CHAPTER 6

Longitudinal Development of Nonsymbolic and Symbolic Magnitude Comparison Skills: Contradictions, Methodologies, and Moving Forward

Xenidou-Dervou, I., Ansari, D., Molenaar, D., van der Schoot, M., & van Lieshout, E. C. D. M. (*manuscript in preparation*). Longitudinal Development of Nonsymbolic and Symbolic Magnitude Processing: Contradictions, Methodologies, and Moving Forward.

ABSTRACT

What developmental roles do nonsymbolic (e.g., dot arrays) and symbolic (i.e., Arabic numerals) magnitude comparison skills play as precursors of children's mathematics achievement? To date, the literature has yielded a patchwork of findings. Notably, different studies have been assessing children of different ages, with different task-formats and domain-general capacities were rarely controlled for. Furthermore, the issue of how children's individual growth rates in these skills relate to their future math achievement had not been previously addressed. We assessed a large sample in kindergarten, grade 1 and grade 2 on two well-known nonsymbolic and symbolic magnitude comparison task-formats: one entails simultaneously presented small numerosities, the other sequentially presented large numerosities. Results demonstrated that symbolic processing underwent larger developmental improvements than nonsymbolic. Kindergarteners' accuracy in all four measures correlated with their future math achievement. With respect to their growth rates, however, only developmental change in the symbolic sequential-large task predicted children's future math achievement. Analyses, where WM capacity and initial IQ were controlled for, revealed how the targeted predictive relationships dynamically change over time. In kindergarten and grade 1, both nonsymbolic and symbolic magnitude processing uniquely predicted future math achievement. In grade 2, however, it was only predicted by symbolic processing. Symbolic magnitude processing was consistently a stronger predictor of future math achievement compared to nonsymbolic but the latter also played an important role. The present study explains and reconciles existing contradictions and proposes a unitary view.

INTRODUCTION

The question of what underlies the development of mathematical achievement has attracted a lot of attention the last decades. The reason is simple: math skills play a prominent role in our cognitive development and life success (e.g., Dougherty, 2003; Reyna & Brainerd, 2007). Numbers are everywhere and they can take many forms: for example, there is the *nonsymbolic* representation consisting of five dots on a screen and the *symbolic* representation of the number "5" in its Arabic form. What both of these representations have in common is the "fiveness" of the numerosities' magnitude. Extensive focus has been placed on the early markers of numerical cognition, particularly on the role that nonsymbolic and symbolic magnitude comparison processing skills play as building blocks of numerical cognition (for reviews see De Smedt, Noël, Gilmore, & Ansari, 2013; Feigenson, Libertus, & Halberda, 2013). Findings so far have been contradictory. In the literature, however, one notices four striking gaps: a) Up to now, studies have been examining only how children's nonsymbolic and symbolic magnitude processing skills predict future math skills but have not assessed how their *individual growth rates* relate to their future math achievement. Performance changes over time, but not necessarily in the same way or in the same rate for all children, b) There is a shortage of longitudinal developmental studies examining whether and how the different magnitude processing *predictors' power dynamically change* from one grade to another. c) *Measures* with fundamentally different design characteristics and numerosity ranges, have been used interchangeably, d) Domain-general capacities such as *working memory* resources are rarely controlled for. The present study strived to fill in these gaps and thereby resolve the existing contradictions.

Nonsymbolic and Symbolic Magnitude Processing

Research has demonstrated that human and non-human primates are born with an ability to estimate and manipulate abstract quantities in nature. The so-called Approximate Number System (ANS; Dehaene, 2011) is thought to be a pre-linguistic cognitive system where magnitudes are represented and processed. The ANS enables humans to compare and manipulate nonsymbolic numerosities already from infancy onwards (for reviews see Dehaene, 2011; Feigenson et al., 2013; De Smedt et al., 2013). Of course, as humans we also develop higher-order mathematical skills with symbols. So, how does this "innate" ability affect the development of our symbolic processing and what predicts the development of mathematical achievement, nonsymbolic, symbolic processing or both? These questions have generated intense scientific debate since they have important theoretical as well as educational implications (e.g., De Smedt et al., 2013; Noël & Rousselle, 2011). Establishing which early cognitive predictors play an important role and how, in the development of math achievement, can inform educational practice, math curricula contents and guide early intervention designs (De Smedt et al., 2013). For example, should educational practise focus on training children's nonsymbolic or symbolic skills or perhaps place different focus at different ages?

One class of theorists supports the viewpoint that symbolic representations of numbers directly map onto ones readily accessible nonsymbolic representations, i.e., the ANS (e.g., Lipton & Spelke, 2005; Piazza & Izard, 2009). In this respect, the ANS is viewed as the cognitive foundation that fosters and enhances the development of general mathematics achievement. This has been a compelling theory and several studies have demonstrated relations between ANS measures and general math achievement (Gilmore, McCarthy, & Spelke, 2010; Inglis, Attridge, Batchelor, & Gilmore, 2011; Libertus, Feigenson, & Halberda, 2011; Starr, Libertus, & Brannon, 2013; for a review see Feigenson et al., 2013). At the same time, however, many studies have failed to find such relations between the ANS and symbolic processing or math achievement (e.g., Bartelet, Vaessen, Blomert, & Ansari, 2014; Holloway & Ansari, 2009; Lyons, Price, Vaessen, Blomert, & Ansari, 2014; Sasanguie, Defever, Maertens, & Reynvoet, 2014; Sasanguie, Göbel, Moll, Smets, & Reynvoet, 2013). Thus, another class of theorists has suggested that symbolic numbers are processed and acquire meaning in a fundamentally different cognitive system (e.g., Lyons, Ansari, & Beilock, 2012). Within this framework, symbolic magnitude processing is viewed as the best predictor of math achievement, not the ANS (De Smedt et al., 2013; Lyons et al., 2014). If symbolic processing does not directly map only onto ones pre-existing nonsymbolic representations, then we would expect them to demonstrate different developmental trajectories and their growth rates over time to be unrelated. To our knowledge, however, the developmental trajectories of nonsymbolic and symbolic magnitude processing have not been previously compared.

As various contradicting results come forth, the roles of nonsymbolic and symbolic magnitude processing remain unclear. Recently a large-scale study proposed that both nonsymbolic and symbolic magnitude processing play a unique role (Xenidou-Dervou, De Smedt, van der Schoot, & van Lieshout, 2013). This study provided proof for aspects of both of the aforementioned theoretical classes and proposed an integrative view on the issue. Using structural equation modelling, the authors revealed that nonsymbolic and symbolic approximation, i.e. the ability to add and to compare large numerosities in an approximate manner, comprised two related but notably distinct cognitive constructs. Also, both nonsymbolic and symbolic approximation correlated with math achievement over and above working memory capacity. However, once structural paths were modelled a specific cognitive pattern emerged: nonsymbolic approximation was no longer directly related to math achievement; symbolic approximation completely mediated its effect. Specifically, symbolic processing predicted math achievement uniquely, above and beyond the ANS and working memory capacity. Working memory played its own prominent role; it predicted performance in both nonsymbolic and symbolic approximation as well as general math achievement. This was, however, a correlational study, which focused only on age group (children in kindergarten). Therefore, the developmental pathways of nonsymbolic and symbolic magnitude processing skills, and how these relate to individual differences in children's math achievement, have not been clarified. In a recent review of findings concerning the relationship between math achievement and numerical magnitude processing (both symbolic and non-symbolic), De Smedt et al. (2013) acknowledged two factors that may give rise to the patchwork of contradictory results that characterise the extant literature: a) The age of the participants assessed, and b) The task-formats used to assess magnitude comparison.

Sources of Contradiction

In order to identify the role that nonsymbolic and symbolic magnitude skills play, longitudinal and developmental studies are clearly necessary. ANS acuity (Halberda & Feigenson, 2008) and symbolic magnitude precision have been shown to increase with age (Holloway & Ansari, 2009; Sasanguie, De Smedt, Defever, & Reynvoet, 2011). Several longitudinal studies have demonstrated ANS acuity before the start of formal school instruction to correlate or be predictive of later math achievement (Gilmore et al., 2010; Libertus et al., 2011; Mazzocco, Feigenson, & Halberda, 2011; Starr et al., 2013). Furthermore, Inglis et al., (2011) found that ANS acuity correlates with mathematical achievement in childhood but not in adulthood.

These studies, however, did not assess symbolic magnitude processing. With cross-sectional designs, Lyons et al. (2014) and Sasanguie, Göbel, et al. (2013) assessed various nonsymbolic and symbolic measures simultaneously across *primary school children* and found no evidence for nonsymbolic magnitude processing being uniquely predictive of arithmetic. Instead, only symbolic numerical processing played a unique role. On the basis of these findings, we expected that the ANS, as a readily accessible system, would play a unique role primarily in kindergarten, before school instruction starts when formal symbolic math learning has not yet initiated (e.g., Gilmore et al., 2010; Xenidou-Dervou et al., 2013). From grade 1 and onwards, however, the predictive role of symbolic processing would take over (Lyons et al., 2014; Sasanguie, Göbel, et al., 2013). Thus, we hypothesized that the predictive roles of nonsymbolic and symbolic magnitude comparison skills would dynamically change over time. To our knowledge, this is the first study, which – due to its large-scale longitudinal design – allowed the examination of whether and how the predictive roles of magnitude comparison skills change across grades and explore how children's individual growth in these skills relates to their future math achievement.

In contrast to our aforementioned hypothesis, however, Bartelet et al. (2014) recently demonstrated that in kindergarten only symbolic magnitude processing skills predicted children's grade 1 mathematics above and beyond nonsymbolic skills. Notably, though, in this study, children's working memory capacities were not controlled for. Also, the tasks used in this study differed on several aspects from certain other kindergarten studies; for example, the numerosities were presented simultaneoulsy, not sequentially (e.g. Gilmore et al., 2010; Xenidou-Dervou et al., 2013). In general, one notable difference across the various studies conducted so far is the task-formats used to assess nonsymbolic and symbolic magnitude processing skills.

