Journal of Experimental Child Psychology 114 (2013) 418–431

Contents lists available at SciVerse ScienceDirect

Journal of Experimental Child Psychology

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

# Approximate number sense, symbolic number processing, or number–space mappings: What underlies mathematics achievement?

Delphine Sasanguie [a,b,*], Silke M. Göbel [c], Kristina Moll [c,d], Karolien Smets [a], Bert Reynvoet [a,b]

[a] Laboratory of Experimental Psychology, KU Leuven, B-3000 Leuven, Belgium
[b] Subfaculty of Psychology and Educational Sciences, KU Leuven Kulak, B-8500 Kortrijk, Belgium
[c] Department of Psychology, University of York, York YO10 5DD, UK
[d] Department of Psychology, University of Graz, A-8010 Graz, Austria

ARTICLE INFO

*Article history:*
Received 18 June 2012
Revised 26 October 2012
Available online 25 December 2012

*Keywords:*
Mathematics achievement
Number comparison
Number line estimation
Individual differences
Number processing development
Different math tests

ABSTRACT

In this study, the performance of typically developing 6- to 8-year-old children on an approximate number discrimination task, a symbolic comparison task, and a symbolic and nonsymbolic number line estimation task was examined. For the first time, children's performances on these basic cognitive number processing tasks were explicitly contrasted to investigate which of them is the best predictor of their future mathematical abilities. Math achievement was measured with a timed arithmetic test and with a general curriculum-based math test to address the additional question of whether the predictive association between the basic numerical abilities and mathematics achievement is dependent on which math test is used. Results revealed that performance on both mathematics achievement tests was best predicted by how well children compared digits. In addition, an association between performance on the symbolic number line estimation task and math achievement scores for the general curriculum-based math test measuring a broader spectrum of skills was found. Together, these results emphasize the importance of learning experiences with symbols for later math abilities.



* Corresponding author at: Subfaculty of Psychology and Educational Sciences, E. Sabbelaan 53, 8500 Kortrijk, Belgium. Fax: +32 56 24 60 52.
*E-mail address:* delphine.sasanguie@kuleuven-kulak.be (D. Sasanguie).

0022-0965/$ - see front matter 
http://dx.doi.org/10.1016/j.jecp.2012.10.012

## Introduction

In numerical cognition, there are two dominant approaches proposing different underlying factors leading to mathematical competence. First, domain-general approaches have proposed that variability in non-numerical skills such as phonological skills, working memory, long-term memory, and visuo-spatial processing underlies individual differences in mathematics achievement (e.g., Geary, 1993; Raghubar, Barnes, & Hecht, 2010). For instance, Passolunghi and Siegel (2004) observed a general working memory deficit in children with difficulties in mathematics. Alternatively, domain-specific approaches have proposed that individual differences in number-specific processes are related to variance in math skills (e.g., Dehaene, Piazza, Pinel, & Cohen, 2003; Piazza et al., 2010). Typically, domain-specific approaches assume an innate capacity for approximate number processing that is shared by human infants and other species (Agrillo, Piffer, Bisazza, & Butterworth, 2012; Cantlon, Platt, & Brannon, 2009). Studies on infants have provided evidence for this approximate number sense (ANS) (Xu & Spelke, 2000). The ANS has been suggested to form the basis of the arithmetic skill of 5-year-old children who had not yet been taught formal arithmetic but could 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 are thought to acquire meaning by being associated with this preexisting nonsymbolic approximate representation (Mundy & Gilmore, 2009). Indeed, in tasks such as magnitude comparison, similar behavioral effects have been found with nonsymbolic and symbolic numbers (Cohen Kadosh, Lammertyn, & Izard, 2008). In the current study, we investigated three domain-specific numerical processes as possible predictors for individual differences in mathematics achievement: acuity of the ANS, performance in symbolic number comparison, and accuracy of number–space mappings of children.

In an ANS task, participants are instructed to decide which of two presented dot arrays contains the larger number of dots. Typically, ratio-dependent performance is observed following Weber–Fechner's law (Fechner, 1860); that is, participants are less accurate in discriminating two numerosities with a smaller ratio (e.g., ratio 1.25, 8 vs. 10 dots) than with a larger ratio (e.g., ratio 2, 8 vs. 16 dots). This ratio-dependent performance pattern can be expressed as a separate Weber fraction ($w$) for each participant. The Weber fraction reflects the minimum change in number that is needed for each participant to perceive a difference in number and, therefore, reflects the acuity of the ANS (Halberda, Mazzocco, & Feigenson, 2008; Piazza, 2010). This individual Weber fraction, as well as the mean accuracy on ANS tasks, has been shown to be related to children's performance on math achievement tests (e.g., Halberda & Feigenson, 2008; Inglis, Attridge, Batchelor, & Gilmore, 2011; Libertus, Feigenson, & Halberda, 2011; Mazzocco, Feigenson, & Halberda, 2011; Piazza et al., 2010). For instance, Inglis et al. (2011) showed that the Weber fraction measured with 7- to 9-year-olds is related to their performance on a math test. However, findings are inconsistent, and studies that did not find a relation between nonsymbolic number comparison and (future) mathematical abilities exist as well (e.g., De Smedt & Gilmore, 2011; Holloway & Ansari, 2009; Sasanguie, De Smedt, Defever, & Reynvoet, 2012; Soltész, Szücs, & Szücs, 2010).

