Developmental Psychology

Basic Numerical Capacities and Prevalence of Developmental Dyscalculia: The Havana Survey

Vivian Reigosa-Crespo, Mitchell Valdés-Sosa, Brian Butterworth, Nancy Estévez, Marisol Rodríguez, Elsa Santos, Paul Torres, Ramón Suárez, and Agustín Lage

Online First Publication, September 12, 2011. doi: 10.1037/a0025356

Developmental Psychology

0012-1649/11/$12.00 DOI: 10.1037/a0025356

Basic Numerical Capacities and Prevalence of Developmental Dyscalculia: The Havana Survey

Vivian Reigosa-Crespo and Mitchell Valdés-Sosa
Cuban Centre for Neuroscience

Brian Butterworth
Institute of Cognitive Neuroscience, University College London

Nancy Estévez, Marisol Rodríguez, and Elsa Santos
Cuban Centre for Neuroscience

Paul Torres
Central Institute of Pedagogical Sciences

Ramón Suárez
National Institute of Hygiene and Epidemiology

Agustín Lage
Cuban Centre for Neuroscience

The association of enumeration and number comparison capacities with arithmetical competence was examined in a large sample of children from 2nd to 9th grades. It was found that efficiency on numerical capacities predicted separately more than 25% of the variance in the individual differences on a timed arithmetical test, and this occurred for both younger and older learners. These capacities were also significant predictors of individual variations in an untimed curriculum-based math achievement test and on the teacher scores of math performance over developmental time. Based on these findings, these numerical capacities were used for estimating the prevalence and gender ratio of basic numerical deficits and developmental dyscalculia (DD) over the grade range defined above (N = 11,652 children). The extent to which DD affects the population with poor ability on calculation was also examined. For this purpose, the prevalence and gender ratio of arithmetical dysfluency (AD) were estimated in the same cohort. The estimated prevalence of DD was 3.4%, and the male:female ratio was 4:1. However, the prevalence of AD was almost 3 times as high (9.35%), and no gender differences were found (male: female ratio = 1.07:1). Basic numerical deficits affect 4.54% of school-age population and affect more boys than girls (2.4:1). The differences between the corresponding estimates were highly significant ($\alpha <$ .01). Based on these contrastive findings, it is concluded that DD, defined as a defective sense of numerosity, could be a distinctive disorder that affects only a portion of children with AD.

Keywords: arithmetic fluency, enumeration, numerical magnitude comparison, developmental dyscalculia, prevalence

Current theories of typical cognitive development postulate that knowledge acquisition is based on a restricted set of core systems, defined as domain-specific representational primitives that lead and constrain the cultural learning (Spelke & Kinzler, 2007). According to Ansari and Karmiloff-Smith (2002) and Butterworth (2005), in the specific case of numbers, the focus of attention has shifted from higher level, school-like arithmetic skills to an analysis of lower level processes, in particular, capacities such as

Vivian Reigosa-Crespo, Department of Developmental Cognitive Neuroscience, Cuban Centre for Neuroscience; Mitchell Valdés-Sosa, Department of Cognitive Neuroscience, Cuban Centre for Neuroscience; Brian Butterworth, Numeracy and Literacy Group, Institute of Cognitive Neuroscience, University College London; Nancy Estévez, Department of Developmental Cognitive Neuroscience, Cuban Centre for Neuroscience; Marisol Rodríguez, Department of Developmental Cognitive Neuroscience, Cuban Centre for Neuroscience Elsa Santos, Department of Software Development, Cuban Centre for Neuroscience; Paul Torres, Department of Education and Quality, Central Institute of Pedagogical Sciences; Ramón Suárez, Department of Biostatistics, National Institute of Hygiene and Epidemiology, Agustín Lage, Department of Mathematics and Statistics, Cuban Centre for Neuroscience.

Marisol Rodríguez is not currently affiliated with any institution.

This research was carried out in the framework of the Cuban National Study for Disabilities. It was supported by the Ministries of Health and Education of Cuba. We thank Marcia Cobas, Vice Minister of Health, who encouraged this work. We appreciate the collaboration of the National Group for Disabilities, the Government, Health, and Education authorities of Centro Habana municipality, the computing students who created the database, and the National Institute of Hygiene and Epidemiology, which provided the facilities for this work. Finally, we are very grateful to the participating children and teachers, who generously gave their time over a 9-month period.

Author contributions are as follows: Study conceived and designed by Vivian Reigosa-Crespo, Mitchell Valdés-Sosa, and Brian Butterworth; study performed by Vivian Reigosa-Crespo, Nancy Estévez, Marisol Rodríguez, Elsa Santos, and Paul Torres; data analyzed by Vivian Reigosa-Crespo, Mitchell Valdés-Sosa, Ramón Suárez, and Agustín Lage; article written by Vivian Reigosa-Crespo, Mitchell Valdés-Sosa, and Brian Butterworth.

Correspondence concerning this article should be addressed to Vivian Reigosa-Crespo, Department of Developmental Cognitive Neuroscience, Cuban Centre for Neuroscience, Ave 25, #15202esq. 158, Cubanacán. Playa, Postal code 11600. Havana, Cuba. E-mail: vivianr@cneuro.edu.cu

estimating, counting, and processing numerical magnitudes. These capacities may function as part of the "starter kit" for understanding numbers and arithmetic, and they are conceived to be executed by a domain-specific and genetically controlled module (Butterworth, 1999, 2010). The functioning of some of these capacities can be observed in preverbal infants (Gelman & Meck, 1983; Wynn, 1992), and even nonhuman animals seem to be capable of estimating numerosities and comparing the size of sets of objects (see Nieder, 2005, for a review). If the domain-specific core systems are indeed involved in acquiring arithmetic skills, then, on one hand, measures of their proficiency should predict individual differences in arithmetic attainment, and on the other hand, at least a subset of the low math achievers should be characterized by deficits in these core capacities. We now examine these two assumptions in more detail.

Evidence from typical development supports the idea that basic numerical capacities predict individual differences in later mathematics achievement. A Finnish longitudinal study by Aunola, Leskinen, Lerkkanen, and Nurmi (2004), identified counting ability at preschool age as a reliable predictor of mathematical achievement in first grade. Similarly, in an Italian longitudinal study by Passolunghi, Vercelloni, and Schadee (2007), counting skills at the beginning of primary school (especially counting as fast as possible from 1 to 10) were identified as a direct precursor to early mathematics learning 6 months later. Moreover, Holloway and Ansari (2009) found that individual differences in the time taken to compare two digits in 6- to 8-year-olds was related to mathematics achievement but not to reading achievement. This relationship was found to be specific to symbolic numerical comparison. In a longitudinal design, De Smedt, Verschaffel, and Ghesquière (2009) found that the size of the individual's symbolic distance effect in Year 1, calculated based on reaction times in a number-comparison task, was predictively related to mathematics achievement in Year 2.

