Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study

Annemie Desoete*, Annelies Ceulemans, Frauke De Weerdt and Stefanie Pieters
University Ghent, Belgium

Background. The ability to compare numbers, as the most basic form of number sense, has been related to arithmetical achievement.

Aims. The current study addressed the predictive value of non-symbolic and symbolic (number word (NW) and Arabic number (AN)) comparison for arithmetics by means of a longitudinal design.

Sample. Sixteen children with mathematical disabilities (MD), 64 low achievers (LA), and 315 typical achieving (TA) children were followed from kindergarten till grade 2.

Method. The association of comparison skills with arithmetical skills in grades 1 and 2 was studied. The performances of MD, LA and TA children were compared.

Results. Regression analyses showed that non-symbolic skills in kindergarten were predictively related to arithmetical achievement 1 year later and fact retrieval 2 years later. AN comparison was predictively related to procedural calculation 2 years later. In grade 2, there was an association between both symbolic tasks and arithmetical achievement. Children with MD already had deficits in non-symbolic and symbolic AN comparison in kindergarten, whereas in grade 2 the deficits in processing symbolic information remained.

Conclusions. The combination of non-symbolic and symbolic deficits represents a risk of developing MD.

Early numeracy
The past decade, individual differences in early numeracy, and in foundations of arithmetic skills have been receiving growing attention (e.g., Dowker, 2008; Durand, Hulme, Larkin, & Snowling, 2005; Koponen, Aunola, Ahonen, & Nurmi, 2007; Krajewski

*Correspondence should be addressed to Annemie Desoete, Dunantlaan 2, B-9000 Ghent, Belgium (e-mail: annemie.desoete@UGent.be).

DOI:10.1348/2044-8279.002002
The current interest in early predictors of MD is encouraged by the hope that, if predictors, determinants, and core deficits can be addressed as key components in remediation programmes, children may not fall further behind (e.g., DiPerna, Lei, & Reid, 2007; Dowker & Sigley, 2010; Gersten, Jordan, & Flojo, 2005) and avoid math or even develop math anxieties (Ashcraft & Moore, 2009).

Until now, research on individual differences in arithmetic has focused on domain-general cognitive functions such as working-memory or executive functions (e.g., Bull, Espy, & Wiebe, 2008; Noel, 2009; Passolunghi & Cornoldi, 2008; Swanson & Kim, 2007) and fluency or processing speed (e.g., Bull & Johnston, 1997; Hecht, Torgersen, Wagner, & Rashotte, 2001). In addition, domain-specific research on early numeracy has focused mainly on the role of Piagetian logical abilities (e.g., Nunes et al., 2007; Stock, Desoete, & Roeyers, 2009a) and on counting knowledge and skills in young children (e.g., Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Gersten, et al., 2005; Hannula, Rässänen, & Lehtinen, 2007; Stock, Desoete, & Roeyers, 2009b). Those studies have shown that applying counting principles is one of the best predictors of arithmetical achievement in first grade, although seriation and classification were also found to be important preparatory arithmetic abilities for the development of proficient arithmetic performance (e.g., Grégoire, 2005; Stock, Desoete, & Roeyers, 2010).

There are several arguments for the claim that number sense growth and trajectories (Berch, 2005; Jordan, Mulhern, & Wylie, 2009) or ‘number magnitude representation’ should be considered as one of the key precursors of arithmetical development, with deficits leading to mathematical disabilities (MD, e.g., Moeller, Neuburger, Kaufmann, Landerl, & Nuerk, 2009; Piazza et al., 2010). In addition, according to the triple code model (Dehaene & Cohen, 1995, 1997; Noel, 2001; Schmithorst & Brown, 2004) there are three types of representations for numbers. Two of them are symbolic and format-dependent: a visual Arabic number (AN) form (e.g., ‘5’) and a verbal word frame with number words (NWs) (e.g., ‘five’), and one is non-symbolic and format-independent: the analogue magnitude representation (e.g., five dots).

The number of studies in the area of ‘magnitude representation’ is growing rapidly (e.g., De Smedt, Verschaffel, & Ghesquière, 2009). However often studies are cross-sectional in nature (e.g., Berteletti, Lucangeli, Piazza, Dehaene, & Zorzi, 2010; Holloway & Ansari, 2009), making predictions on individual differences in arithmetics difficult to make. In addition, in most studies the focus lies on non-symbolic magnitude representation, sometimes in combination with the symbolic representation with ANs (e.g., Mussolin, Mejias & Noel 2010). On basis of such data, it is often unclear whether it is the AN or NWs processing that is important for arithmetic development. Finally, in clinical studies often the control children have no learning disability history (e.g., Mussolin et al., 2009; Mussolin et al., 2010; Piazza et al., 2010) or are age-matched and normally developing children (e.g., Moeller et al., 2009), so the answer on whether children with MD represent a specific and definable impairment or the lower end of the continuum of arithmetical ability cannot be given from such a design. Therefore, the current investigation tried to extend the available studies by means of a longitudinal design that examined the predictive association between non-symbolic and symbolic (AN and NW) comparison before formal school (i.e., in kindergarten) and arithmetic achievement 1 and 2 years later in typical achieving (TA) children, low achievers (LA), and children with MD.
In what follows in this introduction, we present some arguments for the association between non-symbolic and symbolic comparison skills with arithmetical achievement, we propose a definition on mathematical learning disability and summarize what new information this study will provide.

