.975$, $p < .0001$, $\eta_p^2 = .472$), showing that the RTs decreased with increasing distance (see Fig. 1 and Table 2). Post hoc pairwise comparisons demonstrated that trials with a smaller distance were always slower to respond to than trials with a larger distance (all $ps < .0001$), with the exception that no significant difference was found between distances 2 and 4 and between distances 3 and 4. A significant main effect of grade was also found ($F(3,143) = 25.965$, $p < .0001$, $\eta_p^2 = .353$), showing that the kindergartners and first graders were significantly slower than the second- and third graders (all $ps < .0001$), but no significant difference was found between the two youngest groups ($p = .207$), nor between the two oldest groups ($p = .637$). The main effect of stimulus notation was also significant ($F(1,143) = 24.767$, $p < .0001$, $\eta_p^2 = .148$), yielding faster RTs in the non-symbolic task (on average 1739 ms) compared to the mixed notation task (on average 1891 ms). However, a significant interaction between stimulus notation and grade ($F(3,143) = 10.067$, $p < .0001$, $\eta_p^2 = .174$) indicated that only kindergartners were significantly faster in the non-symbolic task (i.e. 1899 ms) compared to the mixed notation task (i.e. 2362 ms) ($p < .0001$), whereas no differences were found between the notation conditions for the other groups (all $ps > .176$). There was also an effect of mathematics achievement ($F(1,143) = 19.508$, $p < .0001$, $\eta_p^2 = .120$), which was embedded in an interaction between stimulus notation and mathematics achievement ($F(1,143) = 4.241$, $p < .050$, $\eta_p^2 = .029$). To further examine this interaction we computed correlations between the median RTs in each task and mathematics achievement, controlling for grade. These correlations showed that the RTs decreased with increasing mathematics achievement, however, the relation was more pronounced in the mixed notation task ($r_{partial}(145) = -.373$, $p < .0001$) than in the non-symbolic task ($r_{partial}(145) = -.232$, $p < .010$). It should be noted that the RTs on the mixed notation task were moderately correlated with the RTs on the non-symbolic task, when grade was partialled out ($r_{partial}(145) = .611$, $p < .0001$). A Hotelling–Williams *t*-test (Williams, 1959) revealed that the correlation between the RTs and mathematics achievement was significantly stronger for the mixed notation task compared to the non-symbolic task ($t(145) = -2.07$, $p < .050$). No other significant effects were observed (all $ps > .134$).

3.2.2. Error rates

A main effect of numerical distance was observed ($F(3,141) = 5.409$, $p < .010$, $\eta_p^2 = .103$). Post hoc pairwise comparisons demonstrated that trials with distance 1 were always responded to less accurately than trials with a larger distance (all $ps < .050$), but no significant differences were found between the distances 2, 3 and 4. A significant main effect of grade was also found ($F(3,144) = 6.942$, $p < .0001$, $\eta_p^2 = .105$), showing that the kindergartners made more errors than the other grades (all $ps < .050$). No significant differences in number of errors were found between the first-, second- and third graders (all $ps > .056$). The significant interaction between stimulus notation and grade ($F(1,143) = 2.677$ $p < .050$, $\eta_p^2 = .053$) indicated that the kindergartners made more errors in the non-symbolic task compared to the mixed notation task ($p < .050$) whereas no significant differences were found between the other groups (all $ps > .095$). No other significant interactions were observed (all $ps > .065$). Finally, no influence of mathematics achievement was found ($F < 1$).

Table 2
Mean RTs (SD) and mean error rates (SD) as a function of numerical distance and grade in the mixed notation task and the non-symbolic task of Experiment 2.

	Numerical distance			
	Distance 1	Distance 2	Distance 3	Distance 4
Reaction times (ms)				
Mixed notation task				
Kindergartners	2523.04 (700.71)	2327.75 (735.56)	2161.04 (489.33)	2436.19 (1095.68)
First graders	2197.45 (488.91)	2081.30 (407.79)	1965.83 (421.60)	1983.54 (358.69)
Second graders	1727.88 (413.25)	1611.03 (375.99)	1543.95 (406.84)	1562.29 (504.76)
Third graders	1649.55 (431.23)	1536.95 (435.75)	1462.29 (403.60)	1478.98 (427.42)
Non-symbolic task				
Kindergartners	2059.63 (552.15)	1906.85 (451.13)	1757.62 (558.36)	1873.13 (605.76)
First graders	2163.43 (586.43)	1942.92 (341.59)	1911.29 (361.35)	1890.14 (565.30)
Second graders	1714.07 (423.48)	1509.74 (441.62)	1465.92 (466.12)	1458.79 473.79)
Third graders	1737.96 (462.51)	1530.80 (502.15)	1451.59 (431.27)	1449.13 (543.92)
Error rates (%)				
Mixed notation task				
Kindergartners	12.38 (.12)	7.62 (.07)	5.15 (.08)	5.77 (.13)
First graders	3.53 (.05)	4.11 (.06)	5.79 (.08)	5.26 (.12)
Second graders	7.11 (.07)	7.82 (.09)	8.61 (.09)	5.92 (.11)
Third graders	7.28 (.06)	6.30 (.07)	7.70 (.10)	4.89 (.10)
Non-symbolic task				
Kindergartners	12.58 (.11)	10.12 (.14)	12.88 (.17)	7.69 (.15)
First graders	5.45 (.07)	3.87 (.05)	2.87 (.06)	4.61 (.14)
Second graders	7.79 (.09)	3.92 (.07)	5.42 (.08)	5.92 (.12)
Third graders	8.59 (.12)	4.41 (.08)	3.98 (.08)	1.63 (.06)

