Predictors for Mathematics Achievement? Evidence From a Longitudinal Study

Delphine Sasanguie[1,2], Eva Van den Bussche[3], and Bert Reynvoet[1,2]

ABSTRACT—Numerical processing has been extensively studied by examining the performance on basic number processing tasks, such as number priming, number comparison, and number line estimation. These tasks assess the innate "number sense," which is assumed to be the breeding ground for later mathematics development. Indeed, several studies have associated children's performance in these tasks with individual differences in mathematical achievement. To date, however, most of these studies have cross-sectional designs. Moreover, the few longitudinal studies either use complex tasks (e.g., story problems) or investigate only one of these basic number processing tasks at a time. In this study, we examine the association between the performance of children on several basic number processing tasks and their individual math achievement scores on a curriculum-based test measured 1 year later. Regression analyses showed that most of the variance in children's math achievement was predicted by nonsymbolic number line estimation performance (i.e., estimating large quantities of dots) and, to a lesser extent, the speed of comparing symbolic numbers. This knowledge about the predictive value of the performance of 5- to 7-year-olds on these markers of number processing can help with the early identification of at-risk children. In addition, this information can guide appropriate educational interventions.

The ability to represent number has proved to be an innate capacity that is shared by human infants and animals (Cantlon & Brannon, 2006; Xu & Spelke, 2000). Behavioral studies have shown that newborns already respond to changes in numbers and that their ability to do this improves with increasing age (Xu & Spelke, 2000). This innate basic number processing ability is also suggested to lie at the basis of the performance of 5-year-olds, who had not yet been taught formal arithmetic but can compare, add, and subtract different dot arrays or sequences of sounds (Barth, Beckmann, & Spelke, 2008). Later, when children are confronted with symbols to represent numbers, these symbols acquire meaning by being associated with this preexisting nonsymbolic representation (Mundy & Gilmore, 2009). Indeed, in basic numerical tasks, such as comparison, similar behavioral effects are found in studies with nonsymbolic and symbolic numbers (Cohen Kadosh, Lammertyn, & Izard, 2008). This resulted in the popular idea that the innate number representation serves as the basis of future mathematic achievement as measured by curriculum-based tasks (e.g., Halberda, Mazzocco, & Feigenson, 2008; but see Rousselle & Noël, 2007, for an alternative account).

Different tasks have been used to study number processing. A first paradigm is *priming* (e.g., Defever, Sasanguie, Gebuis, & Reynvoet, 2011). Here, a target stimulus is preceded by an irrelevant prime stimulus and the numerical distance between prime and target is manipulated. A priming distance effect (PDE) is observed: reaction times (RTs) increase if the numerical distance between the two presented magnitudes (i.e., prime and target) increases. For instance, the digit "4" is processed faster when it is preceded by "3" than by "2" (Reynvoet, De Smedt, & Van den Bussche, 2009). The PDE is explained by overlapping representations between nearby numbers on the mental number line (Verguts, Fias, & Stevens, 2005). On this mental number line, each magnitude is represented as a Gaussian distribution around the corresponding magnitude. This implies that, whenever a magnitude is presented, not only the representation of that specific magnitude will be activated but also, to a lesser extent, the nearby magnitudes, resulting in faster RTs if both numbers are close to one another. Developmental studies

[1]Laboratory of Experimental Psychology, KU Leuven
[2]Subfaculty of Psychology and Educational Sciences, KU Leuven Kulak
[3]Department of Psychology, Vrije Universiteit Brussel

Address correspondence to Delphine Sasanguie, Subfaculty of Psychology and Educational Sciences, Department of Psychology, KU Leuven Kulak, Etienne Sabbelaan 53, BE - 8500 Kortrijk, Belgium; e-mail: delphine.sasanguie@kuleuven-kulak.be

have shown that the PDE can be observed in kindergartners and more importantly, remains stable through development, indicating no differences in representational overlap from late kindergarten onward (Defever et al., 2011).

The second, frequently used task to investigate basic number processing, is *comparison* (e.g., Moyer & Landauer, 1967). In this task, participants have to indicate the larger of two numbers. This typically results in the numerical distance effect: error rates and RTs decrease with increasing distance between both magnitudes. For instance, deciding that 9 is larger than 8 takes more time than performing the same decision on the pair 9 and 6. Originally, the numerical distance effect was also explained by overlapping number representations: When two close numbers are presented, their activated representations will overlap, making it more difficult to differentiate between them. In developmental studies, a decreasing comparison distance effect with increasing age is observed, which is believed to reflect the decreasing overlap between nearby magnitudes (Holloway & Ansari, 2009). Recently, however, several studies argued that the comparison distance effect can alternatively be considered as the outcome of decision processes on numbers (Holloway & Ansari, 2009; Van Opstal, Gevers, De Moor, & Verguts, 2008). A third task is the *number line estimation* task (e.g., Siegler & Booth, 2004). In this task, participants have to indicate the position of a number on an empty external number line. Typically, it is found that, with increasing age, the pattern to map numbers on the line goes from a logarithmic, with larger magnitudes put closer together than smaller magnitudes, toward a linear pattern.