There are developmental, behavioral, and neuro-imaging arguments for the claim that comparison skills are associated with later arithmetic skills. First of all, there is developmental evidence for number sense even in infants, allowing them to see the difference between two sets of items (Berteletti et al., 2010; Mack, 2006; Xu & Arriaga, 2007). This number sense allows children later on, to check the plausibility of their answers on ‘simple’ procedural (P) calculation tasks in number-problem format (e.g., $12 - 6 = \ldots$). In addition, a good level of number sense is important for more complex calculation exercises. The success in complex calculation tasks depends on the knowledge (K) of base-10 structure relationships (e.g., 47 is composed by 4 decades and 7 units), the translation (L) of words into calculation procedures (e.g., ‘9 less than 47 is \ldots’) and the mental (M) representation of problems to prevent ‘blind calculation’ (e.g., 38 is not the answer to ‘47 is 9 less than \ldots’ although one might translate ‘less’ into ‘subtraction’). For more details on the P, K, L, and M-tasks, see Desoete and Roeyers (2005). Moreover, Booth and Siegler (2006) revealed developmental changes on estimation tasks related to individual differences in arithmetic achievement. Finally, number sense is also needed for fact retrieval, since a good understanding of numerical magnitude narrows the range of candidate answers when problems are presented that can be solved by retrieving the answer from semantic memory. Thus, a variety of studies converge to show the crucial role of number sense for procedural calculation and number-fact retrieval (e.g., Barth et al., 2006; Booth & Siegler, 2008; Halberda, Mazzocco, & Feigenson, 2008; Holloway & Ansari, 2009).

Secondly, there is behavioural evidence of problems of children with MD as results of a more imprecise representation of number magnitude (e.g., Mussolin, Meijas, & Noel, 2010; Piazza et al., 2010; Von Aster & Shalev, 2007). Deficits in number sense and quantity-number competencies were found in elementary school children diagnosed with MD (Geary & Hoard, 2005; Geary, Hoard, Byrd-Craven, & DeSoto, 2004). Butterworth and co-workers (Landerl, Bevan, & Butterworth, 2004) explained this deficit with their ‘defective number module’ hypothesis, assuming that MD occur when the basic ability to process numerosity fails to develop normally, resulting in difficulties to understand number concepts and, consequently, in learning numerical information. According to those authors, MD children have a deficit in number sense per se. Consistent with this defective number module hypothesis, Jordan, Hanich, and Kaplan (2003) provided evidence that MD affects also tasks requiring estimating the approximate result of arithmetic problems or showing the quantities standing for the units and the tens in two-digit numbers (Jordan et al., 2003). Nonetheless, Rousselle and Noël (2007) evaluated an alternative explanation with the ‘access deficit hypothesis’ stating that there was no deficit in number sense in SE, since when investigating numerosity processing with no symbolic processing requirement MD children in second grade were only impaired when comparing Arabic numerals (i.e., symbolic number magnitude) but not when comparing collections of sticks (i.e., non-symbolic number magnitude). The authors suggested that children with MD had difficulty in accessing number magnitude from symbols rather than in processing numerosity per se.

Thirdly, neuro-imaging studies have shown that the intra-parietal sulcus which is dedicated to the processing of magnitudes appears to be active during arithmetical tasks.
(Ansara, 2008; Dehaene, Piazza, Pinel, & Cohen, 2003; Kadosh, Lammertyn, & Izard, 2008). Moreover, MD participants exhibited both structural and functional differences in the cerebral areas involved in the processing of this number magnitude (Molko et al., 2003; Mussolin et al., 2010; Price, Holloway, Rasanen, Vesterinen, & Ansari, 2007; Rubinstein & Henik, 2005; Rotzer et al., 2008).

Mathematical disabilities
Despite the growing interest observed over the last few years, research on MD is actually much less advanced than on dyslexia (Grégoire & Desoete, 2009; Rousselle & Noël, 2007). In addition, there remain some difficulties in defining MD (e.g., Mazzocco & Myers, 2003; Murphy, Mazzocco, Hanich, & Early, 2007).

The term mathematical learning disability (MD) refers to a significant degree of impairment in the arithmetical skills (with substantially below performances). In addition, children do not profit from (good) help. This is also referred to as a lack of responsiveness to intervention (RTI, Fuchs et al., 2007; Kavale & Spaulding, 2008). Finally, the problems in MD cannot be totally explained by impairments in general intelligence or external factors that could provide sufficient evidence for scholastic failure.

