Journal of Experimental Child Psychology 150 (2016) 207–226

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

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

The precision of mapping between number words and the approximate number system predicts children's formal math abilities

Melissa E. Libertus [a,b,*], Darko Odic [b,c], Lisa Feigenson [b], Justin Halberda [b]

[a] Department of Psychology and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260, USA
[b] Department of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218, USA
[c] Department of Psychology, University of British Columbia, Vancouver, British Columbia V6T 1Z4, Canada

ARTICLE INFO

Article history:
Received 3 November 2015
Revised 31 May 2016
Available online 24 June 2016

Keywords:
Number estimation
Approximate number system (ANS)
Informal math
Formal math
Individual differences
ANS-number word mapping

ABSTRACT

Children can represent number in at least two ways: by using their non-verbal, intuitive approximate number system (ANS) and by using words and symbols to count and represent numbers exactly. Furthermore, by the time they are 5 years old, children can map between the ANS and number words, as evidenced by their ability to verbally estimate numbers of items without counting. How does the quality of the mapping between approximate and exact numbers relate to children's math abilities? The role of the ANS–number word mapping in math competence remains controversial for at least two reasons. First, previous work has not examined the relation between verbal estimation and distinct subtypes of math abilities. Second, previous work has not addressed how distinct components of verbal estimation—mapping accuracy and variability—might each relate to math performance. Here, we addressed these gaps by measuring individual differences in ANS precision, verbal number estimation, and formal and informal math abilities in 5- to 7-year-old children. We found that verbal estimation variability, but not estimation accuracy, predicted formal math abilities, even when controlling for age, expressive vocabulary, and ANS precision, and that it mediated the link between ANS precision and overall math ability. These findings suggest that variability in the ANS–number word mapping may be especially important for formal math abilities.

* Corresponding author at: Department of Psychology and Learning Research and Development Center, University of Pittsburgh, Pittsburgh, PA 15260, USA.
E-mail address: libertus@pitt.edu (M.E. Libertus).

http://dx.doi.org/10.1016/j.jecp.2016.06.003
0022-0965/

Introduction

Educated adults and children have at least two means of representing and processing numerical information; they can use the approximate number system (ANS), which forms quick, imprecise numerical estimates without relying on serial counting (Dehaene, Dehaene-Lambertz, & Cohen, 1998), or they can use an exact number system that represents numerical information precisely via counting and number words and that is essential for much of school mathematics (Dehaene, 1992).

The ANS and exact number representations can be distinguished in several ways. The ANS is present in humans from birth (Izard, Sann, Spelke, & Streri, 2009) and also has been demonstrated in non-human animals, including fish (Dadda, Piffer, Agrillo, & Bisazza, 2009), rodents (Meck, Church, & Gibbon, 1985), whales and dolphins (Abramson, Hernandez-Lloreda, Call, & Colmenares, 2013), and non-human primates (Cantlon & Brannon, 2006); its presence in these pre-verbal and non-verbal populations shows that the ANS does not rely on language or an understanding of external symbols. Non-verbal ANS comparison tasks—in which participants typically are presented with two arrays of stimuli that differ by varying numerical ratios (e.g., two arrays of dots, two sequences of tones) and must indicate which is more numerous—reveal that the ANS represents number in a noisy continuous fashion along a mental number line, with increasing overlap between neighboring numerical representations as the target quantity grows (Dehaene, 1992; Piazza, Izard, Pinel, Le Bihan, & Dehaene, 2004; van Oeffelen & Vos, 1982). As a result, performance on ANS comparison tasks is ratio dependent; quantities that differ by larger ratios are more easily distinguished than quantities that differ by finer ratios. The precision of ANS representations (i.e., the degree of non-overlap between neighboring representations) is often measured as the Weber fraction (w), which improves over human development (Halberda & Feigenson, 2008; Halberda, Ly, Willmer, Naiman, & Germine, 2012; Libertus & Brannon, 2010; Odic, Libertus, Feigenson, & Halberda, 2013; Xu & Spelke, 2000) and also varies among individuals in a given age group (e.g., Halberda, Mazzocco, & Feigenson, 2008; Libertus, Feigenson, & Halberda, 2011; Starr, Libertus, & Brannon, 2013).

In contrast, the exact number system is uniquely human and is slowly acquired over the course of several years as children learn to use number words to represent numerical concepts (Carey, 2009; Feigenson, Dehaene, & Spelke, 2004; Le Corre & Carey, 2007; Wynn, 1992). The exact number system appears to depend on language given that people living in communities that lack a verbal counting routine lack exact large number concepts (Frank, Everett, Fedorenko, & Gibson, 2008; Gordon, 2004; Pica, Lemer, Izard, & Dehaene, 2004; Spaepen, Coppola, Spelke, Carey, & Goldin-Meadow, 2011). Critically, whereas the ANS lacks the precision to precisely represent large numbers (e.g., to represent exactly 50 as distinct from 49 and 51), exact numbers support such precision.

Despite the differences between the ANS and exact number words, children eventually learn a correspondence between the two systems. The development of this mapping is protracted (Crollen, Castronovo, & Seron, 2011; Le Corre & Carey, 2007; Odic, Le Corre, & Halberda, 2015; Pinheiro-Chagas et al., 2014) and might not be causally linked (Gunderson, Spaepen, & Levine, 2015), but once complete it allows children and adults both to label approximate number representations with exact number words and to produce approximate representations given an exact number word. For example, a 5-year-old child shown a display of 10 dots presented too quickly to count can translate the resulting imprecise ANS representation of approximately 10 dots to arrive at a verbal estimate of "ten," and a child asked to produce a particular number of actions, such as quickly tapping "ten times without counting," will produce an approximately correct number of taps (Odic et al., 2015).

Although all people who use exact number words can form a mapping between these words and the ANS, the ease with which people map between number formats and the quality of this mapping may differ from person to person (Lyons, Ansari, & Beilock, 2012). Much as we can index individual differences in ANS precision using a non-verbal ANS comparison task, we can index individual differences in the mapping between the ANS and exact number words using a verbal estimation task. [1]In a typical

[1] The other common method of measuring the mapping is the number line estimation task (Siegler & Opfer, 2003). However, because this requires an additional mapping between number and physical space, and hence could be influenced by spatial biases (Barth & Paladino, 2011), we rely here on the verbal estimation task.

version of this task, participants are shown some number of dots too quickly to count and must verbally report how many there were. This yields two measures of performance. First, performance reflects verbal estimation *accuracy* (e.g., a participant should produce a number word close to ''twelve'' for 12 dots, ''twenty-three'' for 23 dots, etc.). By plotting participants' verbal responses against the target number, mapping accuracy can be measured through the intercept, slope, and error rate of their responses (i.e., error rate = response – target) (Crollen et al., 2011; Izard & Dehaene, 2008; Odic et al., 2015). If perfectly accurate, participants' average intercept and error rate will be 0 and their slope will be 1 (i.e., for every additional item in the target array, participants should increment their estimate by one number word in their known list). Estimation accuracy depends on several factors, including children's ability to calibrate the mapping between the ANS and number words (Sullivan & Barner, 2014), potential observer biases (Barth & Paladino, 2011), and the range of number words known (Lipton & Spelke, 2005).

Verbal estimation tasks also produce a second measure of performance: estimation variability. Estimation variability reflects the degree to which a participant's verbal responses vary from trial to trial (Cordes, Gelman, Gallistel, & Whalen, 2001; Odic, Im, Eisinger, Ly, & Halberda, 2015). For example, a child who is shown 15 dots on three occasions and reports seeing ''five,'' ''fifteen,'' and ''twenty-five'' will be (on average) perfectly accurate, but his estimation variability will be high compared with a child who says ''fourteen,'' ''fifteen,'' and ''sixteen.'' Although often measured as a coefficient of variance (CV, the standard deviation of all responses for a given stimulus number divided by that number), we previously showed that estimation variability can be more robustly captured through a simple maximum-likelihood estimation method that returns an estimate that we call σ (Odic et al., 2015).[2] Children with a lower σ are less variable/more precise in their verbal estimates.

What is the relation among the ANS, exact number words, and math abilities?

Do individual differences in the mapping between the ANS and exact number words affect math performance? One reason to think so is that the ANS itself appears to be related to math performance. Prior studies that measured ANS representations without requiring any mappings to words or symbols found evidence of a link between individual differences in ANS precision and school math abilities (for a review, see Feigenson, Libertus, & Halberda, 2013). For example, Halberda and colleagues (2008) showed that adolescents' ANS precision in a number task using non-symbolic dot arrays related to their prior performance on standardized symbolic math assessments. Subsequent work revealed that this relation between the ANS and math is present prior to schooling (Bonny & Lourenco, 2013; Libertus et al., 2011) and is maintained throughout adulthood (DeWind & Brannon, 2012; Halberda et al., 2012; Libertus, Odic, & Halberda, 2012; Lourenco, Bonny, Fernandez, & Rao, 2012; Lyons & Beilock, 2011). Furthermore, ANS precision measured during infancy or during the early preschool years also predicts later math performance (Libertus, Feigenson, & Halberda, 2013a; Mazzocco, Feigenson, & Halberda, 2011b; Starr et al., 2013). Although some studies failed to find a link between ANS precision and math ability (e.g., Castronovo & Göbel, 2012; Gilmore et al., 2013; Price, Palmer, Battista, & Ansari, 2012; Sasanguie, De Smedt, Defever, & Reynvoet, 2012), recent meta-analyses support the existence of this link (Chen & Li, 2014; Fazio, Bailey, Thompson, & Siegler, 2014; Schneider et al., 2016).