According to the second assumption, developmental dyscalculia (DD), a congenital and persistent disability in achieving normal levels of arithmetical skills (Shalev, Manor, & Gross-Tsur, 2005) could arise when the specialized capacity, or "number module" (Butterworth, 1999), fails to develop normally with corresponding deleterious effects in the acquisition of higher level math skills. This has been called "the defective number module hypothesis" (see Butterworth, 2005, for an elaboration of this theory). This entails a core cognitive deficit in a sense of numerosity—a sense of the number of objects in a set—that causes poor performance on very simple tasks, such as numerical magnitude comparison and counting small numbers of dots (Butterworth & Reigosa-Crespo, 2007).

Evidence for an association between deficits on these basic numerical capacities and arithmetic skills has been revealed in studies of individuals with known mathematical disabilities. Severe low achievers (3 standard deviations worse than controls in an item-timed arithmetic test) have been shown to perform differently on tasks of number comparison and counting compared with typically developing children (Landerl, Bevan, & Butterworth, 2004). Geary, Hamsom, and Hoard (2000) found small but systematic group differences between first grade low achievers and controls in magnitude comparison tasks, whereas Koontz and Berch (1996) found that children with poor math abilities appeared to be counting to three rather than subitizing in a dot-matching task. Moreover, Torbeyns, Verschaffel, and Ghesquiere (2004) found that low achievers also demonstrated inadequate counting strategies when doing arithmetic. The Finnish longitudinal study by Aunola et al. (2004) showed not only that low mathematical school performance was associated with low counting abilities in preschool but also that these deficits were cumulative already at this point in time. Atypical performance on basic numerical processing (including counting and number comparisons) has also been demonstrated in individuals with Williams syndrome (Paterson, Girelli, Butterworth, & Karmiloff-Smith, 2006), Turner syndrome (Bruandet, Molko, Cohen, & Dehaene, 2004), and chromosome 22q.11.2 deletion (Simon, Bearden, Mc-Ginn, & Zackai, 2005), syndromes in which dyscalculia is present.

In summary, current evidence provides promising supports for an association between low-level numerical capacities and arithmetical skills as well as their impairment. However, these capacities have only been considered systematically as a school entry-level competence. Accordingly, the most of the studies had been focused at an age when children are first being introduced to formal mathematics. As a consequence, there is a lack of research that examines the nature of this relationship along the acquisition of more complex and increasingly sophisticated arithmetic skills.

Also, most of these studies only used a test of arithmetical attainment without time controls. Under this condition, we may not differentiate between children who process numerical information efficiently and those who take long time to process it. Jordan and Montani (1997) suggested that some children with specific math disabilities are able to compensate under untimed conditions because of relatively good verbal or conceptual skills.

In the light of these limitations, the first aim of this study is to examine the relationship between individual differences in basic numerical capacities and the development of arithmetical competence over a broad developmental time (second to ninth grades) using two item-timed capacity tests of the Basic Numerical Battery (BNB; dot enumeration and numerical magnitude comparison) and three measures of arithmetical competence: (a) an untimed computational test based on curriculum by grade, (b) an item-timed test of mental arithmetic (addition, subtraction, and multiplication), and (c) a teacher report about math attainment. A large body of research supports the inclusion of these achievement measures. As was pointed out, the most of the studies described above used untimed test of arithmetical attainment, but evidence from other studies supports the role of the speed in basic calculation for solving most math problems (Gersten, Jordan, & Flojo, 2005; Goldman & Pellegrino, 1987; Hasselbring, Goin, & Bransford, 1988) and suggests that dysfluent calculation is a distinguishing characteristic of children with low math achievement (Barnes et al., 2006; Jordan & Montani, 1997). On the other hand, Hoge and Coladarci (1989), in a comprehensive review, reported a moderate to strong association between teacher judgments and student achievement (median $r = .66$). Consequently, we expect to find an association between both capacities—enumeration and number comparison—and the measures of arithmetic achievement, even for older learners. We hypothesize that basic capacities will contribute significantly to the individual variability on performance for all these convergent measures of arithmetical achievement, although this contribution could be more significant for one measure than others.

We believe that this will provide empirical support for the use of enumeration and numerical magnitude comparison tasks for identifying DD children with a core cognitive deficit in the sense of numerosity leading to impairment along a broad age range. Accordingly, the second aim of the present study is to obtain, for first time, a prevalence estimate of basic numerical deficits and DD based on the entire school-age population from second to ninth grades of a municipality of Havana, Cuba using the BNB tests. We also are interested in testing whether gender differences are present in these disorders.

Here we hypothesize that children with DD exhibit very low arithmetical attainment related to a domain-specific core deficit and that they form a subset of a more extended group of children with arithmetic disabilities. We can test this assumption by comparing prevalence estimate and gender ratio of DD with similar estimates obtained for arithmetical dysfluency (AD) from the same population. We use the term AD instead of the terms mathematical or arithmetical learning disabilities (MLD/ALD), because the AD definition is focused on poor fluency of calculation, a distinguishing characteristic of a more general deficit on mathematical achievement, which demands, even for the early school grades, a complex set of skills to deal with curriculum requirements (e.g., remembering arithmetical procedures and, more generally, the principles and laws of arithmetic). If DD is a subset of a more extended AD group, then the prevalence would be significantly different. Moreover, differences in gender would be considered evidence of the distinctive nature of both disorders. To our knowledge, convincing evidence supporting this assumption has not yet been reported.

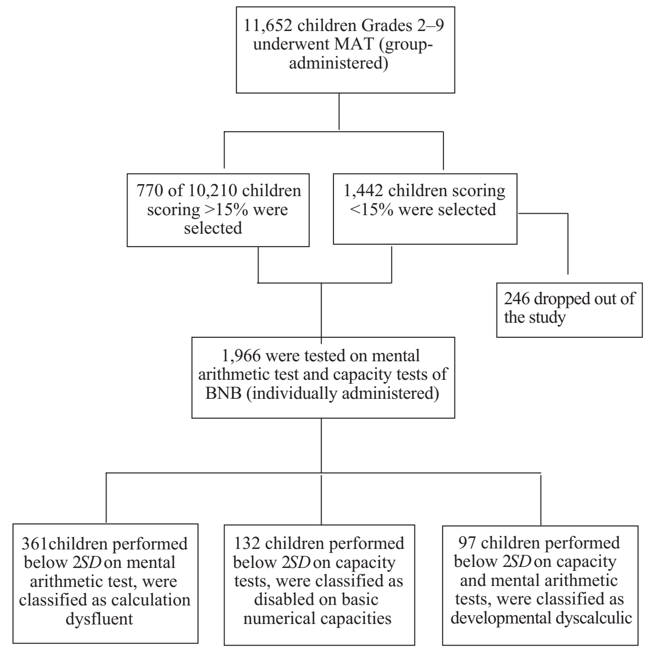

Figure 1. Flow diagram of the two-stage screening. MAT = Mathematics Attainment Test; BNB = Basic Numerical Battery.