4. General discussion

In the current study, we examined the development of symbolic and non-symbolic magnitude representations in elementary school children using a numerical same–different task with symbolic (i.e. digits) and non-symbolic (i.e. arrays of dots) stimuli. In addition, we examined the relationship between the performance on these tasks and mathematics achievement.

One of the aims of this study was to clarify whether children rely on magnitude information when conducting same–different judgments on two digits or whether they base their responses on the physical similarity between the digits. A numerical distance effect was absent, however, a significant effect of physical similarity was observed in this task: subjects responded slower when the physical similarity between the digits was high. These findings replicate and extend the results of Cohen (2009), indicating that the physical similarity influences responses in same–different judgments with digits from kindergarten onwards. Therefore, the present observations make clear that the performance in a same–different task with digits is not an adequate measure to study the development of magnitude representations, as the judgments are based on the physical similarity of the digits and not on their magnitudes. Previous studies relying on this task should therefore keep into account this alternative way of performing the task when interpreting the data (e.g. Duncan & McFarland, 1980).

Despite the inadequacy of the same–different task to study the development of the distance effect when both stimuli are digits, the task does seem appropriate when a matching strategy based on physical similarity is excluded (e.g., mixed notation task, see also Van Opstal & Verguts, 2011). Under these conditions, we found a significant effect of numerical distance. Moreover, crucially for the present purposes, no differences in the size of the distance effect were found over age groups. These findings are consistent with studies using the priming paradigm (Defever, Sasanguie, Gebuis, & Reynvoet, 2011; Reynvoet et al., 2009). In these studies, a priming distance effect was observed that had the same size in all age groups. Like the numerical distance effect observed in same–different judgments, the priming distance effect is assumed to be a reflection of the representational overlap on the mental number line (Van Opstal et al., 2008). Together, the present results and those of previous priming studies suggest that the representational overlap remains stable from kindergarten onwards, or stated differently, that the level of precision of magnitude representations remains constantly over development. The latter is in contrast with the observations in comparison studies (i.e. a decreasing distance effect with increasing age), that has led researchers to conclude that magnitude representations become more precise with increasing age (Holloway & Ansari, 2009). We propose that the results in comparison tasks do not so much reflect the increasing precision of magnitude representations, but that operations or decisions upon these magnitude representations are becoming more refined (see also Van Opstal & Verguts, 2011, for a similar reasoning).

Another goal of this study was to examine whether the performance in a same–different task is related to mathematics achievement. No relationship was found between the numerical distance effect and the performance on a mathematics achievement test. This is not in line with a previous priming study (Defever et al., 2011). In this study, weaker scores on a standardized achievement test were associated with a smaller priming distance effect both for symbolic and non-symbolic stimuli. This finding was interpreted as evidence for a less precise representation of symbolic and non-symbolic magnitudes in children with low scores on a mathematics achievement test. For instance, when children have vague magnitude representations, the representational overlap that is responsible for the priming effect will not differ a lot between the different prime-target distances leading to less increase in RTs with increasing numerical distance (i.e. a small PDE) compared to children with more precise representations. Because both the same–different and priming tasks are supposed to measure magnitude representation, a similar pattern was expected in the current study. More specifically, in this study we expected that vague magnitude representations would result in a slower discrimination performance between two close magnitudes, which would imply a larger distance effect in children with a lower mathematics score (e.g. Holloway & Ansari, 2009). A possible explanation for the differences found in both paradigms may be due to the particular task characteristics. In the priming paradigm, subjects are instructed to compare a digit with a fixed standard and accordingly, need to keep the standard (e.g., '5') in working memory. Previous studies have shown that children with poor mathematics achievement exhibit a deficit in working memory (e.g., De Smedt, Janssen, et al., 2009; Passolunghi, Vercelloni, & Schadee, 2007). In contrast, in the same–different task, no information needs to be retained in memory as both numerosities are presented simultaneously. Consequently, it is possible that a working memory deficit, and not the precision of the magnitude representation, is responsible for the associations observed between the priming distance effect and mathematics achievement. However, this possible explanation is still speculative and needs to be investigated in future studies.