Most practitioners and researchers currently report a prevalence of MDs between 3— and 14% of the school-age population depending on the country of study (Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2005; Dowker, 2005; Shalev, Manor, & Gross-Tsur, 2005). In addition, some authors propose at least a procedural and a semantic memory sub-type within MD (Geary, 1993; 2004; Robinson, Menchetti, & Rogensen, 2002; Temple, 1999). The procedural sub-type would be due to executive dysfunction and characterized by a developmental delay in the acquisition of counting and counting procedures used to solve simple arithmetic problems. The semantic memory sub-type would be due to verbal memory dysfunction and characterized by errors in the retrieval of arithmetic facts (Wilson, Revkin, Cohen, Cohen, & Dehaene, 2006). However, not all studies have found different profiles for these groups (Landerl et al., 2004, Rousselle & Noël, 2007). Moreover, although the criteria for MD seem clear, there are some disagreements on for example the criteria used to define the ‘substantially below’ performances (Geary, 2004; Mazzocco & Myers, 2003). In addition, performances will fluctuate around a cut-point needing repeated testing (Fletcher, Denton, & Francis, 2005; Geary, 2004; Hanley, 2005; Stock et al., 2010). Moreover, there is some disagreement as to whether MD represents a specific and definable impairment or the lower end of the continuum of arithmetical ability. Mazzocco, Devlin, and McKenney (2008) found that children with MD (and a severe form of disability) showed qualitatively different profiles in fact retrieval performances when compared to TA children, whereas the differences between children at the lower end of the continuum (LA with a mild form of disability) and TA children were of a quantitative turn. Geary, Hoard, Byrd-Craven, Nugent, and Numtee (2007) revealed that children with MD (a severe disability) had a severe math cognition deficit and underlying deficit in working memory and speed of processing. The LA groups (with a mild disability) had more subtle deficits in few math domains. Finally, although the criterion of non-RTI (Fuchs, Fuchs, & Prentice, 2004) is an interesting one, some studies suggest that even quite significant arithmetical difficulties are often responsive to interventions targetted at their specific strengths and weaknesses (Baker, Gersten, & Lee, 2002; Dowker & Sigley, 2010; Miller, Butler, & Lee, 1998; Montague, 2008).
Objectives and research questions
In this study, we aim to examine the predictive value of symbolic and non-symbolic comparison skills for individual differences in arithmetical achievement. Within the symbolic comparison, we aim to compare the contribution of ANs and NWs as kindergarten predictor for procedural calculation and numerical facility in grades 1 and 2. Within procedural calculation we investigate differences between the processing of simple and more complex calculation tasks.

In addition, we aim to look for development shifts, as suggested by Booth and Siegler (2006), by analysing the comparison skills in kindergarten (or before the start of formal schooling) as well as in grade 2 (2 years later).

Moreover, the purpose of the current study is to look for specificity and examine kindergarten differences between children at the lower end of the continuum of arithmetical ability or investigate whether children with MD differ from LA. We tested if non-symbolic and symbolic number comparison tasks differentiate MD from LA children and if those tasks can be used as early screeners to identify children with MD.

Finally, it is studied if our data are in line with the ‘defective number’ (Landerl et al., 2004) or ‘access deficit’ (Rousselle & Noël, 2007) hypothesis. According to the defective number module hypothesis we could expect MD children to have problems with all comparison tasks. According to the access deficit hypothesis MD children are supposed to have problems with the symbolic (NW and AN) tasks but not with the non-symbolic comparison tasks.

Method
Participants
This study was carried out in a total group of 395 children (196 boys and 199 girls). All children were Caucasian native Dutch-speaking children living in the Flemish part of Belgium. Three groups of children participated in this study, based on an assessment and consistent achievement on at least two testing points.

Children were retrospectively classified as having MD if they had disabilities non-responsive to remediation and if they scored ≤10th percentile on at least one of the arithmetic achievement tests used to assess procedural or semantic memory disabilities, both in first and second grades (n = 10 boys and 6 girls).

Children, who scored between the 10th and 25th percentiles on at least one of the arithmetical tests, both in first and second grades, were classified as children at the lower end of the continuum of arithmetical ability or as LA (n = 35 boys and 29 girls). Both MD and LA group had a diagnose confirmed by the school psychologist.

The third group consisted of children who scored greater than the 25th percentile on all arithmetic achievement tests in both grades, these children were classified as TA (n = 151 boys and 164 girls).

No significant differences in intelligence were found between the three groups of participants (F(2, 379) = 1.64, p = ns) with a mean IQ of 101.16 (SD = 13.21). In addition no significant differences in socio-economic status derived from the total number of years of parents’ education starting from the beginning of elementary school was found between the AD, LA, and TA groups (Wilks’ lambda = 0.98, F(4, 732) = 1.36, p = ns), with M = 14.96 (SD = 2.40) as mean number of years in education for the mothers and M = 14.56 (SD = 2.88) as mean number of years in education for fathers.
**Materials**

All children were tested in kindergarten (age 5–6) and in grade 2 (age 7–8) on their non-symbolic and symbolic (NW and AN) comparison skills. Moreover, follow-up assessment with two arithmetic tests was conducted in first and second grades and intellectual abilities were tested in second grade.