Method

The study was carried out in Centro Habana, an urban municipality of Havana, Cuba. All schools in the area were included: twenty-seven primary schools and eight junior high schools. A cohort of school-age children (N = 11,652) from second to ninth grades (5,866 boys and 5,786 girls), ages 6.4–17.3 years (M = 10.86 years, SD = 2.36) was evaluated. The cohort included 93% of the municipality school-children. The age and gender distribution, by grade, of the cohort is shown in Table 1. Permission for the study was obtained from the Ministry of Education and the school directors. Informed consent was obtained from all the parents. The study was designed in two stages (see Figure 1) for efficiency and cost-effectiveness (Shrout, Skodol, & Dohrenwend, 1986). In the first stage, a rough assessment of each child's math proficiency was obtained to allow a subsequent stratified sampling of the participants, which ensured one stratum with a richer (and another with a sparser) presence of math disabilities. In the second stage, basic numerical capacities and other variables were assessed.

First Stage

Participants and procedure. In the first stage, the cohort of children underwent a nonstandardized curriculum-based measurement of mathematics attainment (henceforth, MAT) that was given in the classroom without prior notice to all children at the beginning of school year. MAT was group-administered. The population

Table 1
Description of the Cohort by Grade

Grade	n	Age (years) M	SD	Range	% Boys
2	1,317	7.2	0.32	6.4–8.4	50.2
3	1,519	8.3	0.44	7.7–10.5	51.3
4	1,260	9.3	0.45	8–12.7	50.5
5	1,564	10.3	0.51	9.1–13.6	51.3
6	1,493	11.3	0.57	10.6–13	49.8
7	1,502	12.3	0.6	11.7–15.7	48.1
8	1,602	13.4	0.6	12.8–16.7	50.8
9	1,395	14.2	0.55	13.2–17.3	49
Total	11,652	10.8	2.36	6.4–17.3	50.3

was divided in two strata based on MAT score. One stratum included children scoring in the lowest 15% of their grade, who were considered poor arithmetic attainers. This cutoff is restrictive, because we considered that the larger the percentage used as cutoff, the more heterogeneous and more susceptible to environmental factors the sample would be.

Accordingly, 1,442 children passed to second stage. Of these, 246 (129 boys and 117 girls) were unavailable for further testing for several causes (e.g., some children moved away or dropped out because of illness). Therefore, 1,196 children (636 boys and 560 girls), aged 6.9 to 17.3 years (M = 12.3 years, SD = 2.17) and scoring in the lowest 15% in MAT (M score = 1.9, SD = 0.97, range = 0–3) were recruited for the second stage.

The other stratum included a sample of 770 children homogeneously distributed by grade (381 boys and 389 girls), aged 7.1 to 17.3 years (M = 11.2 years, SD = 2.23), scoring above the lowest 15% in MAT (M score = 6.9, SD = 1.22, range = 4–8). These children also passed to the second stage. The inclusion of the latter allowed us to estimate true and false negatives that were included in the analysis of the technical adequacy of MAT as screening tool. This sample was also useful for estimating the prevalence rate of the disorders. It was selected using a stratified random sampling strategy (Pedhazur & Pedhazur-Schmelkin, 1991). The school-age population scoring >15% in MAT (N = 10,210) was first separate by grade and gender and was then divided into strata based on MAT score. Each stratum was initially treated independently. Thus, children within each stratum were randomly selected, and individual estimates (proportions) were obtained. These estimates were then weighted to arrive at an estimate for the population parameters.

In addition, teachers were asked to provide a judgment of the math ability of all the children in both samples using the following scale: 1 (*very poor*), 2 (*poor*), 3 (*moderate*), 4 (*good*) and 5 (*very good*).

Test used in the first stage: MAT. The MAT is a nonstandardized curriculum-based measurement developed by researchers at the Ministry of Education (Bernabeu, M. & León, T., personal communication, 04/16/2003) and employed throughout Cuban schools. MAT comprised eight computational problems by each respective grade (second to ninth). The authors created the measures by selecting problem types representing a proportional sampling of the computation skills within the national curriculum. Total score was up to 8 (one for each problem performed correctly).

Second Stage

Participants and procedure. The second stage was carried out at the end of the school year (9 months later) to identify children with basic numerical deficits, DD and AD. The Basic Numerical Battery (BNB; see below for details) was administered to 1,966 children selected from the first stage (see flowchart in Figure 1). The assessment was conducted in a quiet and illuminated room inside the school. The physical conditions of the evaluation room were similar across participating schools. Each child was seen individually in a single testing session that lasted approximately 20 min. The child sat next to the tester in front of the computer (PC with Pentium 3 processor). The testers were computing teachers who previously received a certificated training in the BNB assessment.

Tests used in the second stage: BNB. BNB is a battery of item-timed computerized tests, with a structure similar to that of the Dyscalculia Screener (Butterworth, 2003). BNB includes two numerical capacity tests: dot enumeration and numerical magnitude comparison and a test of mental arithmetic fluency. Each test included practice trials to ensure the understanding of the instructions. The children always had to give a response by pressing the corresponding key (thus misses were not measurable). Only the keys of the numeric pad (right side of the keyboard) were available for response (except the simple reaction time task).

1. Simple reaction time. Some children are relatively slow at pressing keys in response to any stimuli. The simple reaction time test was designed to evaluate this. This measure was not analyzed by itself. It was considered a baseline measure of processing speed. Accordingly, the reaction times on the following three computer tests described below were adjusted by subtracting simple reaction time from reaction time on each test. Children were asked to press the space bar as soon as they saw a square in the center of display. The interstimulus presentation time was variable (500–1,500 ms). Twenty trials were presented. Five practice trials were given before starting the test. Reaction times were recorded with millisecond precision.

2. Numerical capacity tests.

2.1. Dot enumeration. Randomly arranged dots ranging from 1 to 9 were presented on the computer display. Children were asked to enumerate the quantities and to respond as quickly as they could without making mistakes. Reaction times and errors were recorded by pressing the key corresponding to number of dots enumerated. Eighteen trials were presented altogether, with each number from 1 to 9 being presented twice in a pseudorandom order, with the proviso that no item occurred twice in succession. Five practice trials were given before starting the test. We assume that enumeration can involve at least three strategies: subitizing for numerosities four or fewer; counting for four or more; and a mix of strategies, which varies by individual, depending on both the individual's numerical capacity, age, experience with counting, and so on. The critical point is that both the speed and accuracy of enumeration will index capacity.