Other authors have argued that it is not the representation of numerosities (as measured by discrimination performance) but rather the speed in accessing nonsymbolic magnitude representations from symbols that is crucial for predicting performance on math achievement tests (Rouselle & Noël, 2007). Evidence for this comes from findings with children with mathematical difficulties who perform worse on symbolic number comparison tasks than typically achieving children. In symbolic comparison tasks, the participants need to decide which is the larger number of a pair of single digits (1–9). This results in a distance effect; that is, reaction times are longer when two numerically close numbers (e.g., 8 vs. 9) need to be compared than when two numerically more distant numbers (e.g., 2 vs. 9) need to be compared. In developmental studies, the size of the distance effect decreases with age, and several studies have shown that either the size of the distance effect or average reaction times on symbolic comparison are related to individual differences in mathematics achievement. No such relationship has been reported for comparisons of dot patterns instead of digits (e.g., De Smedt & Gilmore, 2011; Holloway & Ansari, 2009; Landerl & Kölle, 2009; Lonnemann, Linkersdörfer, Heselhaus, Hasselhorn, & Lindberg, 2011; Sasanguie, De Smedt, et al., 2012). This seems to be at odds with studies that have shown a relationship between the precision of the ANS and math achievement. Recently, Noël and Rouselle (2011) proposed a theoretical explanation in which they assumed that children, in addition to their approximate number system, also develop an exact number system for symbolic number processing. This exact number system becomes connected to the approximate representation during development. In this view, an increase in precision (i.e., more fine-tuned representations in the approximate number system) is a direct consequence of the connection to the exact number system. Children with mathematical problems, however, have an early deficit in symbolic magnitude processing that will in turn lead to a less fine-tuned approximate number processing system (i.e., less ANS acuity), especially at a later age (but see Libertus et al., 2011, and Inglis et al., 2011, for ANS associations with math achievement at early ages).

Another explanation for the inconsistent results between studies investigating the influence of symbolic number processing on math achievement and studies investigating the contribution of the ANS to math achievement may lie in the design of both types of studies. An important difference between these studies is that symbolic comparison is contrasted with nonsymbolic comparison of small numbers ranging from 1 to 9. This is done to allow a direct comparison with the symbolic number condition in which digits from the same range are presented (e.g., Holloway & Ansari, 2009). As a result, these designs contain numerosities from below and above the subitizing range (Trick & Pylyshyn, 1994), whereas numerosities from the subitizing range are typically avoided in studies measuring the ANS acuity. This difference in research design might be responsible for diverging results.

Quite independent from these comparison studies, many studies have also demonstrated a relation between the performance on a number line estimation task and mathematics achievement (e.g., Booth & Siegler, 2006, 2008; Sasanguie, De Smedt, et al., 2012; Sasanguie, Van den Bussche, & Reynvoet, 2012; Schneider, Grabner, & Paetsch, 2009; Siegler & Booth, 2004). In this task, participants are instructed to indicate the position of a number on an empty number line. Sasanguie, De Smedt, et al. (2012), for instance, showed that children who are more precise in estimating the corresponding place on the number line of a digit or a dot pattern have higher scores on a math achievement test. When performing this task, it is typically found that with increasing age the pattern to map numbers on the line moves from a logarithmic pattern, with larger magnitudes put closer together than smaller magnitudes, toward a linear pattern (e.g., Booth & Siegler, 2006; Sasanguie, De Smedt, et al., 2012). Dehaene, Izard, Spelke, and Pica (2008) investigated these number–space mappings in the Mundurucu, an Amazonian indigene group with little or no formal education, and observed that "mapping numbers onto space" is a universal intuition that is logarithmic. As a consequence, these authors concluded that the concept of a linear number line is really a cultural invention that fails to develop in the absence of formal education.

Unfortunately, although these studies had the same aim (i.e., explaining individual variance on a math achievement test), none of them contrasted the performance on several of these tasks by incorporating more than one task in the same study. As a result, it remains unanswered which basic numerical process is the most related to performance on a math test. A first attempt was provided by Sasanguie, Van den Bussche, et al. (2012), who included symbolic and nonsymbolic comparison tasks and number line estimation tasks in their study. It was found that mathematical competence was predicted by symbolic number comparison performance and number–space mappings. However, in that study, nonsymbolic comparison was limited to small numerosities. We already mentioned that it is difficult to generalize these results to typical ANS discrimination tasks that do not include numerosities within the subitizing range. Therefore, the current study examined performance of children between 6 and 8 years of age on (a) an ANS discrimination task with large numerosities, (b) a symbolic comparison task, and (c) a symbolic and nonsymbolic number line estimation task. In addition, as a first, we explicitly investigated which of these tasks best predicted future mathematical abilities.

Furthermore, we questioned whether predictive patterns change as a result of different tasks used to assess math achievement. In general, just one of two different math tests is commonly used in these studies: a timed arithmetic test (e.g., the Woodcock–Johnson III Fluency subtest in Holloway & Ansari, 2009) or a more general curriculum-based test (e.g., the Test of Early Mathematical Ability – Second Edition [TEMA-2] in Halberda et al., 2008). Rarely are both tests used at the same time. Therefore, we also examined whether this may have contributed to the conflicting findings. Preliminary evidence for this comes from a study of Holloway and Ansari (2009). They demonstrated that performance on the symbolic comparison task correlates significantly with a timed mathematics fluency subtest but not with an untimed calculation subtest of the Woodcock–Johnson III Tests of Achievement (Woodcock, McGrew, & Mather, 2001). In the current study, therefore, children's math achievement scores were measured 1 year later with both type of tests, and we examined whether the predictive value of the basic number tasks differed depending on the math test. To investigate whether the association between the number processing tasks and mathematics achievement is specific, 1 year later, simultaneously with these math tests, we also administered a curriculum-based spelling test.