The priming effect (Defever et al., 2011), the numerical distance effect (Bugden & Ansari, 2011; Holloway & Ansari, 2009; Mundy & Gilmore, 2009; Sasanguie, De Smedt, Defever, & Reynvoet, 2012), and number line estimation performance (Booth & Siegler, 2006, 2008; Sasanguie et al., 2012; Siegler & Booth, 2004) have all been related to individual differences in mathematical achievement in previous studies. However, some caveats are still present. First, these studies are largely cross-sectional in nature and longitudinal studies are still scarce. De Smedt, Verschaffel, and Ghesquière (2009) provided the first longitudinal evidence that the speed of comparing Arabic digits is related to individual differences in mathematics achievement 1 year later. With regard to the symbolic number line estimation task, Geary et al. (2009) showed that the performance on this task is an important variable in the prediction of later mathematics achievement. It should be noted, however, that these previous studies focused exclusively on only one of the basic number processing tasks described earlier. Furthermore, these studies were not designed to identify the optimal predictor for math achievement.

In this study, we therefore longitudinally examined the relationship between the performance of a single sample of children on multiple indices of basic number processing tasks and their individual mathematics achievement scores on a curriculum-based standardized math test 1 year later. This way, we want to gain insight in which of these markers is the best predictor for later math achievement. The children that participated in this study performed a number priming task, a number comparison task, and a number line estimation task with symbolic and nonsymbolic stimuli. Age groups (i.e., kindergartners, first graders and second graders) were chosen based on previous findings. For example, Siegler and colleagues (e.g., Booth & Siegler, 2006; Siegler & Booth, 2004) showed that first and second graders are the most interesting age groups to investigate, due to their logarithmic-to-linear shift for the 0–100 number lines at that age. In addition, a decreasing distance effect in symbolic and nonsymbolic number comparison has been reported especially in these age groups (Holloway & Ansari, 2009). Furthermore, in line with De Smedt et al. (2009), to get a view on the importance of informal learning experiences kindergartners were included, because this group of children did not receive formal instruction yet at that point.

METHOD

Participants

Participants were recruited from an elementary school in Flanders, Belgium. For the specific details about the participants in the three separate experimental tasks, we refer to our previous studies (Defever et al., 2011; Sasanguie et al., 2012). Subjects that were outliers in one of the three experimental tasks were removed. The final sample of this study therefore consisted of 72 typically developing children, comprising 21 kindergartners (11 males, M_{age} at the initial time of testing = 5.6 years), 25 first graders (5 males, M_{age} at the initial time of testing = 6.7 years), and 26 second graders (10 males M_{age} at the initial time of testing = 7.6 years). None of the children had repeated a grade.

Procedure

Children were tested in March–April 2010, in a quiet room accompanied by two experimenters. Kindergartners were tested in groups of 4 children and first and second graders in groups of 7–10 children. All children first conducted the number line estimation tasks, followed by the comparison tasks and next the priming tasks. For each task, they always first performed the symbolic and afterward the nonsymbolic condition. A short break between the two conditions was provided. After these experiments, the children received a small reward. One year later, around February 2011, the scores on the curriculum-based standardized tests (mathematics and spelling performance) were collected.

Measures

Experimental Tasks

Priming and number comparison were computerized tasks. Number line estimation was administered with pen and paper. Laptops with 14-inch color screens were used, running the Windows XP operating system. Stimulus presentation and the recording of behavioral data (RT and error rates) were controlled by E-prime 1.1 (Psychology Software Tools, http://www.pstnet.com). The symbolic stimuli consisted of Arabic digits ranging from 1 to 9 (Arial font, measuring 5 mm in width and 6 mm in height), whereas the nonsymbolic stimuli comprised arrays of 1–9 dots. The visual properties of the stimuli (e.g., total area occupied and item size) were controlled using the MatLab programme as described in the study by Dehaene, Izard, and Piazza (2005). The dots of a single stimulus were displayed in a circle with a radius of 45 mm.

Number Priming. Prime-target pairs were created by combining primes ranging from 1 to 9 (except 5) and targets 1, 4, 6, and 9, resulting in 32 prime-target pairs. Sixteen trials consisted of a prime and target that were both either smaller or larger than 5 (i.e., congruent trials), whereas the other 16 trials consisted of a prime and target of which one was smaller than 5 and the other larger (i.e., incongruent trials). The 16 congruent trials were presented a second time, and during the second presentation we switched the order of prime and target (prime became target and vice versa) because we did not want the subjects to notice that the numerosities 1, 4, 6, and 9 were always presented as the second numerosity. This resulted in a total of 48 trials (32 congruent and 16 incongruent trials). Participants were instructed to classify both prime and target as either smaller or larger than the standard 5 by pressing the corresponding button. Children were asked to respond as quickly as possible but to avoid making errors. A trial consisted of a fixation-cross (600 ms), the prime (until response), a blank screen (200 ms), the target (until response), and a blank screen (2000 ms). Before each condition, subjects performed five practice trials where feedback on accuracy was provided.