Task-formats used across the literature can differ both on design characteristics as well as numerosity ranges but have nevertheless been used interchangeably. Specifically, in one well-known magnitude comparison task-format, the numerosities to be compared (nonsymbolic or symbolic) are presented *simultaneously* (see for example Figure 1A). This task-format usually entails *small numerosities* within the range of 1 up to 9 (e.g., De Smedt, Verschaffel, & Ghesquière, 2009; Holloway & Ansari, 2008, 2009; Sasanguie, Göbel, et al., 2013; Sasanguie, Van den Bussche, & Reynvoet, 2012; Sekuler & Mierkiewicz, 1977). In contrast, another well-known magnitude comparison task-format entails several *sequential* steps: the child sees a blue (nonsymbolic or symbolic) numerosity dropping down on the left side of the screen, this is then covered by an occluder, and then a comparison red numerosity drops down on the right side of the screen (see Figure 1B). This task-format comprises *large numerosities* ranging for example from 6 up to 70 (e.g., Barth et al., 2006; Barth, La Mont, Lipton, & Spelke, 2005; De Smedt & Gilmore, 2011; Gilmore, McCarthy, & Spelke, 2007; Gilmore et al., 2010; Xenidou-Dervou et al., 2013; Xenidou-Dervou, van Lieshout, & van der Schoot, 2014). Performance across such different task-formats correlate during childhood (Gilmore, Attridge, De Smedt, & Inglis, 2014) but not in adulthood (Gilmore, Attridge, & Inglis, 2011), which might indicate that they have different developmental trajectories. To our knowledge, no previous developmental study has concurrently assessed these different task-formats to look at differences and similarities in their developmental trajectories and relationship to variability in children's developing numerical and mathematical skills.

Another important impediment across the existing literature is that most studies do not control for domain-general capacities, such as working memory (Gilmore et al., 2011; Gullick, Sprute, & Temple, 2011; Xenidou-Dervou et al., 2013). It has recently been demonstrated that nonsymbolic and symbolic magnitude processing call upon different working memory (WM) resources in both task-formats (Friso-van den Bos, Kroesbergen, & van Luit, 2014; Xenidou-Dervou et al., 2013, 2014; in press). WM is a limited-capacity multicomponent cognitive construct, which is responsible for the short-term storage and manipulation of information in an online manner when executing cognitive tasks. According to Baddeley's model (Baddeley, 2003, 2012; Repovs & Baddeley, 2006), WM entails the Phonological Loop (PL), which retains phonological information, the Visuospatial Sketchpad (VSSP), which retains visuospatial information and the Central Executive (CE), which monitors, controls and regulates the processes of the other two components and connects them with one's long-term memory.

WM plays a fundamental role in mathematical achievement (DeStefano & LeFevre, 2004; Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Raghubar, Barnes, & Hecht, 2010). Recently it has been shown that both nonsymbolic and symbolic magnitude processing are related to WM capacity (Friso-van den Bos, Kroesbergen, & van Luit, 2014; Xenidou-Dervou et al., 2013; 2014; in press). The role of the different WM components and their interactions when conducting a given cognitive task depend upon the characteristics of the task and the age of participants (e.g., Friso-van den Bos, van der Ven, Kroesbergen, & van Luit, 2013; Rasmussen & Bisanz, 2005; Simmons, Willis, & Adams, 2012; Xenidou-Dervou, van der Schoot, & van Lieshout, in press). Thus, it becomes clear that in order to identify the *unique* role that performance on a given magnitude processing task-format plays at a given age, one must control for the effect of WM capacities (Hornung, Schiltz, Brunner, & Martin, 2014; Xenidou-Dervou et al., 2013). For the present study we assessed the participants' performance on a wide range of both math specific and math nonspecific WM tasks at each developmental stage (kindergarten, grade 1 and grade 2) in order to control for their effects.

The present study

We aimed to shed light on the conflicting results of studies that focus on nonsymbolic or symbolic magnitude processing as precursors of children's math achievement in different ages. We administered two well-known task-formats of nonsymbolic and symbolic magnitude comparison to a large Dutch-speaking sample in kindergarten, grade 1 and grade 2. We also assessed these children's IQ and WM skills as domain-general control measures and their general mathematical achievement at the end of grade 2. In the Netherlands, formal schooling initiates in grade 1, not earlier. Therefore, this sample is unique in the sense that it allows us to examine the developmental transition from kindergarten to formal schooling. Our aim was to address for the first time the following three research questions:

1. Does performance in the different task-formats (simultaneous-small vs sequential-large) and different stimulus formats (nonsymbolic vs symbolic) develop differently from kindergarten up to grade 2?

Given the fundamental differences between the two task-formats outlined earlier, we hypothesized that performance on these tasks would demonstrate different developmental trajectories (Gilmore et al., 2014, 2011) and children's individual growth rates in these two task-formats would not correlate. Also, we expected that nonsymbolic and symbolic magnitude comparison as two distinct constructs (Lyons et al., 2012; Xenidou-Dervou et al., 2013) would also demonstrate different developmental trajectories and their growth factors would not be related.

2. How does children's performance level in kindergarten (initial status) and their individual developmental growth in the different magnitude comparison measures relate to their future general math achievement?

As outlined earlier, since several studies show performance in the different nonsymbolic and symbolic task-formats to be related to math achievement, we expected that performance level in kindergarten in all magnitude comparison measures would correlate with math achievement at the end of grade 2. However, as symbolic processing seems to be a consistent predictor in primary school ages (De Smedt et al., 2013), we also expected developmental change primarily in symbolic magnitude processing to play a role. Particularly, this was predicted for the task-format that entails large numerosities as children receive symbolic instruction upon entering primary school and their knowledge of large Arabic numerals improves.

3. What predicts math achievement *uniquely* at the end of grade 2, controlling for IQ and WM capacities? Do the roles of the different magnitude processing predictors change across the grades?

We hypothesized that both nonsymbolic (Gilmore et al., 2010; Libertus et al., 2011) and symbolic measures would play a unique role only in kindergarten (Xenidou-Dervou et al., 2013), whereas the symbolic ones would take over from grade 1 and onwards when school instruction starts (De Smedt et al., 2013; Lyons et al., 2014; Sasanguie, et al., 2013).

METHOD

Participants

This data is part of the interlinked MathChild project (http://vu.mathchild.nl/en/home/), during which 444 children from 25 schools in the Netherlands were assessed on a number of measures in kindergarten, grade 1 and grade 2. This large sample's characteristics are representative of the Dutch population (Driessen, Mulder, & Roeeveld, 2012; Xenidou-Dervou et al., 2013). Written consent was acquired from all children's legal guardians. Children identified as extreme outliers, i.e. who scored more than three standard deviations above or below the group mean, in one or more of the present study's measures were removed from the analyses (40 children). Throughout the three years of measurements, 80 children dropped out primarily due to family relocations. At the last measurement wave (see Table 1), the sample consisted of 324 children (M_{age} = 7.99 years, *SD*= 0.33, 176 boys, 148 girls). All children spoke Dutch and 96.6% of them held the Dutch nationality. The sample was acquired from middle- to high- SES environments. 33.6 % of the children's mothers and 26.2% of their fathers had received middle-level applied education (in the Dutch Educational system: MBO). 42.9 % of the mothers and 46.2 % of the fathers attended higher levels of education (in the Dutch Educational system: HBO and higher levels).

Procedure

All participants were tested in quiet settings within their school facilities by trained experimenters with the exception of the general mathematics test (CITO). The CITO ability scores were collected by school staff as part of the usual school tests. The rest of the data of this study comprises a set of tasks administered for the interlinked MathChild project across three testing sessions of approximately 20 mins in kindergarten and across two sessions of 30 min duration in grade 1 and grade 2. Between two sessions, there was a minimum of a day and a maximum of two weeks. Table 1 depicts the timeline of administration of the materials. All experimenters used the same elaborate protocol with instructions for testing administration across all measurements. Kindergarten data on the large-sequential tasks (approximate comparison) have been previously reported in Xenidou-Dervou et al. (2013) and the small-sequential in Friso-van den Bos, Kroesbergen, & van Luit (2014).

Table 1. *Measurement timeline.*

	Kindergarten		Grade 1		Grade 2	
Task	**T1**	**T2**	**T3**	**T4**	**T5**	**T6**
IQ	X					
Dot Matrix, Odd One Out	X		X		X	
Word RF, Word RB		X		X		X
Digit RF, Digit RB		X		X		X
Small-simultaneous	X			X	X	
Large-sequential	X		X		X	
General Math Achievement						X

Note. T1, T3, T5 measurement waves took place in the 1st half of the given academic year (November-December) and T2, T4, T6 in the 2nd half (May-June). The General Math Achievement test (CITO) was administered in June, at the end of grade 2. RF = Recall Forward, RB = Recall Backwards.

Material

All material, apart from the General Math Achievement and IQ tests, were computerized and presented with E-Prime version 1.2 (Psychological Software Tools, Pittsburgh, PA, USA) presented in HP Probook 6550b laptops.

Magnitude Processing

Simultaneous-Small. We administered a nonsymbolic and symbolic task developed on the basis of the widely used "magnitude comparison" task (Friso-van den Bos et al., 2014b;

Holloway & Ansari, 2009; Sekuler & Mierkiewicz, 1977). These tasks entailed 6 practise and 26 testing trials. During testing, no feedback was provided. In each trial, the child saw two numerosities, one on the right and one on the left side of the screen (see Figure 1A). Participants were asked to identify which numerosity was larger by pressing the left or the right response box situated in front of them. In a half of the trials, the larger numerosity was presented on the right side of the screen and in the other half, on the left. Children were instructed to respond as correctly and as fast as possible. Numerosities in these tasks ranged from 1 up to 9. The testing trials included all possible numerical pairs with the absolute distances between the comparison numerosities ranging from 1-4 (distance 1: 8 trials; distance 2: 7 trials; distance 3: 6 trials; distance 4: 5 trials).

The *nonsymbolic* condition started with an alerting beep sound of 100ms followed by a 1500ms warning interval (< >). As depicted in Figure 1A, the dot stimuli were presented in white on a black background, left and right from a yellow asterisk (fixation point). The response interval lasted until an answer was provided or until a maximum of 5000 ms was reached. To prevent the children from counting the dots, the stimuli were only presented for 840 ms. As in previous studies, continuous quantity variables, were controlled for with the methodology developed by Dehaene, Izard and Piazza (2005). According to this methodology, dot diameter was constant in half of the trials whereas in the other half, the size of the total dot surface area was constant. Trial order was randomized. For each continuous quantity variable (constant dot size and constant area) and for each numerosity, there was a pool of 16 different dot patterns. The program chose randomly one of these, so that the individual patterns of the dots were randomized as well. Thus, it is assumed that it is unlikely that the responses could be associated with specific dot patterns instead of quantity.

