DETECTING CHILDREN

Detecting children with arithmetic disabilities from kindergarten:

Evidence from a three year longitudinal study on the role of preparatory arithmetic abilities

Pieter Stock, Annemie Desoete and Herbert Roeyers

Ghent University, Ghent, Belgium

Correspondence concerning this article should be addressed to Annemie Desoete. Dunantlaan 2, B-9000 Ghent, Belgium. Tel: ++32/9/264.94.14; Fax: ++32/9/264.64.89. E-mail: Annemie.Desoete@UGent.be

Abstract

In a three-year longitudinal study 471 children were classified, based on their performances on arithmetic tests in first and second grade, as having persistent arithmetic disabilities (AD), persistent low achieving (LA), persistent typical achieving, inconsistent arithmetic disabilities (DF1) or inconsistent low achieving in arithmetic. Significant differences in the performances on the magnitude comparison in kindergarten (at age 5 to 6) were found between the AD and LA and between the AD and DF1 groups. Furthermore, the percentage of true positive AD children (at age 7 to 8) correctly diagnosed in kindergarten by combination of procedural counting, conceptual counting and magnitude comparison tasks was 87.50%. When composing clinical samples, researchers should pay attention when stipulating restrictive or lenient cut-offs for arithmetic disabilities and select children based on their scores in two consecutive years, since the results of studies on persistent low achievers (LA) or children with inconsistent disabilities (DF1) can not be generalized to children with persistent arithmetic disabilities (AD).

Detecting children with arithmetic disabilities from kindergarten:

Evidence from a three year longitudinal study on the role of preparatory arithmetic abilities

Introduction

Early arithmetic abilities have been found to be the strongest predictor of later school achievement (e.g., DiPerna, Lei, & Reid, 2007; Duncan et al., 2007; Muldoon, Lewis, & Francis, 2007; Stock, Desoete, & Roeyers, in press; Teisl, Mazzocco, & Myers, 2001). Since arithmetical disabilities are persistent (Shalev, Manor, & Gross-Tsur, 2005), it would be interesting to recognize vulnerable young children early in order to prevent children from falling further behind and from developing arithmetic difficulties later on (Coleman, Buysse, & Neitzel, 2006; Gersten, Jordan, & Flojo, 2005; Pasnak, Cooke, & Hendricks, 2006). In addition the first step in a Responsiveness-To-Intervention (RTI) prevention model is determining children who are at risk for developing arithmetic learning disabilities (e.g., Fuchs et al., 2007; Kavale & Spaulding, 2008). Identification of children at risk in kindergarten permits children to participate in prevention services before the onset of substantial deficits (Fuchs et al., 2007).

There is not yet a consensus regarding which of the early math predictors are uniquely associated with early responses to formal math instruction in kindergarten, first and second grade children. Some of the traditional described preparatory arithmetic abilities in kindergartners are the Piagetian logical abilities (Nunes & Bryant, 1996; Piaget, 1965). Besides the Piagetian type tasks, counting knowledge also seems to be associated with early responses to formal math instruction (Baroody, 1992; Frank, 1989; Gersten et al., 2005; Johansson, 2005; Sophian, 1992; Van De Rijt & Van Luit, 1999). Moreover, representation

of number size was found to be involved in numerical competence as well (Jordan, Kaplan, Olah, & Locuniak, 2006). In the present study, we sought to combine the classical Piagetian type tasks with other post-Piagetian domain specific measures such as counting knowledge and insight in magnitude comparison and determine the contribution of intelligence as predictors for children at risk. The purpose of the current study was to examine differences between children with persistent Arithmetic Disabilities (AD), Low Achievers (LA) and Typical Achiever (TA) groups with respect to theoretically relevant predictors and determine the accuracy with which the predictors could be used as a screener to identify children with math difficulties.

First, the main findings on Piagetian type tasks, counting abilities and magnitude comparison as important, although certainly not mutually exclusive, preparatory arithmetic abilities controlling for the role of intelligence will be addressed. Next, we will focus on the implications of the use of different selection criteria in order to define children with arithmetic disabilities. In the last section, the objectives of this study will be specified.

Preparatory Arithmetic Abilities

Piagetian type tasks

Piaget and Szeminska (1941) postulated that four logical abilities are conditional to the development of arithmetic, namely seriation, classification, conservation and inclusion. Seriation is defined as the ability to sort a number of objects based on the differences in one or more dimensions while ignoring the similarities. In contrast, classification is the ability to sort objects based on their similarities in one or more dimensions. Here children have to make abstraction of the differences. In 1959, Piaget and Inhelder stated that the coordination of 'seriation' and 'classification' is needed for the comprehension that '4' is included in '5',

whereas '5' it self is included in '6' (Grégoire, 2005). When children further develop and get older, they use this knowledge to make hierarchical classifications: they learn that numbers are series that contain each other. This is the inclusion principle and it can be seen as the highest form of classification (Piaget & Szeminska, 1941). Once the child is confident in the knowledge that the number of objects in a collection only changes when one or more objects are added or removed, it masters the conservation principle (Piaget & Szeminska, 1941). Piaget (1965) argued that the full development of number comprehension is only possible when the child masters these four logical abilities. However Piaget ignored the importance of counting.

Since the publication of the work of Piaget, several neo-Piagetian researchers questioned the causality of seriation and classification for understanding number (e.g., Dumont, 1994; Lourenço & Machado, 1996) and stated that counting was the best predictor for early arithmetic performances. Nevertheless, other studies revealed that seriation assessed in kindergarten is related to arithmetic achievement in grade 1 (e.g., Grégoire, 2005; Kingsma, 1984; Stock, Desoete, & Roeyers, 200) and grade 2 (Grégoire, 2005), and children adequately solving classification tasks in kindergarten perform better in arithmetic tasks in grade 1 and 2 (Grégoire, 2005). Many studies confirmed that logical abilities are important markers for the development of arithmetic abilities. Even after controlling for differences in working memory, logical abilities in six year-old children remained a strong predictor for arithmetic abilities 16 months later (Nunes et al., 2006).

Post-Piagetian Counting Knowledge

Since the 1980s there is considerable interest in 'counting' as neo-Piagetian milestone in the development of an understanding of number (e.g., Aunola, Leskinen, Lerkkanen, & Nurmi, 2004; Chard, Clarke, Baker, Otterstedt, Braun, & Katz, 2005; Fuson, 1988; Le Corre,

Van de Walle, Brannon, & Carey, 2006). Dumont (1994) hypothesized that counting predicted ordinality (outcome measure of seriation) and cardinality (outcome measure of classification). Moreover, it is obvious that early mathematics involves “counting” (Wynn, 1990). For example in the ‘count all’ or ‘sum’ strategy, the child first counts each collection and then counts the combination of two collections starting from one (i.e., $2+5=_$ 1,2...1,2,3,4,5...1,2,3,4,5,6,7). As practice increases, older children use more effective back-up strategies (Barrouillet & Lepine, 2005). In addition, counting can be seen as the foundation for strategies such as subtraction (Le Fevre et al., 2006) and multiplication (Blöte, Lieffering, & Ouwehand, 2006).

Although a lot of research looked into counting as a unitary ability, Dowker (2005) suggested that counting knowledge consists of procedural and conceptual aspects. *Procedural knowledge* is defined as children’s ability to perform a counting task, for example, when a child can successfully determine that there are five objects in an array (LeFevre et al., 2006). ‘*Conceptual counting knowledge*’ reflects a child’s understanding of why a procedure works or whether a procedure is legitimate. It demonstrates the understanding of the essential counting principles: the stable order principle, the one-one-correspondence principle and the cardinality principle (Gallistel & Gelman, 1990; Gelman, 1990; LeFevre et al., 2006; Wynn, 1992).

Many researchers focused from different theoretical frameworks on the importance of procedural and conceptual counting knowledge in the development of arithmetic abilities (Baroody, 1992; Frank, 1989; Fuchs et al., 2007; Gersten et al., 2005; Johansson, 2005; Le Corre et al., 2006; LeFevre et al., 2006; Sophian, 1992; Van de Rijt & Van Luit, 1999). In their longitudinal study from preschool to second grade, Aunola and colleagues (2004) found that counting ability was the best predictor of the initial level of arithmetic performance. It has been found that children’s basic conceptual understanding of how to count objects and

their knowledge of the order of numbers plays an important role in later arithmetic performance. Mastery of conceptual knowledge allows children to devote attentional resources to more complex arithmetic problem solving (Aunola et al., 2004).

There are a lot of studies on counting skills of subjects with arithmetic disabilities (e.g., Geary, 2004; Jordan, Hanich, & Kaplan, 2003). Dowker (2005) showed that children who had difficulties in any particular aspect of counting had overall below average mathematical performances. Fuchs et al. (2007) included Number Identification/Counting as one of the screening measures for forecasting math disabilities at the end of 2nd grade. In addition, it was shown that toddlers who lacked adequate and flexible counting knowledge went on to develop deficient numeracy skills which resulted in arithmetic disabilities (Aunola et al., 2004; Gersten et al., 2005). Furthermore, Geary, Bow-Thomas and Yao (1992) found that small children with arithmetic disabilities were more likely to make procedural errors in counting and still had considerable conceptual difficulties at the age of six. Desoete and Grégoire (2007) also showed that children with arithmetic disabilities in grade 1 already had encountered problems on numeration in nursery school. They also found some evidence of dissociation of numerical abilities in children with arithmetic disabilities in grade 3. Certain skills appeared to be developed whereas others were not, which made it necessary to investigate them separately and independent of one another. About 13% of the children with arithmetic disabilities still had processing deficits in number sequence and cardinality skills in grade 3. About 67% of these children in grade 3 had a lack of conceptual knowledge. Finally in this field of research Porter (1998) contributed the finding that the acquisition of procedural counting knowledge did not automatically lead to the development of conceptual understanding of counting in children with arithmetic disabilities. Taking into account the complex nature of mathematical problem solving, it may be useful to assess procedural as well as conceptual counting procedures in young children at risk. However most post-

Piagetian researchers ignored the importance of seriation and classification and focused only on counting skills.

Magnitude Comparison as 'Core' deficit?