**Symbolic and non-symbolic comparison skills in kindergarten**

The symbolic and non-symbolic comparison skills were tested with different sub-tests of the TEDI-MATH (Grégoire, Noël, & Van Nieuwenhoven, 2004). The TEDI-MATH has been used (e.g., Wilson et al., 2006) and tested for conceptual accuracy and clinical relevance in previous studies (e.g., Desoete & Grégoire, 2007; Stock et al., 2009b, 2010). The psychometric value was demonstrated on a sample of 550 Dutch-speaking Belgian children from the second year of kindergarten to the third grade of primary school. The TEDI-MATH has proven to be a well validated (Desoete, 2007) and reliable instrument, values for Cronbach’s alpha for the different sub-tests vary between 0.70 and 0.97 (Grégoire et al., 2004).

Non-symbolic magnitude comparison was assessed by comparing a collection of dots. Children were asked where they saw most dots. One point was given for a correct answer. The raw score was converted into a Z-score. The internal consistency of this task was good (Cronbach’s alpha = 0.79).

Symbolic verbal NW comparison was assessed by three kinds of tasks. In the first tasks, children have to judge if a spoken verbal numeral is a NW. In the second tasks, children have to judge if a NW is syntactically correct. In a third task, children have to judge which of two spoken verbal numbers is the larger one. The raw score was converted into a Z-score. The internal consistency of this task was good (Cronbach’s alpha = 0.85).

Symbolic AN comparison was assessed by two kinds of tasks. In the first tasks, children have to judge if a written Arabic symbol is a number. In the second tasks, children have to judge which of two written ANs the larger one is. The raw score was converted into a Z-score. The internal consistency of this task was good (Cronbach’s alpha = 0.87).

**Arithmetical tests in first and second grades**

In order to obtain a complete overview of the arithmetic abilities of children in first and second grades and to test for procedural and semantic memory deficits, two arithmetic tests were used: the revised Kortrijk arithmetic test (Kortrijk Arithmetic Test, KRT-R, Baudonck et al., 2006) and the arithmetic number facts test (tempo test Rekenen, TTR, De Vos, 1992).

The Kortrijk aRithmetic Test revision (KRT-R; Baudonck et al., 2006) is an untimed standardized test on procedural calculations. KRT-R requires that children solve 30 simple calculations (P-tasks) in a number-problem format (e.g., $16 - 12 = \ldots$), and 30 more complex calculations (L, K, C, or M-tasks) often in a word-problem format (e.g., 1 less than 8 is \ldots) in first grade. Children in second grade receive 30 simple calculations (P-tasks) in a number-problem format (e.g., $39 + 60 = \ldots$) and 25 more complex calculations (L, K, C, or M-tasks) often in a word-problem format (e.g., 6 more than 48 is \ldots). The KRT-R results in a score on simple procedural calculations (P-tasks) and a score on complex procedural calculations (L,K, C, and M-tasks). All scores were converted into Z-scores.
KRT-R can be used to test procedural disabilities. The psychometric value of the test has been demonstrated on a sample of 3,246 children. A validity coefficient (correlation with school results) and reliability coefficient (Cronbach’s alpha) of .64 and .94, respectively, were found for second grade.

The arithmetic number facts test (TTR; De Vos, 1992) is a ‘timed’ test consisting of 80 (first grade) or 200 (second grade) arithmetic number fact problems. In first grade, children have to solve as many additions (e.g., \( 5 + 2 = \ldots \)) and subtractions (e.g., \( 6 - 5 = \ldots \)) in 2 min, children in the second grade are presented the same additions and subtractions but also divisions (e.g., \( 2 \times 8 = \ldots \)) and multiplications (e.g., \( 16 \div 4 = \ldots \)) and have 5 min to solve as many items as possible. The TTR is a standardized test that is frequently used in Flemish education as a measure of number-fact retrieval. TTR can be used to assess semantic memory disabilities. The total number of correct items was used as \( Z \)-score for the analyses. The psychometric value of the test has been demonstrated on a sample of 10,059 children in total. Cronbach’s alphas computed for the current study was 0.90. The Guttman split-half coefficient was 0.93; the Spearman–Brown coefficient was 0.95.

**Intelligence**

In order to have an estimation of the intellectual capacities of the child, a short version of the Wechsler Intelligence Scale for Children, third edition (Wechsler, 1991 – WISC-III) was assessed. This is the most recent form in Flanders at that moment. The short version was based on four sub-tests and included both measures for crystallized and fluid intelligence (vocabulary, similarities, block design and picture arrangement; Grégoire, 2001).

**Procedure**

The children were recruited in regular schools. Parents received a letter with the explanation of the research and submitted informed consent in order to participate every year.