The *symbolic* condition was identical to the nonsymbolic, with the key difference that the corresponding Arabic numeral now replaced the dot stimuli. In this condition, the fixation point was now a dot instead of an asterisk in order to prevent possible confusion with the multiplication sign.

Sequential-Large. A nonsymbolic and a symbolic version of the well-known "approximate comparison" tasks were used (Barth et al., 2006; Gilmore, McCarthy, & Spelke, 2007; Gilmore, McCarthy, & Spelke, 2010; Xenidou-Dervou et al., 2013; 2014). We used the version reported in Xenidou-Dervou et al. (2013). These tasks included 6 practise and 24 testing trials. Feedback was only provided during practise. The number of practise trials was reduced to two in grade 1 and grade 2, as children were already familiar with this task-format.

In the *nonsymbolic* version the children were told that Sarah and Peter receive a set of blue and red dots respectively and were asked to respond to the question "Who got more dots? Sarah or Peter?". Within a trial (Figure 1B), the following sequence of events took place: 1) An amount of blue dots appeared and dropped on the right side of the screen next to image of the girl, 2) These were then covered by a grey box, 3) A set of red dots popped up and dropped on the left side of the screen next to the image of the boy. Children were instructed to respond as correctly and as fast as possible by pressing the blue or red response box in front of them. Each animated event lasted 1300 ms and between each event there was a 1200 ms interval. The fast interchange of events prevented counting. The child could respond from the moment the red dots appeared on the screen within a maximum of 7000 ms. Between trials, there was a 300 ms interval. Numerosities ranged from 6 up to 70. The blue array differed from the comparison red array by three ratios: 4:7, 4:6, 4:5 (easy, middle and difficult ratio). There were eight trials for each ratio. In half of the trials the blue array was larger, whereas in the other half the red was larger. Trial order was randomized. To avoid responses being reliant on the physical features of the dots and not quantity per se, dot stimuli followed a commonly used control methodology: Dot size, total dot surface area, total dot contour length and density correlated positively with numerosity whereas array size negatively in half of the trials, in the other half these relations were reversed (see Barth et al., 2006; Gilmore et al., 2010; Xenidou-Dervou et al., 2014).

The *symbolic* version was identical to the nonsymbolic, only now the dot stimuli were replaced by the corresponding Arabic numeral. Children were asked to respond to the question "Who got more stickers, Sarah or Peter?" by pressing the red or the blue response box in front of them (Figure 1B).

General Math Achievement

In the Netherlands, children's progress in primary school is monitored with the administration of the CITO tests. We acquired children's ability scores on the CITO Mathematics tests (in Dutch: CITO Rekenen-Wiskunde), which were assessed at the end of Grade 2 (June). The CITO math tests consist of many problems that cover a wide range of math domains: e.g., numbers and number relations, mental arithmetic (addition, subtraction, multiplication and devision), complex applications (i.e. mostly more than one operation per problem), measurement (e.g. weight, length, time). This series of tests have been demonstrated to have good psychometric properties and high reliability (see Janssen, Verhelst, Engelen, & Scheltens, 2010).

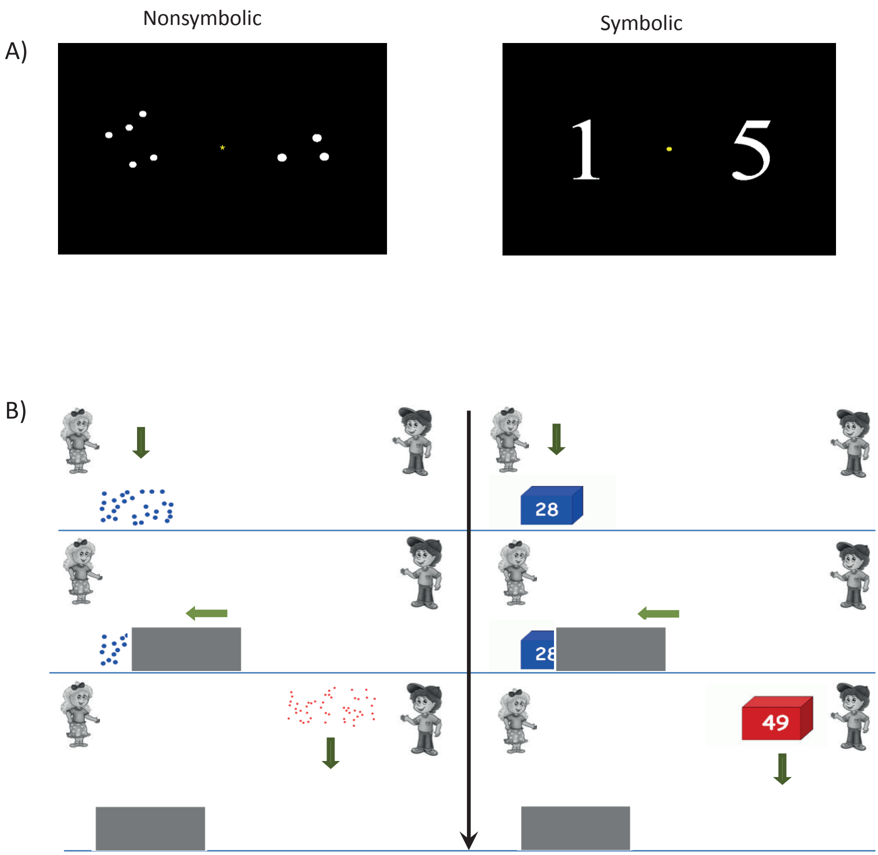

Figure 1. A) Example trials from the nonsymbolic and symbolic simultaneous-small magnitude comparison tasks. B) Example trials from the nonsymbolic and symbolic sequential-large magnitude comparison tasks.

Control measures

IQ. Children's non-verbal intelligence was assessed at the beginning of kindergarten with the Raven's Progressive Matrices (Raven, Raven, & Court, 1998) in a group-testing session. This well-known test entails visual patterns with increasing difficulty. In each trial, a pattern is presented with a missing piece. The participant's task is to identify the missing piece, which will complete the design, out of six pieces. Children's raw scores on this test were used.

Working Memory. We used the Dutch version of six tasks adapted from the Automated WM assessment battery (AWMA; Alloway, 2007; Alloway, Gathercole, Willis, & Adams, 2004) that tap children's capacity on three subcomponents of WM, namely the Phonological Loop, the Visuospatial Sketchpad and the Central Executive component (Baddeley, 2012). We were interested in controlling for all aspects of WM capacity; therefore we used both math-specific, i.e., entailing numbers, and not math-specific WM tasks. Moreover, we assessed both the ability of only retaining visuospatial or phonological information (VSSP and PL, respectively), as well as their interaction with the CE component (Repovs & Baddeley, 2006).

Visuospatial Sketchpad (VSSP). The VSSP component of WM was assessed with the "Cross Matrix". The *Cross Matrix* is identical to the well-known Dot Matrix of the AWMA battery; only in this version dots were replaced with crosses in order to exclude confounding factors with our nonsymbolic tasks that entailed dots. In this task the child saw a 4 x 4 matrix in which a cross appeared and disappeared. The child was instructed to remember the location of the cross and point to the correct box where the cross had previously appeared. A point was awarded for every correct response. After four correctly responded trials, the child was automatically advanced to the next level of difficulty, where one extra cross appeared. The task's levels of difficulty ranged from one up to five series. A correct response necessitated recalling correctly both the location and the right order in which the crosses appeared on the screen. If the child made three errors within one level of difficulty, the task was automatically terminated. The outcome measure entailed the total number of correct responses.

Phonological Loop (PL). Children's PL capacity was assessed with the: "Word Recall Forward" and the "Digit Recall Forward" task. In the *Word Recall*, the child heard a series of unrelated, high frequency words, which had to be later recalled correctly and in the right order. The *Digit Recall* task was the same as the word recall task, only now the child had to recall correctly and in the right order digits instead of words. Scoring and task progression rules were identical to the VSSP tasks.

Central Executive (CE). The CE can be fractionated on the basis of the information that is being manipulated within ones WM (Repovs & Baddeley, 2006). We, therefore, used three tasks to assess the children's CE capacity: the "Word Recall Backwards" (for not math-specific information) and the "Digit Recall Backwards" task (for math-specific phonological information) and the Odd One Out (for visuospatial information). The *Word Recall Backwards* and *Digit Recall Backwards* tasks were similar to the Word Recall Forward and Digit Recall Forward tasks; only now the child was required to recall the words in the reversed order. The *Odd One Out* task started with the child seeing three shapes and was asked to point to the shape, which differed from the other two. The shapes would then disappear from the screen and the child had to point to the location of the previously located odd one out shape. With increasing levels of difficulty the set of presented shapes increased. A response was registered as correct when the child pointed out correctly and in the right order the location of the odd shapes. Task progression rules were identical to the Cross Matrix task.

RESULTS

Table 2 depicts the correlations between accuracy and RT in the four magnitude comparison tasks: nonsymbolic and symbolic sequential-large and simultaneous-small across the 3 years of measurements (kindergarten, grade 1 and grade 2). There was no indication for an accuracy-RT trade-off between these four measures and general math achievement. Also, the correlations between the accuracy and RT data in the sequential-large tasks did not indicate any accuracy-RT trade-off. There was a small indication of such trade-off amongst the data of the simultaneous-small tasks. However, most correlations in this case were relatively small (the highest one was $r = .22$). In large samples such as the current, even small correlations become significant. Therefore, accuracy and RT data were used separately in the subsequent analyses. In Table 2, one notices moderate to large effect sizes (Cohen, 1992) amongst corresponding longitudinal nonsymbolic and symbolic task-formats in both accuracy and RT. Table 3 includes descriptive statistics on math achievement performance and the control measures.