Recently, number sense and representation of number size was found to be also involved in numerical competence (e.g., Berch, 2005; Butterworth, 2005; Gersten et al., 2005; Griffin, 2004; Jordan et al., 2003; Jordan, Kaplan, Olah, & Locuniak, 2006). For a description of the complexity of number sense we refer to Berch (2005) who reviewed the relevant literature and pointed out that processing number sense allows a child to achieve problem-solving from understanding the meaning of numbers to developing strategies; from making number comparison to creating procedures for operating numbers; and integrating her/his knowledge to interpret information. Number sense is involved in 'subitizing' and in 'magnitude comparison'. Subitizing is the rapid apprehension of small numerosity (Kaufman, Lord, Reese, & Volkmann, 1949; Nan, Knösche, & Luo, 2006), while magnitude comparison holds that children have to know which number in a pair is larger (Desoete, Ceulemans, Roeyers, & Huylebroeck, 2008; Gersten & Chard, 1999; Hannula & Lehtinen, 2005; Xu & Spelke, 2000). It was found that the larger the distance between the numbers and the smaller the magnitudes of the numbers, the faster and more accurate the answer on a magnitude comparison task was likely to be (Dehaene, 1997; Dehaene, Bossini, & Giraux, 1993; Gevers, Lammertyn, Notebaert, Verguts, & Fias, 2006; Zhou, Chen, Chen, & Dong, 2008). Furthermore, performances on both magnitude comparison and subitizing tasks improved with increasing age and experience (Laski & Siegler, 2007; Xu, 2003).

Some researchers explain problems of children with arithmetic disabilities as results of a 'core' deficit in 'number sense' a term denoting the ability to picture and manipulate numerical magnitude on an internal number line (e.g., Gersten et al., 2005; Holloway &

Ansari, 2008; Landerl, Bevan, & Butterworth, 2004; Van Aster & Shalev, 2007). Magnitude comparison was found to be an important predictor of variation in arithmetic abilities (Durand, Hulme, Larkin, & Snowling, 2005).

There are far less studies on number sense skills than on counting knowledge of subjects with arithmetic disabilities. However, Reeye and Reynolds (1994) found that 6% of a randomly selected group of children did not subitize (they called them the ‘none’ subitizers). These children were followed in a longitudinal design, revealing that ‘none’ subitizers were slower in reading three digit numbers in comparison to the so called ‘slow’ or ‘fast’ subitizers. These results were repeated the next year as well. The authors however did not report whether or not these ‘non subitizers’ actually developed arithmetic disabilities. The importance of subitizing in arithmetic disabilities was pointed out by Koontz and Berch (1996) who found that children with arithmetic disabilities were slower to process numbers and slower in subitizing tasks, in comparison to children without arithmetic disabilities. This finding was confirmed by Landerl and colleagues (2004) and by Rousselle and Noel (2007) when they found that children with arithmetic disabilities were slower at numerical differentiation in comparison to control groups and that they showed deficits in subitizing. However, not all subjects with arithmetic disabilities were found to have subitizing problems. Desoete and Grégoire (2007) found a severe subitizing deficit in 33% of the 30 average intelligent children of 8 and half years old with a clinical diagnosis of arithmetic disabilities in Flanders. Fischer, Gebhart and Hartnegg (2008) found that between 43% and 79% of the subjects in the age range of 7 to 17 years with arithmetic disabilities performed below the 16-th percentile of the peer control groups on subitizing tasks.

Early Arithmetics skills

Initial arithmetic can be seen as a broad domain of various arithmetical achievement and numerical facility skills (Dowker, 2005). Arithmetical achievement is needed to convert linguistic and numerical information into math equations and algorithms, to understand mathematical concepts and operations and to identify and select appropriate strategies for solving computation and word problems. In addition, by executing arithmetic problems repetitively basic number facts (e.g., $6+2=8$) are retained in long-term memory and ‘automatic’ retrieved if needed. Children with arithmetic disabilities often have problems in the area of automaticity. They lack numerical facility and do not know basic number facts by heart (Geary & Hoard, 2005; Jordan, Levine, & Huttenlocker, 1995).

Intelligence

Often an IQ-assessment serves as an indicator for the general level of achievement (American Psychiatric Association, 2000; Dickerson Mayes, Calhoun, Bixler, & Zimmerman, 2009; Gross-Tsur, Manor, & Shalev, 1996). Recent studies found a correlation of .50 between arithmetical abilities and intelligence (Desoete, 2008; Kort et al., 2002; Ruijsenaars, Van Luit, & Van Lieshout, 2004).

The present study aims to look at the possible group differences in intelligence among children with arithmetic disabilities, low achievers and typical achieving groups not to confound those group differences in the assessment of arithmetic achievement and disabilities.

Criteria in Arithmetic Disabilities Research

Seriation, classification, procedural counting knowledge, conceptual counting knowledge and magnitude comparison as preparatory arithmetic abilities have shown to be promising markers for the early detection of children with arithmetic disabilities. Yet the lack

of a unified set of distinct criteria that describes arithmetic disabilities (Mazzocco, 2001; 2005; Mazzocco & Myers, 2003; Stock et al., 2006; Vaughn & Fuchs, 2003) has profound implications for research that aims to detect children with these disabilities.

The use of divergent selection criteria in the recruitment of research samples based on cutoffs ranging from percentile 3 to even percentile 45 (American Psychiatric Association, 2000; Geary, 2004; Klauer, 1992; Kosc, 1974; Mazzocco, 2001; Stock et al., 2006), may have conflated children with severe and mild forms of arithmetic disability (Geary et al., 2007; Mazzocco et al., 2008). Murphy, Mazzocco, Hanich and Early (2007) revealed differences between children with a severe form of disability (using restrictive criteria), children with a mild form of disability (using lenient criteria) and children with a math performance exceeding the 25th percentile. Children selected with restrictive or lenient criteria showed different arithmetic skills in kindergarten, with the former having more deficient arithmetic abilities than the children with a mild form of disability. Mazzocco et al. (2008) found that children with a severe form of disability showed qualitatively different profiles in fact retrieval performances when compared to typically achieving children, whereas the differences between children with a mild form of disability and typically achieving children were of a quantitative turn. Geary et al. (2007) revealed that children with a severe disability had a severe math cognition deficit and underlying deficit in working memory and speed of processing. The groups with a mild disability had more subtle deficits in few math domains. The variability and wide array of selection criteria may have dispelled a clear and accurate idea of the nature of the disability complicating identification, diagnosis, treatment and remediation of children with arithmetic disabilities (Geary et al., 2007; Hammill, 1990; Mazzocco et al., 2008; Murphy et al., 2007).

Finally, it is important to consider consistency in performance over time (Hanley, 2005; Mazzocco & Myers, 2003). In the 'status model' children are identified as children

with arithmetic disabilities based on an assessment at a single testing point (Fletcher et al. 2005). However, since arithmetic capacities are dynamic, there is much variability in the developmental process (Shalev, 2004) and test scores will fluctuate around a cut-point needing repeated testing (Fletcher et al., 2005). In order to be sure of the persistence of disabilities although remediation took place, in recent studies often children are selected based on their scores in two or more consecutive years. Geary (2004) used the cut-off criterion of the 25th percentile, but warranted that only children who had scores across successive academic years beneath this cut-off may have a diagnosis of Arithmetic Disabilities (AD). The implementation of such a ‘fluctuation model’ and ‘persistence criterion’ is important because it was found that performances in children meeting criteria for AD at one point in time could not be generalized to children who successively meet criteria year after year although adequate instruction took place (Fletcher et al, 2005; Mazzocco, 2001; Murphy et al., 2007). This ‘persistence criterion’ is also included in the recent emphasis on Resistance to Instruction (RTI) for defining AD (e.g., Fuchs et al., 2007). Only when children do not enhance in arithmetic performance although they got special task-specific instruction are defined, according to the RTI-principle as children with AD.

Objectives and Research Questions

Since there is not yet a consensus regarding which type of predictors are uniquely associated with early arithmetic performances, in this study Piagetian and post-Piagetian ideas are combined. In addition magnitude comparison was added since this might be a core deficit in children with arithmetic disabilities. The combination of the classical Piagetian type tasks with other measures of conceptual and procedural knowledge and magnitude comparison, without confounding group differences in intelligence. is a unique and interesting aspect of this study.

The first purpose of the current study was to investigate, in line with Mazzocco & Thompson (2005), kindergarten predictors (at age 5 to 6) of arithmetic performance in grade 1 (at age 6 to 7) and grade 2 (at age 7 to 8). Furthermore, in this three year longitudinal study it was questioned if it was possible to detect children with persistent arithmetic disabilities (AD) in grade 2 based on the preparatory arithmetic abilities in kindergarten.

Moreover in line with Murphy et al. (2007) it is studied whether characteristics of children with arithmetic disabilities vary as a function of the cut-off criterion used to define the disabilities. The research question of whether it was possible to find differences in kindergarten characteristics between groups based on ‘severity’ or between children with persistent arithmetic disabilities (AD), persistent low achieving (LA) and persistent typically achieving (TA) children is addressed.

Finally in line with Fletcher et al. (2005) contrasting the ‘status model’ versus the ‘fluctuation model’, the kindergarten characteristics of children with a persistent arithmetic disabilities based on an assessment at two testing points (AD) and the scores of children with inconsistent arithmetic disabilities (DF1) and clinical scores at one testing point are compared.

Method

Participants

Participants were 471 children (227 boys and 244 girls) who entered the study in the last year of kindergarten (age 5 to 6) and completed data collection in 1st grade (age 6 to 7) and 2nd grade (age 7 to 8). The children had a mean age of 70.02 months ($SD = 4.01$ months) and attended on average 7.42 months ($SD = 1.03$ months) of school in the last kindergarten class when tested the first time. The follow up testing was in grade 1 ($M = 82.02$ months) and

in grade 2 ($M = 94.02$ months). In the present study, the data of 464 of these children (225 boys and 239 girls) were analysed. Only at least average intelligent children were included in the analyses. All children were Caucasian native Dutch-speaking children living in the Flemish part of Belgium. In Belgium children attend kindergarten class for about three years (from the moment they turn 2 year 6 month old) and move to elementary classes in the year they turn six year old.

To separate groups based on ‘potential severity’, three groups of children (AD, LA, TA) participated in this study ($n = 319$), based on an assessment and consistent achievement on at least two testing points (the first testing took place in grade 1; the second testing took place in grade 2). Children were classified in grade 2 retrospectively as having arithmetic disabilities (AD) if they scored $\leq$ the 10th percentile on at least one of the arithmetic achievement tests, both in first and second grade ($n = 10$ boys and 6 girls). Children who scored percentile 11-25 on at least one of the arithmetic achievement tests, both in first and second grade, were classified as low achieving (LA; $n = 14$ boys and 13 girls). The third group consisted of ‘typical achievers (TA; $n = 154$ boys, 165 girls) or children who scored $>$ the 25th percentile on all arithmetic achievement tests in both grades.