Toddlers were tested with TEDI-MATH (Grégoire et al., 2004) in a separate and quiet room. In first and second grade children were tested with TTR (De Vos, 1992) and KRT-R (Baudonck et al., 2006). In addition, children in grade 2 were tested with a short version of the WISC III (Wechsler, 1991) and TEDI-MATH.

The test leaders all received training in the assessment and interpretation of the tests. After completion of the test procedure, all the parents of the children received individual feedback on their children’s results.

**Results**

**Association of preschool measures and tests in grades 1 and 2**

The correlations, controlled for intelligence, between the non-symbolic and symbolic comparison skills in kindergarten and the arithmetical abilities in grades 1 and 2, are presented in Table 1.

There was a very limited relationship between the symbolic and non-symbolic comparison skills in kindergarten. Moreover, the correlations between the performances in kindergarten and the results 2 years later were significant, but very low.
Table 1. Association of preschool measures and grade 1 and 2 tests

<table>
<thead>
<tr>
<th>Kindergarten</th>
<th>Grade 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-symbolic</td>
<td>NW</td>
</tr>
<tr>
<td>Kindergarten</td>
<td></td>
</tr>
<tr>
<td>Non-symbolic</td>
<td>/</td>
</tr>
<tr>
<td>Symbolic NW</td>
<td>0.03</td>
</tr>
<tr>
<td>Symbolic AN</td>
<td>0.06</td>
</tr>
<tr>
<td>Grade 1</td>
<td></td>
</tr>
<tr>
<td>Simple calculation</td>
<td>0.16**</td>
</tr>
<tr>
<td>Complex calculation</td>
<td>0.16**</td>
</tr>
<tr>
<td>Fact retrieval</td>
<td>0.20**</td>
</tr>
<tr>
<td>Grade 2</td>
<td></td>
</tr>
<tr>
<td>Simple calculation</td>
<td>0.01</td>
</tr>
<tr>
<td>Complex calculation</td>
<td>0.02</td>
</tr>
<tr>
<td>Fact retrieval</td>
<td>0.16**</td>
</tr>
</tbody>
</table>

Note. NW, number word; AN, Arabic number. *p ≤ .05; **p ≤ .001.

Prospective prediction from kindergarten to grade 1

Since all variables were normally distributed and did meet the assumptions for multiple regressions, regression analyses were conducted in the sample to evaluate how well the kindergarten abilities predicted procedural calculation in number-problem and word-problem format and numerical facility in grades 1 and 2. Three kindergarten number comparison abilities at ages 5–6 were included simultaneously as predictor variables: comparison of symbolic Arabic numerals, comparison of spoken verbal numerals as symbols and non-symbolic number magnitude comparison. The univariate $F$-tests were Bonferroni-adjusted to control for the number of comparisons.

The linear combination of the kindergarten abilities was significantly related to simple calculations in number-problem format ($F(3, 389) = 3.272, p ≤ .05, R^2 = .03$), complex calculations in word-problem format ($F(3, 389) = 6.159, p ≤ .0005, R^2 = .05$), and to number-fact retrieval ($F(3, 389) = 6.366, p ≤ .0005, R^2 = .05$) in grade 1.

Non-symbolic number magnitude comparison in kindergarten was associated (see Tables 2 and 3) with individual arithmetical performances in grade 1 (at ages 6–7).

Table 2. Prospective predictors for procedural calculation in grade 1

<table>
<thead>
<tr>
<th>Kindergarten (ages 5–6)</th>
<th>Complex calculations grade 1 (ages 6–7)</th>
<th>Simple calculations grade 1 (ages 6–7)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstand. coeff.</td>
<td>$\beta$</td>
</tr>
<tr>
<td>Constant</td>
<td>63.185</td>
<td>47.198</td>
</tr>
<tr>
<td>Arabic numbers</td>
<td>1.673</td>
<td>0.062</td>
</tr>
<tr>
<td>Number-words</td>
<td>3.217</td>
<td>0.119</td>
</tr>
<tr>
<td>Magnitude comparison</td>
<td>4.169</td>
<td>0.154</td>
</tr>
</tbody>
</table>

*p ≤ .01 after Bonferroni adjustment.
Table 3. Prospective predictors for procedural calculation in grade 2

<table>
<thead>
<tr>
<th>Kindergarten (ages 5–6)</th>
<th>Complex calculations grade 2 (ages 7–8)</th>
<th>Simple calculations grade 2 (ages 7–8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstand. coeff.</td>
<td>β</td>
</tr>
<tr>
<td>Constant</td>
<td>0.110</td>
<td>2.443</td>
</tr>
<tr>
<td>Arabic numbers</td>
<td>0.341</td>
<td>0.357</td>
</tr>
<tr>
<td>Number-words</td>
<td>0.054</td>
<td>0.056</td>
</tr>
<tr>
<td>Magnitude comparison</td>
<td>−0.038</td>
<td>−0.040</td>
</tr>
</tbody>
</table>

*p ≤ .01 after Bonferroni adjustment.