Comparing Developmental Trajectories

To compare the developmental trajectories of nonsymbolic and symbolic simultaneous-small and sequential-large magnitude comparison accuracy, we conducted a 3 (Year: kindergarten, grade 1, grade 2) x 2 (Task: simultaneous-small and sequential-large) x 2 (Stimulus: nonsymbolic and symbolic) repeated measures ANOVA. In the case of violation of the assumption of sphericity, degrees of freedom were corrected using Greenhouse-Geisser estimates. As expected, we found a significant Year by Task by Stimulus interaction effect, $F(1.83, 598.48) = 111.05$, $p < .001$ (Figure 2A and 2B). Therefore, the two task-formats and the two stimuli formats demonstrated different developmental trajectories. To unravel the simple effects, two additional analyses were conducted for each task-format (simultaneous-small and sequential-large). For the simultaneous-small tasks, results demonstrated only main effects of Year, $F(1.54, 509.6) = 133.38$, $p < .001$

Table 2. *Correlations Between Nonsymbolic and Symbolic Sequential-Large and Simultaneous-Small Accuracy (1-12) and RT (13-24) across the 3 years as well as General Math Achievement at the end of Grade 2.*

		Math	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	NS SeqL Y1	.24***										
2	NS SeqL Y2	.28***	.32***									
3	NS SeqL Y3	.26***	.25***	.41***								
4	S SeqL Y1	.37***	.20***	.24***	.19***							
5	S SeqL Y2	.50***	ns	.25***	.19***	.38***						
6	S SeqL Y3	.33***	.17**	.19***	.28***	.24***	.32***					
7	NS SimS Y1	.24***	.29***	.27***	.27***	.22***	.28***	.22***				
8	NS SimS Y2	.17**	.12*	.12*	.17**	ns	.14*	.22***	.21***			
9	NS SimS Y3	.12*	.13*	.14*	.14*	ns	.15**	.22***	.18***	.28***		
10	S SimS Y1	.27***	.21***	.13*	.26***	.28***	.34***	.31***	.44***	.18***	.22***	
11	S SimS Y2	.16**	ns	ns	ns	ns	.23***	.13*	.16**	.30***	.19***	.21***
12	S SimS Y3	.13*	ns	.13*	.14*	ns	.12*	.20***	.16**	.20***	.44***	.14*
13	NS SeqL Y1_RT	-.11*	ns	ns	ns	-.23***	-.18***	-.15**	ns	ns	ns	-.14*
14	NS SeqL Y2_RT	-.18***	ns	ns	ns	-.17**	-.21***	-.18**	ns	-.14*	ns	-.14*
15	NS SeqL Y3_RT	-.26***	ns	-.13*	ns	-.22***	-.19***	-.12*	ns	ns	ns	ns
16	S SeqL Y1_RT	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns	ns
17	S SeqL Y2_RT	-.21***	-.13*	ns	ns	-.30***	-.17**	-.17**	ns	ns	ns	-.16**
18	S SeqL Y3_RT	-.38***	ns	-.13*	ns	-.35***	-.32***	-.19***	-.22***	ns	ns	-.18**
19	NS SimS Y1_RT	ns	ns	ns	ns	-.15**	ns	-.17**	.18***	ns	ns	ns
20	NS SimS Y2_RT	-.25***	-.12*	-.17**	ns	-.22***	-.16**	-.26***	ns	.15**	ns	ns
21	NS SimS Y3_RT	-.20***	ns	-.13*	-.12*	-.14*	-.16**	-.17**	ns	ns	.20***	ns
22	S SimS Y1_RT	-.18***	-.13*	ns	ns	-.24***	-.19***	-.12*	ns	ns	ns	.22***
23	S SimS Y2_RT	-.28***	-19***	-.13*	-.13*	-.26***	-.25***	-.22***	-.16**	ns	ns	-.18***
24	S SimS Y3_RT	-.31***	-15**	-.13*	-.17**	-.27***	-.29***	-.25***	-.11*	ns	ns	-.12*

Note. NS = Nonsymbolic, S = Symbolic, SeqL = Sequential-Large, SimS = Simultaneous-Small, Y1 = Kindergarten, Y2 = Grade 1, Y3 = Grade 2, ns = nonsignificant.
*** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$

11	12	13	14	15	16	17	18	19	20	21	22	23
.25***												
ns	ns											
ns	ns	.28***										
ns	ns	.15**	.44***									
ns	ns	.44***	.21***	.15**								
ns	ns	.26***	.57***	.38***	.24***							
-.11*	ns	.23***	.46***	.65***	ns	.42***						
ns	ns	.23***	ns	ns	.26***	ns	ns					
.17**	.11*	ns	.14*	.25***	ns	.15**	.27***	.23***				
ns	.17**	.16**	.15**	.25***	ns	.11*	.33***	.25***	.41***			
ns	ns	.26***	.18***	.16**	.23***	.23***	.24***	.45***	.21***	.21***		
.13*	ns	.23***	.25***	.31***	ns	.33***	.40***	.17**	.58***	.40***	.30***	
ns	.19***	.28***	.22***	.37***	ns	.27***	.49***	.18**	.39***	.61***	.27***	.59***

Table 3. *Means (and SDs) of the Control Measures and the Dependent Variable (Math Achievement).*

	Kindergarten	Grade 1	Grade 2
IQ	21.24 (5.08)		
Dot Matrix	9.97 (2.73)	13.24 (3.00)	15.29 (3.03)
Odd One Out	8.47 (2.88)	11.39 (2.56)	13.27 (2.58)
Word RF	13.89 (2.48)	14.98 (2.65)	15.77 (2.48)
Word RB	4.96 (1.82)	6.14 (2.11)	6.81 (2.26)
Digit RF	14.02 (2.42)	15.83 (2.41)	17.04 (2.48)
Digit RB	4.54 (1.64)	6.20 (2.03)	7.11 (2.32)
Math Achievement			67.06 (14.73)

and Stimulus, $F(1, 330) = 141.28$, $p < .001$ (Figure 2A), indicating that performance in the nonsymbolic and the symbolic task increased in a similar manner across the grades. Inspecting Figure 2A, though, one notices that there was a possible ceiling effect in this task-format. For the sequential-large tasks, results showed significant main effects for Year, $F(1.96, 643.86) = 447.95$, $p < .001$, and Stimulus, $F(1, 328) = 135.61$, $p < .001$ but also an interaction effect, $F(1.92, 631.17) = 135.61$, $p < .001$ (Figure 2B). Thus, as expected, nonsymbolic and symbolic performance in the large-sequential task format demonstrated different developmental trajectories. The symbolic condition underwent larger developmental change than its nonsymbolic counterpart.

The same analyses were conducted with the four measures' RT data. Once again, the 3 x 2 x 3 repeated measures ANOVA showed a significant 3-way interaction: Year by Task by Stimulus, $F(1.81, 594.59) = 7.86$, $p = .001$. For the simultaneous-small task, RT results now demonstrated significant Year, $F(1.59, 526.04) = 309.05$, $p < .001$, and Stimulus, $F(1, 330) = 164.95$, $p < .001$, main effects but this time also the expected Year by Stimulus interaction effect, $F(1.63, 539.34) = 81.22$, $p < .001$ (Figure 2C). For the sequential-large, as in the case of the accuracy data, we found significant Year, $F(1.78, 583.18) = 270.21$, $p < .001$, and Stimulus, $F(1, 328) = 156.1$, $p < .001$, as well as the expected Year by Stimulus interaction effect, $F(1.83, 600.75) = 17.92$, $p < .001$ (Figure 2D). Thus, in line with our hypotheses, our findings confirmed that the two task-formats demonstrate different developmental trajectories. Furthermore, as hypothesized, nonsymbolic and symbolic magnitude processing in both task-formats demonstrated different developmental trajectories.

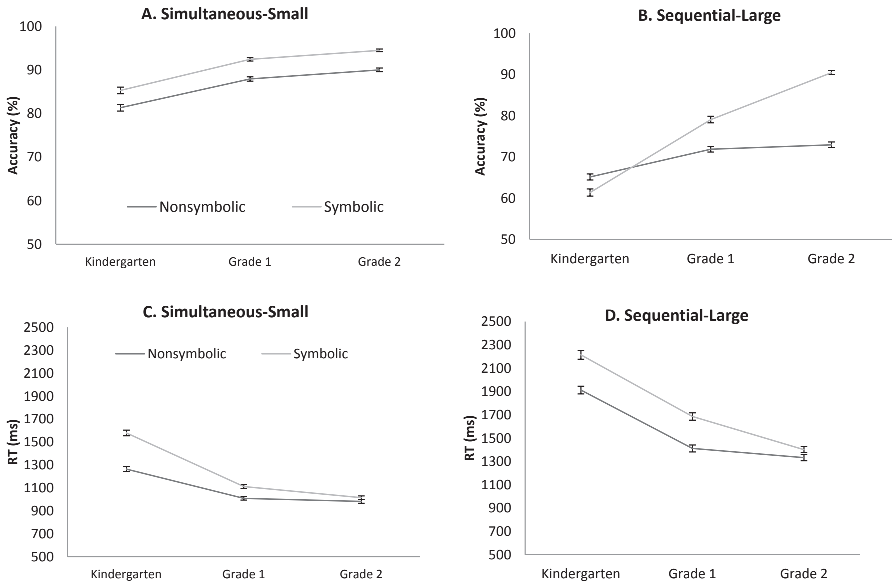

Figure 2. Development in nonsymbolic and symbolic simultaneous-small and sequential-large comparison accuracy (A, B) and RT respectively (C, D) from kindergarten up to grade 2. Nonsymbolic and symbolic magnitude processing demonstrated different developmental trajectories. Also, the two task-formats (simultaneous-small vs sequential-large) had different developmental trajectories.