## Methods

### *Participants*

A total of 92 children in first, second, and third grades (between 6 and 8 years old) were recruited from an elementary school in a middle to higher income neighborhood in Flanders, Belgium. These age groups were chosen because previous developmental studies have shown that many developmental changes occur at this age (e.g., Holloway & Ansari, 2009; Siegler & Booth, 2004). Of the original sample, 8 children were removed from the analyses because they were outliers (i.e., 3 standard deviations above/below the group average: nonsymbolic comparison [$n = 2$], symbolic comparison [$n = 2$], and number line estimation [$n = 2$]) or because they performed at chance level in the nonsymbolic comparison task ($n = 2$). In addition, participants whose performance on the nonsymbolic comparison task resulted in a bad model fit in order to compute the Weber fraction were removed from the analyses ($n = 10$; see Results for details). Finally, 3 children were discarded from the analyses because of missing math achievement or spelling achievement data. The final sample consisted of 71 typically developing children, comprising 19 first graders ($M_{age}$(t1) = 6.6 years, $SD = 0.24$, 10 boys and 9 girls), 24 second graders ($M_{age}$(t1) = 7.7 years, $SD = 0.30$, 7 boys and 17 girls), and 28 third graders ($M_{age}$(t1) = 8.6 years, $SD = 0.24$, 12 boys and 16 girls).

### *Procedure*

The experimental tasks were administered in March 2011 (Time 1, t1). For all experimental measures, the children were tested in a separate room, accompanied by two experimenters, in groups of 7 to 10. All children first performed the symbolic and nonsymbolic number line estimation tasks, followed by the symbolic and nonsymbolic comparison tasks. A short break between each task was provided. After the experiment, the children received a small reward. At the same time (t1), the score on the general curriculum-based test for mathematics was collected for a first time by classroom testing. One year later, around February 2012 (Time 2, t2), the scores on the two standardized math tests (i.e., general curriculum-based test and a timed arithmetic test) and the spelling test were collected by classroom testing.

### *Measures*

#### *Experimental tasks*

*Nonsymbolic comparison task.* The nonsymbolic comparison task was similar to the task used by Piazza, Izard, Pinel, Le Bihan, and Dehaene (2004). Stimuli were pairs of arrays of yellow squares presented in two gray areas (450 $\times$ 450 pixels) on the left- and right-hand sides of the 14-inch screen (see Fig. 1). One array always contained the reference numerosity of 16 squares, whereas the other array contained 6, 8, 10, 12, 13, 14, 15, 17, 18, 19, 20, 22, 24, or 26 squares, resulting in 14 possible pairs. Each of the 14 pairs was repeated 10 times, with each numerosity being presented 5 times on the left-hand side and five times on the right-hand side. The experiment started with 3 training trials before the 140 test trials were given in a pseudo-randomized order. The program (Neurobehavioral Systems, http://www.neurobs.com) randomly allocated the location of the squares within the gray area while running the task. To avoid that performance was based on non-numerical parameters, such as total surface area and item size (Dehaene, Izard, & Piazza, 2005), the surface area between the

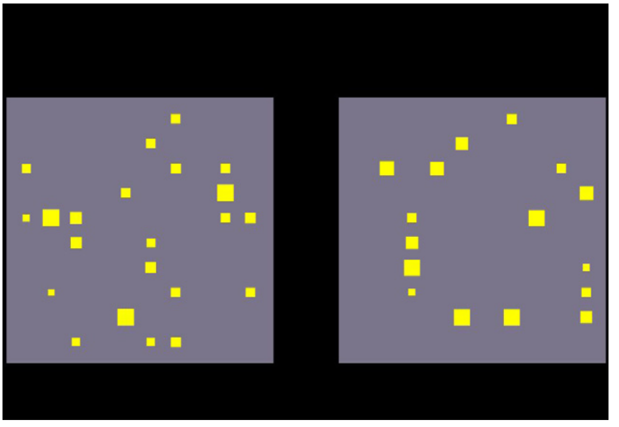

**Fig. 1.** Sample stimulus used in the nonsymbolic comparison task. Participants indicated whether there were more yellow squares in the left or right gray array while visual parameters such as total area and item size were controlled for. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

two arrays was kept constant and mixed-sized item sets were used. The size of squares varied within each array but always contained at least 1/6 of the smallest item size (12 × 12 pixels) and 1/6 of the largest item size (28 × 28 pixels) of squares allowed in the task. The remaining stimuli were created online, with item sizes ranging in between the smallest and largest item sizes. Participants were asked to indicate without counting which of the arrays consisted of more squares by pressing the left or right button on the keyboard. Stimuli remained on the screen until the participants responded, and the next trial was presented after an interstimulus interval of 300 ms.

*Symbolic comparison task.* Stimulus presentation and recording of the data were controlled by E-Prime 1.1 (Psychology Software Tools, http://www.pstnet.com). A trial started with a fixation cross for 600 ms, after which two stimuli appeared. Stimuli were white Arabic digits on a black background (Arial font, 16), simultaneously presented 4.25 cm to the left and right from the center of the screen, and remained on the screen until the children responded. Participants needed to select the larger number by pressing the left or right button on the keyboard. Children were asked to respond as quickly as possible without making errors. The intertrial interval was 1000 ms. The stimulus set consisted of numbers 1 to 9, but only combinations of stimuli with a maximum distance of 5 between both numbers were presented to the children, resulting in 60 trials. This was done because in a previous study (Sasanguie, De Smedt, et al., 2012) reaction times decreased with increasing distance, but from a distance of 5 onward any further decrease was minimal. Five practice trials were included, and feedback on accuracy was provided during practice trials only.

*Number line estimation tasks.* Children were presented with 25-cm-long lines in the center of white A4 sheets (i.e., 29.7 × 21 cm). The 0-to-100 interval was administered in both a symbolic format and a nonsymbolic format. Nonsymbolic stimuli were generated using the MATLAB program as described by Dehaene et al. (2005). The dots of a single stimulus were displayed in a circle with a radius of 45 mm. The end points of the number lines were labeled by an empty circle on the left and a circle with 100 dots on the right in the nonsymbolic condition. In the symbolic condition, the dots were replaced by Arabic digits. The target quantity that needed to be positioned was shown in the center of the sheet, 6 cm above the number line. The quantities used in the experiment (2, 3, 4, 6, 18, 25, 48, 67, 71, and 86) were taken from Siegler and Opfer (2003). The presentation order of the quantities was randomized, and each line was presented on a separate sheet. Children were instructed to put a mark on the line where they thought the quantity needed to be positioned. To ensure that the children were 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 100 [dots]. If here is 0 and here is 100, where would you position this number [quantity]?" Afterward, children went through all sheets at their own pace.