Number Comparison. All numerosities from 1 to 9 were included, but only combinations of stimuli with a maximum distance of 5 were presented. This was done because in a previous study (Sasanguie, Defever, Van den Bussche, & Reynvoet, 2011), we showed that RTs decreased with increasing distance, but from distance 5 onward, a further decrease was minimal. This resulted in a set of 60 experimental trials per condition. On each trial, a fixation cross was presented for 600 ms, after which the two stimuli that had to be compared appeared, one on the left-hand side and one on the right-hand side of the screen. These stimuli remained on the screen until the child responded. The intertrial interval was 1000 ms. Participants had to select the larger of two quantities, by pressing at the side of the largest quantity. Five practice trials were included for each condition and feedback on accuracy was provided.

Number Line Estimation. Children were presented with 25-cm-long lines in the center of white A4 sheets. Two different intervals (0–10 and 0–100) were administered in both symbolic and nonsymbolic formats. The end points of the number lines were labeled on the left by 0 and on the right by either 10 or 100 in the symbolic condition and by an empty circle on the left and a circle with 10 or 100 dots on the right in the nonsymbolic condition. The to-be-positioned quantity was shown in the center of the sheet, 2 cm above the number line. All numbers and dot patterns, except 0 and 10 (100), had to be positioned on the 0–10 interval, whereas for the 0–100 interval quantities were 2, 3, 4, 6, 18, 25, 48, 67, 71, and 86 (corresponding to sets A and B for the same interval used in the study by Siegler & Opfer, 2003). The presentation order of the quantities was randomized and each line was presented on a separate sheet. Children were instructed to mark on the line where they thought that the quantity had to be positioned. To ensure that the child was aware of the interval size, the experimenters took the first number line as an example and pointed to each item on the sheet, while saying: "This line goes from 0 (dots) to 10 (or 100) (dots). If here is 0 and here is 10 (or 100), where would you position this number (quantity)?" Afterward, the children went through all sheets at their own pace. Kindergartners only solved a 0–10 number line task, first graders both a 0–10 and a 0–100 task and second graders only a 0–100 task.

Standardized Tests

Mathematics. Mathematics achievement was assessed with a curriculum-based standardized achievement test for mathematics from the Flemish Student Monitoring System (Dudal, 2000a). This test consists of 60 items covering number knowledge, understanding of operations, (simple) arithmetic, word problem solving, measurement, and geometry. Cronbach's α for this test was .90, .92, and .90 for the first, second and third grades, respectively.

Spelling. The curriculum-based standardized Spelling test of the Flemish Student Monitoring System (Dudal, 2000b) was used to measure children's spelling skills. This test involved the dictation of letters, words, and sentences. Cronbach's α of this test was .94, .90, and .89 for the first, second, and third grades, respectively.

RESULTS

Number Priming

The PDE was examined by analyzing the performance on congruent trials as a function of the distance between prime and the target (see also Reynvoet et al., 2009). A repeated measures analysis of variance (ANOVA) was conducted with distance (three levels) as within-subject factor and grade (three levels) as between-subject factor. To control for individual differences in RTs and to capture the PDE in one measure, a normalized measure of the PDE for each individual was computed. Similar to previous research (Reynvoet et al., 2009), this was done by subtracting the average RT on trials with a prime-target distance of 1 from the average RT on trials with a prime-target distance of 3. The RT difference was then divided by the average RT on trials with a numerical prime-target distance of 1. For the error rates, the average error rate on trials with a prime-target distance of 1 was subtracted from the average error rate on trials with a prime-target distance of 3. A repeated measures ANOVA was conducted on this computed measure of the PDE with stimulus notation (two levels) as within-subject factor and grade (three levels) as between-subject factor. The mean error rates and RTs are presented in Table 1.

In the symbolic condition, the repeated measures analysis on the error rates showed a main effect of prime-target distance, $F(2, 68) = 5.38$, $p < .01$, $\eta_p^2 = .14$, indicating that participants made more errors when distance between prime and target increased. The repeated measures analysis on the RTs revealed a main effect of prime-target distance, $F(2, 68) = 5.52$, $p < .01$, $\eta_p^2 = .14$, indicating that RTs were larger when the distance between prime and target increased. A main effect of grade, $F(3, 69) = 13.12$, $p < .001$, $\eta_p^2 = .28$, was observed, showing that RTs decreased with increasing grade.

In the nonsymbolic condition, the repeated measures analysis on the error rates showed a main effect of prime-target distance, $F(2, 68) = 7.36$, $p < .01$, $\eta_p^2 = .18$, indicating that participants made more errors when distance between prime and target increased. The repeated measures analysis on the RTs revealed a main effect of prime-target distance, $F(2, 68) = 7.45$, $p < .01$, $\eta_p^2 = .18$, indicating that RTs were larger when the distance between prime and target increased. A main effect of grade, $F(3, 69) = 9.23$, $p < .001$, $\eta_p^2 = .21$, was observed, showing that RTs decreased with increasing grade.

The repeated measures analyses with the normalized PDEs on both the error rates and the RTs yielded no main effects of stimulus notation (all Fs < 1), meaning that these PDEs were similar for the symbolic and nonsymbolic notation. Moreover, no main effects of grade (all Fs < 1) were present, suggesting that the PDEs did not change with increasing age.