Relations with General Math Achievement

Initial Status and Developmental Growth

To address our 2nd research question, we built two latent growth models (LGMs) in order to analyse the relation between magnitude processing and math achievement, one with the accuracy scores (Figure 3) and one with the RTs (Figure 4) of the magnitude comparison tasks, using the Mplus analytic software, version 5.1 (Muthén & Muthén, 1998–2011). LGMs refer to a class of models for longitudinal data that can be analysed with SEM and allow us to decompose the data to intercepts and slopes for each individual. This is a sharp contrast to common regressions, which assume a single intercept and slope for the whole sample (Duncam & Duncam, 2009). We specified one Initial Status (IS) latent factor and one Change (C) latent factor for the data of each magnitude comparison task (nonsymbolic and symbolic, sequential-large and simultaneous-small). Our aim was to explore how the IS and C factors on each of these measures related to distant math achievement. The first two loadings on the latent C factor were fixed to constants that correspond to the years of measurement: 0 for kindergarten and 1 for grade 1. The 0 loading for the first measurement (kindergarten) in the C factor is required to ensure that kindergarten is the initial level in the model. Thus, the IS factor is defined based on the kindergarten measurement and its unstandardized loadings were all fixed to 1. As observed in Figure 2, development in the all the magnitude formats did not appear to be strictly linear. Therefore, in our LGM models the third factor loading (i.e., grade 2 performance) in a given C factor was allowed to be freely estimated. Corresponding IS and C factors were specified to covary, i.e., they were free to correlate. Such a covariance would reflect the degree to which initial levels of magnitude processing correlates with subsequent rates of developmental change. Also, the four different IS factors were specified to covary, as well as the four C factors. Lastly, since the two task-formats and two stimulus formats demonstrated different developmental trajectories, we specified the covariances between each nonsymbolic and symbolic version within one type of task-format (simultaneous-small or sequential-large) for each year and between each task-format within one type of stimulus (nonsymbolic or symbolic) for each year.

Figure 3 depicts the baseline latent growth model for the accuracy data, which demonstrated acceptable fit to the data (*RMSEA* = 0.079). Table 4 depicts the means and standard errors in the IS and C factors, where one notices that children improved most in the symbolic sequential-large task over time (C_2 factor). As expected, we found no significant correlation between the four growth (C) factors. In other words, children's individual growth in each of the four task-formats was uncorrelated. The model's results demonstrated that the IS factors of all four tasks significantly correlated with math achievement at the end of grade 2. The symbolic sequential-large magnitude comparison measure had the highest correlation coefficient. Additionally, the developmental change factor in the symbolic sequential-large magnitude measure (C_2) correlated negatively with future math achievement. The negative direction of the association can be explained by the negative correlation between the IS_2 and C_2 factors ($r = -.82, p < .001$), which indicated that children with lower initial levels of symbolic sequential-large performance showed steeper linear increases over time. This was perhaps caused by the ceiling effect observed in grade 2 (skewness value = -1.21). Essentially, this demonstrated that children's individual developmental change in this specific task was related to their future math achievement.

Figure 4 depicts the baseline LGM with the RT data, which demonstrated acceptable fit to the data (*RMSEA* = 0.065). For SEM normalization purposes, we used the logarithmic transformations of the RTs. Results indicated that the variance of the two simultaneous-small change measures (C_3 and C_4) approached 0, i.e., on average there were

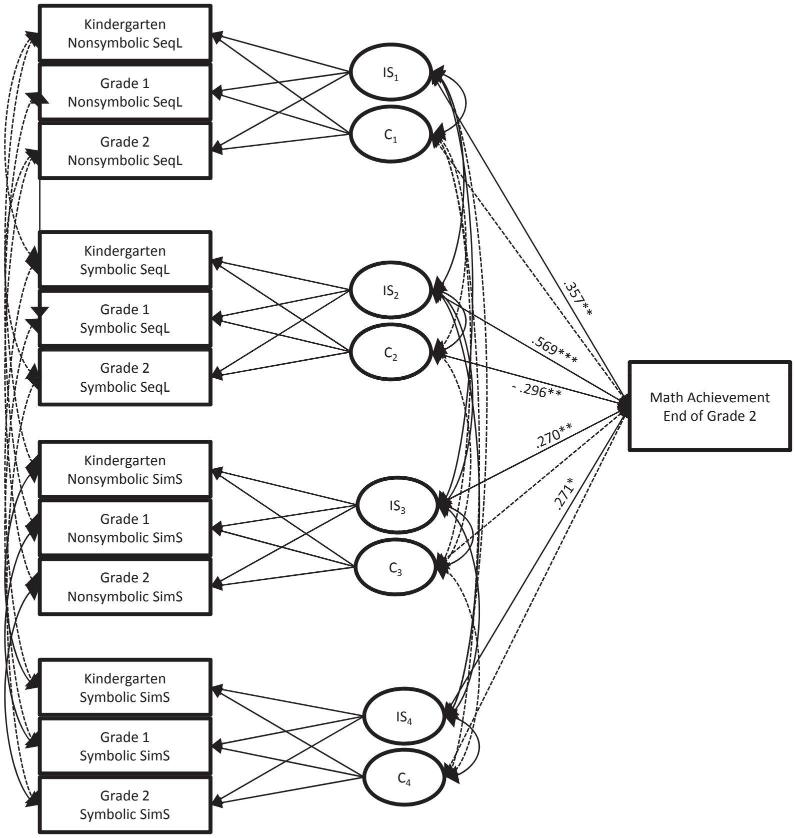

Figure 3. Baseline latent growth model with *accuracy* data on the four magnitude processing tasks. The IS latent factors on all four magnitude comparison measures correlated with math achievement longitudinally. Only development change on the symbolic sequential-large measure (C_2) correlated with math achievement at the end of grade 2. IS = initial status, C = Developmental Change, SeqL = Sequential-Large task format, SimS = Simultaneous-Small task format. Solid lines indicate significant correlations. Dashed lines indicate nonsignificant correlations. Values next to lines represent correlation coefficients. *** $p \leq .001$, ** $p \leq .01$, * $p \leq .05$

no individual differences in the growth of these measures. Therefore, these parameters were fixed to 0. Table 4 shows that children's speed improved in both of the sequential-large tasks (C_1 and C_2 factors). The growth factors, i.e., the rate of decrease in RT, of the nonsymbolic (C_1) and the symbolic (C_2) sequential-large tasks significantly correlated ($r = .34$, $p = .001$), which probably accounts for the common processing speed variance. With respect to the four magnitude comparison measures' relationship to future math achievement, the IS factors of the two simultaneous-small measures correlated with distant math achievement. Specifically, the symbolic simultaneous-small measure had the highest correlation coefficient. The IS factor of the nonsymbolic sequential-large task marginally (*p* = .054) correlated with distant math achievement. Interestingly, this time the IS factor of the symbolic sequential-large task did not correlate with future math achievement. Its growth factor, however, did (C_2), as was the case in the accuracy model. In other words, developmental change, i.e., decrease in RT, from kindergarten till the beginning of grade 2 in the speed with which children compared symbolic magnitudes

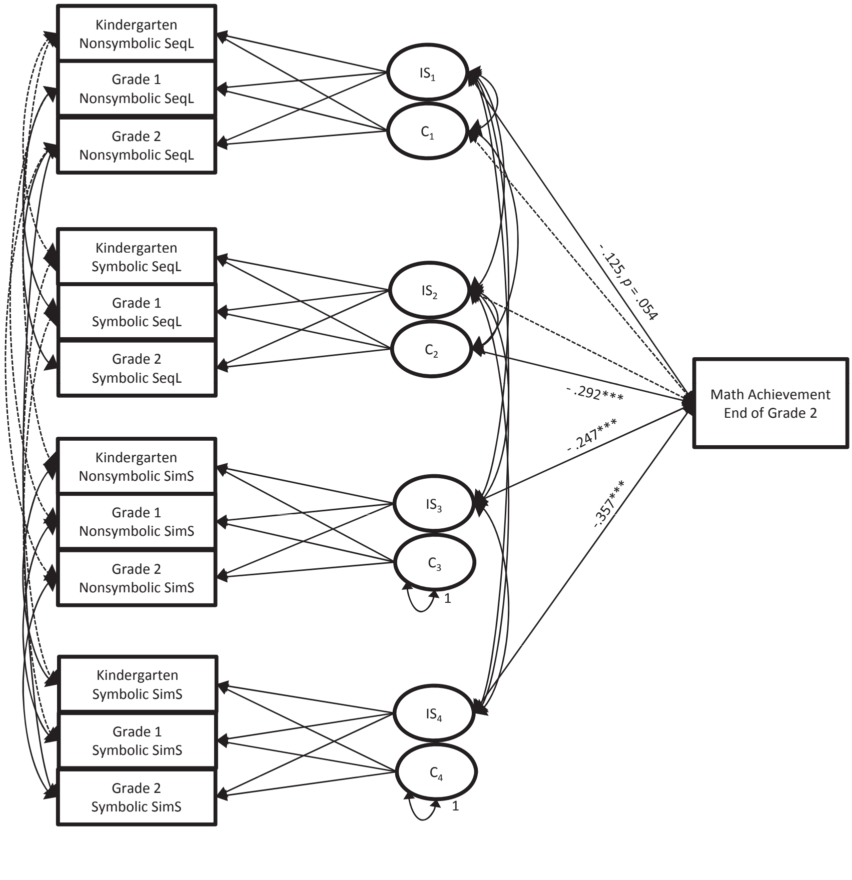

Figure 4. Baseline latent growth model with *RT* data on the four magnitude processing tasks. Kindergarten performance in the two simultaneous-small measures (IS_3 and IS_4) and developmental change only in the symbolic sequential-large task correlated with future math achievement. IS = initial status, C = Developmental Change, SeqL = Sequential-Large task format, SimS = Simultaneous-Small task format. Solid lines indicate significant correlations. Dashed lines indicate nonsignificant correlations. Values next to lines represent correlation coefficients. *** p ≤ .001

in the sequential-large task-format correlated with their math achievement at the end of grade 2. As in the case of the accuracy model, there was a negative correlation between the IS_2 and the C_2 factor (r = -.54, p < .001), which suggests that the longer the child's kindergarten RT was, the steeper his or her RT decline was over time.

It was numerically infeasible to specify structural paths (regressions) on math achievement as this step in the models gave unstable results. Therefore, in the following section we conducted common multiple regressions to identify the unique predictive role of each magnitude comparison skill for each grade.