*Standardized tests*

*Timed arithmetic test.* The TTR (Tempo Test Rekenen; De Vos, 1992) is a test consisting of 200 arithmetic number fact problems presented in five rows (one row with addition, one row with subtraction, one row with division, one row with multiplication, and one mixed problem row). Within each row, the problems increase in difficulty. Children needed to solve as many items as possible within 1 min per row. The TTR is a standardized test that is frequently used in the Flemish education system. The test is identical for each grade. For the analyses, the total number of correct items was transformed into a *z*-score for each grade separately.

*General curriculum-based math achievement test.* The Flemish Student Monitoring System (Dudal, 2000a) consists of 60 items covering number knowledge, understanding of operations, (simple) arithmetic, word problem solving, measurement, and geometry, based on the curriculum per grade. This test specifically focuses on what the children are supposed to have learned during formal math education according to their grade curriculum. Thus, the same constructs are included in each grade but have a different difficulty level. The maximum score is 60, with 1 point for each correct answer. For the analyses, the scores were transformed into a *z*-score for each age group separately. The reliability indexes of Kuder–Richardson (KR 20) for the test are .90, .89, .90, and .89 for the first, second, third, and fourth grades, respectively.

*Curriculum-based spelling achievement test.* The spelling test of the Flemish Student Monitoring System (Dudal, 2000b) was used to measure children's spelling skills. This test involves the dictation of letters, words, and sentences. The maximum score is 60, with 1 point for each correct answer. For the analyses, the scores were transformed into a *z*-score for each age group separately. The reliability index of Kuder–Richardson (KR 20) for this test is .90 for second, third, and fourth grades.

## Results

In this section, we focus on the associations between the experimental tasks and the math achievement tests. The detailed results for each number task are described in the Supplementary material. In sum, all tasks showed the number processing effects typically described in the literature (see Table 1). The nonsymbolic comparison task revealed a ratio effect; children were more accurate when the difference between the two numerosities to be discriminated was larger. Moreover, the computed Weber fractions ($w$) decreased with grade. In the symbolic comparison task, a distance effect was present and children became faster with increasing grade. In line with previous research (e.g., Holloway & Ansari, 2009), a numerical distance effect was calculated for each individual using the following formula. The average of the median reaction times on trials with numerical distances of 4 and 5 was subtracted from the average of the median reaction times on trials with numerical distances of 1 and 2. The difference was then divided by the average of the median reaction times on trials with numerical distances of 4 and 5 to correct for individual differences in reaction times. In the number line estimation tasks, children showed a more linear estimation pattern with increasing age and error rates decreasing with grade for both the symbolic and nonsymbolic number line tasks. Children's estimation accuracies were obtained by computing the percentage of absolute error per child according to the following equation (Siegler & Booth, 2004):

$$\left|\frac{\text{Estimate} - \text{Estimated Quantity}}{\text{Scale of Estimates}}\right|.$$

**Table 1**
Children's performance on the nonsymbolic comparison task, the symbolic comparison task, and the number line estimation tasks, displayed by grade.

| | Nonsymbolic comparison task | Symbolic comparison task | | Number line estimation tasks | |
|---|---|---|---|---|---|
| | Weber fraction (w) | Reaction time (ms) | Accuracy (%) | Symbolic notation mean PAE (%) | Nonsymbolic notation mean PAE (%) |
| First graders | .96 (.71) | 1043.71 (238.33) | 94.41 (5) | 8.63 (5) | 14.16 (5) |
| Second graders | .63 (.40) | 899.31 (159.56) | 96.11 (3) | 7.58 (3) | 13.42 (5) |
| Third graders | .56 (.35) | 784.53 (103.58) | 97.44 (2) | 4.93 (3) | 10.96 (5) |

*Note:* PAE, percentage of absolute error.

For example, if a child was asked to estimate 18 on a 0-to-100 number line and placed the mark at the point on the line corresponding to 30, the percentage of absolute error would be (30–18)/100 or 12%.

### *Correlation analysis*

We examined the relation between the indexes of the different experimental measures (see Table 2). For the nonsymbolic comparison task, the Weber fractions and the mean accuracies were used as indexes. For the symbolic comparison task, we used the median reaction times and the individual distance effects. As indexes for the number line estimation tasks, the symbolic and nonsymbolic mean percentages of absolute error were used. To exclude that correlations were due to differences between grades or general abilities, we controlled for grade and general spelling achievement. There was a high negative correlation between the Weber fraction and the mean accuracies of the children on the nonsymbolic comparison task, $r(67) = -.76$, $p < .001$, indicating that the more accurate children were on this task, the lower their Weber fraction was. There was also an association between reaction times on the symbolic comparison task and mean percentage of absolute error on the symbolic number line estimation task, $r(67) = .37$, $p < .01$, demonstrating that children who were faster at comparing digits were also better at placing digits on an external number line. In addition, a large positive significant correlation was found between both math achievement tests, $r(67) = .59$, $p < .001$.

More important for the current purposes of the study, we also investigated the relation between the individual differences on the experimental measures and the two standardized mathematics achievement scores measured 1 year later. Math achievement measured with the timed arithmetic test correlated negatively with reaction times on the symbolic number comparison task, $r(67) = -.35$, $p < .01$, indicating that children with a high score on this timed math test were faster in comparing digits. Other correlations were not significant. When math achievement was measured with the general curriculum-based math test, significant correlations were found between mathematics achievement and symbolic comparison, $r(67) = -.37$, $p < .01$. In addition, math achievement also correlated with performance on the symbolic number line estimation task, $r(67) = -.40$, $p < .01$, and the nonsymbolic number line estimation task, $r(67) = -.25$, $p = .04$, suggesting that children with a high score on the curriculum-based math test were more accurate in placing numbers on a number line.