To compare the ‘status model’ versus ‘fluctuation model’ we explored our dataset for children with fluctuating test scores and within-subject inconsistency on poor arithmetic achievement across the primary school age years. To be accepted in our sample as children with fluctuating scores, children had to belong to the group of ‘severe arithmetic difficulties’ (DF1) or to the group of ‘mild arithmetic difficulties’ (DF2). The DF1-group was a group of 26 boys and 39 girls performing $\leq$ the 10th percentile in one grade (so they would have been classified as AD children based on an assessment at a single testing point) and above the 25th percentile in the other grade (so they belonged to the TA group based on this single testing).

The DF2 group was the group of 21 boys and 16 girls who scored between pc 11-25 in one grade and above the 25th percentile in the other grade.

A socio-economic status was derived from the total number of years of scholarship of the parents (starting from the beginning of elementary school), with a mean of 14.84 years ($SD = 2.41$ years) for mothers and 14.48 years ($SD = 2.85$ years) for fathers.

< Insert Table 1 about here >

No significant differences in SES were found between the AD, LA, TA, DF1 and DF2 groups (see Table 1).

Materials

All children were tested in kindergarten on their preparatory arithmetic abilities. Follow-up assessment with two arithmetic tests was conducted in first and second grade and intellectual abilities were tested in second grade.

Preparatory Arithmetic Abilities in Kindergarten

All preparatory arithmetic abilities were tested with different subtests of the TEst for the Diagnosis of Mathematical Competences (TEDI-MATH; Grégoire, Noel, & Van Nieuwenhoven, 2004). The TEDI-MATH is an individual assessment battery that was constructed in order to detect arithmetic disabilities. The manual suggests to use scores < pc 25 as clinical cut-off scores for children 'at risk'. Important basic works for the construction of the test were the theory on logic thinking of Piaget (Piaget, 1965; Piaget & Szeminska, 1941), the research on counting of Gelman and Gallistel (1978) and Fuson and colleagues (e.g., Fuson, 1988) and the research work of Geary (e.g., Geary, 1994) on the development of

arithmetic abilities. The TEDI-MATH has been tested for conceptual accuracy and clinical relevance in previous studies (e.g., Desoete & Grégoire, 2007; Stock et al., 2007). 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, 2006; 2007a & b) and reliable instrument, values for Cronbach's Alpha for the different subtests vary between .70 and .97 (Grégoire et al., 2004). The predictive value has been established in a longitudinal study of 82 children from kindergarten till grade (Desoete & Grégoire, 2007) and on 240 children assessed in grade 1, 2 or 3 with TEDI-MATH and reassessed two years later with arithmetic tasks (Desoete, 2007). In addition the Flemish data were confirmed with similar data from the French speaking part of Belgium and France (Desoete, Roeyers, Schittekatte, & Grégoire, 2006).

Counting knowledge included procedural and conceptual knowledge of counting.

Procedural knowledge of counting. Procedural knowledge of counting was assessed using accuracy in counting numbers, counting forward to an upper bound (e.g., 'count up to 6'), counting forward from a lower bound (e.g., 'count from 3'), counting forward with an upper and lower bound (e.g., 'count from 5 up to 9' – see Appendix A). One point was given for a correct answer without helping the child. The task included 13 items; the maximum total score was 14 points. The total row item scores were summed and converted to z-scores. In addition percentile scores were computed to link the scores of our sample to the scores of the stratificated normative sample and to obtain an easier clinical interpretation of the findings. The internal consistency of this task was good (Cronbach's Alpha = .73).

Conceptual knowledge of counting. Conceptual knowledge of counting was assessed with judgments about the validity of counting procedures. Children had to judge the counting of linear and random patterns of drawings and counters (see Appendix A). To assess the abstraction principle, children had to count different kinds of objects that were presented in a

heap. Furthermore, a child who counts a set of objects is asked ‘How many objects are there in total?’, or ‘How many objects are there if you start counting with the leftmost object in the array?’. When children have to count again to answer, they do not gain any points, as this is considered to represent good procedural knowledge, but a lack of understanding of the counting principles of Gelman and Gallistel (1978). One point was given for a correct answer with a correct motivation (e.g., you did not add objects so the number of objects has not changed). The maximum total score was 13 points. The total row item scores were summed and converted to z-scores. In addition percentile scores were computed to link the scores of our sample to the scores of the stratificated normative sample. The internal consistency of this task was good (Cronbach’s Alpha = .85).

Logical abilities. Logical abilities were assessed using two different tasks (see Appendix A). Children had to seriate numbers (e.g., ‘Sort the (six) cards from the one with the fewest trees to the one with the most trees’). The maximum score was 3 points. Children had to make groups of nine cards in order to assess the classification of numbers (e.g., ‘Make (three) groups with the cards that go together’). The maximum score was 3 points. The total row item scores were summed and converted to z-scores. In addition percentile scores were computed to link the scores of our sample to the scores of the stratificated normative sample. The internal consistency of the two tasks were good with Cronbach’s alpha of .68, and .73 respectively.

Magnitude comparison. Magnitude comparison (see Appendix A) was assessed by comparison of dot sets (e.g., 4 dots versus 6 dots or 7 dots versus 2 dots). Children were asked where they saw most dots. One point was given for a correct answer. As the task included 6 items, the maximum score was 6 points. The total row item scores were summed and converted to z-scores in order to analyze the results. In addition percentile scores were

computed to link the scores of our sample to the scores of the stratificated normative sample. The internal consistency of this task was good (Cronbach's Alpha = .79).

Arithmetic Tests in First and Second Grade

In order to have a full sight on the arithmetic abilities of children in first and second grade, two arithmetic tests were used: The Kortrijk Arithmetic Test Revised (Kortrijkse Rekentest Revision, KRT-R, Baudonck et al., 2006) and the Arithmetic Number Facts Test (Tempo Test Rekenen, TTR, De Vos, 1992).

The Kortrijk Arithmetic Test Revision (Kortrijkse Rekentest Revision, KRT-R; Baudonck et al., 2006) is a standardized test on arithmetical achievement which requires that children solve 30 mental arithmetics (e.g., ' $16-12 = _$ ') and 30 number knowledge tasks (e.g., '1 more than 3 is $_$ ') in first grade; children in second grade received 30 and 25 tasks respectively. The KRT-R is frequently used in Flemish education as a measure of arithmetic achievement (e.g., Desoete, Roeyers, & De Clercq, 2004; Desoete & Grégoire, 2007). The test results in a score for mental computation, number system knowledge and a total score. The raw item scores were converted to percentile scores. The psychometric value of the KRT-R has been demonstrated on a sample of 3,246 children. A validity coefficient (correlation with school results) and reliability coefficient (Cronbach's alpha) of .50 and .92 respectively were found for first grade.

The Arithmetic Number Facts Test (Tempo Test Rekenen, TTR; De Vos, 1992) is a numerical facility test consisting of 80 (first grade) or 200 (second grade) arithmetic number fact problems. In first grade children receive a form with two subtests: one subtest with 40 additions (e.g., ' $2+3 = _$ ') and one subtest with 40 subtractions (e.g., ' $8-3 = _$ '). Children have to solve as many additions as possible in one minute. After that, they receive another minute to solve as many of the 40 subtractions as possible. From second grade a form with five

subtests is used. The first subtest requires addition (e.g., '2+3'= _), the second subtraction (e.g., '8-3=_'), the third multiplication (e.g., '5x9=_'), the fourth division (e.g., '15:3=_'), the fifth mixed exercises (additions, subtractions, multiplications and division through each other). Children have to solve as many items as possible in five minutes; they can work one minute on every subtest. The TTR is a standardized test that is frequently used in Flemish education as a measure of early arithmetic acquisition. The total number of correct items were summed and converted to percentile scores. The psychometric value of the TTR has been demonstrated on a sample of 10,059 children in total (Ghesquière & Ruijsenaars, 1994). Cronbach's alphas computed for the current study was .90. The Guttman Split-Half Coefficient was .93; the Spearman-Brown coefficient was .95.

Intelligence

In order to have an estimation of the intellectual capacities of the children, 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 the moment. The short version is based on four subtests and includes both measures for crystallized and fluid intelligence (Vocabulary, Similarities, Block Design and Picture Arrangement; Grégoire, 2001). The mean intelligence of the children was $M = 100.54$ ($SD = 13.37$). No significant differences in intelligence were found between the AD, LA, TA, DF1 and DF2 groups (see Table 1).

Procedure

The children were recruited in 33 schools. The schools were randomly selected

with schools in a city and out of the city and representing all three Types of schools in Belgium: Catholic schools, schools organized by the federal government, and schools organized by the cities or provinces (Desoete, Roeyers, & De Clercq, 2004). Parents received a letter with the explanation of the research and submitted informed consent in order to participate every year.

Children were tested during school time in a separate and quiet room. Children were tested individually in kindergarten. The total duration of the individual testing was about 40 minutes. Tests were assessed in the same order and in the same period of the school year. In first and second grade the children were assessed on arithmetic abilities (with KRT-R) and on numerical facility (with TTR). The TTR was presented first and after that the KRT-R. The short version of the WISC III was assessed individually in second grade.

The test leaders all received training in the assessment and interpretation of the tests. For every subtest, instructions and scoring rules were explained orally. In order to guarantee reliability of the assessment, each tester had to test one child and score the protocol in advance. This protocol was analyzed and corrected by the main researcher of the study. The test-protocols were not included in the analyses of this study. All responses were entered on an item-by-item basis into SPSS. A second and third scorer independently re-entered all protocols, with 99.9% agreement. After completion of the test procedure, all the parents of the children received individual feedback on the results of their children.

Results

Research Question One: How well do kindergarten (age 5 to 6) performances model the arithmetic abilities and numerical facility in Grade 1 (age 6 to 7) and Grade 2 (age 7 to 8).

The research question on the proportion of variance in arithmetic skills that could be modelled by the performances in kindergarten is answered with a correlation among the kindergarten predictors and a regression analyses. The correlation among the kindergarten variables are reported in Table 2.

< Insert Table 2 about here >

Intelligence correlated significantly with almost all measures except with classification in kindergarten. The correlation between the Piagetian tasks in kindergarten was $r = .112$ ($p = .01$). Moreover procedural counting and conceptual counting skills correlated significantly in kindergarten $r = .296$ ($p < .0005$). In kindergarten also magnitude comparison correlated significantly with procedural counting $r = .156$ ($p = .001$), seriation $r = .102$ ($p < .05$) and classification $r = .093$ ($p < .05$).

To address the question on how well the kindergarten predictors (at age 5 to 6) model arithmetic achievement, regression analyses were conducted on the 1nd and 2nd grade outcome (see Table 3).