Table 4. *Unstandardized Means and Standard Errors (SE) of the IS and C latent factors in the Accuracy (Figure 3) and RT (Figure 4) Latent Growth Models.*

Measure	Accuracy	RT
IS_1	15.46 (0.16)	2.91 (0.02)
C_1	1.70 (0.19)	-0.33 (0.02)
IS_2	14.51 (0.19)	3.04 (0.02)
C_2	4.23 (0.21)	-0.28 (0.02)
IS_3	20.90 (0.18)	2.50 (0.02)
C_3	1.83 (0.20)	--
IS_4	21.76 (0.18)	2.73 (0.02)
C_4	2.22 (0.19)	--

Note. -- = fixed parameters (see results section); IS = Initial Status; C = Developmental Change; subscripts: 1 = nonsymbolic sequential-large, 2 = symbolic sequential-large, 3 = nonsymbolic simultaneous-small, 4 = symbolic simultaneous-small.

6

Unique Predictive Power Across Grades

For our 3rd research question, we sought to identify how the unique predictive power of the different nonsymbolic and symbolic magnitude processing skills changes across grades when predicting distant math achievement. Therefore, we conducted a series of multiple linear regression analyses, one for each year, controlling for initial IQ and performance in the PL, VSSP and CE WM tasks in the respective year. One regression analysis for each grade was conducted by entering all variables in one step. This was done separately for magnitude comparison accuracy and RT data. In the case of the accuracy scores, F-tests indicated that all models significantly explained variance in grade 2 general math achievement: kindergarten, $F(12, 298) = 15.12$, $p < .001$, Adj. $R^2 = 0.35$, grade 1, $F(12, 299) = 17.16$, $p < .001$, Adj. $R^2 = 0.38$ and grade 2, $F(12, 293) = 13.46$, $p < .001$, Adj. $R^2 = 0.33$. Table 5 depicts the models' results for accuracy. In kindergarten and grade 1, both nonsymbolic and symbolic sequential-large magnitude processing uniquely predicted distant math achievement above and beyond WM skills and IQ (Table 5). We further compared the nonsymbolic and symbolic regression coefficients and found that symbolic sequential-large was a better predictor of math achievement both in kindergarten ($p = .000$) as well as in grade 1 ($p = .011$) than the nonsymbolic one. In grade 2, out of the four magnitude comparison predictors assessed in the beginning of grade 2, only symbolic sequential-large performance explained unique variance in math achievement at the end of the grade.

Also in the case of the RTs, *F*-tests showed that all models significantly explained variance in grade 2 general math achievement: kindergarten, $F(12, 298) = 10.89$, $p < .001$, Adj. $R^2 = 0.28$, grade 1, $F(12, 299) = 12.95$, $p < .001$, Adj. $R^2 = 0.34$ and grade 2, $F(12, 293) = 15.69$, $p < .001$, Adj. $R^2 = 0.37$. Table 6 shows the regression models' results for magnitude comparison RT. In this case, only the symbolic magnitude comparison predictors reached significance. Speed in comparing small numerosity digits in kindergarten, and large numerosities in grade 1 and grade 2 uniquely predicted distal math achievement above and beyond nonsymbolic magnitude processing speed, WM resources and initial IQ.

DISCUSSION

The present study strived to shed light on the role that nonsymbolic and symbolic magnitude comparison skills play in the transition from kindergarten to formal schooling (grade 1 and grade 2) and to reconcile existing contradicting findings in the literature (for reviews see De Smedt et al., 2013; Feigenson et al., 2013). For the first time, a single large sample of children was assessed on two different well-known nonsymbolic and symbolic magnitude comparison task-formats simultaneously from kindergarten through to grade 2.

We found that symbolic magnitude processing undergoes larger developmental changes than nonsymbolic. Also, the two different magnitude comparison task-formats (small simultaneously and large sequentially presented numerosities) demonstrated different developmental pathways. Children's accuracy growth rates in the two stimulus-formats of magnitude processing both within the sequential-large and simultaneous-

Table 5. *Accuracy (%) Per Year on the Magnitude Comparison Measures Predicting Math Achievement at the End of Grade 2, While Controlling for WM capacity on the Given Year and initial IQ.*

	Kindergarten		Grade 1		Grade 2	
Predictors	β	*p*	β	*p*	β	*p*
Age	*-.15*	*.002*	-.04	.336	.03	.552
IQ	*.19*	*.000*	*.17*	*.001*	*.25*	*.000*
Word RF	-.04	.561	.07	.238	-.01	.918
Word RB	*.12*	*.022*	.04	.393	*.12*	*.043*
Digit RF	*.19*	*.002*	.06	.301	*.14*	*.033*
Digit RB	*.13*	*.013*	.04	.400	.11	.072
Dot Matrix	.05	.310	*.14*	*.008*	.07	.228
Odd One Out	.03	.889	.08	.127	.11	.053
Nonsymbolic SimS	.10	.069	.04	.368	.02	.735
Symbolic SimS	.02	.753	.03	.644	.03	.525
Nonsymbolic SeqL	***.12***	***.019***	***.10***	***.028***	.08	.114
Symbolic SeqL	***.24***	***.000***	***.33***	***.000***	***.18***	***.001***

Note. General math achievement at the end of grade 2 was the dependent variable for all models. Models' adjusted R^2 and predictors' standardized beta coefficients (β) are reported. Significant predictors italicized. Significant magnitude comparison predictors in bold. SeqL = Sequential-Large, SimS = Simultaneous-Small.

Table 6. *Average RT (ms) Per Year on the Magnitude Comparison Measures Predicting Math Achievement at the End of Grade 2, While Controlling for WM capacity on the Given Year and initial IQ.*

	Kindergarten		Grade 1		Grade 2	
Predictors	β	*p*	β	*p*	β	*p*
Age	-.10	.058	-.04	.427	.03	.581
IQ	*.24*	*.000*	*.21*	*.000*	*.23*	*.000*
Word RF	-.01	.837	.06	.328	.10	.870
Word RB	*.14*	*.020*	.05	.397	*.11*	*.042*
Digit RF	*.20*	*.003*	*.12*	*.041*	*.14*	*.028*
Digit RB	*.14*	*.009*	.10	.065	*.13*	*.028*
Dot Matrix	.08	.140	*.15*	*.009*	.05	.353
Odd One Out	.02	.668	.11	.056	.12	*.029*
Nonsymbolic SimS	.07	.219	-.10	.638	.05	.409
Symbolic SimS	***-.16***	***.004***	-.06	.350	-.12	.054
Nonsymbolic SeqL	-.05	.394	-.06	.284	.01	.939
Symbolic SeqL	-.01	.836	***-.13***	***.033***	***-.25***	***.000***

Note. General math achievement at the end of grade 2 was the dependent variable for all models. Models' adjusted R^2 and predictors' standardized beta coefficients (β) are reported. Significant predictors italicized. Significant magnitude comparison predictors in bold. SeqL = Sequential-Large, SimS = Simultaneous-Small.

small task-formats were uncorrelated. Also, children's growth in the two task-formats both within the nonsymbolic and symbolic stimulus-format were uncorrelated. Our latent growth model analyses with accuracy data showed that kindergarteners' performance in all magnitude processing measures predicted children's future math achievement. With the RT data, however, we found that kindergarteners' speed only in the two simultaneous-small (nonsymbolic and symbolic) task-formats correlated with distant math achievement. Noticeably, though, with both accuracy and RT data, only individual developmental growth rates in comparing symbolic magnitudes in the sequential-large task-format correlated with distant math achievement. Our multiple regression findings revealed how the predictive roles of nonsymbolic and symbolic magnitude comparison skills dynamically change across grades. In particular, accuracy in both nonsymbolic and symbolic sequential-large comparison played a unique role in *kindergarten* and *grade 1* in predicting distant math achievement beyond WM capacities and initial IQ. Even though both predicted math achievement at these ages, our results showed that symbolic sequential-large magnitude comparison was a stronger predictor than nonsymbolic in both cases. By *grade 2,* only the ability to compare large digits in sequential steps significantly predicted math achievement longitudinally. With respect to magnitude comparison speed, we found that only symbolic processing RT uniquely predicted distant math achievement: small numerosities in kindergarten and larger numerosity ranges in grades 1 and 2. It should be highlighted that we found different patterns of results in the predictive relations of the four magnitude comparison measures with math achievement on the basis of the data that was used: accuracy or RT. In the following section, we discuss more elaborately our findings from a theoretical and methodological standpoint and their educational implications.

The fact that we share a cognitive ability with other species – the "innate" ability to estimate abstract quantities in nature, i.e., the ANS – generates a lot of questions. How does this evolutionary ancient system relate to our ability as humans to use symbols to represent quantities? It has often been assumed that symbolic representations directly map onto our readily accessible nonsymbolic representations (Lipton & Spelke, 2005; Mundy & Gilmore, 2009; Piazza & Izard, 2009). If that were the case, then symbolic and nonsymbolic magnitude processing would demonstrate similar developmental trajectories and children's individual growth rates in these measures would correlate. Our results, however, with two different magnitude comparison task-formats revealed that this is not the case. Nonsymbolic and symbolic magnitude processing demonstrated different developmental pathways, supporting the assumption that they are two distinct cognitive systems: one for representing and manipulating abstract magnitudes and one for representing and manipulating symbolic information (Lyons et al., 2012; Xenidou-Dervou et al., 2013). In other words, even though nonsymbolic and symbolic magnitude processing correlates, symbolic representations do not necessarily map only onto nonsymbolic representations. The ANS seems to influence symbolic processing in the initial stages of development, before formal schooling (Xenidou-Dervou et al., 2013) but with the start of formal school instruction, symbolic processing takes its own developmental path improving at a faster rate than nonsymbolic. Thus, education, age and experience appear to modulate symbolic magnitude processing more than nonsymbolic magnitude processing.

So, how does the ANS relate to the development of mathematics achievement? The existing literature has brought forth many contradicting results in this respect (De Smedt et al., 2013) and one source of contradiction is the task-formats used to assess magnitude processing. But before addressing what relates to math achievement and how, we examined whether these measures actually assess the same underlying constructs. One well-known task format includes small numerosities (1-9), which are presented simultaneously, whereas another task-format includes sequential steps and large numerosities (6 to 70). As expected, we found that performance and RT in these two task-formats demonstrated different developmental trajectories and children's individual growth rates in these measures were uncorrelated. Both of these findings constitute a warning against their interchangeable usage within the literature. Subsequently, we examined their relation to distant math achievement. Latent growth modelling revealed that accuracy rate in kindergarten in all of these magnitude comparison measures (nonsymbolic and symbolic, sequential-large and simultaneous-small) correlated with math achievement at the end of grade 2. With both accuracy and RT data though, it was evident that children's individual growth only in the *symbolic sequential-large* measure significantly related with children's future math achievement. As a next step, future research should identify the factors that potentially affect children's individual developmental growth rates in this magnitude comparison skill.