Prospective prediction from kindergarten to grade 2

The linear combination of the kindergarten abilities was significantly associated to simple calculations ($F(3, 393) = 12.114, p ≤ .0005, R^2 = 0.09$) and to complex calculations ($F(3, 393) = 20.303, p ≤ .0005, R^2 = 0.13$) assessed in grade 2 (at ages 7–8).

Symbolic Arabic numeral comparison in kindergarten was associated with procedural calculation skills in grade 2 (see Table 3).

Kindergarten abilities were also significantly predictively associated with number-fact retrieval in grade 2 (at ages 6–7), $F(3, 387) = 4.737, p ≤ .005$. $R^2$ was 0.04. Especially non-symbolic comparison was beneficial for semantic fact retrieval in grade 2 (see Table 4).

Concurrent predictions within grade 2

The linear combination of the magnitude, AN and NW comparison skills in grade 2 was significantly related to simple calculations ($F(3, 379) = 33.504, p ≤ .0005, R^2 = 0.21$), complex calculations ($F(3, 379) = 37.876, p ≤ .0005, R^2 = 0.23$), and to number-fact retrieval ($F(3, 377) = 14.227, p ≤ .0005, R^2 = 0.10$) tested at the same moment.

Both symbolic (Arabic numeral as well as verbal NW) comparison skills were associated with arithmetic achievement in grade 2 (see Table 5).

Table 4. Prospective predictions for fact retrieval in grades 1 and 2

<table>
<thead>
<tr>
<th>Kindergarten (age 5–6)</th>
<th>Fact retrieval grade 1 (ages 6–7)</th>
<th>Fact retrieval grade 2 (ages 7–8)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstand. coeff.</td>
<td>β</td>
</tr>
<tr>
<td>Constant</td>
<td>0.214</td>
<td>4.552</td>
</tr>
<tr>
<td>Arabic numbers</td>
<td>0.102</td>
<td>.107</td>
</tr>
<tr>
<td>Number-words</td>
<td>0.018</td>
<td>.019</td>
</tr>
<tr>
<td>Magnitude comparison</td>
<td>0.171</td>
<td>.181</td>
</tr>
</tbody>
</table>

*p ≤ .01 after Bonferroni adjustment.
Mathematical learning disabilities in kindergarten

Table 5. Concurrent predictors for procedural calculation in grade 2

<table>
<thead>
<tr>
<th>Comparison skills</th>
<th>Complex calculations</th>
<th>Simple calculations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unstand. coeff. β T p</td>
<td>Unstand. coeff. β T p</td>
</tr>
<tr>
<td>Constant</td>
<td>0.064 1.524 .128</td>
<td>0.042 1.004 .316</td>
</tr>
<tr>
<td>Arabic numbers</td>
<td>0.141 0.150 2.869 .004*</td>
<td>0.181 0.200 3.774 .000*</td>
</tr>
<tr>
<td>Number-words</td>
<td>0.364 0.389 7.398 .000*</td>
<td>0.292 0.322 6.051 .000*</td>
</tr>
<tr>
<td>Magnitude comparison</td>
<td>0.001 0.001 0.019 .985</td>
<td>0.010 0.011 0.237 .813</td>
</tr>
</tbody>
</table>

*p ≤ .01 after Bonferroni adjustment.

Group differences in MD, LA, TA children

Differences in kindergarten

A multi-variate analysis of variance (MANOVA) was conducted to investigate kindergarten differences between the children with a MD, LA and TA peers on three dependent variables: AN comparison, verbal number comparison, and magnitude comparison (assessed in kindergarten).

The MANOVA was significant on the multi-variate level, Wilks' lambda = 0.894, F (6, 778) = 7.434, p ≤ .0005, partial η² = 0.05. The means and standard deviations of the dependent variables for the three groups are shown in Table 6.

As can be concluded from Table 6, post hoc follow-up analyses (see indexes in Table 6) revealed that TA and LA performers were better than MD performers on the comparison of ANs in kindergarten. No significant differences were found between MD, TA, and LA on the comparison of NWs in kindergarten. All three performance groups also differed on magnitude comparison tasks in kindergarten. TA problem solvers were better than LA problem solvers and LA problem solvers did better than MD problem solvers on magnitude comparison tasks in kindergarten.

Differences in grade 2

A second MANOVA was conducted to investigate differences between the MD, LA, and TA groups on three dependent variables: AN comparison, verbal number comparison, and magnitude comparison (assessed in grade 2).