< Insert Table 3 about here >

It was first tested whether the arithmetic abilities in *grade 1* could be modelled by classical Piagetian type tasks and with other measures of procedural and conceptual counting knowledge and magnitude comparison tests in kindergarten. The linear combination of kindergarten predictors (at age 5 to 6) was significantly related to arithmetic reasoning

assessed in grade 1 (at age 6 to 7) with KRT-R, $F(5, 458) = 20.423$; $p = .0005$. R^2 was .184; meaning that about 18% of the variance in arithmetic reasoning in grade 1 could be modelled by assessing the performances of children one year earlier in kindergarten. Especially the Piagetian seriation and classification tasks and the neo-Piagetian procedural counting and magnitude comparison tasks were beneficial as kindergarten predictors for arithmetic reasoning in grade 1. The results also suggested that 8% of the variance in numerical facility, assessed in first grade (at age 6 to 7) with TTR could be modelled by assessing the preparatory arithmetic abilities in kindergarten (at age 5 to 6), $F(5, 458) = 8.066$; $p = .0005$. R^2 was .082. Especially the Piagetian classification task was beneficial as kindergarten predictor.

It was further tested whether the arithmetic abilities in *grade 2* could be modelled by kindergarten predictors. The combination of kindergarten predictors (at age 5 to 6) was significantly related to arithmetic reasoning assessed in grade 2 (at age 7 to 8) with KRT-R, $F(5, 458) = 10.775$; $p = .0005$. R^2 was .106. Especially the classical Piagetian type seriation task and the neo-Piagetian procedural counting tasks were beneficial as kindergarten predictors for arithmetic reasoning in grade 2. The kindergarten predictors (at age 5 to 6) were also significantly related to numerical facility assessed in grade 2 (at age 7 to 8) with TTR, $F(5, 458) = 4.408$; $p = .001$. R^2 was .047. Especially procedural counting was beneficial as kindergarten predictor for numerical facility in grade 2.

Research Question two: Group Differences between persistent arithmetic ability groups (AD, LA, TA).

The research question of whether it was possible to find differences between groups of children with ‘persistent’ arithmetic performances based on ‘potential severity’ without confounding group differences in intelligence, is addressed with a multivariate analysis of

covariance (MANCOVA). Using the Bonferroni method, each ANOVA was tested at the .008 level correcting for experiment-wise error rate.

The MANCOVA was conducted on intelligence as covariate, group (persistent arithmetic disabilities or AD, persistent low achievers or LA, persistent typically achievers or TA) as independent variable and procedural counting knowledge, conceptual counting knowledge, seriation, classification, magnitude comparison (at age 5 to 6) as dependent variables. The MANCOVA was significant on the multivariate level, for intelligence ($F(5, 336) = 12.921; p = .0005$) but also for the group (AD, LA, TA) $F(10, 672) = 6.265, p = .0005$. The means and standard deviations of the dependent variables for the three groups are shown in Table 4.

< Insert Table 4 about here >

On the univariate level, post hoc analyses revealed that the TA group (a -index) performed significantly better than the AD group (b -index) on seriation and magnitude comparison tasks (at age 5 to 6) and significantly better than the LA group (b -index) on seriation but not on magnitude comparison tasks (a -index). LA children (b -index) did not differ significantly from AD children (b -index) on seriation. However, LA children (a -index) did significantly better than AD children (b -index) on magnitude comparison tasks.

Research Question Three: Prediction of Group Membership within persistent arithmetic ability groups.

Our next research question was whether it was possible to predict ‘persistent’ arithmetic disabilities (AD) in grade 2 (at age 7 to 8) based on the kindergarten predictors assessed two years earlier. This research question was answered with four analyses. Firstly

the research question was addressed by conducting a discriminant analyses on kindergarten tests to investigate the overall accurateness of the predicted classifications in the AD, LA or TA groups (at age 7 to 8). Secondly, the contribution of each kindergarten predictor uniquely was reported. Thirdly, with ROC curves the group membership (AD versus non-AD) was predicted and the sensitivity and specificity of the predictors was described. Finally, in order to obtain an easier clinical and sample-independent interpretation, it was investigated if the kindergarten clinical cut-off scores (at age 5 to 6), based on the stratificated normative sample of the TEDI-MATH, are consistent with the sample-dependent analyses with z-scores.

Firstly, the Fisher's linear discriminant function was used to investigate the accurateness of the predicted classifications in the AD, LA and TA groups. The overall Wilks' lambda was significant, $\Lambda = .831$, $\chi^2 (10, N = 357) = 65.263$, $p = .0005$, indicating that overall the kindergarten predictors differentiated among the AD, LA and TA group. In Table 4 the standardised weights of the predictors are presented. Based on these coefficients, magnitude comparison and seriation demonstrated the strongest relationships with the general arithmetic achievement of the children. The means on the discriminant function were consistent with this interpretation. The TA children did better on the preparatory arithmetic abilities. Based on the scores for these kindergarten predictors, 88.0% was classified correctly into the AD, LA or TA group in grade 1 and 2, whereas 88.0% of the cross-validated grouped cases were classified correctly. However only 18.8% of the AD children no LA children and 99.0% of the TA children were classified correctly. Table 5 gives an overview of the classification results based on this discriminant function.

< Insert Table 5 about here >

In order to take into account chance agreement, a kappa coefficient was computed and obtained a value of .17, indicating a weak prediction.

To see how well each predictor uniquely predicts group membership, the contribution of each predictor uniquely is reported.

The overall Wilks' lambda for the Piagetian tasks was significant, namely $\Lambda = .929$, $\chi^2(2, N = 358) = 26.256$, $p = .0005$ for seriation (with 100% of the TA children but non of the LA and AD children classified correctly) and $\Lambda = .982$, $\chi^2(2, N = 358) = 6.577$, $p = .037$ for classification respectively (with 100% of the TA children but non of the LA and AD children classified correctly), indicating that overall the Piagetian-kindergarten predictors differentiated among the AD, LA and TA group.

The overall Wilks' lambda for the neo-Piagetian counting tasks and the magnitude comparison tasks were significant, namely $\Lambda = .971$, $\chi^2(2, N = 357) = 10.365$, $p = .006$ for procedural counting knowledge (with 100% of the TA children but non of the LA and AD children classified correctly), $\Lambda = .981$, $\chi^2(2, N = 357) = 6.901$, $p = .032$ for conceptual counting knowledge (with 100% of the TA children but non of the LA and AD children classified correctly), $\Lambda = .919$, $\chi^2(2, N = 357) = 30.119$, $p = .0005$ for number comparison tasks (with 99% of the TA children, non of the LA children and 18.8% of the AD children classified correctly), indicating that overall the neo-Piagetian-kindergarten predictors differentiated among the AD, LA and TA group.

In addition, Table 6 gives an overview of the Receiver Operating Characteristic (ROC) curves based on the 'standardized scores' as coordinates to determine the sensitivity and specificity to identify AD versus non-AD children in elementary school. Magnitude comparison was the most sensitive task, whereas procedural knowledge had the highest specificity as kindergarten predictor.

< Insert Table 6 about here >

Finally, in order to obtain an easier clinical and sample-independent interpretation, it was investigated if the use of clinical cut-off scores in kindergarten (at age 5 to 6) were consistent with the previous analyses on z-scores. The percentage of true positive AD children (at age 7 to 8) correctly detected in kindergarten (at age 5 to 6) by the combination of seriation, classification, procedural counting, conceptual counting and magnitude comparison predictors was 87.50%. There were 18.25% false positives or AD children not failing at age 5 to 6 on the kindergarten measures but developing AD (at age 7 to 8) in elementary school. None of the AD children in elementary school failed on the Piagetian seriation or classification tasks as preschooler in our sample. There were 12.50% of the AD children in elementary school with a clinical score on neo-Piagetian procedural counting tasks assessed in kindergarten. In addition, conceptual counting assessed in kindergarten detected 31.25% of the AD children in elementary school. Finally 43.75% of the AD children in elementary school already had a clinical score on magnitude comparison tasks in kindergarten.

In addition the clinical specificity was 89.40% in the TA group. In our dataset 0.15% of the TAs children scored below the clinical cutoff on procedural counting tasks in preschool and 5.95% scored below this cut off on conceptual counting tasks. No false negatives were found on seriation, classification and magnitude comparison tasks in kindergarten.

Research Question Four: Do children with persistent or inconsistent arithmetic abilities differ in kindergarten skills?

The research question of whether it was possible to find differences between children with persistent or inconsistent performances groups without confounding group differences in intelligence is addressed with a MANCOVA. We were especially interested in the differences between AD (persistent arithmetic disabilities) and DF1 (inconsistent arithmetic disabilities). In addition we were also interested in the differences between the TA (persistent typical achievers) and DF2 children (inconsistent low achievers)

From all participants 64.23% children appeared to have stable or persistent arithmetic abilities whereas 35.77% children had inconsistent abilities (DF1 and DF2) with fluctuating arithmetic test scores.

A MANCOVA was conducted to investigate the differences between the AD, TA, DF1 and DF2 group on five dependent variables: procedural counting knowledge, conceptual counting knowledge, seriation, classification, magnitude comparison (assessed in kindergarten) and intellectual abilities as covariate. The MANCOVA was significant on the multivariate level for the covariate ($F(5, 411) = 12.506, p = .0005$) but also for the group $F(15, 1134.991) = 4.341, p = .0005$. The means and standard deviations of the dependent variables for the two instable groups are shown in Table 7.

< Insert Table 7 about here >

Post hoc analyses, using the Bonferroni method, revealed that the children with inconsistent arithmetic disabilities (DF1) performed significantly better than the children with persistent arithmetic disabilities (AD) on neo-Piagetian magnitude comparison tasks (a -index versus b -index). The children with persistent typical achievement (TA) were significantly better than the children with inconsistent low achievement (DF2) on Piagetian classification tasks (a -index versus b -index).

Discussion

Several cognitive antecedents have been suggested as factors that play a role in the development of initial arithmetic performance and eventually as early markers for arithmetic disabilities. In 1941, Piaget postulated that logical abilities are conditional for the development of arithmetic (Piaget & Szeminska, 1941). However, up till now, the debate on the value of seriation and classification as Piagetian abilities remains unsolved (e.g., Grégoire, 2005; Lourenço & Machado, 1996; Nunes et al., 2006; Stock et al., 2007, in press). Besides these Piagetian logical abilities, neo-Piagetian researchers focused on the importance of procedural and conceptual counting knowledge in the development of arithmetic performance (e.g., Aunola et al., 2004). Finally since Landerl, Bevan, and Butterworth (2004) suggested that the core problem of arithmetic disabilities might be a deficit in number sense, it might be interesting to explore if magnitude comparison can be used as early marker for arithmetic disabilities (e.g., Durand et al., 2005). Since there is not yet a consensus regarding which of these type of kindergarten predictors are uniquely associated with early responses to formal arithmetic instruction in first and second grade, this study used a combination of Piagetian seriation and classification tasks with other measures of conceptual and procedural counting knowledge as well as a magnitude comparison task. Based on the possible group differences in intelligence among children that might present a confound in the assessment of arithmetic performance, in this study IQ was added as covariate. The combination of these measures using a large sample of children makes this study relatively rare in this type of research.