2.2. Numerical magnitude comparison. Children were presented with two digits (1–9) on the computer, one to the left and one to the right of the screen, and they were asked to compare the magnitude of numbers from left to right (e.g., $5 < 7$, $7 > 5$). The numerical distance between pairs was manipulated (distances 1–8). The response keys were "1" for "<," "2" for "=," and "3" for ">." Thirty-six trials were presented in a pseudorandom order. Five practice trials were given before starting the test. Reaction times and errors were recorded.

3. Mental arithmetic. Fifteen simple additions, 15 subtractions, and 15 multiplications were presented in three separate blocks. All involved single-digit numbers from 2 to 9, excluding 0 and 1, since number facts involving 0 and 1 can be solved by application of a rule rather than calculation or retrieval. No ties (e.g., $3 + 3$, 5×5) were presented, and items were not repeated. Items were presented on the computer screen in the form "$2 + 4$." Two practices trials were given before the start of each block. Children were asked to type in the answer as quickly as they could without making any mistake. Reaction time (RT) was measured with the first key stroke. Errors were also recorded. Second graders did not receive the multiplication block, because at the time of the assessment, they were starting to learn the multiplication tables.

The median is usually considered a trimming procedure for excluding spuriously fast or slow reaction times from the analysis. It allows one to obtain better estimation of central value with less variability than does the mean (Ulrich & Miller, 1994). Consequently, median reaction times for correct responses in dot enumeration and numerical magnitude comparison were calculated. The medians were adjusted, subtracting each from the median of the simple RT for that participant (adjRTs). Then an *efficiency measure* (EM) for each test was calculated by diving adjRTs by the proportion of hits (EM = adjRT/Hits). For the mental arithmetic test, the EM scores for each operation (addition, subtraction, and multiplication) were calculated for each child. The mean of these medians for each child was then used as a measure of efficiency on the mental arithmetic test overall. As in the Landerl et al. (2004) study, these two measures (RT and proportion of hits) were used because it had been noted that children with low numeracy tend to adopt strategies that produce generally accurate answers but extremely long RT latencies (see also Jordan & Montani, 1997); or they would simply guess quickly, leading to inaccurate answers but short RT latencies. Note that higher EM scores represent worse performance.

Individual *Z*-score for each test was calculated using the mean and standard deviation (*SD*) of the residuals of the regressions of EMs as a function of age, estimated from the normative sample. Residuals are differences between the observed values and the corresponding values that are predicted by the model and, thus, represent the variance that is not attributable to age. The operational criterion for classifying children with AD was a *Z*-score < 2 *SDs* in the mental arithmetic test. The operational criterion for classifying children with basic numerical deficits was a *Z*-score < 2 *SDs* in at least one of the two capacity tests (dot enumeration and numerical magnitude comparison). Finally, DD was operationally defined as *Z*-score < 2 *SDs* in the mental arithmetic test and in at least one of the two capacity tests.

Results

Basic Numerical Capacities and Arithmetical Competence

All children assessed with both tests (MAT and BNB) were included in the subsequent analysis. First, a one-way analysis of variance (ANOVA) was conducted, with the EMs of the enumeration and comparison tests as dependent variables and grade as the independent variable. Means and standard error for each EM by grade are shown in Table 2. The correlations in Table 2 between the EMs of both tests were positive and significant across grades.

We found a significant effect of grade. Earlier grades showed longer EMs than later grades for both enumeration, $F(7, 1958) = 133.11$, $p < .0001$, and comparison tasks, $F(7, 1958) = 80.313$, $p < .0001$. This reflects that each grade differed from other grades in EMs, except for older grades (sixth through eighth grades for enumerating and seventh through ninth grades for comparing), as demonstrated by post hoc Bonferroni-corrected *t* tests ($\alpha = .05/28 = .001785$).

We also tested whether efficiency for enumeration and number comparison predicted individual variations in math performance. Because EMs did not fit a normal distribution (the data were skewed to the right), a logarithmic transformation was performed (logEMs).

First, a correlational analysis across all grades between the outcome measures (MAT scores and teacher scores) and the predictors (logEMs of capacity tasks) was performed, including grade as independent variable. Significant negative partial correlations were found between teacher scores and the logEMs of enumeration (–.11, $p < .01$) and number comparison (–.17, $p < .01$) as well as between the logEMs of capacity tests and MAT scores (–.13 and –.09, $p < .001$, respectively, for enumeration and comparison).

The second analysis was performed in two ways: by grade and collapsed across grades. In both cases, multiple regressions were performed. In the models, logEMs calculated for dot enumeration and number comparison were defined simultaneously as continuous predictor variables. MAT scores (0–8) and teachers' opinion scores ("very poor" to "very good") were defined separately as dependent variables. The models assumed an ordinal multinomial

Table 2
Efficiency on Numerical Capacity Tests by Grade

	Dot enumeration			Numerical magnitude comparison			
Grade	*M*	*SE*	95% CI	*M*	*SE*	95% CI	*R*
2	4,504	214.8	4,073.3–4,934.6	5,791.6	874.8	4,038.4–7,544.9	.38**
3	3,484.2	84.8	3,316.2–3,652.3	2,954.9	171.0	2,616.2–3,293.6	.46**
4	2,992.1	76.2	2,840.8–3,143.4	2,317	122.9	2,073.2–2,560.8	.22*
5	2,713.1	57.8	2,598.8–2,827.4	2,152.4	103.8	1,947.4–2,057.4	.46**
6	2,435.5	59.6	2,317.6–2,553.5	1,755.1	80.8	1,595.2–1,915.1	.48**
7	2,389.2	67.7	2,295.3–2,523.1	1,742.1	104.5	1,535.2–1,949.1	.41**
8	2,236.3	23.6	2,189.8–2,282.7	1,607.6	23.8	1,560.7–1,654.4	.37**
9	2,117.9	27.4	2,064.1–2,171.7	1,485	29.1	1,427.8–1,542.2	.32**

Note. CI = confidence interval; R = Pearson product-moment correlation coefficient between efficiency on dot enumeration and numerical magnitude comparison.
* $p < .05$. ** $p < .01$.

distribution of dependent variables because they may be ordered as categories.

In the analysis by grade, we found that logEMs of dot enumeration were significant predictors of MAT scores for all grades except fourth grade, $W(1, 8)$ between 6.63 and 53.18, $p < .01$. LogEMs of number comparison were also significant predictors of MAT scores for all grades except second grade, $W(1, 8)$ between 3.98 and 36.5, $p < .05$. When the teacher's judgment was analyzed, the efficiency in both capacity tests was significant predictor for all grades without exception, $W(1, 4)$ between 3.75 and 9.11, $p < .05$, for dot enumeration; $W(1, 4)$ between 3.82 and 14.1, $p < .05$, for number comparison.