Number Comparison

In this task, RTs were adjusted to reflect both speed and accuracy of performance by combining the RTs and error rates using the formula RT/(1 − error). This way, the RTs remain unchanged with 100% accuracy and increase in proportion with the number of errors. This adjustment was made to control for speed–accuracy tradeoffs (see also Simon et al., 2008). The effect of distance (DE) was examined by conducting a repeated measures ANOVA with distance as within-subject factor (five levels) and grade (three levels) as between-subject factor on children's adjusted RTs. The adjusted RTs are shown in Table 2. To examine the DE in more detail, we computed the size of the DE for each child by calculating the slope of a regression in which distance predicted the adjusted RTs. This slope should be negative because the distance effect predicts a negative relationship between distance and RT. The size of the slope reflects the DE, with steeper slopes indicating larger distance effects. One-sample t-tests and a one-way ANOVA on the slopes were executed (see Sasanguie et al., 2012, for full details).

In the symbolic condition, analyses revealed a significant main DE on the RTs, $F(4, 66) = 14.80$, $p < .001$, $\sigma_p^2 = .47$. There was also a main effect of grade, $F(2, 69) = 34.02$, $p < .001$, $\sigma_p^2 = .50$, indicating that the RTs decreased with increasing grade.

Table 1
Mean Error Rates, Mean Reaction Times (RTs), and Corresponding (Standard Deviations) for the Three Distances of the Number Priming Task, per Notation and per Grade

	Number priming task					
	Symbolic notation			*Nonsymbolic notation*		
	d1	*d2*	*d3*	*d1*	*d2*	*d3*
Error rates (%)						
Kindergartners	0.04 (0.06)	0.11 (0.14)	0.09 (0.10)	0.03 (0.07)	0.10 (0.11)	0.11 (0.14)
First graders	0.05 (0.06)	0.04 (0.06)	0.09 (0.09)	0.03 (0.09)	0.05 (0.08)	0.07 (0.11)
Second graders	0.05 (0.08)	0.02 (0.05)	0.10 (0.10)	0.03 (0.05)	0.06 (0.09)	0.10 (0.13)
RT (ms)						
Kindergartners	1224.10 (369.14)	1371.95 (313.48)	1339 (313.56)	1138.52 (262.21)	1188.55 (351.71)	1217.45 (258.56)
First graders	1096.96 (302.66)	1142.50 (377.12)	1248.06 (330.46)	958.60 (239.57)	980.34 (226.57)	999.26 (179.97)
Second graders	897.71 (207.01)	957.61 (286.81)	949.92 (215.31)	883.63 (186.34)	904.63 (208.41)	977.94 (198.43)

Table 2
Mean Adjusted Reaction Times and Corresponding (Standard Deviations) for the Five Distances of the Number Comparison Task, per Notation and per Grade

	Number comparison task				
	d1	*d2*	*d3*	*d4*	*d5*
Symbolic					
Kindergartners	2122.41 (898.54)	1942.11 (692.36)	1915.16 (751.52)	1887.03 (825.82)	1677.55 (600.51)
First graders	1427.58 (356.64)	1300.88 (357.61)	1167.50 (251.14)	1163.62 (278.34)	1101.04 (255.64)
Second graders	1080.36 (207.47)	991.48 (198.47)	933.79 (167.92)	875.78 (136.57)	863.12 (171.43)
Nonsymbolic					
Kindergartners	2017.43 (432.14)	1595.67 (389.05)	1450.95 (360.10)	1363.36 (328.68)	1292.74 (343.04)
First graders	1760.21 (450.92)	1314.42 (294.67)	1197.95 (315.45)	1108.89 (309.35)	1094.55 (326.78)
Second graders	1349.43 (300.95)	1095.54 (174.32)	939.04 (170.92)	938.63 (187.69)	853.00 (141.89)

In the nonsymbolic condition, analyses revealed a significant main DE on the RTs, $F(4, 66) = 87.48$, $p < .001$, $\sigma_p^2 = .84$. There was also a main effect of grade, $F(2, 69) = 20.49$, $p < .001$, $\sigma_p^2 = .37$, indicating that the RTs decreased with increasing grade.

One-sample t-tests revealed that the slopes were significantly different from zero for all grades, in both conditions: kindergarten symbolic, $t(20) = -2.52$, $p = .02$, and nonsymbolic slope, $t(20) = -11.77$, $p < .001$; first grade symbolic, $t(24) = -7.34$, $p < .001$, and nonsymbolic slope, $t(24) = -9.07$, $p < .001$; and second grade symbolic, $t(25) = -8.89$, $p < .001$, and nonsymbolic slope, $t(25) = -12.83$, $p < .001$. A one-way ANOVA revealed that the size of the distance effect decreased with increasing grade level for the nonsymbolic, $F(2, 69) = 4.07$, $p = .02$, task, but not for the symbolic ($F < 1$) one.