Children were grouped based on their arithmetic ability into one of five groups: persistent arithmetic disabilities (AD), persistent low achievement (LA), persistent typical achievement (TA), inconsistent arithmetic disabilities (DF1), and inconsistent low achievement (DF2). In line with Murphy et al. (2007) the persistent AD and LA groups were separated based on ‘potential severity’. AD children corresponded to the restrictive selection criteria of an arithmetic disability scoring successively $\leq$ the 10th percentile on arithmetic tests, whereas LA children were low achieving in arithmetic scoring successively between percentile 11-25. TA children had age appropriate achievement in grade 1 and 2, scoring $>$ percentile 25 on arithmetic tests. Moreover in line with Fletcher et al. (2005) and Murphy et al. (2007) the presented study separated ‘persistent’ and ‘inconsistent’ achievers, addressing the ‘status models’ that attempt to identify children based on an assessment at one single testing point (DF1, DF2) and the ‘fluctuating model’ or the results of children classified based on their arithmetic achievement scores for at least 2 years (AD, TA). DF1 children had ‘severe but non-persistent arithmetic difficulties’ performing $\leq$ the 10th percentile in one grade and above the 25th percentile in the other grade. The DF2 children had ‘mild but non-persistent arithmetic difficulties’ scoring between pc 11-25 in one grade and above the 25th percentile in the other grade. The DF1 and DF2 children may be excluded from prior studies that use persistence as a basis for inclusion in arithmetic ability groups. In short, we explored if the kindergarten characteristics of children varied as a function of the cut-off (AD versus LA) and persistence (AD versus DF1) as a basis for inclusion in arithmetic ability groups.

This study revealed that nearly one-fifth of the variance in arithmetic reasoning in *grade 1* (age 6 to 7) could be explained by assessing the performances of children in kindergarten (age 5 to 6) two year earlier. Especially the Piagetian seriation and classification tasks and the neo-Piagetian procedural counting and magnitude comparison tasks were beneficial as kindergarten (age 5 to 6) predictors. Nearly one-tenth of the variance in

numerical facility in first grade (age 6 to 7) could be modelled by assessing the kindergarten performances. Especially the Piagetian classification task was a significant predictor.

In addition one-tenth of the variance in arithmetic abilities and about one-twentieth of the numerical facility in *grade 2* (age 7 to 8) could be modelled by kindergarten (age 5 to 6) performances two years earlier. Especially the classical Piagetian type seriation tasks and the neo-Piagetian procedural counting tasks were beneficial to model the variance in arithmetic reasoning, whereas the neo-Piagetian procedural counting task was beneficial as kindergarten predictor for the variance in numerical facility.

These findings stress the need for a model on the development of arithmetic abilities that includes seriation and classification as Piagetian logical abilities, but also procedural counting knowledge and magnitude comparison as neo-Piagetian insights (age 5 to 6). This model could serve as a framework for a better understanding of the development of arithmetic abilities in grade 1 and 2. The longer the follow-up period, the weaker the prediction from kindergarten (age 5 to 6) performances, but the assessment of preparatory arithmetic abilities in kindergarten seemed still to have a predictive value for the prediction of arithmetic abilities even two years later. However the studies in grade 1 and 2 have not consistently pointed out one and the same predictor. It makes it clear that it is not good practice to look for a single deficient arithmetic ability. These conclusions indicate that it is important to build models that include the several markers for arithmetic development and then investigate the interactions between those components.

Moreover in line with studies of Geary et al. (2007) and Murphy et al. (2007) the present study highlighted differences in kindergarten skills based on the definition of arithmetic disabilities. Concerning performances on the Piagetian seriation tasks in kindergarten (age 5-6) no significant differences were found between the AD and LA children, but both groups performed significantly worse on seriation tasks than TA children

in kindergarten. However, there were significant differences between the AD and LA groups on magnitude comparison task in kindergarten, with children with AD having less developed preparatory arithmetic abilities than the LA-children. The significant differences in the kindergarten performances between AD and LA children on the magnitude comparison task strengthen the idea that this might be one of the core deficits in children with arithmetic disabilities (Butterworth, 2005; Gersten et al., 2005; Holloway & Ansari, 2008; Landerl et al., 2004). This supports the fact that results of low achievers (LA) cannot automatically be applied to children with arithmetic disabilities (AD), implying major consequences for the selection of samples in future research. It is hypothesized that children with arithmetic disabilities (AD) show qualitatively different skills than children who are low achieving in arithmetic (LA), and this hypothesis is confirmed by the recent findings of Mazzocco, Devlin, and Mc Kenny (2008) in older children and by studies of Geary et al. (2007). They also found that children with arithmetic disabilities showed qualitatively different profiles in fact retrieval performances when compared to typically achieving children, whereas the differences between low achieving and typically achieving children were of a quantitative turn. Since the criteria used to define children with arithmetic disabilities (AD) in the current study approach criteria used in clinical practice, this also implicates that clinicians have to be careful with conclusions of scientific studies that used a more lenient criterion.

The third purpose of this study was to classify children as persistent typical achiever (TA), persistent low achiever (LA) or children with persistent arithmetic disabilities (AD) based on the kindergarten preparatory abilities. The study revealed that seven out of eight children in grade 2 (at age 7 to 8) could be classified correctly into the arithmetic disabilities, low achieving or typically achieving group based on their kindergarten preparatory arithmetic abilities assessed two years earlier (at age 5 to 6). However less than one fifth of the children with arithmetic disabilities, no low achieving children but nearly all typically achievers could

be classified correctly. All kindergarten tasks contributed to the general prediction, but only the number comparison tasks contributed to a correct classification of children with arithmetic disabilities in grade 2 (age 7 to 8). These results showed that it was easier to screen the children who are 'not' at-risk than to detect the at-risk children based on their kindergarten abilities. The Receiver Operating Characteristic (ROC) curves revealed that magnitude comparison was the most sensitive kindergarten task, whereas procedural knowledge had the highest specificity as kindergarten predictor. Moreover, based on clinical cut-off scores, seven out of eight children with arithmetic disabilities (at age 7 to 8) could already be correctly detected in kindergarten (at age 5 to 6). One out of eight AD-children failed on neo-Piagetian procedural counting tasks and the conceptual counting tasks detected about one out of three AD-children in grade 2 already in kindergarten (age 5 to 6). Finally more than two fifth of the AD-children (age 7 to 8) already had a clinical score on magnitude comparison tasks in kindergarten (age 5 to 6). In addition the clinical specificity of the five kindergarten predictors was about nine tenth. No false negatives were found on magnitude comparison as kindergarten task. These results underline that it is important to include magnitude comparison measures but also procedural and conceptual counting tasks in assessment batteries that aim to prospectively detect kindergartners (age 5 to 6) who are at risk for arithmetic disabilities (AD) in the first two years of elementary schools.

The fourth purpose of this study was to look for differences in kindergarten performances (age 5 to 6) between children with 'persistent' or 'inconsistent' arithmetic abilities. In line with the 'status model' our data revealed that about two thirds of the children had a persistent arithmetic achievement (AD, LA, TA). However, in line with the 'fluctuation model' (Fletcher et al., 2005), there was a big amount of instability in the identification of class members because test scores fluctuated with repeated testing in on third of the children (DF1, DF2). Similar fluctuations were described by Fletcher et al. (2005).

Moreover, the group with inconsistent arithmetic disabilities (DF1) differed significantly on magnitude comparison tasks but not on seriation tasks in kindergarten (age 5 to 6) from the group with persistent arithmetic disabilities (AD). In addition, the children with inconsistent low achievement (DF2) differed significantly from the persistent typical achievers (TA) on classification but not on seriation tasks in kindergarten. These results showed that different conclusions could be drawn when using children with inconsistent (DF1) and persistent (AD) arithmetic disabilities together in one group. The use of DF1 and AD children together in one sample may conflate children with severe arithmetic disabilities (AD) and children with fluctuating arithmetic performances. Moreover, the use of DF1 children challenges intervention studies, since the spontaneous fluctuations in arithmetic performances might mask or overemphasize differences between children with arithmetic disabilities and average achievers. Also the use of DF2 children in the recruitment of clinical control samples might be a problem.

The results of the current study have to be interpreted with care since some other possible powerful predictors for arithmetic disabilities were not taken into account. Geary, Bailey and Hoard (in press) assessed speed and accuracy with which children can identify and process quantities. Based on children's performance in first grade, they were able to detect two out of three children identified as having arithmetic disabilities at the end of third grade. Perhaps the inclusion of the speed factor was responsible for the better classification results obtained compared to our ROC-analyses. In addition, several authors stressed the importance of executive functions (e.g., Mazzocco & Kover, 2007; Van der Sluis, de Jong, & van der Leij, 2007), working memory (e.g., Bull, Espy, & Wiebe, 2008; Geary & Widaman, 1992; Passolunghi, Mammarella, & Altoe, 2008; Passolunghi & Siegel, 2004; Ricken & Fritz, 2006) and attention (Marzocchi, Lucangeli, De Meo, Fini, & Cornoldi, 2002) in the development of arithmetic (dis)abilities. It even might be so that working memory is related

to some of the indicators proposed in the study as well. Future research is needed on speed and accuracy, EF and working memory in the differentiation of AD, LA and DF1 children. Finally, context variables such as home and school environment, learning packages and parental involvement (e.g., Reusser, 2000) should be included in order to obtain a complete overview of the arithmetic development of these children. These limitations indicate that only a part of the picture was investigated, so the results of the study have to be interpreted with care. Yet the study included an in-depth assessment of the arithmetic performances of the children during three consecutive years. Since few large-scale studies have been done (Porter, 1998), the size of the group of children that was assessed in this study strengthens the generalizability of the results. Finally, the use of lenient as well as restrictive criteria and the inclusion of an assessment of the intellectual capacities of the children further empower this study.