If having more precise ANS representations is linked to better symbolic math performance, this raises the question of whether the *mapping* between the ANS and exact number words predicts additional variance in math abilities. Better mappings could plausibly provide children with a more accurate or reliable sense of the approximate magnitude corresponding to a number word or symbol, thereby improving mathematical intuitions. Indeed, children's exact number abilities—especially those related to their emerging understanding of exact number symbols—have also been shown to be important for math achievement. For example, Holloway and Ansari (2009), Brankaer, Ghesquiere, and De Smedt (2014), and Sasanguie and colleagues (Sasanguie, Defever, Maertens, & Reynvoet, 2014; Sasanguie et al., 2012) found that children's performance on a symbolic number comparison task (e.g., showing participants two Arabic numerals and asking them to quickly judge which is greater) correlated with

[2] As shown in the Results Section, all of our results remain identical, although less precise, if we analyze CV instead of σ.

math abilities. Similarly, Kolkman, Kroesbergen, and Leseman (2013) found that kindergarteners' performance on a symbolic number comparison task and a symbolic number line estimation task (but not on a non-symbolic number comparison task or a non-symbolic number line estimation task) predicted math performance in first grade. Other studies have found that children with math learning disabilities show significantly worse performance on symbolic number tasks than children without math learning disabilities (De Smedt & Gilmore, 2011; Iuculano, Tang, Hall, & Butterworth, 2008; Landerl & Kölle, 2009; Rousselle & Noel, 2007). These findings have sometimes been taken to indicate that the mapping of number words and symbols to their meanings is critical for math achievement—perhaps even more important than the ANS itself for the math skills tested (for a review, see De Smedt, Noel, Gilmore, & Ansari, 2013). However, some have recently questioned whether symbolic number comparison tasks actually reflect any use of approximate number representations. Lyons, Nuerk, and Ansari (2015) suggested that ratio effects in symbolic tasks (e.g., taking longer to respond that 19 is more than 17 than to respond that 19 is more than 11) do not reflect a mapping to approximate number representations but instead stem from more general issues of symbolic processing (e.g., frequency effects of particular number symbols). If so, then previously observed correlations between ratio effects in symbolic number tasks and math performance might not reveal anything about the role of the mapping between the ANS and number words in mathematical thinking.

One way to address this issue is to use tasks that more directly probe the mapping between the ANS and exact number words. The verbal estimation task described earlier requires explicit translation between non-symbolic quantities and symbolic quantities and, as such, may be better suited to asking whether individual differences in the ANS–number word mapping relate to math abilities. However, previous studies using verbal estimation tasks have yielded mixed results with regard to math performance. Some verbal estimation studies have found a significant relation between ANS–number word mapping and math ability in children. For example, Mazzocco, Feigenson, and Halberda (2011a) found that in a sample of ninth graders individual differences in ANS precision and verbal estimation independently explained variability in children's standardized math test performance. Similarly, Pinheiro-Chagas and colleagues (2014) found that 10-year-old children's ANS and verbal estimation performance each correlated with calculation skills, with verbal estimation partly mediating the relation between ANS and calculation. Mundy and Gilmore (2009) found that the accuracy of 6- to 8-year-old children's mappings between non-symbolic quantities and number words predicted a small but significant amount of variance in math performance. In addition, Bartelet, Vaessen, Blomert, and Ansari (2014) found that kindergarteners' verbal number estimation skills predicted their arithmetic performance in first grade, whereas their efficiency on a non-symbolic ANS comparison task did not. However, other studies found no evidence of this association; Lyons, Price, Vaessen, Blomert, and Ansari (2014) observed no reliable link between success at estimating the numerosity of a dot array and math ability in first through sixth graders. In addition, other studies found mixed evidence within a single sample; in a study by Castronovo and Göbel (2012), adults' verbal estimation error rates correlated with math abilities, but their estimation variability did not. Together, these findings do not offer a coherent picture of the relation among the ANS, verbal estimation, and mathematics.

At least two possibilities may help to explain these divergent findings. First, verbal estimation abilities may be more important for some math skills than for others. In preschool and early elementary school instruction, mathematics is often described as composed of informal and formal components (Baroody, 1987; Baroody & Wilkins, 1999). Informal math knowledge is acquired from children's everyday experiences with collections or events in their environment. For example, preschool-aged children's informal math abilities include counting, comparing number words to determine which is greater, and determining the answers to simple arithmetic problems using tokens or fingers. In contrast, formal math knowledge requires an understanding of symbols and mathematical conventions and, hence, is learned through explicit instruction. Children's formal math abilities include reading and writing Arabic numerals, understanding the place value system, and recalling memorized addition, subtraction, and multiplication facts.

The distinction between formal and informal aspects of mathematics plays an important role in interpreting the correlations between ANS precision and math ability. In a recent study, Libertus, Feigenson, and Halberda (2013b) found that 3- to 7-year-old children's ANS precision predicted their informal math abilities, but not their formal math abilities, up to 2 years later. In adults, Lourenco and colleagues (2012) showed that ANS precision correlated with performance on the Woodcock–Johnson (WJ) calculation subtest and KeyMath geometry subtest but not with performance on the WJ math fluency, applied math problems, and quantitative concepts subtests (but see Inglis, Attridge, Batchelor, & Gilmore, 2011, for a failure to find a link between any WJ math subtest and ANS acuity in adults). In contrast, Holloway and Ansari (2009) found that children's performance on a symbolic number comparison task, but not on a non-symbolic ANS comparison task, significantly correlated with performance on the WJ math fluency subtest and marginally correlated with the calculation subtest. Taken together, these findings suggest that the link between ANS precision and exact number representations does not play a uniform role across all aspects of mathematics. Verbal estimation abilities may likely be more closely related to those aspects of math that heavily rely on understanding the meaning of number words and symbols (i.e., formal math abilities).

A second potential explanation for the discrepancies in the literature is that only some indexes of verbal number estimation may be linked to math abilities. In particular, the accuracy of the mapping between the ANS and exact number words may play a distinct role from the variability in this mapping. If, for example, children use the mapping between number words and the ANS to perform quick mental arithmetic (accessing approximate representations to estimate the answers to math problems that use exact number symbols), a highly variable mapping would lead to many more errors over time even if the mapping is (on average) accurate. Similarly, a highly variable mapping may lead to a low-confidence internal signal, leading children to blindly guess or to use other non-optimal strategies (e.g., see Odic, Hock, & Halberda, 2014). As a result, individual differences in the variability of children's mappings between ANS representations and exact number words may affect math abilities differently than estimation accuracy.

Goals of the current study

The goals of the current study were twofold. First, we aimed to identify whether children's symbolic math abilities depend on both verbal estimation accuracy (intercept, slope, and error rate) and verbal estimation variability (σ) even when controlling for ANS precision. Second, we aimed to investigate the degree to which children's ANS precision and verbal estimation accuracy and variability are differentially linked to informal versus formal math abilities. To this end, we tested 5- to 7-year-old children on a verbal number estimation task (to measure their accuracy through intercepts, slopes, and error rates and to measure their variability through σ), a non-symbolic ANS comparison task (to measure their ANS precision or w), and a standardized assessment of informal and formal math abilities. We focused on this age range because previous work shows that 5-year-olds have established a mapping between the ANS and number words (Le Corre & Carey, 2007; Odic et al., 2015) and because 5- to 7-year-olds are at the point of transitioning from mostly informal math skills to more formal ones.

Method

Participants

We tested 51 children (23 girls; average age = 6 years 8 months, range = 5 years 1 month to 8 years 0 months). Data from 2 additional children were excluded because the children were unable to complete all three tasks. Most children came from families of middle to high socioeconomic status, all from the Baltimore, Maryland, area in the eastern United States. Parents provided informed written consent prior to their children's participation, and all children received a small gift (e.g., toy, book) to thank them for their participation.

Materials and procedure

Children were tested in the laboratory and completed three tasks, all in the same order: a verbal Number Estimation task, then a standardized assessment of math ability (Test of Early Mathematics Ability [TEMA-3]; Ginsburg & Baroody, 2003), and finally an ANS Comparison task. The tasks were administered in this order so that our primary measure of interest—performance in the Number Estimation task—would not be biased by children first completing another numerical task that used either number words or dot arrays. We administered the ANS Comparison task last so that it would not immediately follow the Number Estimation task (because both tasks involved dot arrays and, therefore, might lead to decreased engagement). In addition, parents completed an assessment of children's expressive vocabulary (the Developmental Vocabulary Assessment for Parents [DVAP]; Libertus, Odic, Feigenson, & Halberda, 2015). It took children approximately 5 to 10 min to complete the Number Estimation task, 20 to 30 min to complete the TEMA-3, and 5 to 10 min to complete the ANS Comparison task.