Although no relationship was found between the numerical distance effect and mathematics achievement, an interaction between stimulus notation and mathematics achievement was observed in the reaction time analysis of Experiment 2. A significant association between the reaction times and mathematics achievement was observed in both the mixed notation task and the non-symbolic task but the relationship was significantly stronger in the mixed notation task. It could be argued that this association is due to general processing speed. We were not able to evaluate this as no speeded control task was included in this study. However, it does not seem likely that it is only general processing speed that caused the association with mathematics achievement, given that the mathematics achievement test is an untimed measure of children's mathematical ability and given the stronger correlation in the mixed notation task. The only difference between the tasks was that in the mixed notation task a symbolic processing requirement came into play. This seems to indicate that predominantly the ability to link a symbol with its corresponding magnitude is related to mathematics achievement, more than magnitude processing per se. This observation is in line with the findings of previous studies in which low achievers showed problems with the processing of symbols (De Smedt & Gilmore, 2011; Rousselle & Noël, 2007; Sasanguie, De Smedt, Defever, & Reynvoet, 2011).

To conclude, our results demonstrated that the numerical same–different task can be used to investigate the development of magnitude representations. However, one has to be careful with interpreting the results in this task when using digits, because even young children seem to base their responses on the physical similarity between the digits instead of their numerical value. Once a physical similarity strategy is excluded (e.g., with a mixed notation), a numerical distance effect was present. More importantly, the numerical distance effect was of equal size in all age groups, suggesting a stable magnitude representation from late kindergarten onwards. Individual differences in the numerical distance effect were not related to mathematics achievement. However, children performing worse on the mathematics achievement test, were also slower to match a digit and a dot pattern suggesting that these children might have problems with linking a symbol with its corresponding magnitude.

Acknowledgments

The authors thank all children, parents and schools that took part in this study. This research was supported by the Grant G.0451.10 from the Fund for Scientific Research-Flanders and by the Research Fund K.U. Leuven.

Appendix A

SD 1	SD 2	Phys.	SD 1	SD 2	Phys.
1	2	2.19	5	6	4.69
1	3	1.81	5	7	2.38
1	4	2.50	5	8	3.38
1	5	1.44	5	9	3.19
1	6	1.44	6	1	1.13
2	1	2.13	6	2	1.94
2	3	3.44	6	3	2.88
2	4	2.50	6	4	1.81
2	5	3.00	6	5	3.94
2	6	2.63	6	7	1.81
2	7	2.88	6	8	5.13
3	1	1.63	6	9	5.38
3	2	3.06	7	2	2.56
3	4	2.19	7	3	1.88
3	5	3.25	7	4	2.13
3	6	3.13	7	5	1.69
3	7	2.19	7	6	1.69
3	8	5.06	7	8	1.50
4	1	2.44	7	9	1.81
4	2	2.00	8	3	4.81
4	3	1.88	8	4	1.88
4	5	2.38	8	5	2.81
4	6	1.94	8	6	4.63
4	7	2.44	8	7	1.75
4	8	1.56	8	9	5.50
4	9	1.69	9	4	2.06
5	1	1.25	9	5	2.75
5	2	2.56	9	6	5.25
5	3	2.81	9	7	1.75
5	4	2.13	9	8	4.88

Note. SD 1 = digit presented at the left side of the screen. SD 2 = digit presented at the right side of the screen. Phys = mean physical similarity rating.

References

Barth, H., La Mont, K., Lipton, J., & Spelke, E. (2005). Abstract number and arithmetic in preschool children. *PNAS*, *102*, 14116–14121.

Bugden, S., & Ansari, D. (2011). Individual differences in children's mathematical competence are related to the intentional but not automatic processing of Arabic numerals. *Cognition*, *118*, 35–47.

Campbell, J. I. D., & Clarck, J. M. (1988). An encoding-complex view of cognitive number processing: Comment on McCloskey, Sokol, and Goodman (1986). *Journal of Experimental Psychology. General*, *117*, 204–214.

CLB Haacht (2005). *TOETERS: Toets voor taal- en rekenvoorwaarden, schrijfmotoriek en observatie van de werkhouding* (2e druk). februari 2005.

CLB Haacht, & Department of Psychology of the Lessius Hogeschool (2005). *TOETERS: Toets voor taal- en rekenvoorwaarden, schrijfmotoriek en observatie van de werkhouding.*

Cohen, D. J. (2009). Integers do not automatically activate their quantity representation. *Psychonomic Bulletin & Review*, *16*, 332–336.