However important implications and challenges for future research can be withdrawn. First of all, researchers should be careful with selection criteria when composing research samples. Restrictive cut-off criteria (scores $\leq$ the 10th percentile) and the implementation of the Resistance To Instruction criterion (clinical scores in at least two consecutive years although task specific instruction was given between both measuring points) seem indicated. Secondly, we can only detect kindergarten predictors of arithmetic performance in grade 1 and 2 with a hybrid model combining Piagetian type tests, neo-Piagetian counting tasks and number comparison screening in kindergarten. We propose to integrate seriation and classification in a model that also includes procedural and conceptual counting knowledge and magnitude comparison. It is recommended that we should not try to assess ‘one’ kindergarten predictor, but rather look for a set of markers when we aim to model variance in later arithmetic abilities. Moreover, researchers should be aware that a large part of the variance in arithmetic abilities cannot yet be explained. Additional research is needed on

larger context variables and on other underlying factors as well as on the Resistance to Instruction (RTI) to build such a model. The third and perhaps most important educational challenge for the future is to accommodate the need for intervention studies (Gersten et al., 2005). Early intervention can remediate difficulties and alter children's arithmetic development (Fuchs & Fuchs, 2001), but if we really want to screen children on their early arithmetic abilities in order to prevent the development of arithmetic disabilities, it will be important to focus on assessments that are directly related to instruction. Only response to instruction studies can sufficiently serve that goal (Fletcher, Coulter, Reschly, & Vaught, 2004).

In conclusion, some of the variance in arithmetics could be explained in the three-year longitudinal study by the kindergarten performances on Piagetian type tasks as well as on neo-Piagetian procedural counting and magnitude comparison tasks. Moreover, significant differences in the performances on magnitude comparison can be found between children with persistent arithmetic disabilities (AD) and persistent low achievers (LA), but also between children with persistent (AD) and inconsistent (DF1) arithmetic disabilities. Furthermore, based on the clinical cut-off scores on number comparison, procedural counting and conceptual counting tasks in kindergarten (at age 5-6) seven out of eight kindergartners at risk for persistent arithmetic disabilities (AD) in grade 1 and 2 could be detected.

References

- American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders. Fourth edition, Text Revision*. Washington, DC: Author.
- Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental Dynamics of Math Performance From Preschool to Grade 2. *Journal of Educational Psychology, 96*, 699-713.
- Baroody, A. J. (1992). Remedying common counting difficulties. In J. Bideaud, C. Meljac, & J.-P. Fischer (Eds), *Pathways to number. Children's developing numerical abilities* (307-324). Hillsdale, NJ: Erlbaum.
- Barrouillet, P., & Lepine, R. (2005). Working memory and children's use of retrieval to solve addition problems. *Journal of Experimental Child Psychology, 91*, 183-204.
- Baudonck, M., Debusschere, A., Dewulf, B., Samyn, F., Vercaemst, V., & Desoete, A. (2006). *De Kortrijkse Rekentest Revision KRT-R. [The Kortrijk Arithmetic Test Revision KRT-R]*. Kortrijk: CAR Overleie.
- Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities, 38*, 333-339.
- Blöte, A. W., Lieffering, L. M., & Ouweland, K. (2006). The development of many-to-one counting in 4-year-old children. *Cognitive Development, 21*, 332-348.
- Bull, R., Espy, K.A., & Wiebe, S.A. (2008). Short-term memory, working memory and executive functioning in preschoolers: Longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology, 33*, 205-228.
- Butterworth, B. (2005). The development of arithmetical abilities. *Journal of Child Psychology and Psychiatry, 46*, 3-18.
- Chard, D.J., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using

- measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention*, 30, 3-14.
- Coleman, M.R., Buysse, V., & Neitzel, J. (2006). *Recognition and response. An Early Intervening system for Young children at-risk for Learning Disabilities. Research synthesis and recommendations*. UNIC FPG Child Development Institute.
- Dehaene, S. (1997). *The Number Sense*. Oxford University Press. Dehaene, S., Bossini, S., & Giraux, P. (1993). The Mental Representation of Parity and Number Magnitude. *Journal of Experimental Psychology: General*, 122, 371-396.
- Desoete, A. (2006). Validiteitsonderzoek met de TEDI-MATH [Validity research on the TEDI-MATH]. *Diagnostiek-wijzer*, 9, 140-157.
- Desoete, A. (2007a). *Students with mathematical disabilities in Belgium: from definition, classification and assessment to STICORDI devices*. (pp. 181-222). In T.E. Scruggs & M.A. Mastropieri (Eds.), *Advances in Learning and Behavioral Disabilities*, Vol. 20. International Perspectives Amsterdam & Oxford: Elsevier Press..
- Desoete, A. (2007b). De TEDI-MATH. De plaats van de Tedi-Math in de diagnostiek van dyscalculie in Vlaanderen. [The value of the Tedi-Math in the assessment of mathematical learning disabilities in Flanders]. *Caleidoscoop*, 19 (6), 6-19.
- Desoete, A. (2008). Co-Morbidity in Mathematical Learning Disabilities: Role or Exception? *The Open Rehabilitation Journal*, 1, 15-26.
- Desoete, A., Ceulemans, A., Roeyers, H., & Huylebroeck, A. (2008). Subitizing or counting as possible screening variables for learning disabilities in mathematics education or learning? *Educational Research Review*, doi:10.1016/j.edurev.2008.11.003.
- Desoete, A., & Grégoire, J. (2007). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences*, 16, 351-367.

- Desoete, A., Roeyers, H., & De Clercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities*, 37, 50-61.
- Desoete, A., Roeyers, H., Schittekatte, M., & Grégoire, J. (2006). Dyscalculie-gevoelige kennis en vaardigheden in het basisonderwijs in Vlaanderen, Wallonië en Frankrijk. [Dyscalculia-sensitive knowledge and skills in elementary education in the Flemish and French speaking part of Belgium and in France]. *Pedagogische Studieën*, 83, 105-121.
- De Vos, T. (1992). *TTR. Tempotest rekenen [Arithmetic number fact test]*. Lisse: Swets & Zeitlinger.
- Dickerson Mayes, S., Calhoun, S.L., Bixler, E.O., & Zimmerman, D.N. (2009). IQ and neuropsychological predictors of academic achievement. *Learning and Individual Differences*, 19, 238-241.
- DiPerna, J.C., Lei, P-W, & Reid, E.E. (2007). Kindergarten Predictors of Mathematical Growth in the Primary Grades: An Investigation Using the Early Childhood Longitudinal Study – Kindergarten Cohort. *Journal of Educational Psychology*, 99, 369-379.
- Dowker, A. (2005). *Individual Differences in Arithmetic. Implications for Psychology, Neuroscience and Education*. New York: Psychology Press.
- Dumont, J.J. (1994). *Leerstoornissen I. Theorie en model (nieuwe uitgave) [Learning disabilities I. Theory and model (new edition)]*. Rotterdam: Lemniscaat.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Egel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, 43, 1428-1446.
- Durand, M., Hulme, C., Larkin, R., & Snowling, M. (2005). The cognitive foundations of

- reading and arithmetic skills in 7- to 10-year olds. *Journal of Experimental Child Psychology*, 91, 113-136.
- Fischer B, Gebhardt, C., & Hartnegg, K. (2008). Subitizing and visual counting in children with problems acquiring basic arithmetic skills. *Optom Vis Dev* 39, 24-29.
- Fletcher, J.M., Coulter, W.A., Reschly, D.J., & Vaughn, S. (2004). Alternative approaches to the definition and identification of learning disabilities. Some questions and answers. *Annals of Dyslexia*, 54, 304-331.
- Fletcher, J.M., Denton, C., & Francis, D.J. (2005). Validity of Alternative Approaches for the Identification of Learning Disabilities : Operationalizing Unexpected Underachievement. *Journal of Learning Disabilities*, 38, 545-552.
- Frank, A. R. (1989). Counting skills: A foundation for early mathematics. *Arithmetic Teacher*, 37, 14-17.
- Fuchs, L.S., & Fuchs, D. (2001). Principles for the prevention and intervention of mathematics difficulties. *Learning Disabilities Research and Practice*, 16, 85-95.
- Fuchs, L.S., Fuchs, D., Compton, D.L., Bryant, J.D., Hamlett, C.L., & Seethaler, P.M. (2007). Mathematics Screening and Progress Monitoring at First Grade: Implications for Responsiveness to Intervention. *Exceptional Children*, 73, 311-330.
- Fuson, C.K. (1988). *Children's counting and concepts of number*. New York: Springer-Verlag.
- Gallistel, C. R., & Gelman, R. (1990). The what and how of counting. *Cognition*, 34, 197-199.
- Geary, D.C. (1994). *Children's Mathematical Development*. Washington, D.C.: American Psychological Association.
- Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, 37, 4-15.

- Geary, D. C., Bailey, D. H., & Hoard, M. K. (in press). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: the Number Sets Test. *Journal of Psychoeducational Assessment*.
- Geary, D. C., Bow-Thomas, C. C., & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, 54, 372-391.
- Geary, D.C., & Hoard, M.K. (2005). Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. In J.I.D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 253-268). New York: Psychology Press.
- Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, 78, 1343-1359.
- Gersten, R., Jordan, N.C., & Flojo, J.E. (2005). Early identification and intervention for students with mathematics difficulties. *Journal of Learning disabilities*, 38, 293-304.
- Geary, D. C., & Widaman, K. F. (1992). Numerical cognition: On the convergence of componential and psychometric models. *Intelligence*, 16, 47-80.
- Gelman, R. (1990). First principles organize attention to and learning about relevant data: number and animate-inanimate distinctions as examples. *Cognitive Science*, 14, 79-106.
- Gelman, R., & Galistel, C. R. (1978). *The child's understanding of number*. Cambridge, MA: Harvard University Press.
- Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *Journal of Special Education*, 44, 18-28.
- Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and Intervention for

- Students With Mathematics Difficulties. *Journal of Learning Disabilities*, 38, 293-304.
- Gevers, W., Lammertyn, J., Notebaert, W., Verguts, T. & Fias, W. (2006). Automatic response activation of implicit spatial information: Evidence from the SNARC effect *Acta Psychologica*, 122, 221-233.
- Ghesquière, P., & Ruijsenaars, A. (1994). *Vlaamse normen voor studietoetsen Rekenen en technisch lezen lager onderwijs. [Flemish norms for schooltests on Mathematics and Technical Reading in elementary school]*. Leuven: K.U.L.-C.S.B.O.
- Grégoire, J. (2001). *L'Evaluation clinique de l'intelligence de l'enfant : théorie et pratique du WISC-III*. (2nd ed.) (pp. 259). Sprimont : Mardaga.
- Grégoire, J. (2005). Développement logique et compétences arithmétiques. Le modèle piagétien est-il toujours actuel ? In M. Crahay, L. Verschaffel, E. De Corte & J. Grégoire. *Enseignement et apprentissage des mathématiques*. (pp.57-77). Bruxelles : De Boeck.
- Grégoire, J., Noel, M., & Van Nieuwenhoven (2004). *TEDI-MATH*. TEMA: Brussel/Harcourt: Antwerpen.
- Griffin, S. (2004). Building number sense with Number Worlds: a mathematics program for young children. *Early Childhood Research Quarterly*, 19, 173-180.
- Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: prevalence and demographic features. *Developmental Medicine and Child Neurology*, 38, 25-33.
- Hammill, D. D. (1990). On defining learning disabilities: An emerging consensus. *Journal of Learning Disabilities*, 23, 74-84.
- Hanley, T.V. (2005). Commentary on early identification and interventions for students with mathematical difficulties: Make sense – *Do the math*. *Journal of Learning Disabilities*, 38, 346-349.