Number Estimation task ("How many?")

Children were told that they would see briefly flashed displays of dots on a computer screen and should say how many dots appeared in each flash (Fig. 1). Unlike in the ANS Comparison task (see below), children were asked to produce a number word on each trial. The experimenter initiated each trial when children appeared to be attentive (i.e., quietly looking at the screen) and recorded their response by pushing a button as soon as they answered. In addition, the experimenter wrote down all responses on a piece of paper, and the computer recorded children's responses in an audio file. The stimuli were arrays of 5, 8, 12, or 18 white dots on a gray background, and each array was presented for 2000 ms on a 13-inch laptop screen (i.e., long enough for comfortable viewing but too short for serial counting). We chose these stimulus numerosities to ensure that the arrays triggered ANS representations rather than non-cardinal representations of individual objects (as is often the case for arrays containing 4 or fewer items) and so that the numerosities fell within the range of known number words for all of our participants (i.e., numbers <20). To verify that children in fact knew the number words for all target values, children were asked to count up to the biggest number they knew; all children reliably counted to at least 22. We avoided choosing stimulus numerosities that were too close to each other to be reliably distinguishable; hence, the numerosities all had a ratio of at least 1.5 from the next nearest numerosity. The dots were heterogeneous in size and were drawn from a uniform distribution of 20 to 50 pixels. Finally, although we presented only four different stimulus numerosities, we expected that the inherent noisiness of the ANS would lead children to experience more variability between the stimuli (i.e., children would be unaware that they had been shown just four different numerosities).

Children received four practice trials containing three and four dots, followed by 32 test trials (8 trials for each numerosity, presented in random order). After each 2000-ms target stimulus, a blank screen remained visible until children responded. Regardless of the correctness of their response, children always received neutral positive feedback after each practice and test trial (e.g., "Great!" "Awesome!"). We used neutral positive feedback rather than accurate feedback in order to avoid discouragement and because previous studies showed that feedback can modify estimation abilities within a single experimental session (Izard & Dehaene, 2008). Because we were interested in individual differences in children's existing mappings, we avoided providing feedback that might change children's responses.

Standardized mathematics assessment

To assess children's mathematics abilities, we administered Form A of the TEMA-3 (Ginsburg & Baroody, 2003), which has been normed for children between the ages of 3 years 0 months and 8 years 11 months. The TEMA-3 is composed of 72 items divided into two broad categories. One category tests informal math abilities such as numbering skills (e.g., verbally counting the number of objects on a page), number comparison facility (e.g., determining which of two spoken number words is larger), informal calculation (e.g., solving word problems with the aid of tokens or fingers), and informal number concepts (e.g., the cardinality principle, i.e., knowing that the last number in a count sequence is the number of items in the set). The other TEMA-3 category tests formal math abilities such as numeral literacy (e.g., reading and writing Arabic numerals), mastery of number facts (e.g., retrieving addition, subtraction, and multiplication facts), calculation skills (e.g., solving mental and written addition and subtraction problems), and number concepts (e.g., answering how many tens are in one hundred). Although the authors of the TEMA have not published factor analyses of TEMA items,

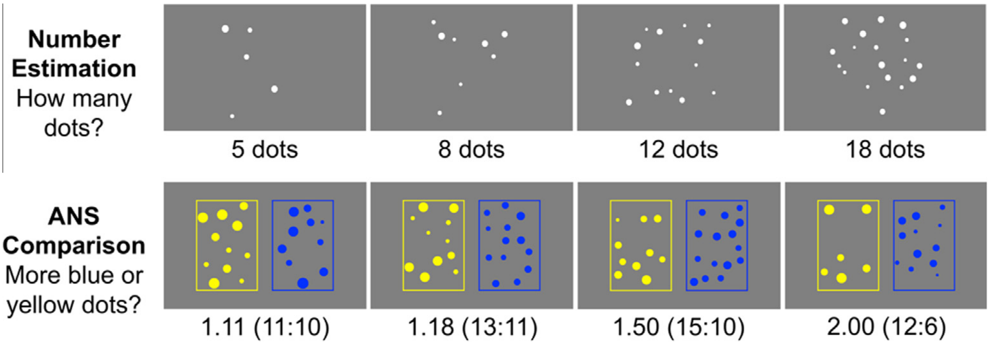

Fig. 1. Sample stimuli for the Number Estimation task (top panel) and the ANS Comparison task (bottom panel).

subsequent studies found that children's performance was better fit by a model containing multiple subtypes of mathematical abilities as opposed to a single math ability score (Ryoo et al., 2015).

We administered the TEMA-3 using the standardized procedure; that is, the first item administered was determined by children's age, after which we administered items consecutively and stopped after children incorrectly answered five items in a row. Items before the age-defined starting point were administered only if children did not initially succeed on five consecutive items. No feedback was given.

ANS Comparison task ("Which is more?")

To measure the precision of children's ANS representations, we administered a version of Panamath (the Psychophysical Assessment of Number-sense Acuity; http://panamath.org; Halberda & Ly, 2016)—an approximate number discrimination task. Children were shown arrays of spatially separated blue and yellow dots on a 13-inch laptop screen and were asked to indicate whether more of the dots were blue or yellow (Fig. 1). The experimenter initiated each trial when children appeared to be attentive (i.e., quietly looking at the screen). Each stimulus array was visible for 2000 ms and was followed by a blank screen that remained until children gave a verbal response (e.g., "yellow"). The experimenter immediately pressed the corresponding key on an external keyboard (e.g., "y" for "yellow"). The computer provided positive feedback (e.g., "That's right!" "Fantastic!") or negative feedback (e.g., "Oh, that's not right"). We provided accurate feedback in this task in accordance with the methods used in previous studies of ANS discrimination (e.g., Halberda & Feigenson, 2008; Halberda et al., 2008; Libertus et al., 2011; Libertus et al., 2013a). Children were familiarized to the task and given feedback on four practice trials, during which the experimenter provided additional verbal feedback to ensure that children understood the task and were motivated to participate.

Following these practice trials, children completed 56 test trials. On each of these trials, the presented numerosities were drawn randomly from one of seven numerical ratio bins: 1.08, 1.11, 1.15, 1.18, 1.25, 1.5, or 2.0 (with the absolute number of dots in each collection varying between 5 and 22, such that a trial with, e.g., 5 yellow dots vs. 10 blue dots would contribute to the 2.0 ratio bin). On half of the trials the yellow dots were more numerous, and on the other half the blue dots were more numerous. Orthogonally, the dots in each array varied in size in order to discourage children from using dot size as a cue. The default radius of the dots was 60 pixels, and the maximum variability in size between the dots was ±35%. Previous research has found that non-numerical aspects of an array can affect numerical performance (Clayton, Gilmore, & Inglis, 2015; Gebuis & Reynvoet, 2012; Smets, Sasanguie, Szücs, & Reynvoet, 2015). As such, on half of the trials the two arrays were equated for individual dot size (i.e., the average size of the dots in each collection was equal), and on the other half the cumulative surface area of the blue dots and the yellow dots was equated.

Expressive vocabulary

We measured children's expressive vocabulary using the DVAP (Libertus et al., 2015). The DVAP questionnaire contained a list of 212 nouns, verbs, and adjectives taken from Form A of the Peabody Picture Vocabulary Test (PPVT-4; Dunn & Dunn, 2007), and parents were asked to indicate which of the listed words they had heard their children say. Previous work has established the DVAP as a stable and accurate indicator of children's expressive verbal knowledge that correlates with children's PPVT-4 performance (Libertus et al., 2015). Parents completed the DVAP either before their children's testing session began or, if they did not accompany their children into the testing room, during the testing session.

Results

We first report the data from each task, comparing performance in our sample with performance previously reported in the literature. Subsequently, we examine the relations among the three tasks. Where the underlying distribution of scores was found to be left or right skewed, we used Spearman rank-order coefficients instead of Pearson coefficients (Inglis & Gilmore, 2014). We report the correlations that were identified a priori as our primary interest in the main text, but a full overview of all zero-order correlations can be found in the Appendix. Finally, although all participants were tested in the Number Estimation task first, the standardized mathematics assessment second, and the ANS Comparison task third, for ease of explication we report results from the Number Estimation and ANS Comparison tasks first.