When the grades were collapsed in the analysis, we found that the logEMs of both capacity tests were significant predictors of MAT scores, $W(1, 8) = 10.96$, $p = .0009$, for dot enumeration and $W(1, 8) = 7.13$, $p = .008$, for number comparison. Note in Figure 2 that low and moderate math performance (score range 0–2 and 3–5, respectively, on MAT test) were associated with poor efficiency in enumeration and comparison. There was no significant difference between them according to Bonferroni-corrected t tests ($\alpha = .05/3 = .0166$), whereas high math performance (score range 6–8 on MAT test) was associated with better efficiency in capacity tests. Significant differences in efficiency between high math performance and low/moderate math performance were found. LogEMs of the number comparison test also predicted teacher opinion about math achievement, $W(1, 4) = 13.9$, $p = .0002$, but in this case, the prediction of the dot enumeration test was only marginally significant, $W(1, 4) = 4.4$, $p = .056$. Figure 2 shows fall-off patterns, which means that the lower the category assigned by the teacher, the worse the EMs are on capacity tasks. No significant differences between "very poor," "poor," and "moderate" categories and between "good" and "very good" categories, were found with respect to efficiency on enumeration. However, a significant difference was found between these two groupings of categories according to Bonferroni-corrected t tests ($\alpha = .05/3 = .0166$). With respect to efficiency on number comparison, the pattern of differences between categories of teacher' expectancy was quite similar to that described for enumeration, except that, in

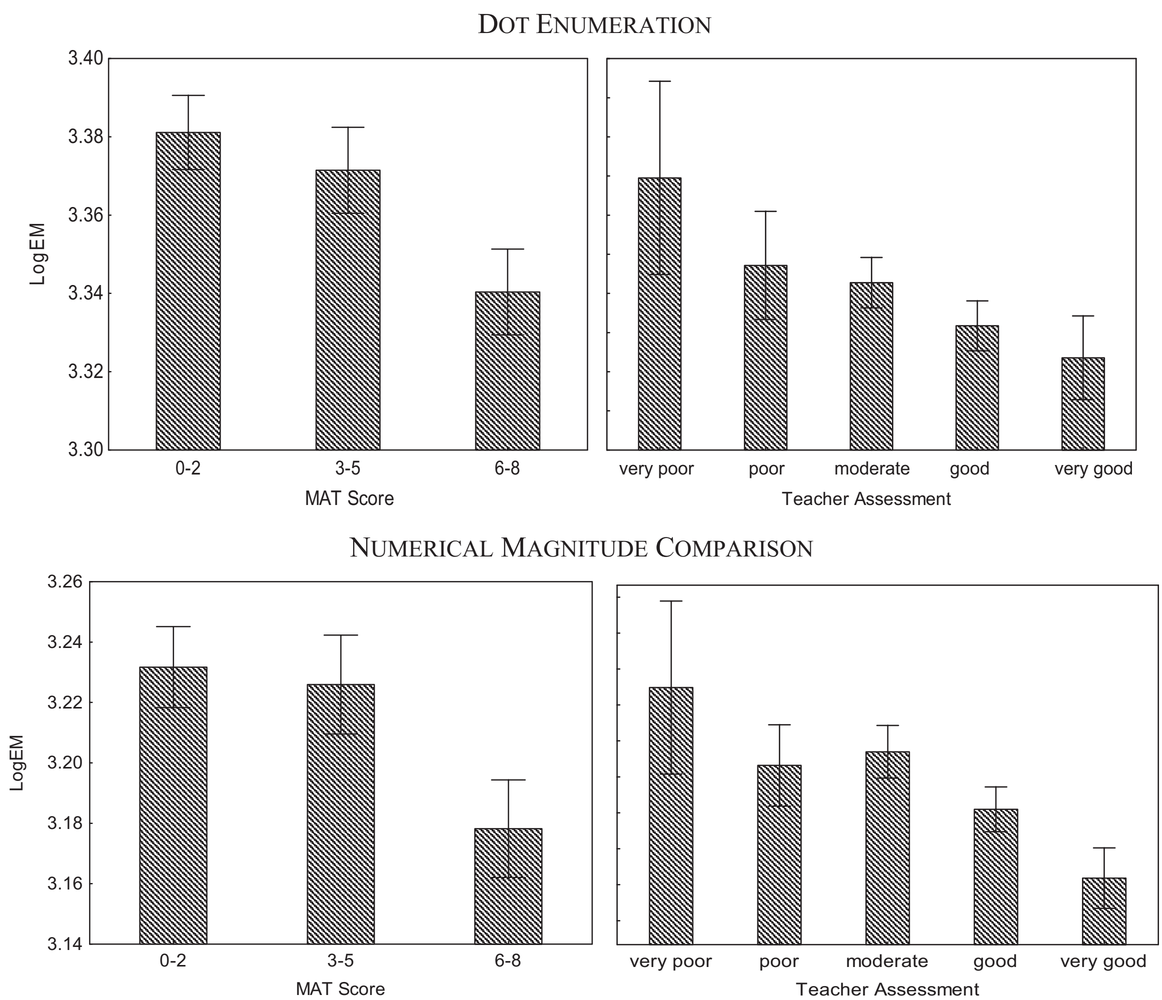

Figure 2. Mean efficiency in dot enumeration and numerical magnitude comparison tests related to mathematics attainment (MAT) scores and teacher's opinion of math achievement. Efficiency is an inverse measure; lower numbers indicate better performance. Error bars represent 95% confidence intervals.

this case, there was a significant difference between “good” and “very good” categories.

A further question to address was the unique contribution of enumeration and comparison with the individual variability on mental arithmetic fluency during the development. That is, the relationship of one numerical capacity and the efficiency of mental calculation while controlling for the other numerical capacity. To examine this, we performed several regression analyses by grade. In the models, logEMs of mental arithmetic by operation (addition, subtraction, and multiplication) and a composite logEM on the mental arithmetic overall were defined as dependent variables separately in each model. For each analysis, logEMs in enumeration and comparison were considered simultaneously as independent variables. The bivariate correlations between efficiency on mental arithmetic and efficiency on enumeration by grade were highly significant (range = .39–.65, $p < .01$) and were also significant between mental arithmetic and number comparison (range = .50–.61, $p < .01$).

The analysis of partial correlations showed that enumeration and number comparison capacities were significantly independent predictors of the calculation efficiency (overall and by operation) in all the grades tested (see Tables 3 and 4). As a trend, partial correlations were stronger for overall efficiency than for efficiency by operation. In-depth analysis revealed some interesting issues. In a similar fashion, enumeration and number comparison separately accounted the individual differences in efficiency for addition and subtraction facts. However, the individual variance in multiplication efficiency was better explained by number comparison than by enumeration. In fact, there was no significant contribution of dot enumeration until fifth grade. Focusing the attention on the younger children, we found that enumeration showed higher partial correlations with addition and subtraction than with comparison for second graders. This pattern was inverted for third and fourth graders. For older children, similar contribution of enumeration and comparison to calculation efficiency was found.