- Hannula, M. M., & Lehtinen, E. (2005). Spontaneous focussing on numerosity and mathematical skills of young children. *Learning and Instruction, 15*, 237-256.
- Holloway, I. D., & Ansari, D. (2008). Mapping numerical magnitudes into symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology*, doi:10.1016/j.jecp.2008.04.001
- Johansson, B. S. (2005). Number-word sequence skill and arithmetic performance. *Scandinavian Journal of Psychology, 46*, 157-167.
- Jordan, H.C., Hanich, L.B., & Kaplan, B. (2003). A Longitudinal Study of Mathematical Competencies in Children with Specific Mathematics Difficulties Versus Children with Comorbid Mathematics and Reading Difficulties. *Child Development, 74*, 834-850.
- Jordan, N. C., Kaplan, D., Olah, L. N., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child development, 77*, 153-175.
- Jordan, N.C., Levine, S.C., & Huttenlocher, J. (1995). Calculation abilities of young children with different patterns of cognitive functioning. *Journal of Learning Disabilities, 28*, 53-67.
- Kavale, K.A., & Spaulding, L.S. (2008). Is response to Intervention Good Policy for Specific Learning Disability? *Learning Disabilities Research & Practice, 23*, 169-179.
- Kaufman, E.L., Lord, M.W., Reese, T.W., & Volkman, J. (1949). The discrimination of visual number. *American Journal of Psychology, 62*, 498-525.
- Kingma, J. (1984). Traditional intelligence, Piagetian Tasks, and Initial Arithmetic in kindergarten and primary-school grade-one. *Journal of Genetic Psychology, 145*, 49-60.

- Klauer, K. J. (1992). In Mathematik mehr leistungsschwache Madchen, im Lesen und Rechtschreiben mehr leistungsschwache Junden? *Zeitschrift fur Entwicklungspsychologie und Padagogische Psychologie*, 26, 48-65.
- Koontz, K.L., & Berch, D.B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited arithmetic learning disabled children. *Mathematical Cognition*, 2, 1-23.
- Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities*, 7, 46-59.
- Kort, W., Schittekatte, M., Compaan, E. L., Bosmans, M., Bleichrodt, N., Vermeir, G., Resing, W. & Verhaeghe, P. (2002). *Handleiding WISC-III^{NL} [Manual of the WISC-III^{NL}]*. The Psychological Corporation. NIP Dienstencentrum: Londen, Amsterdam.
- Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: a study of 8-9 year old students. *Cognition*, 93, 99-125.
- Laski, E. V., & Siegler, R. S. (2007). Is 27 a big number? Correlational and causal connections among numerical categorization, number line estimation and numerical magnitude comparison. *Child Development*, 78, 1723-1743.
- Le Corre, M., Van de Walle, G., Brannon, E. M., & Carey, S. (2006). Re-visiting the competence/performance debate in the acquisition of the counting principles. *Cognitive Psychology*, 52, 130-169.
- Le Fevre, J.-A., Smith-Chant, B. L., Fast, L., Skwarchuk, S.-L., Sargla, E., Arnup, J. S., Penner-Wilger, M., Bisanz, J., & Kamawar, D. (2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through Grade 2. *Journal of Experimental Child Psychology*, 93, 285-303.
- Lourenço, O., & Machado, A. (1996). In Defense of Piaget's Theory: A Reply to 10 Common Criticisms. *Psychological Review*, 103, 143-164.
- Marzocchi, G. M., Lucangeli, D., De Meo, T., Fini, F., & Cornoldi, C. (2002). The

- Disturbing Effect of Irrelevant Information on Arithmetic Problem Solving in Inattentive Children. *Developmental Neuropsychology*, 21, 73-92.
- Mazzocco, M. M. M. (2001). Math learning disability and math LD subtypes: Evidence from studies of Turner syndrome, Fragile X syndrome and Neurofibromatosis type 1. *Journal of Learning Disabilities*, 34, 520-533.
- Mazzocco, M.M.M. (2005). Challenges in Identifying Target Skills for Math Disability Screening and Intervention. *Journal of Learning Disabilities*, 38, 318-323.
- Mazzocco, M.M.M., Devlin, K.T., & McKenney, S.J. (2008). Is it a fact? Timed arithmetic performance of children with mathematical learning disabilities (MLD) varies as a function of how MLD is defined. *Developmental Neuropsychology*, 33, 318-344.
- Mazzocco, M. M. M., & Kover, S. T. (2007). A Longitudinal Assessment of Executive Function Skills and Their Association with Math Performance. *Child Neuropsychology*, 13, 18-45
- Mazzocco, M.M.M., & Myers, G.F. (2003). Complexities in Identifying and Defining Mathematics Learning Disability in the Primary School-Age Years, *Annals of Dyslexia*, 53, 218-253.
- Mazzocco M.M.M., & Thompson RE. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research and Practice*, 20, 142-155.
- Muldoon, K. P., Lewis, C., & Francis, B. (2007). Using cardinality to compare quantities: the role of social-cognitive conflict in early numeracy. *Developmental science*, 10, 694-711.
- Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B., & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, 40, 458-478.

- Nan, Y., Knösche, T.R., & Luo, Y-J. (2006). Counting in everyday life: discrimination and enumeration. *Neuropsychologica*, 44, 1103-1113.
- Nunes, T., & Bryant, P. (1996). *Children doing mathematics*. Oxford: Blackwell Publishers.
- Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, A., Gardner, A., & Carraher, J. (2006). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental psychology*, 00, 1-21.
- Pasnak, R., Cooke, W. D., & Hendricks, C. (2006). Enhancing academic performance by strengthening class-inclusion reasoning. *The Journal of Psychology*, 140, 603-613.
- Passolunghi, M. C., Mammarella, I. C., & Altoe, G. (2008). Cognitive abilities as precursors of the early acquisition of mathematical skills during first through second grades. *Developmental Neuropsychology*, 33, 229-250.
- Passolunghi, M.C., & Siegel, L.S. (2004). Working memory and access to numerical information in children with disability in mathematics. *Journal of Experimental Child Psychology*, 88, 348-367.
- Piaget, J. (1965). *The child's conception of number*. New York: Norton.
- Piaget, J., & Inhelder, B. (1959). *La genèse des structures logiques élémentaires. Classification et sériation [The development of basic logical structures. Classification and seriation]*. Neuchâtel France: Delachaux and Niestlé.
- Piaget, J., & Szeminska, A. (1941). *La genèse du nombre chez l'enfant*. Neuchâtel : Delanchaux et Niestlé.
- Porter, J. (1998). The understanding of counting in children with severe learning difficulties and nursery children. *British Journal of Educational Psychology*, 68, 331-345.
- Reeye, R.A., & Reynolds, F.J. (2004, July). *The nature of young children's mathematical (dis)abilities*. Poster presented at the International Society of the Study of Behavioral Development. Ghent.

- Reusser, K. (2000). Success and failure in school mathematics: effects of instruction and school environment. *European Child & Adolescent Psychiatry*, 9, II/7-II/26.
- Ricken, G., & Fritz, A. (2006). Working memory functions in children with different arithmetical performance at kindergarten. *Psychologie in Erziehung Und Unterricht*, 53, 263-274.
- Rousselle, L., & Noël, M.P. (2007). Basic numeric skills in children with mathematics learning disabilities : a comparison of symbolic vs non-symbolic number magnitude processing. *Cognition*, 102, 361-395.
- Ruijsenaars, A.J.J.M., Van Luit, H., & E.C.D.M. Van Lieshout (Eds.) (2004). *Rekenproblemen en dyscalculie. [Arithmetic problems and dyscalculia]*. Lemniscaat: Rotterdam.
- Shalev, R.S. (2004). Developmental Dyscalculia. *Journal of Child Neurology*, 10, 766-771.
- Shalev, R. S., Manor, O., & Gross-Tsur, V. (2005). Developmental dyscalculia: a prospective six-year follow-up. *Developmental Medicine & child Neurology*, 47, 121-125.
- Sophian, C. (1992). Learning about numbers: Lessons for mathematics education from preschool number development. In J. Bideaud, C. Meljac & J.-P. Fischer (Eds), *Pathways to number. Children's developing numerical abilities* (pp. 19-40). Hillsdale, NJ: Erlbaum.
- Stock, P., Desoete, A., & Roeyers, H. (2006). Focussing on mathematical disabilities: a search for definition, classification and assessment (pp. 29-62). In Soren V. Randall (Ed.) *Learning Disabilities New Research Nova Science: Hauppauge, NY*.
- Stock, P., Desoete, A., & Roeyers, H. (2007). Early markers for arithmetic difficulties. *Educational and Child Psychology*, 24, 28-39.
- Stock, P., Desoete, A., & Roeyers, H. (in press). Predicting Arithmetic Abilities: The Role of

Preparatory Arithmetic Markers and Intelligence. *Journal of Psychoeducational Assessment*

- Teisl, J.T., Mazzocco, M.M.M., & Myers, G.F. (2001). The utility of Kindergarten Teacher Ratings for Predicting Low Academic Achievement in First Grade. *Journal of Learning Disabilities*, 34, 286-293.
- Van de Rijt, B. A. M., & Van Luit, J. E. H. (1999). Milestones in the development of infant numeracy. *Scandinavian Journal of Psychology*, 40, 65-71.
- Van der Sluis, S., de Jong, P. F., & van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, 35, 427-449
- Vaughn, S., & Fuchs, L.S. (2003). Redefining learning disabilities as inadequate response to instruction: The promise and potential problems. *Learning Disabilities Research & Practice*, 18, 137-143.
- Von Aster, M.G. & Shalev, R.S. (2007) Number development and developmental dyscalculia. *Developmental Medicine & Child Neurology*, 49 , 868–873.
- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children (3rd ed.)*. San Antonio: Psychological Corporation.
- World Health Organisation, 1992. *The ICD-10 Classification of mental and behavioural disorders. Clinical descriptions and diagnostic guidelines*. Geneva: WHO.
- Wynn, K. (1990). Children's understanding of counting. *Cognition*, 36, 155-193.
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, 24, 220-251.
- Xu, F. (2003). Numerosity discrimination in infants: Evidence for two systems of representations. *Cognition*, 89, B15-B25.
- Xu, F., & Spelke, E.S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74, 1-11.