Number Line Estimation

Children's estimation accuracies were obtained by computing the percent absolute error per child, according to the following equation (Siegler & Booth, 2004):

$$\left|\frac{\text{Estimate} - \text{estimate quantity}}{\text{Scale of estimates}}\right|$$

For example, if a child was asked to estimate 18 on a 0–100 number line and placed the mark at the point on the line corresponding to 30, the percentage of absolute error (PAE) would be (30 − 18)/100 or 12%. The mean PAE per grade and notation are shown in Table 3. A univariate ANOVA on the mean PAE was executed. We analyzed the pattern of estimates of the children in more detail by fitting linear and logarithmic functions for group means and each individual child (Siegler & Opfer, 2003). For group means, a paired t-test was conducted on the mean R^2 linear and mean R^2 logarithmic for each age group. If the t-test indicated a significant difference between both R^2s, the best fitting model (linear or logarithmic) was attributed to the group (see Sasanguie et al., 2012, for full details).

Table 3
Mean Percent Absolute Error (PAE) and the Corresponding Standard Deviations of the Number Lines, per Notation and per Grade

	Number line estimation task
Symbolic	
Kindergartners	0.17 (0.05)
First graders	0.15 (0.06)
Second graders	0.08 (0.04)
Nonsymbolic	
Kindergartners	0.18 (0.08)
First graders	0.18 (0.05)
Second graders	0.14 (0.07)

In the symbolic condition, for the 0–10 interval, the model with the highest R^2 was linear for both the kindergartners ($R^2_{lin} = .69$) and the first graders ($R^2_{lin} = .84$) and differed significantly from the logarithmic fit in both grades, $R^2_{log} = .53$; $t(20) = 5.49$, $p < .001$ and $R^2_{log} = .68$; $t(24) = 10.86$, $p < .001$, respectively. For the 0–100 interval, the fit of the logarithmic model was better for the first graders ($R^2_{log} = .80$), but not significantly different from the linear fit ($R^2_{lin} = .77$; $t < 1$). For the second graders, the fit of the linear model was the best ($R^2_{lin} = .90$) and differed significantly from the logarithmic fit, $R^2_{log} = .86$; $t(25) = 2.31$, $p = .03$.

In the nonsymbolic condition, for the 0–10 interval, the model with the highest R^2 was linear for both the kindergartners ($R^2_{lin} = .76$) and the first graders ($R^2_{lin} = .84$) and differed significantly from the logarithmic model in both grades, $R^2_{log} = .66$; $t(20) = 4.99$, $p < .001$ and $R^2_{log} = .68$; $t(24) = 18.71$, $p < .001$, respectively. For the 0–100 interval, the fit of the logarithmic model was significantly better for the first graders ($R^2_{log} = .88$) and significantly different from the linear fit, $R^2_{lin} = .76$; $t(24) = -4.74$, $p < .001$. In second graders, the linear model provided a better fit ($R^2_{lin} = .80$), however, not significantly different from the logarithmic fit ($R^2_{log} = .79$; $t < 1$).

A univariate ANOVA on the mean PAE did not reveal a significant difference between the kindergartners and the first graders for the symbolic, $F(1, 44) = 1.01$, $p = .32$, $\sigma^2_p = .02$, and nonsymbolic ($F < 1$) 0–10 number line conditions. Accuracy on the symbolic, $F(1, 49) = 7.07$, $p = .01$, $\sigma^2_p = .13$, and nonsymbolic, $F(1, 49) = 8.28$, $p < .01$, $\sigma^2_p = .14$, 0–100 tasks increased with grade.

Correlation Analyses

We explored the correlation between the different markers of number processing that were significantly associated with mathematics achievement in our previous studies (i.e., the normalized PDE in case of number priming, the slope and the mean adjusted RTs in case of number comparison, and the mean PAEs in case of number line estimation) and mathematics achievement measured 1 year later. Hereby, we controlled for grade differences because RTs speed up with grade. To be able to compare all children in the different grades, the raw math scores were transformed to *z*-scores per grade. To investigate whether the associations between the experimental tasks and mathematics achievement was specific to mathematics, we also integrated the performance of the children on a curriculum-based test of spelling as a control variable in the analyses. The correlations (controlled for grade) between the different markers of number processing and mathematics achievement measured 1 year later are displayed in Table 4.

Mathematics achievement was negatively correlated with the RTs of the symbolic comparison task, $r(69) = -.31$, $p = .01$ (Figure 1) and the mean PAE of both the symbolic, $r(69) = -.35$, $p < .001$ (Figure 2) and the nonsymbolic number line tasks, $r(69) = -.48$, $p < .001$ (Figure 3), indicating that children with a high score on the math achievement test were faster in comparing symbolic magnitudes and were more accurate in the number line estimation tasks. These correlations remained significant when spelling ability was controlled for, RTs symbolic comparison: $r(68) = -.32$, $p = .01$; mean PAE symbolic: $r(68) = -.38$, $p < .001$ and mean PAE nonsymbolic: $r(68) = -.50$, $p < .001$.