Prevalence and Gender Differences for AD, Basic Numerical Deficits, and DD

The criterion of AD was fulfilled by 361 children. Of them, 306 scored below 15% and 55 scored above 15% in MAT. Using the formula shown in Table 5, we estimated the prevalence of AD in the entire school-age population as 9.35%. The criterion of basic numerical deficits was satisfied by 132 children: one hundred one scored below 15% and 31 scored above 15% in MAT. The estimated prevalence was 4.54%. Finally, 97 children fulfilled the criterion of DD. Of them, 74 scored below 15% and 23 scored above 15% in MAT. We estimated that the prevalence of DD in the entire school-age population was 3.4%. Notice that 35 of 132 children (26%) of those identified as poor performers on basic numerical capacity tasks did not perform poorly in mental arithmetic. Significant differences between the estimates of the disorders were found ($p < .0001$).

On the other hand, the analysis of gender differences in children with AD showed no significant preponderance of boys with respect to girls (m:f ratio = 1.07:1). Conversely, the gender difference was significant for children with basic numerical deficits (m:f ratio = 2.4:1), $\chi^2(1, N = 132) = 73.3$, $p < .00001$. However, at the higher end of efficiency on basic numerical capacities, gender differences were not found (m:f ratio = 1:1). Interestingly, the preponderance of boys with DD was twice that of those with only basic numerical deficits (m:f ratio = 4:1), χ^2 $(1, N = 97) = 103.9$, $p < .00001$.

Effectiveness of MAT and BNB as Screening Tools of DD

The evaluation of the technical adequacy of MAT and the capacity tests of BNB as screening tools of DD is crucial with respect to a criterion measure, or “gold standard” (GS). As was previously pointed out, there is agreement that poor fluency on calculation is a distinguishing feature of arithmetical disorders. For this reason, mental arithmetic efficiency was used here as the GS.

Table 3
Correlations of Performance on Mental Arithmetic With Efficiency on Dot Enumeration After Controlling for Efficiency on Numerical Magnitude Comparison

	Composite measure		Addition		Subtraction		Multiplication	
Grade	r_p^2	t	r_p^2	t	r_p^2	t	r_p^2	t
2	.52**	4.45	.54**	4.65	.43**	3.42	na	na
3	.16	1.75	.19*	1.99	.19*	1.99	.09	0.95
4	.37**	4.02	.16	1.6	.14	1.39	.09	0.85
5	.37**	5.07	.33**	4.4	.27**	3.43	.21**	2.62
6	.37**	4.41	.26**	2.87	.29**	3.28	.25**	2.81
7	.40**	5.07	.30**	2.98	.26**	2.95	.18*	2.04
8	.38**	10.22	.37**	9.64	.27**	6.92	.22**	5.58
9	.41**	11.16	.38**	10.04	.31**	8.2	.25**	6.43
Overall	.423*	20.684	.40**	18.73	.34**	15.68	.25**	11.04

Note. na = not assessed. Composite measures represent the efficiency on the mental arithmetic test overall. r_p^2 values represent the proportion of the variance in mental calculation accounted for by dot enumeration when controlling for numerical magnitude comparison. t values represent the distance, measured in units of standard errors, between the obtained correlation and the null hypothesis of no correlation. p values represent the probability of obtaining the observed correlation in a sample of data by random chance when there is truly no relation in the population.
* $p < .05$. ** $p < .01$.

Table 4
Correlations of Performance on Mental Arithmetic With Efficiency on Numerical Magnitude Comparison After Controlling for Efficiency on Dot Enumeration

	Composite measure		Addition		Subtraction		Multiplication	
Grade	r_p^2	t	r_p^2	t	r_p^2	t	r_p^2	t
2	.25*	1.92	.29*	2.15	.14	0.98	na	na
3	.40**	4.61	.24**	2.62	.30**	3.26	.21*	2.16
4	.59**	7.41	.41**	4.45	.44**	5.00	.46**	5.07
5	.29**	3.81	.29**	3.75	.22**	2.85	.24**	3.07
6	.40**	4.79	.24**	2.62	.30**	3.39	.19*	2.03
7	.38**	4.82	.44**	5.44	.38**	4.56	.24**	2.77
8	.35**	9.37	.24**	6.02	.25**	6.38	.22**	5.52
9	.40**	10.71	.31**	8.06	.27**	6.76	.29**	7.29
Overall	.422*	20.663	.35**	16.14	.32**	14.85	.28**	12.2

Note. na = not assessed. Composite measures represent the efficiency on the mental arithmetic test overall. r_p^2 values represent the proportion of the variance in mental calculation accounted by numerical magnitude comparison when controlling for dot enumeration. t values represent the distance, measured in units of standard errors, between the obtained correlation and the null hypothesis of no correlation. p values represent the probability of obtaining the observed correlation in a sample of data by random chance when there is truly no relation in the population.
* $p < .05$. ** $p < .01$.

Therefore, presence/absence of AD was considered as criterion measure of the effectiveness of the screening. The cutoff for the GS was 2 *SD* below the mean for EMs on the composite measure of mental arithmetic test (as was defined for AD in the Method section). Four indices were taken into account to evaluate the effectiveness of MAT and capacity tests of BNB: sensitivity, specificity, negative predictive value, and positive predictive value. Each index was calculated using values from the outcome matrices of the screening tools respect to the GS. This means that true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) must all be taken into account (see Table 6).

Table 6 also summarizes the four indices for MAT and capacity tests of BNB. The first screener, MAT, exhibited high sensitivity (.85) and negative predictive value (.93) but low specificity (.45) and very low positive predictive value (.26). Conversely, the second screener showed poor sensitivity (.27) and very high specificity (.98). Consequently, a positive result in BNB was itself fair at confirming DD (predictive positive value was .74), and a negative result was very good at reassuring that a child does not have DD (negative predicted value was .86).

Discussion

Basic Numerical Capacities and Arithmetical Competence

It has been assumed that the manipulation of numerical representation in basic numerical processing serve as a cognitive precursor for the development of complex mathematical skills (Dehaene, 1997; Gallistel & Gelman, 1992; Geary, 1995). To address this question, the current study examined whether the efficiency in comparing symbolic numerical magnitudes and in enumerating sets of dots is related to individual differences in children's math achievement (including calculation fluency) and to the teachers' opinions about math ability. Initial analysis revealed that efficiencies in dot enumeration and number comparison were specific and significant predictors of individual differences in math attainment and teacher judgment. A remarkable finding was that these core numerical capacities improved with age and contributed significantly to math attainment across all grades studied (including adolescence). Contrary to previous consensus, this finding demonstrates the plausibility of the hypothesis that quantitative knowl-

Table 5
Proportions of Children With Arithmetic Dysfluency, Basic Numerical Deficits, and Developmental Dyscalculia Based on Two-Stage Screening Assessment

		AD		Numerical deficit		DD	
Total population	Sample[i]	N	Proportion	N	Proportion	N	Proportion
10,210 (a) > 15% in MAT	770	55 (c)	.07	31 (e)	.04	23 (g)	.03
1,442 (b) ≤ 15% in MAT	1,196	306 (d)	.26	101 (f)	.084	74 (h)	.062

Note. AD = arithmetical dysfluency; DD = developmental dyscalculia; MAT = mathematics attainment. (a) = population scoring above the lowest 15% in MAT; (b) = population scoring in the lowest 15% in MAT. The proportions (c) through (h) were calculated by taking the actual sample tested with the Basic Numerical Battery (BNB). The prevalence of AD was calculated with the formula (a × c) + (b × d)/(a + b). The prevalences of the other conditions were calculated substituting the proportions (c) and (d) with the corresponding (e), (f), (g), and (h), respectively, in the formula.
[i] The actual sample tested with BNB.