Zhou, X., Chen, C., Chen, C., & Dong, Q. (2008). Holistic or compositional representation of two-digit numbers? Evidence from distance, magnitude, and SNARC-effects in a number-matching task. *Cognition*, 106, 1525-1536.

Table 1

Descriptive information based on the subsamples

	AD	LA	TA	DF1	DF2	F
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	(4, 459)
SES: Number of years of scholarship of the mother	14.86 (2.44)	14.27 (2.60)	14.94 (2.35)	14.38 (2.48)	15.14 (2.59)	1.179 <i>p</i> = .319
SES: Number of years of scholarship of the father	13.57 (2.77)	14.12 (3.44)	14.65 (2.91)	14.30 (2.58)	14.03 (2.19)	0.996 <i>p</i> = .409
Intelligence (TIQ)	96.12 (13.99)	100.44 (16.28)	101.41 (12.75)	97.96 (15.22)	99.67 (15.22)	1.399 <i>p</i> = .233

Note. AD = Persistent Arithmetic Disabilities Group; LA = Persistent Low Achieving Group;

TAs = Persistent Typically Achieving Group; DF1 = Inconsistent Arithmetic Disability

group; DF 2 = Inconsistent Low Achievers

Table 2

Correlations among kindergarten predictors and intelligence

	Conceptual				Magnitude
	Intelligence	Counting	Seriation	Classification	Comparison
Procedural Counting	.290**	.296**	.197**	.087	.156**
Conceptual Counting	.219**	-	.231**	.185**	.076
Seriation	.241**	-	-	.112*	.102*
Classification	-.003	-	-	-	.093*
Magnitude Comparison	.116*	-	-	-	.093*

Note. ** $p < .0005$. * $p < .05$

Table 3

Prediction of Arithmetic Abilities in Grade 1 and 2 from Preparative Abilities

Kindergarten abilities (age 5 to 6)	Arithmetic abilities Grade 1 (age 6 to 7)				Numerical facility (age 6 to 7)			
	Unstand. Coeff.	β	t	P	Unstand. Coeff.	β	T	p
Constant	59.088		48.076	.000	9.097		35.558	.000
Procedural counting	4.871	.164	3.626	.000*	.531	.090	1.884	.060
Conceptual counting	3.045	.100	2.178	.030	.146	.024	.499	.618
Seriation	6.125	.210	4.769	.000*	.677	.117	2.513	.012
Classification	4.433	.157	3.609	.000*	.909	.163	3.529	.000*
Magnitude Comparison	3.958	.121	2.805	.005*	.656	.101	2.217	.027
Kindergarten abilities (age 5 to 6)	Arithmetic abilities Grade 2 (age 7 to 8)				Fast fact retrieval Grade 2 (age 7 to 8)			
	Unstand Coeff	β	t	p	Unstand Coeff	β	t	p
Constant	62.391		49.184	.000	14.437		25.875	.000
Procedural counting	4.957	.168	3.548	.000*	1.735	.138	2.827	.005
Conceptual counting	3.637	.109	2.275	.023	.812	.063	1.276	.203

Seriation	3.637	.125	2.723	.007*	1.296	.105	2.207	.028
Classification	2.554	.091	1.999	.046	-.962	-.081	-1.717	.087
Magnitude Comparison	1.879	.058	1.281	.201	-.275	-.020	-.428	.669

Note. * $p \leq .008$ after Bonferonni adjustment

Table 4

Means, Standard Deviations, F-Values and Standardised Canonical Discriminant Function

Coefficients of the Predictors for the Three Groups of Persistent Achievers

Kindergarten					Standard.
Predictors	AD	LA	TA	F	canonical
	M (SD)	M (SD)	M (SD)	(2, 340)	discriminant
					function
					coefficients
Procedural Counting	-0.47 (0.92)	0.16 (0.96)	0.28 (0.92)	3.735*	.206
Conceptual Counting	-0.38 (0.82)	0.24 (0.74)	0.24 (0.91)	2.785*	.120
Seriation	-0.58 _b (1.10)	-0.52 _b (1.08)	0.24 _a (0.88)	11.920**	.543
Classification	-0.43 (0.98)	0.29 (1.05)	0.19 (0.96)	3.336*	.135
Magnitude Comp.	-0.76 _b (1.77)	-0.05 _a (1.11)	0.19 _a (0.52)	15.395**	.690

Note. AD = Persistent Arithmetic Disabilities Group; LA = Persistent Low Achieving Group;

TAs = Persistent Typically Achieving Group; Comp = comparison * $p < .05$ ** $p \leq .008$ after

Bonferroni adjustment. TA children (_a) were better than AD children (_b) on seriation and

magnitude comparison tasks.. TA children were better than LA children on seriation tasks (_{ab})

but not on magnitude comparison tasks (_{aa}). LA children were better than AD children on

magnitude comparison tasks (_{ab}) but not on seriation tasks (_{bb}).

Table 5

*Percentages of Observed and Predicted Group Membership of the Persistent Achievers
Based on the Discriminant Function*

		Predicted Group Membership		
		AD	LA	TA
Observed Group Membership	AD	18.8 %	0 %	81.3 %
	LA	11.1 %	0 %	88.9 %
	TA	1.0 %	0 %	99.0 %

Note. AD = Persistent Arithmetic Disabilities Group; LA = Persistent Low Achieving Group;
TAs = Persistent Typically Achieving Group

Table 6

ROC Curve

	Area under the Curve (AUC)	Standard Error (SE)	Confidence Interval (CI)	<i>p</i>
Procedural Counting	.721	.053	.617-.824	.003
Conceptual Counting	.717	.061	.597-.837	.003
Seriation	.707	.075	.561-.853	.005
Classification	.645	.074	.501-.789	.051
Magnitude Comparison	.725	.071	.586-.865	.002

Table 7

Means, Standard Deviations and F-Values for the Two Groups of Inconsistent Achievers (DF1, DF2) compared with Persistent Achievers (AD and TA children)

	AD	DF1	DF2	TA	F
	M (SD)	M (SD)	M (SD)	M (SD)	(3, 415)
Proc. Counting	-0.47 (0.92)	0.38 (0.96)	0.24 (0.89)	0.28 (0.92)	3.265*
Conc. Counting	-0.38 (0.82)	0.34 (0.93)	0.39 (0.87)	0.24 (0.91)	2.166*
Seriation	-0.58 _b (1.10)	-0.14 _b (1.02)	0.19 _a (0.88)	0.24 _a (0.88)	4.852**
Classification	-0.43 _b (0.98)	0.11 (1.01)	-0.48 _b (0.89)	0.19 _a (0.96)	6.887**
Magnitude Comp.	-0.76 _c (1.77)	-0.17 _b (1.22)	-0.15 (1.17)	0.19 _a (0.52)	9.790**

Note. Proc.= Procedural; Conc.= Conceptual. Comp.=Comparison; AD = Persistent Arithmetic Disabilities Group; DF1 = Inconsistent Arithmetic Disability Group; DF 2 = Inconsistent Low Achievers; TA = Persistent Typically Achieving Group * $p < .05$ ** $p \leq .0005$. AD (_b)children did not differ significantly from DF (_b) children on seriation tasks, but they had worse magnitude comparison skills (_c). The TA (_a) did not differ significantly from DF2 (_a) on seriation and magnitude comparison skills. TA (_a) children were better than DF2 (_b)children on classification tasks.

Appendix A: Subtests and examples of test-items of the TEDI-MATH

Subtest	Content and example of item
1 Procedural knowledge (13 items, max 14 points)	<ul style="list-style-type: none"> - Counting as far as possible (without help=2 points; with starting help=1 point) - Counting forward to an upper bound (<i>“up to 9”</i>) - Counting forward to an upper bound (<i>“up to 6”</i>) - Counting forward from a lower bound (<i>“from 3”</i>) - Counting forward from a lower bound (<i>“from 7”</i>) - Counting forward from a lower bound to an upper bound (<i>“from 5 up to 9”</i>) - Counting forward from a lower bound to an upper bound (<i>“from 4 up to 8”</i>) - Counting forward (<i>“5 steps starting at 8”</i>) - Counting forward (<i>“6 steps starting at 9”</i>) - Count backward (<i>“from 7”</i>) - Count backward (<i>“from 15”</i>) - Count by step (<i>by 2</i>) - Count by step (<i>by 10</i>)
2. Conceptual knowledge (7 items, max 13 points)	<ul style="list-style-type: none"> - Counting linear pattern of items (<i>“how many rabbits are there? How many rabbits are there in total?”, and “how many rabbits are there if you start counting with this one. Why?”</i>). max 3 points - Counting linear pattern of items (<i>“how many lions are there? How many lions are there in total?”, and “how many lions have I hidden. Why?”</i>). max 3 point - Counting random pattern of items (<i>“how many turtles are there? How many turtles are there in total?”</i>). max 2 points - Counting random pattern of items (<i>“how many sharks are there? How many sharks are there in total?”</i>). max 2 points - Counting a heterogeneous set of items (<i>“how many animals are there in total?”</i>). max 1 point - Understanding of the cardinal (<i>“Can you put as much counters as there are on this paper?”</i>) max 1 point - Understanding of the cardinal (<i>“How many hat do I have in my hand, when all the snowman had a hat on this picture?”</i>) max 1 point
3. Logical operations on numbers (6 items, max 6 points)	<ul style="list-style-type: none"> -Seriation <i>“Sort the cards form the one with fewer trees to the one with the most trees. forgot this card” and “Can you put this card in the correct order?” and “I give you carts with numbers now. Do the same as with the trees. Start with the cart with smallest number and go on with the other carts”</i> max 3 points -Classification : <i>“Make groups with the cards that go together. Can you put them together in another way?” and “Make groups with these cards that go together”</i> max 3 points

4. Estimation of the size

(6 items, 6 points)

Comparison of dot sets (subitising): in preschool and grade 1

Where do you have most dots? Here or here? Show me.

1 dot versus 3 dots

3 dots versus 2 dots

4 dots versus 6 dots

7 dots versus 2 dots

7 dots versus 12 dots

15 dots versus 8 dots (see example)