### *Development of the associations over grades*

We also examined whether the associations between the experimental tasks and mathematics achievement changed over grades. Therefore, we performed analyses of covariance with the indexes of the experimental tasks (see Table 1) as dependent variables, grade as a factor, and mathematics achievement and the interaction between grade and mathematics achievement as covariates. If the interaction between grade and mathematics achievement were significant, this would indicate that for performance on the experimental task, the association with mathematics achievement changed over grades. These analyses were conducted for both measures of mathematics achievement.

For the timed arithmetic test, the grade by mathematics achievement interaction was significant for reaction times on the symbolic number comparison task, $F(3,67) = 3.43$, $p = .02$, $\eta_p^2 = .13$. To examine this interaction in more detail, Pearson correlation coefficients were calculated per grade. Although these associations revealed a similar pattern in all grades, the largest association was found in first graders: $r(16) = -.44$, $p = .07$ for first graders, $r(21) = -.35$, $p = .10$ for second graders, and $r(25) = -.35$, $p = .08$ for third graders.

For the curriculum-based math test, the grade by mathematics achievement interaction was significant for reaction times on the symbolic number comparison task, $F(3,67) = 3.52$, $p = .02$, $\eta_p^2 = .14$, performance on the symbolic number line estimation task, $F(3,67) = 3.08$, $p = .03$, $\eta_p^2 = .12$, and performance on the nonsymbolic number line estimation task, $F(3,67) = 2.98$, $p = .04$, $\eta_p^2 = .12$. Pearson correlation coefficients were calculated per grade to investigate the pattern of associations. For reaction times on the symbolic comparison task, associations were $r(16) = -.33$, $p = .19$ for first graders, $r(21) = -.53$, $p < .01$ for second graders, and $r(25) = -.26$, $p = .20$ for third graders, indicating that the association is the largest for second graders. For performance on the symbolic number line estimation task, the associations per grade showed a highly similar pattern: $r(16) = -.42$, $p = .08$ for first graders, $r(21) = -.48$, $p = .02$ for second graders, and $r(25) = -.39$, $p = .04$ for third graders. However, for mean performance on the nonsymbolic number line estimation task, Pearson correlations showed only a negative association, indicating that children with higher math scores made fewer errors in the case of second graders, $r(21) = -.37$, $p = .08$, and third graders, $r(25) = -.44$, $p = .02$, but not in the case of first graders, $r(16) = .18$, $p = .49$. The other interactions were not significant, indicating that there were no grade differences in the correlations between the other experimental tasks and mathematics achievement.

## *Regression analysis*

To examine which of these experimental measures predicted unique variance in both mathematics achievement scores, two hierarchical multiple regression analyses were conducted with each math test as a dependent variable and the experimental measures as predictors. Blocks of independent variables were added in a stepwise procedure. This method allowed us to control for nonexperimental variables (i.e., grade and general abilities) before investigating the unique contribution of the experimental measures to the variance in mathematics achievement. Grade was entered in Step 1 after

**Table 2**

Partial correlations between the diverse indexes of the experimental measures and between these measures and mathematics achievement measured 1 year later with the two math tests (see frame), controlling for grade and spelling achievement ($N = 71$).

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| 1. Standardized score on timed arithmetic test | | | | | | | |
| 2. Standardized score on general curriculum-based math achievement test | .59*** | | | | | | |
| 3. Nonsymbolic comparison Weber fraction | −.17 | −.17 | | | | | |
| 4. Nonsymbolic comparison mean accuracy | .14 | .09 | −.76*** | | | | |
| 5. Symbolic comparison median RT | −.35** | −.37** | .09 | .06 | | | |
| 6. Symbolic comparison DE | .07 | .06 | −.05 | .04 | −.04 | | |
| 7. Symbolic 0-to-100 line mean PAE | −.24 | −.40** | .16 | −.08 | .37** | .06 | |
| 8. Nonsymbolic 0-to-100 line mean PAE | −.19 | −.25* | −.01 | .01 | .01 | .14 | .17 |

*Note:* RT, reaction time; DE, distance effect; PAE, percentage of absolute error.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

recoding this variable in two dummies by means of effect recoding (i.e., Dummy 1: Grade 1 = 1, Grades 2 and 3 = 0; Dummy 2: Grade 2 = 1, Grades 1 and 3 = 0). To control for general abilities, spelling performance, which should also be mediated by general abilities, was entered into the model in Step 2. Finally, in Step 3, the experimental measures were added.

The model for the timed arithmetic test accounted in total for 21% of the variance (see Table 3). After controlling for differences in spelling achievement (11% of variance), reaction time on the symbolic comparison task was the only significant predictor for the timed arithmetic test.

The model for the general curriculum-based math test explained in total 26% of the variance (see Table 4). After controlling for spelling achievement (7% of variance), reaction time on the symbolic comparison task and performance on the symbolic number line estimation task were the only significant predictors.

Finally, we conducted a regression analysis with the scores on the curriculum-based math test measured at t1 also incorporated into the model (in Step 3) before putting the experimental measures in the regression model. This way, it can be examined whether the predictive associations between the basic number processing tasks and math achievement at t2 (see Table 4) remain after controlling for math achievement at t1. Because we collected scores only on the curriculum-based test at t1 (and not on the timed test), this additional regression analysis was performed only with the curriculum-based test at t2 as the dependent variable.

This model explained in total 54% of the variance for the general curriculum-based math test (see Table 5). After controlling for spelling achievement (7% of variance), math achievement measured at t1 explained 51% of the variance in math achievement measured at t2. The number processing indexes entered in the fourth step were no longer significant predictors.