Number Estimation task

Children's performance on the verbal Number Estimation task is shown in Fig. 2 (left panel). Before analysis, we removed any responses that were more than 2.5 standard deviations from children's average response across all trials (<1% of trials). We found that children showed a linearly increasing relation between the target numerosity and the number word produced, $F(3, 150) = 127.21$, $p < .001$; linear contrast $F(1, 50) = 155.50$, $p < .001$. Children's responses also exhibited linearly increasing variability, $F(3, 150) = 17.62$, $p < .001$; linear contrast $F(1, 50) = 21.78$, $p < .001$. Both of these patterns are considered signatures of the ANS, confirming that children did not count in our task but instead used the mapping between the imprecise ANS and exact number words (Barth, Starr, & Sullivan, 2009; Crollen et al., 2011; Huntley-Fenner, 2001; Izard & Dehaene, 2008).

The mathematical model of verbal number estimation predicts that children's responses will track with slope/intercept and that their variability will be normally distributed with standard deviation σ. To estimate the slope, intercept, and σ of each child's verbal number estimates, we used R-PsiMLE software (Odic et al., 2015). This method maximizes the parameters that fit the normal distribution CDF (cumulative distribution function), with slope, intercept, and target number determining the mean of the distribution and σ, proportional to the mean, determining the standard deviation:

$$L(\alpha, \beta, \sigma | Resp, Num) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi(\alpha + \beta * Num_i * \sigma)^2}} \exp\left(-\frac{(Resp_i - \alpha + \beta * Num_i)^2}{2(\alpha + \beta * Num_i * \sigma)^2}\right).$$

Here, *Num* corresponds to the number of dots presented (i.e., either 5, 8, 12, or 18), α corresponds to the intercept of the linear regression, β corresponds to the slope, σ corresponds to the linearly increasing scalar variability (equivalent to CV), and *Resp* corresponds to children's number word response. R-PsiMLE software (freely available online at http://panamath.org/psimle) uses maximum-likelihood estimation to simultaneously estimate all three parameters of interest. Separately, we also calculated the average error rate for each child by subtracting the target numbers from the actual responses and averaging them. Positive error rates indicate over-estimation, and negative error rates indicate under-estimation.[3] Our measure of estimation variability, σ, correlated

[3] Two other measures—percentage absolute error (PAE) and r^2 of best fit line (linearity)—are also sometimes used to examine verbal estimation performance (Booth & Siegler, 2008; Siegler & Opfer, 2003). However, these two measures combine accuracy and variability into a single measure. For example, if a child were (on average) perfectly accurate but highly variable, the PAE would be substantially higher than 0 because the positive and negative errors would not cancel out in the average. Similarly, r^2 values are affected by both the fit of the line (accuracy) and the variability surrounding that line (variability). Although these measures are well suited to some questions concerning number estimation, they are not well suited to the goal of examining accuracy and variability separately.

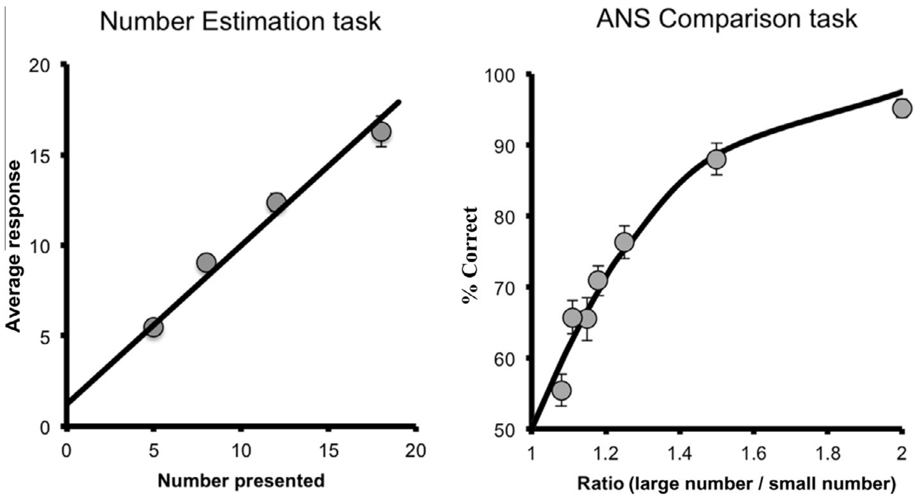

Fig. 2. Performance in the verbal Number Estimation task (left) and the non-symbolic ANS Comparison task (right). Error bars in the Number Estimation graph depict standard errors of the mean and show an expected linear increase in the number word produced as the number of target dots increased. The best fitting linear function had a slope of 0.9 and an intercept of 1.22. Error bars in the ANS Comparison task graph depict standard errors of the mean, and the line reflects the best fitting psychophysical model with $w = 0.25$.

strongly with the more traditionally used CV measure (relation of σ to CV; $r = .96$, $p < .001$). Because the PsiMLE method yields better reliability in estimates (Odic et al., 2015), we used σ throughout our analyses. However, the pattern of results was nearly identical when we analyzed CV instead.

Data from all children were included in our analyses. We found that children in our sample were (on average) accurate in their verbal estimates, with roughly equal numbers of children producing over-estimates versus under-estimates. Across children, the average β (i.e., slope) was 0.90 ($SD = 0.53$), the average α (i.e., intercept) was 1.22 ($SD = 3.61$), and the average error rate was 0.05 ($SD = 2.77$). The average estimation variability (i.e., σ) was 0.26 ($SD = 0.12$), slightly higher than the typical adult average of 0.22 (Cordes et al., 2001).

ANS Comparison task

Performance on the ANS Comparison task is shown in Fig. 2 (right panel). Preliminary analyses revealed no significant differences in accuracy for trials in which individual dot size was equated ($M = 74\%$, $SEM = 2\%$) and in which cumulative surface area was equated ($M = 73\%$, $SEM = 1\%$), $t(50) = 0.94$, $p = .35$, and no correlation between age and the difference in performance on these two trial types, $r(51) = .005$. Thus, we collapsed across these trial types in our further analyses. The Spearman–Brown corrected random split-half reliability for the ANS Comparison task was 0.69, in line with previous estimates (Libertus et al., 2012).

As shown in Fig. 2 and as predicted by Weber's law, children's accuracy at judging the greater of the two arrays decreased as the ratio between the numerosities approached equality, $F(6, 342) = 50.48$, $p < .001$. To determine each individual child's Weber fraction (w), an estimate of the ANS precision, we fit each child's responses over all 56 trials with a widely used psychophysical model (1) (cf. Green & Swets, 1966; Halberda & Feigenson, 2008; Halberda et al., 2008; Pica et al., 2004):

$$\text{expected accuracy} = \Phi\left(\frac{ratio - 1}{w\sqrt{ratio^2 + 1}}\right). \quad (1)$$

In this model, *ratio* is the ratio between the presented numerosities (larger number/smaller number), w is the Weber fraction, and Φ is the standard cumulative distribution function of a Gaussian distribution. The best fitting w parameter was found via non-linear least squares. The model assumes that the underlying representations are distributed along a continuum of Gaussian random variables. An important implication of this model is that the two numerosities presented on each trial will often yield similar and overlapping representations. In other words, as the ratio of two quantities becomes increasingly close (i.e., approaches 1.0), the Gaussian representations of the numerosities will increasingly overlap and children will have greater difficulty in determining which array is more numerous, resulting in decreased accuracy. A greater w indicates greater imprecision in children's ANS representations. Data from 1 child could not be fitted because the child performed at chance across all ratios.

Collapsing across all of the ages we tested, we found that children had an average w of 0.25 (SD = 0.15), a value consistent with previous estimates for children of this age in tasks using similar methods (Halberda & Feigenson, 2008; Piazza et al., 2010).

Standardized mathematics assessment

To measure children's overall math abilities, we calculated their raw math scores following the TEMA-3 procedures. We distinguished children's informal versus formal mathematics abilities using the item categorization given in the TEMA-3 examiner's manual (Ginsburg & Baroody, 2003). As described above, the standardized administration of the TEMA-3 requires starting and ending with different items in the test sequence depending on children's age and performance. To account for the resulting difference in the total number of items children completed, and for the number of informal versus formal math items administered to each child, we averaged children's scores (such that children were scored as having either a 0 or 1 for each administered item) and multiplied these averages by 100 to create separate percentage scores for informal and formal mathematics skills. Only data from children who were tested with at least three items in a given category were included in order to avoid including informal or formal math scores that were based on very few items. In addition, because children started at different points according to their age, the resulting informal and formal math scores are, in practice, roughly controlled for age at time of testing.

Children's average raw TEMA-3 score was 43.06 (SD = 14.34), which corresponds to an average standardized score of 107.14 (SD = 14.11), roughly half a standard deviation above the expected population mean. Children solved an average of 66.09% informal math problems (SD = 17.24%) and 44.72% formal math problems (SD = 16.53%). The higher rate of correctly solved informal problems is expected given that these problems tend to be easier than formal problems.

Expressive vocabulary

To measure children's expressive vocabulary, we summed the number of words that parents had marked on the DVAP questionnaire as having heard their children say (as described by Libertus et al., 2015). Children's average score on the expressive vocabulary questionnaire was 130.51 (SD = 26.65). This is somewhat higher than the average score reported in previous investigations using the DVAP (Libertus et al., 2011; Libertus et al., 2015); this is not surprising given that the children in this sample were (on average) older than those in the previous studies.