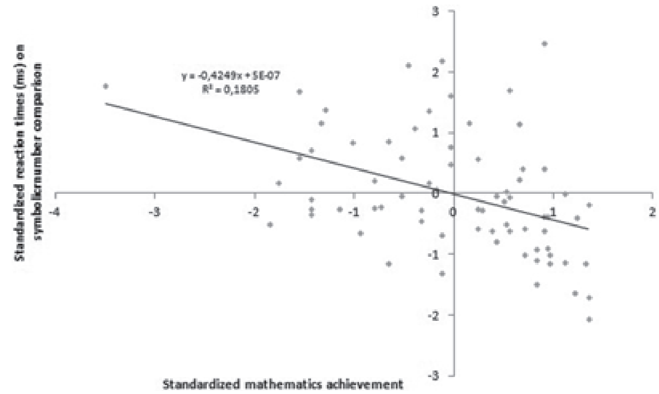

Fig. 1. Scatter plot showing the significant negative correlation between the standardized mean adjusted reaction times on the symbolic number comparison task and the standardized mathematics achievement scores from the curriculum-based standardized math test 1 year later.

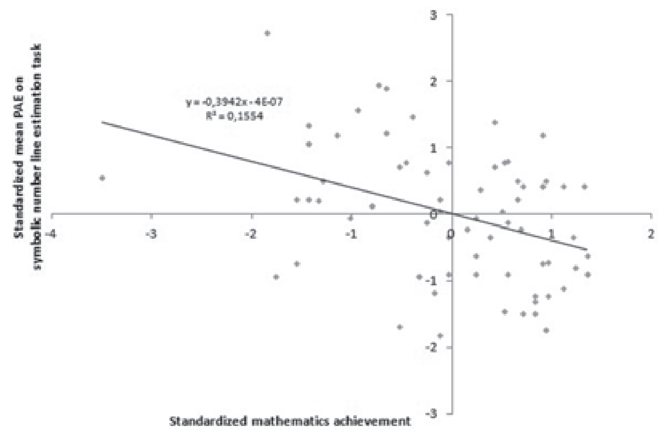

Fig. 2. Scatter plot showing the significant negative correlation between the standardized mean PAE on the symbolic number line estimation task and the standardized mathematics achievement scores from the curriculum-based standardized math test 1 year later.

Table 4
Partial Correlations Between the Diverse Indices of Magnitude Representation and Mathematics Achievement 1 Year Later, Controlled for Grade

	Standardized mathematics achievement
Priming	
Priming distance effect symbolic	0.22
Priming distance effect nonsymbolic	−0.00
Magnitude comparison	
Comparison slope symbolic	0.08
Comparison slope nonsymbolic	−0.12
Comparison mean adjusted reaction times symbolic	−0.31*
Comparison mean adjusted reaction times nonsymbolic	−0.18
Number line estimation	
Mean PAE symbolic	−0.35**
Mean PAE nonsymbolic	−0.48***

*$p < .05$. **$p < .01$. ***$p < .001$.

Regression Analyses

To examine whether priming, number comparison, and number line estimation predicted unique variance in the mathematics achievement scores, hierarchical multiple regression analyses were conducted. Five blocks of independent variables were added in a stepwise procedure. In Step 1, the variable "grade" was included to control for grade differences in RT,

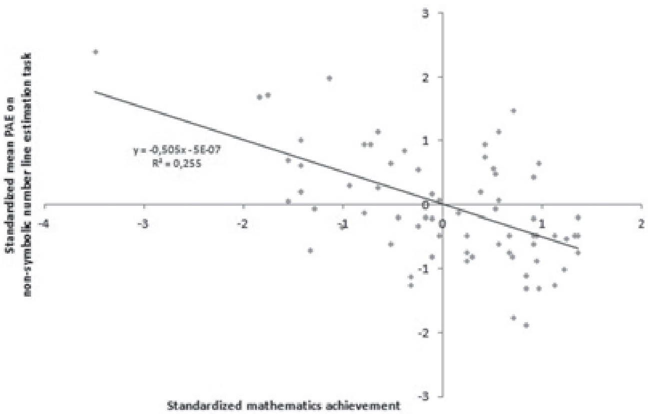

Fig. 3. Scatter plot showing the significant negative correlation between the standardized mean PAE on the nonsymbolic number line estimation task and the standardized mathematics achievement scores from the curriculum-based standardized math test 1 year later.

after recoding this variable in two dummies by means of effect recoding. Spelling performance was entered into the model in Step 2. In Step 3, both the symbolic and the nonsymbolic PDEs were included to determine the unique influence of the magnitude representation level after controlling for grade and spelling performance. Step 4 involved the entrance of the slopes and mean adjusted RTs of the comparison tasks, to examine the unique contribution to the variance of possible decision processes. Finally, in Step 5, the mean PAE of the symbolic and the nonsymbolic number line estimation tasks were included. This way, it was possible to gain insight in the contribution of the number–space interaction to mathematics achievement variance.