Table 6
Effectiveness Indices of MAT and Capacity Tests of BNB Using Efficiency on Mental Arithmetic as the Gold Standard

Effectiveness indices	First screener MAT	Second screener BNB[a]
True negative	715	1570
False negative	55	264
True positive	306	97
False positive	890	35
Sensitivity	0.85	0.27
Specificity	0.45	0.98
Positive predictive value	0.26	0.74
Negative predictive value	0.93	0.86

Note. MAT = mathematics attainment; BNB = Basic Numerical Battery.
[a] Only dot enumeration and numerical magnitude comparison tests.

edge is not only an entry-level competency but that it also affects the formal learning of arithmetic until middle education. Further research is necessary to determine whether math education enhances the quantitative knowledge and the extent to which tertiary factors can affect both.

We also found that numerical capacities were more strongly associated with the math fluency measures than with untimed and indirect measures of math attainment. Partial correlation analysis (see Tables 3 and 4) indicates that the relationship between the efficiency on basic capacities and calculation fluency exists from the second grade and continues until the older grades. This relationship was substantially proved not only for overall calculation efficiency but also for efficiency by operation (addition, subtraction, and multiplication) separately. Moreover, a significant proportion of the variance in mental calculation (overall and by operation) was accounted for by one capacity variable when controlling for the other, and this effect occurred for each grade. This unique contribution of each numerical capacity was found despite the high correlation observed between them.

In-depth analysis reveals subtle differences with respect to the contribution of enumeration and number comparison on basic operations during development. As a trend, efficiency on enumeration accounted more significantly for the individual variability in addition and subtraction, although it also explained performance on multiplication. Moreover, efficiency on number comparison better explained individual differences in multiplication.

In specifically analyzing the first grade, some interesting patterns arise. Compared with previous findings describing the acquisition of arithmetic skills (see Butterworth, 2005, for comprehensive review), in our study, enumeration seems to play a more significant role in addition and subtraction for younger children, whereas manipulating magnitudes is more instrumental in mastering multiplication facts. However, the influence of enumeration gradually diminished and was not significant in fourth graders but became significant again in fifth to ninth grades. On the other hand, the manipulation of magnitudes became highly significant to mental calculation in third grade and continued in a similar fashion until adolescence. The enhanced relationship between basic capacities and mental arithmetic in older grades could be an effect of the influence of formal instruction on numerical capacities, but further research is necessary to confirm this.

Another important contribution for understanding the development of numerical cognition was provided by this study. The specific influence of each numerical capacity on calculation fluency strongly suggests that during development, the arithmetic fact retrieval is not a straightforward retrieval process but, rather, involves the activation of numerical magnitudes (Butterworth, Zorzi, Girelli, & Jonckheere, 2001) and counting strategies (Geary, 1993, 2004; Jordan, Hanich, & Kaplan, 2003).

The fact that capacities as basic as comparing two single-digit numbers and enumerating small sets contribute specifically to predicting variance in more complex mathematical tasks, even for students in older grades, speaks to the importance of these basic numerical processes for mathematical understanding over developmental time.

As highlighted in the Method section, none of the measures of mathematical ability (MAT, the teacher assessment, and the mental arithmetic test) contained number comparison or enumeration. Hence, the correlations revealed in this study are not simply relationships between similar measures but, instead, could be indicating that similar processes are engaged during enumeration and symbolic numerical magnitude comparison and tasks measuring math achievement. Furthermore, it could not be argued that the correlation between basic capacities and mental arithmetic is simply due to both tasks being speeded, because the overall RTs were adjusted by individual differences in processing speed (see Method section). Moreover, a significant relationship between the efficiency on basic numerical capacities and MAT was also found. MAT was an untimed test of children's arithmetical competence. If the relationships reported here were fully accounted for by speed of processing, no correlation between the efficiency on basic numerical processing and MAT should have been found.

As a final consideration, here we examined a large cohort as a single group twice—at the beginning of the school year with MAT and at the end of the school year with BNB—rather than comparing different groups of participants (e.g., average and poor math achievers), but it was not properly a longitudinal design. Consequently, although the correlations between numerical capacities and math achievement were well established here, their implications must be taken with some caution, because the findings do not prove a direct causal relationship between them or the direction of causality. Longitudinal studies can be invaluable in helping clarify the direction of causal relationships in typical development and for highlighting the role that early impairments may play in causing later development.

In conclusion, the current results support the existence of a relationship between individual differences in basic numerical processing and arithmetical competence over developmental time. This finding is in line with recent reports. Holloway and Ansari (2009) found similar results using the slope of numerical distance obtained from a number comparison tasks in typically developing children, whereas De Smedt et al. (2009), in a longitudinal design found that the size of the individual's distance effect, calculated on the basis of reaction times, was predictively related to mathematics achievement in first grade children. However, as we have shown here, using partial correlations, number comparison alone is less predictive than enumeration and comparison together. This suggests that the two tests index somewhat different, and possibly overlapping, capacities. It may be, as both Holloway and Ansari (2009) and Rousselle and Noël (2007) proposed, that one of these capacities is the mapping between number concepts and number symbols (digits), and the other is the capacity to represent the number concepts themselves. In some learners, only one of these capacities will be impaired, whereas in others, both capacities are impaired. This is a subject for future research. In the meanwhile, it would be sensible to use both capacity tests in assessing learners.

Prevalence and Gender Differences in AD, Basic Numerical Deficits, and DD

The present study was based on the research principles for epidemiological studies of learning disabilities indicated by Shalev (2007). As a starting point, DD was precisely defined as a persistent and serious math disability due to a core deficit in the "number module," and accordingly, an operational criterion to classifying DD was established. Moreover, a sufficiently large and representative population was screened, and the DD children were identified using a two-stage design involving two measurements. Consequently, here we reported the first prevalence estimates of deficits in basic numerical capacities per se (4.53%) and DD focused on selective deficits in "number module" (3.4%) in a general school-age population.