Correlations between verbal number estimation and non-symbolic ANS precision

To ask whether the precision of children's ANS was related to the accuracy of their verbal number estimation independent of children's age, we examined the relation between w and the slope, intercept, and error rates of children's verbal number estimates when controlling for age at the time of testing. These analyses revealed no significant correlation between w and slope ($r_s = -.26$, $p = .10$) or error rate ($r_s = -.10$, $p = .52$) but revealed a significant correlation between w and intercept ($r_s = .38$, $p < .01$). We next asked whether ANS precision was related to the variability of children's verbal number estimates. Here, we found a significant correlation between w and σ ($r_s = .56$, $p < .001$)

(Fig. 3); more imprecise approximate number representations (i.e., higher *w* scores) were associated with more variability in the number words children produced. This finding suggests that variability in number estimation is partly determined by the ANS.

Relation between non-symbolic ANS precision and mathematical ability

In replication of previous work (Bonny & Lourenco, 2013; Libertus et al., 2011; Libertus et al., 2013a), we observed a significant correlation between children's *w* and their TEMA-3 scores when controlling for age ($r_s = -.51$, $p < .01$) (Fig. 4A); children with more imprecise approximate number representations performed more poorly on the standardized math test. We further replicated a divergence in the relation between *w* and formal versus informal math abilities. Consistent with the findings of Libertus and colleagues (2013b), we found that *w* significantly related to informal math abilities even when controlling for age ($r_s = -.36$, $p < .05$) (Fig. 5A), but *w* showed no significant relation to formal math abilities ($r_s = -.06$, $p = .71$) (Fig. 6A). This asymmetry motivated us to probe for a similar divergence in the link between informal versus formal math abilities and their relation to verbal number estimation.

Relation between verbal number estimation and mathematical ability

To ask whether children's performance in the verbal Number Estimation task was related to their math abilities, we first examined whether number estimation accuracy (slope, intercept, and error rates) correlated with individual differences in TEMA-3 performance when controlling for children's age. We found no significant correlation between overall TEMA scores and any of the three measures of estimation accuracy (intercept: $r = -.04$, $p = .79$; slope: $r = .004$, $p = .98$; error rate: $r = -.04$, $p = .81$). This lack of a relation held for both informal math abilities (intercept: $r = -.16$, $p = .33$; slope: $r = .10$, $p = .51$; error rate: $r = .02$, $p = .93$) and formal math abilities (intercept: $r = .17$, $p = .29$; slope: $r = -.15$, $p = .34$; error rate: $r = -.07$, $p = .67$). Together, these results suggest that the accuracy of children's verbal number estimation—at least in the range of number words that they reliably know—does not relate to their informal or formal math abilities.

We next examined the relation between math performance and our second measure of verbal number estimation: mapping variability. Here, we found a robust correlation between estimation variability (measured by σ) and overall TEMA-3 scores even when controlling for children's age ($r = -.44$, $p < .01$) (Fig. 4B). And even when controlling for age, estimation variability was marginally correlated with informal math ability ($r = -.30$, $p = .054$) (Fig. 5B) and significantly correlated with formal math ability ($r = -.34$, $p < .05$) (Fig. 6B).

To assess whether estimation variability was a significant predictor of children's math performance above and beyond differences in age, expressive vocabulary, and ANS precision, we conducted three hierarchical linear regression analyses: one for overall TEMA-3 scores, one for informal TEMA-3 scores, and one for formal TEMA-3 scores. In the first step, we entered age and expressive vocabulary (DVAP) as variables. In the next step, we added ANS precision (*w*). In the final step, we entered estimation variability (σ) to test whether σ explained a significant amount of variability in math performance above and beyond the predictors entered in the first two steps. As can be seen in Table 1, σ was a significant predictor of children's overall math performance even when controlling for age, expressive vocabulary, and *w*. Similarly, σ was a significant predictor of children's formal math performance even when controlling for age, expressive vocabulary, and *w* (see Table 2). In contrast, σ was not a significant predictor of children's informal math performance, and neither was *w* (see Table 3).

Finally, to test whether verbal number estimation, specifically estimation variability (σ), mediated the relation between ANS precision (*w*) and overall math performance, we conducted a mediation analysis following the approach outlined by Preacher and Hayes (2008). The overall regression model, including both ANS precision and σ as predictors and both age and expressive vocabulary scores as covariates, was significant ($R^2 = .67$, $p < .001$). The direct path between ANS precision and overall math ability was no longer significant when the indirect path via verbal number estimation was added ($p = .09$) (see Fig. 7), and the 95% confidence interval ranged from -35.75 to -0.34, suggesting a significant mediation effect ($p < .05$).

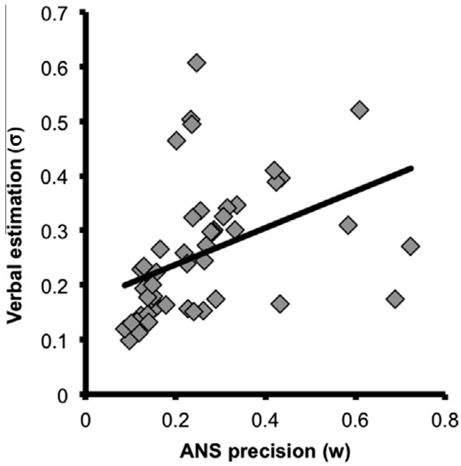

Fig. 3. Scatterplot depicting the relation between ANS precision (w) and verbal number estimation variability (σ). Children with greater ANS precision (i.e., lower w) tended to show less variable verbal estimation performance.

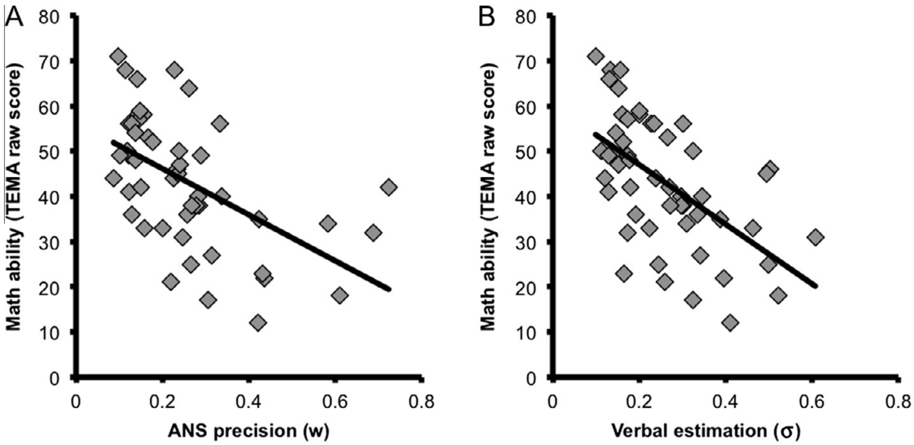

Fig. 4. Scatterplots depicting the relation between overall math ability (TEMA-3 raw scores) and ANS precision (A) and verbal number estimation (B). Children with greater ANS precision (i.e., lower w) or lower verbal estimation variability (i.e., lower σ) tended to have greater overall math ability.

Discussion

By their fifth birthday, children have two ways of representing number: intuitive non-verbal representations produced by the ANS and exact number words and symbols whose meanings are acquired gradually, usually over the preschool years. Although each of these systems has been shown to independently relate to early mathematics ability, the role of the mapping between the two systems (e.g., as measured by the accuracy and variability of children's verbal number estimations) has remained controversial.

Here, we report three main findings. First, we demonstrate that only some aspects of verbal number estimation correlate with early mathematics ability. We found that whereas children's estimation accuracy (indexed by slopes, intercepts, and error rates) did not correlate with their math abilities, the

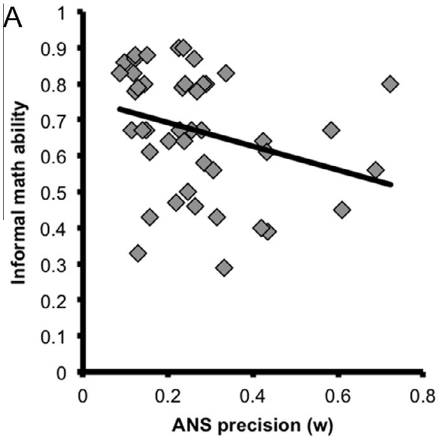

Fig. 5. Scatterplots depicting the relation between informal math ability and ANS precision (A) and verbal number estimation (B). Children with greater ANS precision (i.e., lower w) or lower verbal estimation variability (i.e., lower σ) tended to have greater informal math ability.