## Discussion

In this study, we investigated which basic numerical processes (i.e., numerosity discrimination, symbol processing, or number line estimation) assessed in Grades 1 to 3 was most related to individual differences in math achievement 1 year later. In addition, we wanted to examine whether the type of math achievement test (timed arithmetic vs. untimed, general curriculum-based) has an impact on the predictive power of the basic numerical tasks. The typical results on nonsymbolic and symbolic comparison tasks and on symbolic and nonsymbolic number line estimation tasks were obtained. We then investigated whether the children's performances on these tasks were related to their math achievement scores. The math achievement scores of the children on both math achievement tests were predicted by symbolic number processing.

Children who were fast at comparing digits performed better on the math tests. In addition, performance on the curriculum-based math test was also predicted by the accuracy of placing digits on an external number line. When the math achievement scores measured at t1 were incorporated into the regression analysis, the number processing measures no longer contributed to the variance in math achievement measured at t2. This indicates that the indexes of the experimental tasks do not predict later math achievement on top of math achievement measured 1 year earlier. This is not surprising given the high correlation between the scores on the math achievement tests at t1 and t2, $r(69) = .72$, $p < .001$. However, the predictive power of the experimental indexes as presented in Table 4 are still interesting because they provide a more refined picture of the basic representations and processes that are associated with later math achievement.

The current findings demonstrate that symbolic number comparison is clearly associated most with the children's mathematics achievement measured 1 year later. This association was particularly strong in the youngest groups and was somewhat reduced in older children, a finding that is consistent with previous studies (e.g., Holloway & Ansari, 2009; Sasanguie, De Smedt, et al., 2012). Younger children (i.e., first and second graders) are in the process of mapping symbolic digits onto preexisting nonsymbolic representations (Barth et al., 2008; Mundy & Gilmore, 2009), which might explain why symbolic number comparison is more sensitive to explaining individual differences in math achievement at this age. It is unlikely that this association is the result of an underlying general processing speed mechanism because the association was present for both the timed arithmetic test and the

**Table 3**
Hierarchical regression analysis predicting mathematics achievement 1 year later measured with a timed arithmetic test.

| Variable | Standardized $\beta$ | $t$ | $p$ |
|---|---|---|---|
| Grade_dummy1 | .00 | 0.00 | 1.00 |
| Grade_dummy2 | .00 | 0.00 | 1.00 |
| $F_{Change} < 1$ | | | |
| Spelling achievement | .38 | 3.39 | .001 |
| $F_{Change}$ $(1,67) = 11.48$, $p < .01$, $R^2 = .15$ | | | |
| Nonsymbolic comparison Weber fraction | −.02 | −0.10 | .923 |
| Nonsymbolic comparison mean accuracy | .13 | 0.75 | .456 |
| Symbolic comparison mean RT | −.36 | −2.59 | .012 |
| Symbolic comparison DE | .08 | 0.71 | .482 |
| Symbolic 0-to-100 line mean PAE | −.08 | −0.63 | .532 |
| Nonsymbolic 0-to-100 line mean PAE | −.18 | −1.59 | .118 |
| $F_{Change}(6,61) = 2.50$, $p = .03$, $R^2 = .17$ | | | |

*Note:* RT, reaction time; DE, distance effect; PAE, percentage of absolute error.

**Table 4**
Hierarchical regression analysis predicting mathematics achievement 1 year later measured with a general, curriculum-based math test.

| Variable | Standardized $\beta$ | $t$ | $p$ |
|---|---|---|---|
| Grade_dummy1 | .00 | 0.00 | 1.00 |
| Grade_dummy2 | .00 | 0.00 | 1.00 |
| $F_{Change} < 1$ | | | |
| Spelling achievement | .33 | 2.86 | .006 |
| $F_{Change}(1,67) = 8.18$, $p < .01$, $R^2 = .11$ | | | |
| Nonsymbolic comparison Weber fraction | −.08 | −0.45 | .656 |
| Nonsymbolic comparison mean accuracy | .03 | 0.15 | .885 |
| Symbolic comparison mean RT | −.30 | −2.25 | .028 |
| Symbolic comparison DE | .08 | 0.78 | .439 |
| Symbolic 0-to-100 line mean PAE | −.27 | −2.14 | .037 |
| Nonsymbolic 0-to-100 line mean PAE | −.22 | −1.97 | .054 |
| $F_{Change}(6,61) = 3.93$, $p < .01$, $R^2 = .25$ | | | |

*Note:* RT, reaction time; DE, distance effect; PAE, percentage of absolute error.

**Table 5**
Hierarchical regression analysis predicting mathematics achievement 1 year later measured with a general, curriculum-based math test, controlling for math achievement at t1.

| Variable | Standardized $\beta$ | $t$ | $p$ |
|---|---|---|---|
| Grade_dummy1 | .00 | 0.00 | 1.00 |
| Grade_dummy2 | .00 | 0.00 | 1.00 |
| $F_{Change} < 1$ | | | |
| Spelling achievement | .33 | 2.86 | .006 |
| $F_{Change}(1,67) = 8.18$, $p < .01$, $R^2 = .11$ | | | |
| Math achievement at t1 | .67 | 7.86 | .000 |
| $F_{Change}(1,66) = 61.80$, $p < .001$, $R^2 = .43$ | | | |
| Nonsymbolic comparison Weber fraction | −.15 | −1.08 | .286 |
| Nonsymbolic comparison mean accuracy | −.03 | −0.23 | .816 |
| Symbolic comparison mean RT | −.13 | −1.18 | .244 |
| Symbolic comparison DE | .05 | 0.60 | .553 |
| Symbolic 0-to-100 line mean PAE | −.07 | −0.64 | .522 |
| Nonsymbolic 0-to-100 line mean PAE | −.16 | −1.87 | .066 |
| $F_{Change}(6,60) = 1.54$, $p = .18$, $R^2 = .06$ | | | |

*Note:* RT, reaction time; DE, distance effect; PAE, percentage of absolute error.