Table 6. Kindergarten and grade 2 skills for the three groups of achievers

<table>
<thead>
<tr>
<th>Kindergarten skills</th>
<th>MD M (SD)</th>
<th>LA M (SD)</th>
<th>TA M (SD)</th>
<th>F (2, 391)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic numbers</td>
<td>−1.02b (1.03)</td>
<td>−0.16a (1.06)</td>
<td>0.08a (0.93)</td>
<td>10.2465**</td>
</tr>
<tr>
<td>Number-words</td>
<td>−0.32 (0.98)</td>
<td>−0.12 (0.90)</td>
<td>0.05 (1.01)</td>
<td>1.644</td>
</tr>
<tr>
<td>Magnitudes</td>
<td>−0.97b (1.61)</td>
<td>−0.22b (1.29)</td>
<td>0.09a (0.85)</td>
<td>10.558**</td>
</tr>
<tr>
<td>Grade 2 skills</td>
<td></td>
<td></td>
<td></td>
<td>F (2, 377)</td>
</tr>
<tr>
<td>Arabic numbers</td>
<td>−0.45b (0.72)</td>
<td>−0.34 (0.89)</td>
<td>0.09a (1.01)</td>
<td>6.721*</td>
</tr>
<tr>
<td>Number-words</td>
<td>−0.30b (1.03)</td>
<td>−0.29 (0.89)</td>
<td>0.08a (1.00)</td>
<td>4.345*</td>
</tr>
<tr>
<td>Magnitudes</td>
<td>−0.42 (1.04)</td>
<td>−0.07 (1.01)</td>
<td>0.04 (0.99)</td>
<td>1.735</td>
</tr>
</tbody>
</table>

Note. MD, mathematical disabilities group; LA, low achieving group; TA, typically achieving group; *p < .01, **p ≤ .0005 (abc) post hoc indexes p < .05.
Table 7. Concurrent predictors for number-fact retrieval in grade 2

<table>
<thead>
<tr>
<th></th>
<th>Unstand. coeff.</th>
<th>$\beta$</th>
<th>$t$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.045</td>
<td>1.045</td>
<td>.297</td>
<td></td>
</tr>
<tr>
<td>Arabic numbers</td>
<td>0.184</td>
<td>0.211</td>
<td>3.719</td>
<td>.000*</td>
</tr>
<tr>
<td>Number-words</td>
<td>0.123</td>
<td>0.141</td>
<td>2.464</td>
<td>.014*</td>
</tr>
<tr>
<td>Magnitude comparison</td>
<td>0.037</td>
<td>0.043</td>
<td>0.841</td>
<td>.401</td>
</tr>
</tbody>
</table>

*p ≤ .01 after Bonferroni adjustment.

The MANOVA was significant on the multi-variate level, Wilks' lambda = 0.957, $F(6, 750) = 2.808, p ≤ .01$, partial $\eta^2 = 0.02$. The means and standard deviations of the dependent variables for the three groups are shown in Table 6.

As can be concluded from Table 6, post hoc follow-up analyse (see indexes in Table 6) revealed that TA performers were better than MD performers on the comparison of ANs and NWs in grade 2. No significant differences were found between LA children and the TA or MD children on the comparison of magnitudes in grade 2.

Discussion

Several cognitive skills have been suggested as key precursors for arithmetical achievement and eventually as early markers for MD (e.g., Mussolin et al., 2010; Piazza et al., 2010; Stock et al., 2010). Nevertheless, relatively few studies have examined the predictive value of symbolic and non-symbolic (NW and AN) comparison together for individual differences in specific arithmetical abilities (namely simple and complex procedural calculation and numerical facility) by means of a longitudinal design, including typically and non-typically developing children. Moreover, from developmental perspective (Booth & Siegler, 2006), it was studied if the impact of symbolic and non-symbolic comparison changes from kindergarten to grade 2. Finally, differences were examined between children with MD, LA, and TA groups.

Regression analyses showed that non-symbolic skills in kindergarten were predictively related to arithmetical achievement 1 year later and to fact retrieval 2 years later. AN comparison skills were predictively related to procedural calculation 2 years later. In grade 2, there was a concurrent association between both symbolic tasks and arithmetical achievement. Easy and complex calculation tasks seem to be elaborated in the same manner. The assessment of non-symbolic comparison in grade 2 did not provide additional longitudinal information on procedural calculation and fact retrieval and a shorter test (without such tasks) may be administered reducing costs to administer and improve scores because participants are less fatigued.

Our kindergarten findings might indicate a developmental shift from depending on a non-symbolic approximate representation of magnitude in grade 1 to a more precise and complex symbolic representation in grade 2. The understanding of approximate magnitudes might aid children’s early arithmetic development in grade 1, when dealing with calculations up till 20. However, a more precise visual AN representation seems associated with multi-digit calculation procedures up till 100 and insight in base-10 structure relationships in grade 2. Moreover, the analyses in grade 2 reveal that both
non-symbolic representations for numbers are associated with arithmetical achievement. This could indicate that, in addition to the visual AN representation (needed for multi-digit calculation), also the verbal word frame is associated with procedural calculation (depending on stored addition and multiplication tables).