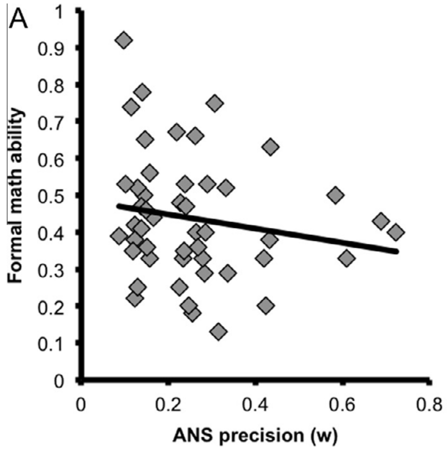

Fig. 6. Scatterplots depicting the relation between formal math ability and ANS precision (A) and verbal number estimation (B). Children with lower verbal estimation variability (i.e., lower σ) tended to have greater formal math ability, but no relation was observed between ANS precision and formal math ability.

variability in their estimates (indexed by σ) did. Second, we found that this variability in children's mappings between the ANS and number words (σ) uniquely correlated with formal math abilities. σ was a significant predictor of formal math abilities (but not informal math abilities) even when controlling for ANS precision, age, and expressive vocabulary. Third, σ mediated the relation between children's ANS precision and their overall math ability. As we elaborate below, these findings suggest that the link between the ANS and number words may uniquely contribute to children's math skills.

One of the primary motivations for our investigation was the mixed and sometimes contradictory findings on the relation between verbal number estimation and math performance (Castronovo & Göbel, 2012; Lyons et al., 2014; Mazzocco et al., 2011a; Mejias, Mussolin, Rousselle, Gregoire, & Noel, 2012; Mejias & Schiltz, 2013; Mundy & Gilmore, 2009; Pinheiro-Chagas et al., 2014). Our results suggest at least two reasons for these discrepancies. First, the relation between number estimation and math appears to be restricted to *estimation variability*; estimation accuracy showed no relation

Table 1
Summary of hierarchical linear regression analyses for variables predicting overall math ability.

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*
Age	.04	.005	.75***	.03	.005	.64***	.03	.005	.53***
Expressive vocabulary (DVAP)	.02	.06	.03	.01	.05	.03	.05	.05	.10
ANS acuity (w)				−24.16	9.54	−.25*	−16.54	9.35	−.17
Estimation variability (σ)							−32.06	11.70	−.28**
R^2	.58			.63			.68		
F for R^2 change	32.00***			6.42*			7.52**		

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

Table 2
Summary of hierarchical linear regression analyses for variables predicting formal math ability.

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*
Age	<.001	<.001	.20	<.001	<.001	.16	<.001	<.001	.02
Expressive vocabulary (DVAP)	.001	.001	.11	.001	.001	.11	.001	.001	.20
ANS acuity (w)				−.10	.18	−.09	.01	.18	.01
Estimation variability (σ)							−.47	.23	−.34*
R^2	.07			.07			.15		
F for R^2 change	1.56			0.30			4.12*		

* $p < .05$.

Table 3
Summary of hierarchical linear regression analyses for variables predicting informal math ability.

	Model 1			Model 2			Model 3		
Variable	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*	*B*	*SE B*	*ß*
Age	<.001	<.001	.25	<.001	<.001	.14	<.001	<.001	.03
Expressive vocabulary (DVAP)	−.001	.001	−.14	−.001	.001	−.14	<.001	.001	−.07
ANS acuity (w)				−.30	.19	−.26	−.21	.20	−.18
Estimation variability (σ)							−.36	.24	−.26
R^2	.06			.11			.16		
F for R^2 change	1.24			2.46			2.22		

with math in our analyses. Second, when controlling for ANS precision, estimation variability related to *formal math skills* but not to informal math skills. This suggests that future studies might clearly delineate these components of math ability, especially across development.

Our results also confirm that non-verbal ANS comparison and verbal number estimation tasks share some common cognitive resources—including, critically, the degree of precision of the underlying ANS representations themselves (Ansari, Donlan, & Karmiloff-Smith, 2007). This is evidenced by the correlation between ANS precision (w) and σ. However, our results also show that estimation variability (σ) measures more than just ANS precision. In our regression analyses, σ was a significant predictor of formal math ability even when controlling for ANS precision, age, and vocabulary. We conclude that the variability in σ unaccounted for by ANS precision is not just measurement noise. We suggest instead that it may reflect shifting calibration in children's mappings (e.g., children may be unable to maintain a stable association between the ANS and number words) (Sullivan & Barner,

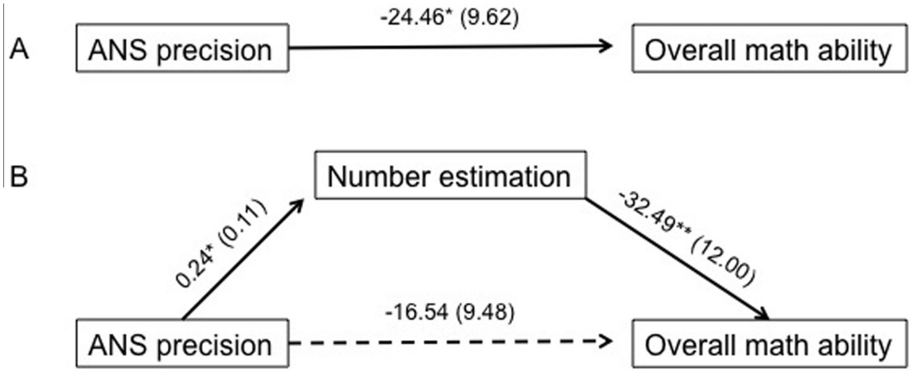

Fig. 7. Mediation model for the relation between ANS precision and overall math ability by number estimation when controlling for age and expressive vocabulary scores. Unstandardized path coefficients for the direct effects of ANS precision on overall math ability (A) and for the indirect effects via verbal number estimation (B) are shown. Standard errors are given in parentheses. *$p < .05$; **$p < .01$.

2014) or may reflect the effects of children's precision in being able to categorize the continuous ANS representation into a single number word (Odic et al., 2015). As an example, children may vary in whether their mental category for the word "ten" neatly aligns with only the section of the ANS mental number line coding for 10 items or whether it bleeds over onto nearby ANS representations. As a result, children may be (on average) correct in their verbal estimates but very imprecise in what they consider "ten" to be. Future studies could investigate whether verbal number estimation is unique in this regard or whether other number estimation tasks, including symbol production (in which children are required to write a number in response to a non-symbolic stimulus), matching tasks (in which children match numerosities to Arabic numerals or vice versa) (Brankaer et al., 2014), and non-symbolic numerosity production tasks (in which children produce an approximate number of objects or actions to match a given number symbol) yield similar results.

Our findings also suggest that, in line with previous investigations, symbolic number skills may mediate the relation between the ANS and math abilities. For example, previous studies found that verbal number abilities mediated the relation between ANS precision and math achievement in preschoolers and elementary school-aged children (Pinheiro-Chagas et al., 2014; van Marle, Chu, Li, & Geary, 2014). Similarly, performance on a symbolic number comparison task mediated the relation between ANS precision and various aspects of math achievement in third graders (Price & Fuchs, 2016). Here, we found evidence of a similar relation; σ mediated the relation between ANS precision and overall math ability. This mediated relation makes sense given that σ reflects the precision of the ANS plus additional variability in the mapping between the ANS and number words.

Finally, we note that in previous work (Libertus et al., 2013b), ANS precision (w) correlated with informal math abilities but not with formal math abilities; this association remained significant even when controlling for children's age and non-verbal IQ. In the current work, we found that ANS precision was no longer a significant predictor of children's informal math abilities when controlling for age and expressive vocabulary. It is possible that the lack of an observed relation between ANS precision and informal math performance in the current study was due to the smaller sample size relative to previous work (Libertus et al., 2013b). Indeed, the qualitative pattern we observed here mirrored that in our previous investigation; ANS precision was a better predictor of children's informal math abilities than of their formal math abilities (see Tables 2 and 3). Future studies should continue to explore the predictive value of ANS precision for different subtypes of math abilities.

Possible explanation for the link between number estimation variability and formal math abilities

It is, of course, important to note that the observed relation between the variability in children's number estimation and their formal math abilities is correlational and that, therefore, any suggestions about the mechanisms linking them will be speculative. However, one possible explanation may be children's *confidence in their number estimation abilities.* Children with low variability in their mappings between the ANS and exact number words can rapidly form fairly precise estimates of numerical magnitude (i.e., can reliably narrow in on a small region of the ANS mental number line) in response to a number word or symbol. This mapping precision may cause children to experience a sense of confidence in their "gut sense" of number when interacting with number words and symbols, leading them to make fewer guesses when engaging in formal mathematics (Hacker, Dunlosky, & Graesser, 1998). Recently, work studying children's non-verbal numerical approximation abilities has demonstrated the importance of confidence for ANS comparison in young children (Halberda & Odic, 2014; Odic et al., 2014). Accessing the approximate magnitudes associated with exact number symbols may be less important for children when they engage in informal math processing (as when using tokens or fingers to solve word problems or engage in rote counting); hence, individual differences in mapping variability may specifically affect formal math processing.