Results of Model 1 (Table 5) revealed grade ($\beta = .44$, $p = .010$), standardized spelling performance ($\beta = .35$, $p < .001$), mean adjusted symbolic RT ($\beta = -.61$, $p = .010$), and the mean PAE of the nonsymbolic number line estimation ($\beta = -.38$, $p = .001$) as significant and unique predictors of math achievement. As shown in Table 5, spelling performance predicted 11% of the variance, $F_{\text{Change}}(1, 68) = 8.61$, $p = .005$, $R^2 = .07$. In Step 4, decision processes (i.e., number comparison) added significantly to the model, $F_{\text{Change}}(4, 62) = 2.85$, $p = .031$, $R^2 = .18$. Also the number–space interactions added in Step 5 predicted 20% additional variance, $F_{\text{Change}}(2, 60) = 11.74$, $p < .001$, $R^2 = .39$. The magnitude representation level reflected in the priming effects added in Step 3 was not significant.

In a second model (Model 2, Table 6), standardized mathematics achievement measured in the same year in which the children performed the experimental tasks was additionally entered into the model in a sixth block. This was done to examine the contribution of the predictors if additionally controlled for the math achievement score of the previous year. The results of Model 2 showed that the additional incorporation of mathematics achievement of the previous year into the model did not change the results pattern observed in Model 1. However, this additional variable turned out to be significant ($\beta = .32$, $p = .002$), leading to an extra 8% of explained variance, $F_{\text{Change}}(1, 59) = 10.50$, $p = .002$, $R^2 = .47$.

DISCUSSION

The current longitudinal study aimed to investigate the relationship between the performance of a single sample of children on multiple indices of basic number processing tasks and their individual mathematics achievement scores on a curriculum-based standardized math test 1 year later. This way, we wanted to investigate which of these markers are good predictors for later arithmetic performance.

Results revealed that decision processes—especially decisions on symbolic numbers—and an adequate mapping of

Table 5
Hierarchical Regression Analysis Predicting Mathematics Achievement 1 Year Later

Step	*Independent variables*	*Standardized* β	ΔR^2	*Adjusted* R^2
1	Grade_dummy1	.44*	.00	.03
	Grade_dummy2	−.04		
2	Spelling ability	.35***	.01**	.00
3	PDE symbolic	.05	.04	.09
	PDE nonsymbolic	−.10		
4	Slope symbolic	−.14	.03*	.18
	Mean adjusted RT symbolic	−.61*		
	Slope nonsymbolic	−.12		
	Mean adjusted RT nonsymbolic	.22		
5	Mean PAE symbolic	−.25	.20***	.39
	Mean PAE nonsymbolic	−.38**		

Note. Standardized βs from the last step in the analyses are displayed.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Table 6
Hierarchical Regression Analysis Predicting Mathematics Achievement 1 Year Later

Step	*Independent variables*	*Standardized* β	ΔR^2	*Adjusted* R^2
1	Grade_dummy1	.32	.00	−.03
	Grade_dummy2	−.03		
2	Spelling ability	.33***	.11**	.07
3	PDE symbolic	.01	.04	.09
	PDE nonsymbolic	−.19		
4	Slope symbolic	−.11	.13*	.18
	Mean adjusted RT symbolic	−.43		
	Slope nonsymbolic	−.15		
	Mean adjusted RT nonsymbolic	.13		
5	Mean PAE symbolic	−.18	.20***	.39
	Mean PAE nonsymbolic	−.37**		
6	Math achievement year 1	.32**	.08**	.47

Note. Standardized βs from the last step in the analyses are displayed.
*$p < .05$. **$p < .01$. ***$p < .001$.

numbers onto space were the strongest predictors for individual differences in mathematics achievement scores. Children who performed well at comparing digits and who were more precise at placing a (symbolic or) nonsymbolic number on an external number line had a higher score on a curriculum-based math achievement task measured 1 year later. Importantly, these longitudinal correlations held even when controlling for math achievement at the time the number processing measures were tested.

Many previous studies already demonstrated that particularly the performance on the symbolic tasks is related with mathematics (e.g., De Smedt et al., 2009; Holloway & Ansari, 2009; Sasanguie et al., 2012). On the basis of these findings, it is suggested that mathematics achievement is related to how well children could access the corresponding magnitude from the symbol rather than with quantity processing per se (e.g., Rousselle & Noël, 2007). The current finding is strongly in line with this assumption. However, it could be argued that the performance on the symbolic comparison task reflects individual processing speed and that the latter also affects performance in the mathematical achievement test. However, if this was the case, the same would hold for the nonsymbolic number comparison performance, but no correlation was present between the performance on this task and math achievement. Moreover, our mathematical test is an untimed standardized test, implying that if the relation between RTs and mathematics achievement were fully accounted by processing speed, no correlation is expected.

No association has been found between the performance on the nonsymbolic number comparison task and mathematics achievement. Although this finding is in line with findings from previous cross-sectional studies with small numbers from 1 to 9 (e.g., Holloway & Ansari, 2009; Sasanguie et al., 2012), this contrasts with findings from other studies including large numbers (e.g., Halberda & Feigenson, 2008; Inglis, Attridge, Batchelor, & Gilmore, 2011). Therefore, we think that the differences in the number ranges chosen in these studies explain the different results.