Cohen Kadosh, R., Brodsky, W., Levin, M., & Henik, A. (2008). Mental representation: What can pitch tell us about the distance effect? *Cortex*, *44*, 470–477.

De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first graders with mathematical difficulties. *Journal of Experimental Child Psychology*, *108*, 278–292.

De Smedt, B., Janssen, R., Bouwens, K., Verschaffel, L., Boets, B., & Ghesquière, P. (2009). Working memory and individual differences in mathematics achievement: A longitudinal study from first grade to second grade. *Journal of Experimental Child Psychology*, *103*, 186–201.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 469–479.

Defever, E., Sasanguie, D., Gebuis, T., & Reynvoet, B. (2011). Children's representation of symbolic and non-symbolic magnitude examined with the priming paradigm. *Journal of Experimental Child Psychology*, *109*, 174–186.

Dehaene, S., & Akhavein, R. (1995). Attention, automaticity, and levels of representation in number processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 314–326.

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

Deloof, G. (2005). *LVS-VCLB Leerlingvolgsysteem. Wiskunde Toetsen 6–-Basisboek [Student monitoring system. Mathematics Tests 6–-manual].* Leuven, Belgium: Garant.

Dudal, P. (2000). *LVS-VCLB Leerlingvolgsysteem. Wiskunde Toetsen 1–-Basisboek [Student monitoring system. Mathematics Tests 1–-manual].* Leuven, Belgium: Garant.

Dudal, P. (2001). *LVS-VCLB Leerlingvolgsysteem. Wiskunde Toetsen 2–-Basisboek [Student monitoring system. Mathematics Tests 2–-manual].* Leuven, Belgium: Garant.

Dudal, P. (2002). *LVS-VCLB Leerlingvolgsysteem. Wiskunde Toetsen 3–-Basisboek [Student monitoring system. Mathematics Tests 3–-manual].* Leuven, Belgium: Garant.

Duncan, E. M., & McFarland, C. E. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory and Cognition*, *8*, 612–622.

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

Ganor-Stern, D., & Tzelgov, J. (2008). Across-notation automatic numerical processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*, 430–437.

Gebuis, T., & Reynvoet, B. (2011, November 14). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology. General*, doi:10.1037/a0026218 Advance online publication.

Goldfarb, L., Henik, A., Rubinsten, O., Bloch-David, Y., & Gertner, L. (2011). The numerical distance effect is task dependent. *Memory and Cognition*, *8*, 1508–1517.

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development in children's development of number comparison. *Developmental Science*, *11*, 644–649.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, *103*, 17–29.

Izard, V., Sann, C., Spelke, E., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*, 10382–10385.

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgments of numerical inequality. *Nature*, *215*, 1519–1520.

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and non-symbolic representations of number. *Journal of Experimental Child Psychology*, *103*, 490–502.

Mussolin, C., Mejias, S., & Noël, M. P. (2010). Symbolic and non-symbolic number comparison in children with and without dyscalculia. *Cognition*, *115*, 10–25.

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development*, *22*, 165–184.

Reynvoet, B., De Smedt, B., & Van den Bussche, E. (2009). Children's representation of symbolic magnitude: The development of the priming distance effect. *Journal of Experimental Child Psychology*, *103*, 480–489.

Rousselle, L., & Noël, M. P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic versus non-symbolic number magnitude processing. *Cognition*, *102*, 361–395.

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2011). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, *136*, 73–80.

Sasanguie, D., Defever, E., Van den Bussche, E., & Reynvoet, B. (2011). The reliability of and the relation between non-symbolic numerical distance effects in comparison, same–different judgments and priming. *Acta Psychologica*, *136*, 73–80.

Soltész, F., Szücs, D., & Szücs, L. (2010). Relationships between magnitude representation, counting and memory in 4- to 7-year-old children: A developmental study. *Behavioural and Brain Functions*, *6*, 13, doi:10.1186/1744-9081-6-13.

Trick, L., & Pylyshyn, Z. (1993). What enumeration studies can show us about spatial attention: Evidence for limited capacity preattentive processing. *Journal of Experimental Psychology. Human Perception and Performance*, *19*, 331–351.

Van Opstal, F., Gevers, W., De Moor, W., & Verguts, T. (2008). Dissecting the symbolic distance effect: Priming and comparison distance effects in numerical and nonnumerical orders. *Psychonomic Bulletin & Review*, *15*, 419–425.

Van Opstal, F., & Verguts, T. (2011). The origins of the numerical distance effect: The same different task. *Journal of Cognitive Psychology*, *23*, 112–120.

Verguts, T., & van Opstal, F. (2005). Dissociation of the distance effect and size effect in one-digit numbers. *Psychonomic Bulletin & Review*, *12*, 925–930.

Williams, E. J. (1959). The comparison of regression variables. *Journal of the Royal Statistical Society: Series B*, *21*, 395–399.