Another possible explanation for the relation between verbal estimation variability and math performance is the potential shared need to inhibit irrelevant information. In the verbal Number Estimation task, representing the approximate numerosity of a stimulus array requires inhibiting irrelevant perceptual information, including cumulative surface area, density, and convex hull of the array (Szücs, Nobes, Devine, Gabriel, & Gebuis, 2013). Inhibitory control also appears to play a role in children's math performance (Blair & Razza, 2007; Clark et al., 2013; St Clair-Thompson & Gathercole, 2006). Thus, it is plausible that variability in inhibitory control may contribute to the link among performance on ANS comparison tasks, number estimation tasks, and math ability. Although this has not been directly examined for number estimation tasks, previous studies examining the role of inhibitory control in explaining the link between ANS precision and math ability have yielded mixed results (Fuhs & McNeil, 2013; Gilmore et al., 2013; Keller & Libertus, 2015; Libertus et al., 2013); hence, further work is needed.

We also note that the understanding of exact number words or symbols may influence math ability via routes other than just the mapping between ANS representations and number words, as we assessed here. Previous work by Lyons and Beilock (2011) showed that in adults symbolic ordering skills (e.g., determining whether three Arabic numerals are presented in numerically increasing order) are a mediator between ANS precision and math ability. Thus, future research could determine whether symbolic ordering and verbal number estimation skills independently mediate the relation between ANS precision and math ability. Alternatively, symbolic ordering and verbal number estimation skills may be independently associated with distinct aspects of mathematics and/or different mediators may be at play at different stages of development.

Conclusion

We found that the precision of 5- to 7-year-old children's mapping between the approximate number system and exact number words correlated with their math abilities. More specifically, the variability in the ANS–number word mapping predicted formal math abilities, even when controlling for ANS precision, age, and verbal abilities, and mediated the relation between ANS precision and overall math abilities. These findings demonstrate that verbal number estimates—and especially the variability in producing these estimates—reflect additional cognitive processes beyond ANS precision and that these additional processes play an important role in math performance. Future studies are necessary to replicate and extend these findings to more clearly delineate the relation between ANS precision, number estimation, and subcategories of informal and formal mathematical abilities across developmental time. This work will further our understanding of the mechanisms that link basic math cognition with the variety of skills required by abstract mathematical thought.

Acknowledgments

We thank Rebecca Zhu for help with testing participants, and all families and children for their participation. This research was supported by National Institute of Child Health and Human Development

(NICHD) Grant R01 HD057258 to L.F. and J.H. Creation of the Panamath software was supported by National Science Foundation (NSF) Grant DRL0937675 to J.H. The authors declare no conflicts of interest.

Appendix

Pearson correlation results for all measures of interest when controlling for age at the time of testing.

	Estimation accuracy (intercept)	Estimation accuracy (slope)	Estimation accuracy (error rate)	Estimation variability (σ)	ANS precision (w)	Expressive vocabulary (DVAP)
1. Overall math	−.04	.004	−.04	−.44**	−.51**	.02
2. Informal math	−.16	.10	.02	−.30†	−.36*	−.12
3. Formal math	.17	−.15	−.07	−.34*	−.06	.09
4. Estimation accuracy (intercept)	–	−.92***	−.57***	−.04	.38*	.10
5. Estimation accuracy (slope)		–	.84***	.23	−.26	.005
6. Estimation accuracy (error rate)			–	.38*	−.10	.14
7. Estimation variability (σ)				–	.56***	.31*
8. ANS precision (w)					–	.04
9. Expressive vocabulary (DVAP)						–

Note. All correlations involving ANS precision (w) are Spearman rank-ordered correlations because the distribution of the Weber fraction scores was slightly skewed. Correction for multiple correlations using the correction by Benjamini and Hochberg (1995) shifts the significance level to $p < .009$.

† $p < .06$.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

References

Abramson, J. Z., Hernandez-Lloreda, V., Call, J., & Colmenares, F. (2013). Relative quantity judgments in the beluga whale (*Delphinapterus leucas*) and the bottlenose dolphin (*Tursiops truncatus*). *Behavioural Processes, 96*, 11–19.

Ansari, D., Donlan, C., & Karmiloff-Smith, A. (2007). Typical and atypical development of visual estimation abilities. *Cortex, 43*, 758–768.

Baroody, A. J. (1987). *Children's mathematical thinking: A developmental framework for preschool, primary, and special education teachers*. New York: Teachers College Press.

Baroody, A. J., & Wilkins, J. L. M. (1999). The development of informal counting, number, and arithmetic skills and concepts. In J. V. Copley (Ed.), *Mathematics in the early years* (pp. 48–65). Washington, DC: National Association for the Education of Young Children.

Bartelet, D., Vaessen, A., Blomert, L., & Ansari, D. (2014). What basic number processing measures in kindergarten explain unique variability in first-grade arithmetic proficiency? *Journal of Experimental Child Psychology, 117*, 12–28.

Barth, H., & Paladino, A. M. (2011). The development of numerical estimation: Evidence against a representational shift. *Developmental Science, 14*, 125–135.
Barth, H., Starr, A. B., & Sullivan, J. (2009). Children's mappings of large number words to numerosities. *Cognitive Development, 24*, 248–264.
Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society B: Methodological, 57*, 289–300.
Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development, 78*(2), 647–663.
Bonny, J. W., & Lourenco, S. F. (2013). The approximate number system and its relation to early math achievement: Evidence from the preschool years. *Journal of Experimental Child Psychology, 114*, 375–388.
Booth, J. L., & Siegler, R. S. (2008). Numerical magnitude representations influence arithmetic learning. *Child Development, 79*, 1016–1031.
Brankaer, C., Ghesquiere, P., & De Smedt, B. (2014). Children's mapping between non-symbolic and symbolic numerical magnitudes and its association with timed and untimed tests of mathematics achievement. *PLoS ONE, 9*(4), e93565.
Cantlon, J. F., & Brannon, E. M. (2006). Shared system for ordering small and large numbers in monkeys and humans. *Psychological Science, 17*, 401–406.
Carey, S. (2009). *The origin of concepts*. Oxford, UK: Oxford University Press.
Castronovo, J., & Göbel, S. M. (2012). Impact of high mathematics education on the number sense. *PLoS ONE, 7*(4), e33832.
Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica, 148*, 163–172.
Clark, C. A., Sheffield, T. D., Wiebe, S. A., & Espy, K. A. (2013). Longitudinal associations between executive control and developing mathematical competence in preschool boys and girls. *Child Development, 84*(2), 662–677.
Clayton, S., Gilmore, C., & Inglis, M. (2015). Dot comparison stimuli are not all alike: The effect of different visual controls on ANS measurement. *Acta Psychologica, 161*, 177–184.
Cordes, S., Gelman, R., Gallistel, C. R., & Whalen, J. (2001). Variability signatures distinguish verbal from nonverbal counting for both large and small numbers. *Psychonomic Bulletin & Review, 8*, 698–707.
Crollen, V., Castronovo, J., & Seron, X. (2011). Under- and over-estimation: A bi-directional mapping process between symbolic and non-symbolic representations of number? *Experimental Psychology, 58*, 39–49.
Dadda, M., Piffer, L., Agrillo, C., & Bisazza, A. (2009). Spontaneous number representation in mosquitofish. *Cognition, 112*, 343–348.
De Smedt, B., & Gilmore, C. (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., Noel, M. P., Gilmore, C., & Ansari, D. (2013). How do symbolic and non-symbolic numerical magnitude processing relate to individual differences in children's mathematical skills? A review of evidence from brain and behavior. *Trends in Neuroscience and Education, 2*(2), 48–55.
Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1–42.
Dehaene, S., Dehaene-Lambertz, G., & Cohen, L. (1998). Abstract representations of numbers in the animal and human brain. *Trends in Neuroscience, 21*, 355–361.
DeWind, N. K., & Brannon, E. M. (2012). Malleability of the approximate number system: Effects of feedback and training. *Frontiers in Human Neuroscience, 6*. http://dx.doi.org/10.3389/fnhum.2012.00068.
Dunn, L. M., & Dunn, D. M. (2007). *Peabody Picture Vocabulary Test-Fourth Edition*. San Antonio, TX: Pearson.
Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123*, 53–72.
Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*, 307–314.
Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links between the intuitive sense of number and formal mathematics ability. *Child Development Perspectives, 7*(2), 74–79.
Frank, M. C., Everett, D. L., Fedorenko, E., & Gibson, E. (2008). Number as a cognitive technology: Evidence from Piraha language and cognition. *Cognition, 108*, 819–824.
Fuhs, M. W., & McNeil, N. M. (2013). ANS acuity and mathematics ability in preschoolers from low-income homes: Contributions of inhibitory control. *Developmental Science, 16*, 136–148.
Gebuis, T., & Reynvoet, B. (2012). The interplay between nonsymbolic number and its continuous visual properties. *Journal of Experimental Psychology: General, 141*, 642–648.
Gilmore, C., Attridge, N., Clayton, S., Cragg, L., Johnson, S., Marlow, N., ... Inglis, M. (2013). Individual difference in inhibitory control, not non-verbal number acuity, correlate with mathematics achievement. *PLoS ONE, 8*(6), e67374.
Ginsburg, H. P., & Baroody, A. J. (2003). *Test of Early Mathematics Ability-Third Edition*. Austin, TX: Pro-Ed.
Gordon, P. (2004). Numerical cognition without words: Evidence from Amazonia. *Science, 306*, 496–499.
Green, D., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: John Wiley.
Gunderson, E. A., Spaepen, E., & Levine, S. C. (2015). Approximate number word knowledge before the cardinal principle. *Journal of Experimental Child Psychology, 130*, 35–55.
Hacker, D. J., Dunlosky, J., & Graesser, A. C. (1998). *Metacognition in educational theory and practice*. Mahwah, NJ: Lawrence Erlbaum.
Halberda, J., & Ly, R. (2016). Panamath: The Psychophysical Assessment of Number–sense Acuity. Manuscript in preparation.
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., Ly, R., Willmer, J., Naiman, D., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 11116–11120.
Halberda, J., Mazzocco, M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455*, 665–668.