When addressing the question of non-typically developing children, our results do not validate the ‘defective module hypothesis’ (Butterworth, 2005) nor the ‘access deficit hypothesis’ (Rousselle & Noël, 2007). Part of the results are in line with Butterworth (2005) because the MD children in our sample already had difficulties in the non-symbolic comparison tasks in kindergarten. However, the present findings also indicate in line with Rousselle and Noël (2007) that older children with MD (in grade 2) no longer significantly differed from LA and TA peers on accuracy in non-symbolic comparison tasks but that they only significantly differed on symbolic comparison accuracy. Moreover, there seems to be a developmental shift leading to individual differences on symbolic tasks in kindergarten and on non-symbolic tasks in grade 2. However, it should be noted that we tested untimed accuracy or precision and not the fluency or speed as measure of children’s understanding of numerical magnitude. Fluency might be associated with a reduction of working memory load when doing arithmetic. Durand et al. (2005) revealed that the general speed of comparing numbers accounted for unique variance in individual differences in mathematics achievement. Perhaps if we looked at fluency (instead of accuracy), the differences on NWs in kindergarten and magnitude comparison in grade 2 between MD and TA would also be significant.

Our findings also reveal that MD should be considered as a specific and definable impairment and not the lower end of a continuum of arithmetical ability. In line with Geary et al. (2007), Mazzocco et al. (2008) and Stock et al. (2010) children with MD and children who were low achieving on arithmetic tests had different profiles. ‘Children with MD’ on the one hand already had significantly deficits in accuracy on non-symbolic and symbolic AN comparison tasks in kindergarten. In grade 2 they still had a deficit in accuracy on both symbolic comparison tasks. ‘LA’ on the other hand had a mild problem on non-symbolic magnitude comparison tasks but no lack of accuracy on symbolic comparison tasks in kindergarten. In grade 2 no significant differences were found between low achieving and TA children on any of the comparison tasks.

These results should be interpreted with care, since there are several limitations to the present study. Firstly as already mentioned, we only tested accuracy in number comparison skills. Additional research is needed on the fluency with which magnitude information is available. This should be done by timed tasks (as did De Smedt et al., 2009; Geary, Bailey, & Hoard, 2009) to capture how quick children decide which of two dot sets, ANs or NWs is larger. Secondly, we did not differentiate between quantities in our tasks. However, according to Weber’s Law the representation becomes increasingly imprecise as numbers get larger (Noel, 2001). Moreover, in a recent study Mussolin et al. (2010) revealed that children with MD had especially higher error rates when discriminating close numerical quantities and they were more sensitive than controls to continuous dimensions such as surface area or density. Future studies should examine these relationships more in detail. Such studies are currently being conducted in infants (Ceulemans, Desoete, & Roeyers, 2010). Thirdly the results of the current study need to be interpreted with care, since other possible even more powerful predictors for MD were not taken into account. Several authors stressed the importance of counting (Gersten et al., 2005; Stock et al., 2009b) executive functions (e.g., Andersson, 2008; Mazzocco & Kover, 2007; Van der Sluis, De Jong, & Van der Leij, 2007), working memory (e.g., Alloway, Gathercole, Adams, Willis, Eaglen, & Lamont,
2005; Passolunghi, Mammarella, & Altoe, 2008) and attention (Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002) in the development of mathematical (dis)abilities. Studies on executive functions of children with MD are currently being analysed (De Weerdt, Stock, Desoete, & Roeyers, 2009). Fourthly, it should be pointed out that arithmetic and its early precursors might have may components (Dowker, 2005; 2008; Jordan et al., 2009) and that it is therefore likely that MD are not homogeneous (Iuculano, Tang, Hall, & Butterworth, 2008; Von Aster, 2000). Finally, context variables such as home and school environment and expectations (e.g., Brady & Woolfson, 2008; Flouri, 2006; Rubie-Davies, 2010), learning packages (e.g., Van Steenbrugge, Valcke, & Desoete, 2009) and parental involvement (e.g., Reusser, 2000) should be included in order to obtain a complete overview of the arithmetical development of these children. These limitations indicate that only a part of the picture was investigated, so additional studies should focus on these aspects.

Nevertheless this study was longitudinal in nature, allowing us to determine whether individual accuracy differences in kindergarten (before the start of formal mathematics education) on symbolic and non-symbolic comparison tasks can predict later individual differences in arithmetical achievement in grades 1 and 2. It seems clear that the choice of the task matters and that the prediction depends on the age of the children and the aim of the assessment. ‘When’ and ‘what’ you test is what you get. The accuracy in non-symbolic magnitude comparison in kindergarten was predictively related to early arithmetic in grade 1. The accuracy in symbolic AN comparison tasks in kindergarten was predictively associated to the procedural calculation skills in grade 2. Both symbolic comparison skills assessed in grade 2 were associated with procedural calculation at the same moment. If the aim is to screen for non-typically developing children, children with MD already seem to have deficits in non-symbolic and symbolic AN comparison in kindergarten, whereas in grade 2 the lack of accuracy in processing symbolic information remains.

References


Krajewski, K., & Schneider, W. (2009). Early development of quantity to number-word linkage as a precursor of mathematical school achievement and mathematical difficulties: Findings from a four-year longitudinal study. Learning and Instruction, 19, 513–526.