Developmental studies so far examining concurrently nonsymbolic and symbolic magnitude processing had only used task-formats that presented the numerosities simultaneously – with small or large numerosities, had had cross-sectional designs or only assessed children's performance at two time-points (Bartelet et al., 2014; Hornung et al., 2014; Lyons et al., 2014; Sasanguie et al., 2014). To our knowledge, this is the first study examining nonsymbolic and symbolic processing simultaneously from kindergarten through grade 2 that addressed the issue of children's initial status and developmental change (growth). A fundamental observation in behavioural sciences is that performance changes over time. Children's performance improves with age, experience and education but not all of them change in the same way or in the same rate. As much as children's early performance in a certain skill (initial status) can be an important predictor of their later achievements, their potential for growth in this skill and the consequences of this growth is equally – or even more – important. Interventions can potentially target and improve the rates of change in these specific skills.

Our latent growth model analyses examined the children's (intra)individual change over time in the four different magnitude comparison measures (nonsymbolic and symbolic, sequential-large and simultaneous-small). With the accuracy data, we saw that children's initial status in all four measures predicted their general math achievement three years later. More importantly, however, our results were the first to demonstrate that children's individual growth rates in the symbolic sequential-large magnitude comparison also had a significant effect on their future math achievement. This growth, however, was negatively correlated with kindergarten performance in this task. It seemed that children who initially performed poorly on this task showed steeper increases in their performance compared to higher performing children. Also, the negative association between the growth factor and future math achievement (Figure 3), indicated that poor performers in this task did not reach the same level as the higher performers. This was evident in the multiple regression analyses: High performance in the symbolic sequential-large task during the last measurement (beginning of grade 2) was still predictive of future math achievement. Developmental change also in the RT data of this task demonstrated the same pattern of results. There was a negative association between children's kindergarten RT and their developmental decrease in RT. This decrease in RT over time correlated negatively with children's future math achievement (Figure 4). So, it appears that once again the slowest performing children in this task did not catch up over time with the children who responded faster from the beginning (kindergarten). As demonstrated in the multiple regression analyses, children's RT in the symbolic sequential-large task in grade 2 was highly predictive of their future achievement.

The unique predictive role that each magnitude comparison skill plays at each developmental stage became clear when controlling for domain-general capacities, i.e., WM abilities at the given age and initial IQ, confirming and reconciling the various contradicting results across the literature. With the accuracy data, we found that *both* nonsymbolic and symbolic sequential-large magnitude comparison played a unique role in predicting distant math achievement in kindergarten and grade 1 (Gilmore et al., 2010; Hornung et al., 2014; Xenidou-Dervou et al., 2013). By grade 2 though, only symbolic sequential-large magnitude comparison explained unique variance in math achievement longitudinally (Lyons et al., 2014). This confirms the assumption that the ANS plays a unique role in the development of math achievement at the initial stages of development and gradually lessens that role with education (Inglis et al., 2011). With respect to the speed of comparing numerosities (RT), only the symbolic measures explained individual differences in math achievement longitudinally. The simultaneous-small symbolic measure played a unique role in kindergarten and the sequential-large in grades 1 and 2, demonstrating a shift of roles from small numerosities to large numerosities with the start of formal schooling.

Interestingly, as in previous studies, we found no evidence for the nonsymbolic simultaneous-small measure uniquely predicting math achievement at any stage with either accuracy or RT data (Bartelet et al., 2014; Holloway & Ansari, 2009; Lyons et al., 2014; Sasanguie et al., 2011, 2014, 2013). Kindergarten performance in this task did correlate with distant math achievement in the latent growth model but in the regression models where domain-general capacities were controlled for, it demonstrated no effect. This result highlights that research in this domain should take a closer look at what exactly is being measured with the different task-formats. Perhaps other cognitive processes underlie children's performance in the simultaneous-small magnitude comparison task-format to a higher extent than the ANS.

As mentioned earlier, the two task-formats differed on the basis of both numerosity ranges and design characteristics. The existing literature suggests that different WM resources are employed when estimating different nonsymbolic numerosity ranges. For example, Xenidou-Dervou et al. (2014) demonstrated that kindergarteners' nonsymbolic numerosity processing with sequential steps and large quantities (6-70) necessitates CE processing, whereas the VSSP did not play a role. The authors concluded that for this task-format the nonsymbolic quantities are mentally represented as condensed whole arrays within the kindergarteners' CE. In another study, though, the same task-format but this time with smaller numerosities (1-9) was found to be predicted only by the kindergarteners' VSSP capacity (Xenidou-Dervou et al., in press). Xenidou-Dervou et al. (in press) suggest that smaller nonsymbolic quantities can be represented by the kindergarteners' readily accessible mental model in a one-by-one manner (Huttenlocher, Jordan, & Levine, 1994;

Rasmussen & Bisanz, 2005). Thus, it is plausible that the VSSP mediates the effect of the nonsymbolic simultaneous-small comparison task, which would explain why performance in this task did not play a unique role in predicting math achievement at any developmental stage beyond WM capacity. It should also be noted that nonsymbolic task-formats, which entail numerosities from 1 up to 9 also call upon subitizing processes (processes that allow instantaneously the exact determination of numerosities up to four) since many of their trials include one up to four numerosities (Revkin, Piazza, Izard, Cohen, & Dehaene, 2008).

In their review, however, De Smedt et al. (2013) point out that a distinction across the contradicting studies with respect to the role of the ANS could not be based solely on the numerosity ranges. So, perhaps the sequential versus simultaneous distinction plays an important role too in conjunction with the numerosity range. One may intuitively expect that the use of an occluder in the sequential tasks (see Figure 1) may impose more WM load than the simultaneous task-format. However, our results do not support that assumption as performance in this task-format explained unique variance above and beyond all the different WM capacities at each developmental stage. Perhaps the use of the occluder affects performance in a difference manner: it may force children to store the condensed whole array in their CE (Xenidou-Dervou et al., 2014) and extract a global magnitude code from this configuration in order to compare it with the second quantity, which becomes subsequently visible on the screen. In contrast, with the simultaneous presentation of the quantities, children are "invited" to look and compare the dots there and then, which may allow them to use visual search strategies with intermediate WM storage when looking back and forth between the two stimulus displays. Research has indicated that different visuospatial, phonological or transcoding processes may take place during numerosity processing depending on the familiarity with the task format, its level of difficulty and the stimulus used (Xenidou-Dervou et al., in press). Since the simultaneous-small tasks appeared to be easy for the children, it could be that they represented the nonsymbolic stimuli in their mental model in a one by one manner, or transcoded the nonsymbolic stimuli to the symbolic code, or by phonologically storing symbolic information as it has been evidenced in small-numerosity nonsymbolic and symbolic mental addition (Rasmussen & Bisanz, 2005; Xenidou-Dervou et al., in press). These are of course tentative assumptions and further research is necessary to identify the underlying mechanisms involved in the different task-formats and numerosity ranges. Nevertheless, our findings suggest that the sequential presentation of large numerosities taps the ANS or the symbolic system in a more direct manner compared to the simultaneous-small task-formats. They do so even above and beyond the effects of math-specific or domain-general WM capacities and IQ. We showed that the effect of simultaneous-small magnitude comparison was minimal having controlled for the children's WM capacity. Future research should address the aforementioned assumptions by examining the different strategies and WM components employed when executing nonsymbolic or symbolic magnitude comparison in the different task-formats.

Our findings also empirically verified, for the first time, the assumption introduced by De Smedt et al.'s review (2013) that accuracy and RT yield different patterns of results, which can explain the contradictory findings reported across the literature (De Smedt et al., 2013; Gilmore et al., 2014, 2011). The unique predictive role of the ANS became evident with accuracy - not speed - in both the latent growth models and the regression analyses. The unique predictive role of symbolic processing, however, was evident with both accuracy and RT data. In general, symbolic magnitude processing was a stronger predictor of children's future math achievement overall. (De Smedt et al., 2013). This was evident across both our latent growth models and the multiple regression analyses. Actually, our findings reveal that initially at the kindergarten level, symbolic *small-numerosity RT* (Figure 4 & Table 6) and symbolic *large-numerosity accuracy* (Figure 3 & Table 5) both play important roles. From *grade 1 and onwards*, however, focus should be placed on the improvement of children's both accuracy and speed in symbolic *large-numerosity* magnitude comparison (Tables 5 & 6). Noticeably, symbolic sequential-large magnitude accuracy and RT often demonstrated comparative regression coefficients to initial IQ, which highlights even more its importance as a building block for the development of general mathematics achievement (see Tables 5 & 6).

This study's findings bring forth important implications for educational assessment and practice. Firstly, we demonstrated that symbolic magnitude processing improves at a faster rate than nonsymbolic from kindergarten through grade 2. This indicates that age, education or experience modulate this type of processing more than nonsymbolic magnitude processing (the ANS). Also, children's individual developmental growth rates in this skill were found to be an important predictor of their future math achievement. Since education can modulate this skill, our results suggest that instruction should target its development in order for low performers to eventually catch up with children who demonstrate higher symbolic magnitude processing skills. Secondly, ANS performance appeared to also play an important role, although to a lesser extent and ANS screening seemed to be more effective with sequentially presented large quantities. Thirdly, we showed that at different developmental stages, different skills should be targeted. Our results suggested that before and at the start of school instruction, assessment and interventional designs should target both nonsymbolic and symbolic magnitude processing. After that though, focus should be placed on children's speed and accuracy in comparing digits. It should be noted that the relationship between development in the symbolic sequential-large task and math achievement must not imply that later math achievement is a result of this ability. It is also possible that the increase of numerical knowledge due to increased general math achievement may contribute to the development in this task. Nevertheless, as simple as it may seem to adults to compare digits, targeting the development of this skill from very early on appears to be of primary importance, as it appears to have comparative predictive effects of later mathematics achievement as domain-general capacities such as WM and IQ (Tables 5 and 6).