Halberda, J., & Odic, D. (2014). The precision and internal confidence of our approximate number thoughts. In D. C. Geary, D. B. Berch, & K. Mann Koepke (Eds.). *Evolutionary origins and early development of number processing* (pp. 305–333). Waltham, MA: Academic Press.
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.
Huntley-Fenner, G. (2001). Children's understanding of number is similar to adults' and rats': Numerical estimation by 5–7-year-olds. *Cognition, 78*, B27–B40.
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.
Inglis, M., & Gilmore, C. (2014). Indexing the approximate number system. *Acta Psychology, 145*, 147–155.
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.
Izard, V., & Dehaene, S. (2008). Calibrating the mental number line. *Cognition, 106*, 1221–1247.
Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences of the United States of America, 106*, 10382–10385.
Keller, L., & Libertus, M. (2015). Inhibitory control may not explain the link between approximation and math abilities in kindergarteners from middle class families. *Frontiers in Psychology, 6*. http://dx.doi.org/10.3389/fpsyg.2015.00685.
Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction, 25*, 95–103.
Landerl, K., & Kölle, C. (2009). Typical and atypical development of basic numerical skills in elementary school. *Journal of Experimental Child Psychology, 103*, 546–565.
Le Corre, M., & Carey, S. (2007). One, two, three, four, nothing more: An investigation of the conceptual sources of the verbal counting principles. *Cognition, 105*, 395–438.
Libertus, M. E., & Brannon, E. M. (2010). Stable individual differences in number discrimination in infancy. *Developmental Science, 13*, 900–906.
Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science, 14*, 1292–1300.
Libertus, M. E., Feigenson, L., & Halberda, J. (2013a). Is approximate number precision a stable predictor of math ability? *Learning and Individual Differences, 25*, 126–133.
Libertus, M. E., Feigenson, L., & Halberda, J. (2013b). Numerical approximation abilities correlate with and predict informal but not formal school mathematics abilities. *Journal of Experimental Child Psychology, 116*, 829–838.
Libertus, M. E., Odic, D., Feigenson, L., & Halberda, J. (2015). The Developmental Vocabulary Assessment for Parents (DVAP): A novel tool to measure vocabulary size in 2- to 7-year-old children. *Journal of Cognition and Development, 16*, 442–454.
Libertus, M. E., Odic, D., & Halberda, J. (2012). Intuitive sense of number correlates with math scores on college-entrance examination. *Acta Psychologica, 141*, 373–379.
Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development, 76*, 978–988.
Lourenco, S. F., Bonny, J. W., Fernandez, E. P., & Rao, S. (2012). Nonsymbolic number and cumulative area representations contribute shared and unique variance to symbolic math competence. *Proceedings of the National Academy of Sciences of the United States of America, 109*, 18737–18742.
Lyons, I. M., Ansari, D., & Beilock, S. (2012). Symbolic estrangement: Evidence against a strong association between numerical symbols and the quantities they represent. *Journal of Experimental Psychology: General, 141*, 635–641.
Lyons, I. M., & Beilock, S. L. (2011). Numerical ordering ability mediates the relation between number-sense and arithmetic competence. *Cognition, 121*, 256–261.
Lyons, I. M., Nuerk, H. C., & Ansari, D. (2015). Rethinking the implications of numerical ratio effects for understanding the development of representational precision and numerical processing across formats. *Journal of Experimental Psychology: General, 144*, 1021–1035.
Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in Grades 1–6. *Developmental Science, 17*, 714–726.
Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011a). Impaired acuity of the approximate number system underlies mathematical learning disability. *Child Development, 82*, 1224–1237.
Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011b). Preschoolers' precision of the approximate number system predicts later school mathematics performance. *PLoS ONE, 6*(9), e23749.
Meck, W. H., Church, R. M., & Gibbon, J. (1985). Temporal integration in duration and number discrimination. *Journal of Experimental Psychology: Animal Behavior Processes, 11*, 591–597.
Mejias, S., Mussolin, C., Rousselle, L., Gregoire, J., & Noel, M. P. (2012). Numerical and nonnumerical estimation in children with and without mathematical learning disabilities. *Child Neuropsychology, 18*, 550–575.
Mejias, S., & Schiltz, C. (2013). Estimation abilities of large numerosities in kindergartners. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00518.
Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology, 103*, 490–502.
Odic, D., Hock, H., & Halberda, J. (2014). Hysteresis affects approximate number discrimination in young children. *Journal of Experimental Psychology: General, 143*, 255–265.
Odic, D., Im, H. Y., Eisinger, R., Ly, R., & Halberda, J. (2015). PsiMLE: A maximum-likelihood estimation approach to estimating psychophysical scaling and variability more reliably, efficiently, and flexibly. *Behavior Research Methods*, 1–18.
Odic, D., Le Corre, M., & Halberda, J. (2015). Children's mappings between number words and the approximate number system. *Cognition, 138*, 102–121.
Odic, D., Libertus, M. E., Feigenson, L., & Halberda, J. (2013). Developmental change in the acuity of approximate number and area representations. *Developmental Psychology, 49*, 1103–1112.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., ... Zorzi, M. (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.
Pica, P., Lemer, C., Izard, V., & Dehaene, S. (2004). Exact and approximate arithmetic in an Amazonian indigene group. *Science, 306*, 499–503.
Pinheiro-Chagas, P., Wood, G., Knops, A., Krinzinger, H., Lonnemann, J., Starling-Alves, I., ... Haase, V. G. (2014). In how many ways is the approximate number system associated with exact calculation? *PLoS ONE, 9*(11), e111155.
Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods, 40*, 879–891.
Price, G. R., & Fuchs, L. S. (2016). The mediating relation between symbolic and nonsymbolic foundations of math competence. *PLoS ONE, 11*(2), e0148981.
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.
Rousselle, L., & Noel, 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.
Ryoo, J. H., Molfese, V. J., Brown, E. T., Karp, K. S., Welch, G. W., & Bovaird, J. A. (2015). Examining factor structures on the Test of Early Mathematics Ability–3: A longitudinal approach. *Learning and Individual Differences, 41*, 21–29.
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., Defever, E., Maertens, B., & Reynvoet, B. (2014). The approximate number system is not predictive for symbolic number processing in kindergarteners. *Quarterly Journal of Experimental Psychology, 67*, 271–280.
Schneider, M., Beeres, K., Coban, L., Merz, S., Schmidt, S. S., Stricker, J., et al (2016). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*. http://dx.doi.org/10.1111/desc.12372.
Siegler, R. S., & Opfer, J. E. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science, 14*, 237–243.
Smets, K., Sasanguie, D., Szücs, D., & Reynvoet, B. (2015). The effect of different methods to construct non-symbolic stimuli in numerosity estimation and comparison. *Journal of Cognitive Psychology, 27*, 310–325.
Spaepen, E., Coppola, M., Spelke, E., Carey, S., & Goldin-Meadow, S. (2011). Number without a language model. *Proceedings of the National Academy of Sciences of the United States of America, 108*, 3163–3168.
St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59*, 745–759.
Starr, A. B., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences of the United States of America, 110*, 18116–18120.
Sullivan, J., & Barner, D. (2014). Inference and association in children's early numerical estimation. *Child Development, 85*, 1740–1755.
Szücs, D., Nobes, A., Devine, A., Gabriel, F. C., & Gebuis, T. (2013). Visual stimulus parameters seriously compromise the measurement of approximate number system acuity and comparative effects between adults and children. *Frontiers in Psychology, 4*. http://dx.doi.org/10.3389/fpsyg.2013.00444.
van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science, 17*, 492–505.
van Oeffelen, M. P., & Vos, P. G. (1982). A probabilistic model for the discrimination of visual number. *Perception & Psychophysics, 32*, 163–170.
Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology, 24*, 220–251.
Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition, 74*, B1–B11.