In addition, mapping numbers onto space also predicted a large amount of variance in mathematics achievement. Although the results of the regression analysis revealed that only the nonsymbolic number line performance was a significant contributor, also the correlation between the performance in the symbolic number line estimation task and mathematics achievement was significant and the regression coefficient tended to go in the same direction. These findings show the importance of number–space mappings for mathematical ability. Indeed, Kucian et al. (2011) already reported that the ability of children to solve arithmetical problems improved when they practiced their spatial abilities of positioning a number on a number line with the game "Rescue Calcularis."

Finally, no predictive association was observed between mathematics achievement and the performance on the priming tasks. This suggests that the amount of representational overlap of the magnitude representation is not a predictive factor for math achievement performance.

Together, these results lead to the conclusion that mathematical competence is, on one hand, related to efficient symbolic decision making. These culturally determined processes highlight the importance of formal (e.g., classroom activities) and informal (e.g., number board games) learning episodes in which numerals have to be associated with their numerical meaning. On the other hand, mathematical competence is also predicted by the ability of mapping numbers onto space. Again, this ability is a direct reflection of what children "learn" in a certain culture, in this case, the Western one. Mathematical performance in 5- to 7-year-old children thus especially seems to be determined by cultural factors.

Knowledge about the predictive value of these cognitive markers of number processing can help in dealing with the early identification of children at risk for mathematical difficulties. In addition, this information can also guide appropriate educational intervention. More specifically, our data suggest that intervention programs should focus on connecting symbols with their meaning and on mapping numbers onto space. Most educational intervention programs of which the effectiveness has been examined so far included the association of numerical symbols with their meaning as an important instructional principle (e.g., Griffin, 2004; Wilson et al., 2006). In addition, some intervention programs also focused on training with mapping numbers onto space trough games (e.g., Kucian et al., 2011) or via a linear number-board game (e.g., Ramani & Siegler, 2008). Intervention studies with these programs seem promising, but it might be interesting to investigate the impact of these programs on mathematical achievement on larger samples.

Acknowledgments —The authors thank all children, parents, and the school that took part in this study. In addition, we thank Bert De Smedt for the very helpful discussion on the data analysis procedure and Titia Gebuis for her thoughtful comments. This research was supported by the Research Fund KU Leuven.

REFERENCES

Barth, H., Beckmann, L., & Spelke, E. (2008). Nonsymbolic, approximate arithmetic in children: Abstract addition prior to instruction. *Developmental Psychology*, *44*, 1466–1477.

Booth, J. L., & Siegler, R. S. (2006). Developmental and individual differences in pure numerical estimation. *Developmental Psychology*, *41*, 189–201.

Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development*, *79*, 1016–1031.

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.

Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science*, *17*, 401–406.

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

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., Izard, V., & Piazza, M. (2005). *Control over non-numerical parameters in numerosity experiments*. Unpublished manuscript. Retrieved November 16, 2009, from www.unicog.org

Dudal, P. (2000a). *Leerlingvolgsysteem. Rekenen: Toetsen 1-2-3. Basisboek.* Leuven, Belgium: Garant.

Dudal, P. (2000b). *Leerlingvolgsysteem. Spelling: Toetsen 1-6.* Leuven, Belgium: Garant.

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.

Geary, D.C., Bailey, D.H., Littlefield, A., Wood, P., Hoard, M.K., & Nugent, L. (2009). First-grade predictors of mathematical learning disability: A latent class trajectory analysis. *Cognitive Development*, *24*. doi:10.1016/j.cogdev.2009.10.001

Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early Childhood Research Quarterly*, *19*, 173–180.

Halberda, J., & Feigenson, L. (2008). Developmental change in the acuity of the "Number Sense": The approximate number system in 3-, 4-, 5-, and 6-year-olds and adults. *Developmental Psychology*, *44*, 1457–1465.

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature*, *455*, 665–668. doi:10.1038/nature07246

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.

Inglis, M., Attridge, N., Batchelor, S., & Gilmore, C. (2011). Non-verbal number acuity correlates with symbolic mathematics achievement: But only in children. *Psychonomic Bulletin Review*, *18*, 1222–1229.

Kucian, K., Grond, U., Rotzer, S., Henzi, B., Schönmann, C., Plangger, F., et al. (2011). Mental number line training in children with developmental dyscalculia. *NeuroImage.* doi: 10.1016/j.neuroimage.2011.01.070

Moyer, R. S., & Landauer, T. K. (1967). Time required for judgements 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.

Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development*, *79*, 375–394.

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.

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

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

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, *30*, 344–357.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428–444.

Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237–243.

Simon, T. J., Takarae, Y., DeBoer, T., McDonald-McGinn, D. M., Zackai, E. H., & Ross, J. L. (2008). Overlapping numerical cognition impairments in children with chromosome 22q11.2 deletion or Turner syndromes. *Neuropsychologia*, *46*, 82–94.

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.

Verguts, T., Fias, W., & Stevens, M. (2005). A model of exact small-number representation. *Psychonomic Bulletin & Review*, *12*, 66–80.

Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, *74*, B1–B11.

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race", an adaptive computer game for remediation of dyscalculia. *Behavioural and Brain Functions*, *2.* doi:10.1186/1744-9081-2-20