untimed general curriculum-based math test. Processing symbolic numbers seems to be important for later mathematical skills. When digits need to be compared, they must first be associated with their corresponding magnitude representation. In a next step, a decision would need to be made on the activated representations, which typically results in a distance effect (Moyer & Landauer, 1967; Restle, 1970). Previous studies have indeed reported a relationship between the distance effect in symbolic comparison and mathematics achievement, indicating that the decisional aspect is problematic in digit comparison (e.g., De Smedt, Verschaffel, & Ghesquière, 2009; Holloway & Ansari, 2009). It is possible that the larger distance effect observed in children with mathematics difficulties is due to a more imprecise representation of Arabic numbers (e.g., Holloway & Ansari, 2009). However, an association between the distance effect and mathematics achievement was absent in the current study, which is in line with other studies (e.g., De Smedt & Gilmore, 2011; Landerl, Fussenegger, Moll, & Willburger, 2009; Sasanguie, Van den Bussche, et al., 2012). Taking a closer look at the studies that reported an association between the distance effect and mathematics achievement, it becomes clear that the amount of variance that could be explained by the size of the symbolic distance effect was generally relatively smaller than the variance explained by general reaction times. In addition, associations between the distance effect and mathematics achievement were limited to younger participants, at least in some studies (e.g., Holloway & Ansari, 2009). Therefore, we argue that it is not an imprecise representation but rather a delayed activation of the corresponding magnitude of a symbol that is responsible for the association between symbolic comparison and mathematics achievement. This is in line with the so-called "access-deficit hypothesis" proposed by Rouselle and Noël (2007).

In contrast to some previous studies (e.g., Halberda & Feigenson, 2008; Inglis et al., 2011; Libertus et al., 2011; Piazza et al., 2010), neither the accuracy nor the Weber fraction derived from the nonsymbolic comparison task was related to math achievement, suggesting that the innate approximate representation of numbers does not underlie later mathematics achievement in 6- to 8-year-old children. Recently, Price, Palmer, Battista, and Ansari (2012) put forward several possibilities for the absence of a relation between ANS acuity and math achievement in some studies (e.g., Holloway & Ansari, 2009; Iuculano, Tang, Hall, & Butterworth, 2008; Price et al., 2012). They proposed that differences between studies, such as sample size, age group, type of arithmetic test, and nature of the ANS metric, might be responsible for the inconsistent findings. However, none of these explanations seems to account for the absence of an association between ANS acuity and math achievement in the current study. The sample size and age range were similar to those in studies having found a relationship between the ANS and math achievement (e.g., Halberda & Feigenson, 2008; Inglis et al., 2011). Moreover, we used different metrics of ANS performance as well as different math tests. Soltész et al. (2010) suggested that the visual cues of the number stimuli (e.g., surface, diameter) might play a significant role in the variance observed in the outcomes of nonsymbolic number studies. These researchers showed that 4-year-olds have the tendency to respond to the visual properties of the stimuli instead of number when they need to decide which stimulus contains more items. Such biases were also observed in older children (Inglis et al., 2011) and adults (Gebuis & Reynvoet, 2012b; Gilmore, Attridge, & Inglis, 2011). Furthermore, the methods used to control the visual cues in studies might, thus, account for the discrepancies in the results (Gebuis & Reynvoet, 2012a). More research focusing on subtle differences in the procedure, such as stimuli used (but also presentation rate, etc.), is required to unravel why an association between the ANS acuity and math achievement in children sometimes is found and sometimes is not found.

Although we have considered the observed associations between experimental tasks and math ability as evidence in favor of a delayed access from symbols to their corresponding quantity in participants with lower math achievement scores (see also Mundy & Gilmore, 2009), an alternative explanation cannot be excluded. Recently, some researchers have argued in favor of two independent representations: an approximate representation for large numerosities and an exact representation for symbols (e.g., Lyons, Ansari, & Beilock, 2012; Noël & Rouselle, 2011). Both representations need to be related to one another in the process of acquisition of symbols (Noël & Rouselle, 2011). According to Lyons et al. (2012), the approximate nonsymbolic system and the acquired exact symbolic system gradually diverge again, resulting in two rather distinct representations during adulthood. According to this theoretical framework, the current results show that the exact symbolic system explains variance in individual mathematics achievement at least during the primary school years. However, our results cannot distinguish between either accounts with one magnitude representation onto which symbols are mapped or accounts with distinct representations for nonsymbolic stimuli and symbols.

Another goal of this study was to clarify whether the type of math test that is used to assess math achievement may have led to diverging results between studies. A strong positive correlation between both math tests was observed. Therefore, it is not surprising that the results of the current study seem to suggest more similarities than differences if another type of math test is used. Both performance on a timed arithmetic test, which was identical for each grade, and on a more general curriculum-based test, adjusted to the level that needs to be acquired by the children in each grade in terms of geometry, measurement, word problem solving, and so forth, were associated with reaction times on the symbolic comparison task. Individual differences on the curriculum-based test were also explained by performance on the symbolic number line estimation task, which in turn can be explained by the fact that the curriculum-based math test measures a broader spectrum of skills that are, among others, more related to mapping digits onto space (e.g., spatial abilities, order aspects, left and right, before and after) than the arithmetic test. The observation that performance on the general curriculum-based math test was better predicted by the number–space mapping of symbols than that of dot patterns is also evidence in favor of symbolic number processing as the main contributor to individual variance in mathematics achievement (i.e., the "access deficit hypothesis"; Rouselle & Noël, 2007). In particular, this association between number–space mappings and math achievement was stronger in the oldest groups (i.e., second and third graders). This might be due to the fact that a 0-to-100 number range was used in all age groups. Not all children in the first grade are completely familiar with this number range, leading to more random variance in number line positioning in the youngest group and a smaller association with mathematics achievement (e.g., Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Siegler & Booth, 2004).