General Conclusions

The present study explains and reconciles contradictory findings in the literature regarding the roles of the ANS and symbolic magnitude processing during children's transition from kindergarten to official schooling. It provided proof for certain aspects of both existing theoretical classes in this domain and proposes an integrative view. The ANS, as an intuitive nonverbal, evolutionary system, plays a unique role in the development of general mathematics achievement until the first year of formal schooling. It relates to our ability to use symbols at these early developmental stages, however symbolic representations do not necessarily map only onto the ANS. Symbols are represented and manipulated in a different cognitive system. Until the first year of formal schooling, both the ANS and symbolic processing play their own unique role in explaining individual differences in later general math achievement. However, since by default math education focuses on using symbols to conduct complex math, from grade 2 and onwards only symbolic magnitude processing plays a unique role. With age, education or experience symbolic magnitude processing improves at a faster rate than nonsymbolic and children's individual growth rates in this skill significantly predicts their future general math achievement. Overall, symbolic magnitude processing arises as a consistent unique developmental predictor of children's mathematical achievement above and beyond WM capacities and IQ.

REFERENCES

Alloway, T. P. (2007). Automated Working Memory Assessment. London, UK: Pearson Assessment. (Translated and reproduced by permission of Pearson Assessment).

Alloway, T. P., Gathercole, S. E., Willis, C., & Adams, A. (2004). A structural analysis of working memory and related cognitive skills in young children, *87*, 85–106. doi:10.1016/j.jecp.2003.10.002

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews. Neuroscience*, *4*, 829–39. doi:10.1038/nrn1201

Baddeley, A. (2012). Working memory: theories, models, and controversies. *Annual Review of Psychology*, *63*, 1–29. doi:10.1146/annurev-psych-120710-100422

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*, *117C*, 12–28. doi:10.1016/j.jecp.2013.08.010

Barth, H., La Mont, K., Lipton, J., Dehaene, S., Kanwisher, N., & Spelke, E. (2006). Non-symbolic arithmetic in adults and young children. *Cognition*, *98*, 199–222. doi:10.1016/j.cognition.2004.09.011

Barth, H., La Mont, K., Lipton, J., & Spelke, E. S. (2005). Abstract number and arithmetic in preschool children. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 14116–21. doi:10.1073/pnas.0505512102

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*, 278–92. doi:10.1016/j.jecp.2010.09.003

De Smedt, B., Noël, 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*, 1–8. doi:10.1016/j.tine.2013.06.001

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*, 469–79. doi:10.1016/j.jecp.2009.01.010

DeStefano, D., & LeFevre, J. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology*, *16*, 353–386. doi:10.1080/09541440244000328

Dougherty, C. (2003). Numeracy, literacy and earnings: evidence from the National Longitudinal Survey of Youth. *Economics of Education Review*, *22*, 511–521. doi:10.1016/S0272-7757(03)00040-2

Duncam, T. E., & Duncam, S. C. (2009). The ABC's of LGM: An introductory guide to latent variable growth curve modeling. *Social & Personality Psychology Compass*, *3*, 979–991. doi:10.1111/j.1751-9004.2009.00224.x.The

Feigenson, L., Libertus, M. E., & Halberda, J. (2013). Links Between the Intuitive Sense of Number and Formal Mathematics Ability. *Child Development Perspectives*, *7*, 74–79. doi:10.1111/cdep.12019

Friso-van den Bos, I., Kroesbergen, E. H., & van Luit, J. E. H. (2014). Number sense in kindergarten children: Factor structure and working memory predictors. *Learning and Individual Differences*, *33*, 23–29. doi:10.1016/j.lindif.2014.05.003

Friso-van den Bos, I., van der Ven, S. H. G., Kroesbergen, E. H., & van Luit, J. E. H. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, *10*, 29–44. doi:10.1016/j.edurev.2013.05.003

Gilmore, C., Attridge, N., De Smedt, B., & Inglis, M. (2014). Measuring the approximate number system in children: Exploring the relationships among different tasks. *Learning and Individual Differences*, *29*, 50–58. doi:10.1016/j.lindif.2013.10.004

Gilmore, C., Attridge, N., & Inglis, M. (2011). Measuring the approximate number system. *The Quarterly Journal of Experimental Psychology*, *64*, 2099–2109.

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2007). Symbolic arithmetic knowledge without instruction. *Nature*, *447*(7144), 589–91. doi:10.1038/nature05850

Gilmore, C. K., McCarthy, S. E., & Spelke, E. S. (2010). Non-symbolic arithmetic abilities and mathematics achievement in the first year of formal schooling. *Cognition*, *115*, 394–406. doi:10.1016/j.cognition.2010.02.002

Gullick, M. M., Sprute, L. a., & Temple, E. (2011). Individual differences in working memory, nonverbal IQ, and mathematics achievement and brain mechanisms associated with symbolic and nonsymbolic number processing. *Learning and Individual Differences*, *21*, 644–654. doi:10.1016/j.lindif.2010.10.003

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*, 1457–65. doi:10.1037/a0012682

Holloway, I. D., & Ansari, D. (2008). Domain-specific and domain-general changes in children's development of number comparison. *Developmental Science*, *11*, 644–9. doi:10.1111/j.1467-7687.2008.00712.x

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. doi:10.1016/j.jecp.2008.04.001

Hornung, C., Schiltz, C., Brunner, M., & Martin, R. (2014). Predicting first-grade mathematics achievement: the contributions of domain-general cognitive abilities, nonverbal number sense, and early number competence. *Frontiers in Psychology*, *5*, 272. doi:10.3389/fpsyg.2014.00272

Huttenlocher, J., Jordan, N. C., & Levine, S. C. (1994). A mental model for early arithmetic. *Journal of Experimental Psychology. General*, *123*, 284–96.

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*, 1222–9. doi:10.3758/s13423-011-0154-1

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*, 1292–300. doi:10.1111/j.1467-7687.2011.01080.x

Lipton, J. S., & Spelke, E. S. (2005). Preschool children's mapping of number words to nonsymbolic numerosities. *Child Development*, *76*, 978–88. doi:10.1111/j.1467-8624.2005.00891.x

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*, 635–41. doi:10.1037/a0027248

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1-6. *Developmental Science*, 1–13. doi:10.1111/desc.12152

Mazzocco, M. M. M., Feigenson, L., & Halberda, J. (2011). Preschoolers ' Precision of the Approximate Number System Predicts Later School Mathematics Performance, *6*, 1–8. doi:10.1371/journal.pone.0023749

Mundy, E., & Gilmore, C. K. (2009). Children's mapping between symbolic and nonsymbolic representations of number. *Journal of Experimental Child Psychology*, *103*, 490–502. doi:10.1016/j.jecp.2009.02.003

Noël, M.-P., & Rousselle, 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. doi:10.3389/fnhum.2011.00165

Piazza, M., & Izard, V. (2009). How humans count: numerosity and the parietal cortex. *The Neuroscientist : A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *15*, 261–73. doi:10.1177/1073858409333073

Raghubar, K. P., Barnes, M. a., & Hecht, S. a. (2010). Working memory and mathematics: A review of developmental, individual difference, and cognitive approaches. *Learning and Individual Differences*, *20*, 110–122. doi:10.1016/j.lindif.2009.10.005

Rasmussen, C., & Bisanz, J. (2005). Representation and working memory in early arithmetic. *Journal of Experimental Child Psychology*, *91*, 137–157. doi:10.1016/j.jecp.2005.01.004

Repovs, G., & Baddeley, a. (2006). The multi-component model of working memory: explorations in experimental cognitive psychology. *Neuroscience*, *139*, 5–21. doi:10.1016/j.neuroscience.2005.12.061

Revkin, S. K., Piazza, M., Izard, V., Cohen, L., & Dehaene, S. (2008). Does subitizing reflect numerical estimation? *Psychological Science*, *19*, 607–14. doi:10.1111/j.1467-9280.2008.02130.x

Reyna, V. F., & Brainerd, C. J. (2007). The importance of mathematics in health and human judgment: Numeracy, risk communication, and medical decision making. *Learning and Individual Differences*, *17*, 147–159. doi:10.1016/j.lindif.2007.03.010

Sasanguie, D., De Smedt, B., Defever, E., & Reynvoet, B. (2011). Association between basic numerical abilities and mathematics achievement. *The British Journal of Developmental Psychology*, *30*, 344–57. doi:10.1111/j.2044-835X.2011.02048.x

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*, 37–41. doi:10.1080/17470218.2013.803581

Sasanguie, D., Göbel, S. M., Moll, K., Smets, K., & Reynvoet, B. (2013). Approximate number sense, symbolic number processing, or number-space mappings: What underlies mathematics achievement? *Journal of Experimental Child Psychology*, *114*, 418–31. doi:10.1016/j.jecp.2012.10.012

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. doi:10.1111/j.1751-228X.2012.01147.x

Sekuler, R., & Mierkiewicz, D. (1977a). Children's Judgments of Numerical Inequality. *Child Development*, *48*, 630. doi:10.2307/1128664

Simmons, F. R., Willis, C., & Adams, A.-M. (2012). Different components of working memory have different relationships with different mathematical skills. *Journal of Experimental Child Psychology*, *111*, 139–55. doi:10.1016/j.jecp.2011.08.011

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, 1–6. doi:10.1073/pnas.1302751110

Xenidou-Dervou, I., De Smedt, B., van der Schoot, M., & van Lieshout, E. C. D. M. (2013). Individual differences in kindergarten math achievement: The integrative roles of approximation skills and working memory. *Learning and Individual Differences*, *28*, 119–129. doi:10.1016/j.lindif.2013.09.012

Xenidou-Dervou, I., van Lieshout, E. C. D. M., & van der Schoot, M. (2014). Working memory in nonsymbolic approximate arithmetic processing: a dual-task study with preschoolers. *Cognitive Science*, *38*, 101–27. doi:10.1111/cogs.12053

Xenidou-Dervou, I., van der Schoot, M., & van Lieshout, E. C. D. M. (in press). Working Memory and Number Line Representations in Single-digit Addition: Approximate versus Exact, Nonsymbolic versus Symbolic. *The Quarterly Journal of Experimental Psychology*.