To conclude, the current findings revealed that children's performance on an ANS task was not associated with their math achievement score, implying that the innate representation of number is not predictive of children's later math achievement. In contrast, symbolic number processing seems to predict math achievement of 6- to 8-year-olds measured 1 year later. When math achievement was measured with a curriculum-based math test, the accuracy of placing digits on an external number line had additional predictive value. These findings suggest that children who have mathematical difficulties struggle with delayed activation from the corresponding magnitude from a symbol or a deficient exact representational system for symbols. Thus, focusing on symbols and the translation to their corresponding representation is very important in learning contexts because this process seems so crucial for later math achievement.

## Acknowledgments

This research was supported by the Research Fund K.U. Leuven. The authors thank all children, parents, and schools that took part in this study. In addition, special thanks go to Giovanni Delaere for checking the manuscript for grammatical errors.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.jecp.2012.10.012.

## References

Agrillo, C., Piffer, L., Bisazza, A., & Butterworth, B. (2012). Evidence for two numerical systems that are similar in humans and guppies. *PLoS One, 7*, 1923–2012.

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

Berteletti, I., Lucangeli, D., Piazza, M., Dehaene, S., & Zorzi, M. (2010). Numerical estimation in preschoolers. *Developmental Psychology, 46*, 545–551.

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.
Cantlon, J. F., Platt, M. L., & Brannon, E. M. (2009). Beyond the number domain. *Trends in Cognitive Sciences, 13*, 83–91.
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.
Dehaene, S., Izard, V., & Piazza, M. (2005). *Control over non-numerical parameters in numerosity experiments.* Unpublished manuscript. Available from http://www.unicog.org/docs/DocumentationDotsGeneration.doc.
Dehaene, S., Izard, V., Spelke, E., & Pica, P. (2008). Log or linear? Distinct intuitions of the number scale in Western and Amazonian indigene cultures. *Science, 320*, 1217–1220.
Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506.
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., 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.
De Vos, T. (1992). *TTR: Tempotest rekenen*. Lisse, The Netherlands: Swets & Zeitlinger.
Dudal, P. (2000a). *Leerlingvolgsysteem, Rekenen: Toetsen 1-2-3, Basisboek*. Leuven, Belgium: Garant.
Dudal, P. (2000b). *Leerlingvolgsysteem, spelling: Toetsen 1–6*. Leuven, Belgium: Garant.
Fechner, G. T. (1860). *Elemente der Psychophysik*. Wiesbaden, Germany: Breitkopf & Hartel.
Geary, D. C. (1993). Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin, 114*, 345–362.
Gebuis, T., & Reynvoet, B. (2012a). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General, 141*, 642–648.
Gebuis, T., & Reynvoet, B. (2012b). The role of visual information in numerosity estimation. *PLoS One, 7*, e37426.
Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *Quarterly Journal of Experimental Psychology, 64*, 2099–2109.
Halberda, J., & Feigenson, L. (2008). Set representations required for the acquisition of the "natural number" concept. *Behavioral and Brain Sciences, 31*, 655–656.
Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with math achievement. *Nature, 455*, 665–668.
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.
Iuculano, T., Tang, J., Hall, C. W. B., & Butterworth, B. (2008). Core information processing deficits in developmental dyscalculia and low numeracy. *Developmental Science, 11*, 669–680.
Landerl, K., Fussenegger, B., Moll, K., & Willburger, E. (2009). Dyslexia and dyscalculia: Two learning disorders with different cognitive profiles. *Journal of Experimental Child Psychology, 103*, 309–324.
Landerl, K., & Kölle, C. (2009). Typical and atypical development of basic numerical skills in elementary school. *Journal of Experimental Child Psychology, 103*, 546–565.
Libertus, M., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292–1300.
Lonnemann, J., Linkersdörfer, J., Heselhaus, V., Hasselhorn, M., & Lindberg, S. (2011). Relations between balancing and arithmetic skills in children: Evidence of cerebellar involvement? *Journal of Neurolinguistics, 24*, 592–601.
Lyons, I. M., Ansari, D., & Beilock, S. L. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General, 141*, 635–641.
Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers' precision of the approximate number system predicts late school mathematics performance. *PLoS One, 6*, e23749.
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.
Noël, M.-P., & Rouselle, L. (2011). Developmental changes in the profiles of dyscalculia: An explanation based on a double exact-and-approximate number representation model. *Frontiers in Human Neuroscience, 5*, 165.
Passolunghi, M. C., & Siegel, L. S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology, 88*, 348–367.
Piazza, M. (2010). Neurocognitive start-up tools for symbolic number representations. *Trends in Cognitive Sciences, 14*, 542–551.
Piazza, M., Facoetti, A., Trussardi, A., Berteletti, I., Conte, S., Lucangeli, D., et al (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*, 33–41.
Piazza, M., Izard, V., Pinel, P., Le Bihan, D., & Dehaene, S. (2004). Tuning curves for approximate numerosity in the human intraparietal sulcus. *Neuron, 44*, 547–555.
Price, G. R., Palmer, D., Battista, C., & Ansari, D. (2012). Nonsymbolic numerical magnitude comparison: Reliability and validity of different task variants and outcome measures, and their relationship to arithmetic achievement in adults. *Acta Psychologica, 140*, 50–57.
Raghubar, K., Barnes, F., & Hecht, S. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences, 20*, 110–122.
Restle, F. (1970). Speed of adding and comparing numbers. *Journal of Experimental Psychology, 91*, 191–205.
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., De Smedt, B., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology, 30*, 344–357.

Sasanguie, D., Van den Bussche, E., & Reynvoet, B. (2012). Predictors for mathematics achievement? Evidence from a longitudinal study. *Mind, Brain, and Education, 6*, 119–128.
Schneider, M., Grabner, R. H., & Paetsch, J. (2009). Mental number line, number line estimation, and mathematical achievement: Their interrelations in Grades 5 and 6. *Journal of Educational Psychology, 101*, 359–372.
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.
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.
Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80–102.
Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock–Johnson III tests of achievement*. Itasca, IL: Riverside.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1–B11.