We also obtained for first time a prevalence estimate of AD. This estimate was at 9.4% in the same school-age population. AD was occurred three times more frequently than DD in the population studied. This finding supports the assumption that DD could be a subset of a more extended arithmetical disability group. That seems plausible because DD, as defined here, is related to a very selective deficit in basic numerical capacities, whereas AD may reflect a variety of cognitive disabilities related to inadequate counting-based and retrieval-based strategies from long-term memory (Geary, 1993; Jordan et al., 2003; Jordan & Montani, 1997) as well as other factors, such as a lack of exposure of arithmetic facts, a low confidence criterion, or low IQ score (Geary et al., 2000). Math anxiety or attention disorders could also be related, but they have not been fully explored.

Note that traditional studies of prevalence of MLD have focused on tests of mathematical attainment that demand, even for the early school grades, a complex set of skills to deal with curriculum requirements (e.g., recognizing and understanding numerals, number words, and operation symbols; remembering arithmetical facts, arithmetical procedures, and more generally, the principles and laws of arithmetic). Moreover, mathematical attainment depends significantly of many factors, including quality of teaching, behavioral problems, anxiety, and assistance at school. Therefore, attainment tests probably capture a heterogeneous sample of learners who could be bad in math for many different reasons. The prevalence estimates based on the math attainment measurement range from 3% to 14% (*Mdn* $\approx$ 6%; Badian, 1983; Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2005; Desoete, Roeyers, & De Clercq, 2004; Gross-Tsur, Manor, & Shalev, 1996; Kosc, 1974; Lewis, Hitch, & Walker, 1994). The samples of children with MLD identified in these studies are probably more comparable to our AD group than to the DD group we define here.

It is interesting that 26% of children with basic numerical deficits did not perform poorly in arithmetic. One plausible explanation is that they have surmounted their inherent disabilities by compensating for defective capacities in numerical cognition, most likely through formal instruction in the first grades (Locuniak & Jordan, 2008). Over time these children may eventually perform above the low average range in arithmetical achievement using laborious strategies, and consequently, they may not fully qualify as having arithmetical disabilities (Mazzocco, 2007).

As far as we know, the present study is the first to examine the gender ratios of AD focused on calculation dysfluency and DD, defined as arithmetical disability associated with basic numerical deficits. In most of the previous prevalence studies, boys and girls had similar probabilities of having poor math abilities. However, we found a high preponderance of boys with respect to girls at the lower end of efficiency in the basic numerical capacities. Boys were two times more likely to suffer this deficit than girls. If these children were also at the lower end of arithmetical fluency, then boys were four times more likely to exhibit the deficits than girls.

On the other hand, we found no significant differences between genders at the higher end of efficiency in enumeration and number comparison. This finding is in line with the research and literature to date that supports the assertion that boys and girls with typical development, do not differ greatly in terms of numerical performance during the elementary or secondary school years (Royer & Walles, 2007).

In summary, findings related to prevalence and gender characteristics in DD and AD lead us to conclude that DD involves a subset of a more extended group of arithmetically disabled children, and consequently, both disorders seems to be distinctive in nature. To our knowledge, convincing verifications of these assumptions have not been reported before now. These are important findings for those responsible for scholastic remediation programmes. If the basic capacities for understanding numerosities are weak or defective fluency on calculation is present, these should form the focus of a training strategy, rather than rote learning of number bonds and other arithmetical facts (Butterworth & Yeo, 2004: Wilson et al., 2006).

Effectiveness of MAT and BNB as Screening Tools of DD

The evaluation of the technical adequacy of MAT as screener showed that it was sensitive (85% of children with AD were correctly identified by MAT) but imprecise (only 45% of children without AD were correctly identified by MAT as being disorder-free). Consequently, the positive predictive value of MAT was extremely low (.26). That means that 74 of every 100 children designated as being "at risk" will be able, contrary to predictions, to demonstrate adequate calculation fluency in subsequent school years.

On the other hand, the capacity test of BNB exhibited very poor sensitivity (.27). That means that 73% of children with dysfluency on calculation did not suffer from disorders in basic numerical capacities. This finding is consistent with theoretical assumptions and empirical evidence that suggest that DD, as was defined in this article, could be a specific disorder that affects a fraction of the entire population with arithmetical disabilities. Meanwhile, the precision of the capacity tests of BNB was significantly incremented with respect to MAT. Of all the children without calculation dysfluency, 98% had originally scored adequately in the capacity tests. Accordingly, the positive predictive value dramatically varied from .26 for MAT to .74 for the capacity tests. Capacity tests thus increased the chance that those who test positive do indeed have the disorder. Thus, 74% of those children considered to be at risk as consequence of defective basic numerical capacities later performed poorly on mental arithmetic test. Moreover, the negative predictive values were high for the two stages, and consequently, there was a low chance of over-identifying children who will later have poor outcomes. This is critical, because subsequent actions to serve misidentified individuals can result in reduced opportunities for learning or growth, overuse of programming resources, and increased stress among family members or support personnel.

Recently two small-scale studies using basic numerical tools as screeners of math abilities have been reported (Geary, Bailey, & Hoard, 2009; Locuniak & Jordan, 2008). The Number Sets Test developed by Geary et al. (2009) was assessed for a sample of 223 children. This tool correctly identified about nine out of 10 children (88%) who were not at risk for mathematical disabilities and correctly identified two out of three children (66%) diagnosed as math disabled in third grade, whereas Locuniak and Jordan (2008) found that a number sense screening of 198 kindergarten children, using "at-risk" versus "not-at-risk" criteria, successfully ruled out 84% of the children who subsequently exhibited adequate calculation fluency and positively identified 52% of the children who later showed fluency difficulties. Notice that these screening tools showed similar negative predictive values but lower positive predictive values than the capacity tests of BNB.

The acceptable predictive rate of capacity tests of BNB for detecting the risk of struggling with calculation supports the hypothesis that defective numerical capacities are related to deleterious effects on the subsequent performance on higher level arithmetic tasks (e.g., Ansari & Karmiloff-Smith, 2002; Butterworth, 2005; Butterworth & Reigosa-Crespo, 2007). Partial tests of this hypothesis have used numerosity comparison but not enumeration, and accuracy but not time, so a measure of numerosity-processing efficiency could not be calculated (e.g., Halberda, Mazzocco, & Feigenson, 2008; Piazza et al., 2010). Speed of numerical magnitude comparison alone has been used as predictor (De Smedt et al., 2009; Holloway & Ansari, 2009; Rousselle & Noël, 2007) but not in prevalence studies.

A separate analysis of the technical adequacy of MAT and capacity test of BNB is insufficient, because these screening tools were implicated in a two-stage screening design. Consistent with established criteria for multistage screening (Glover & Albers, 2007), we found that MAT, as the first stage measurement, exhibited high sensitivity, which is crucial for ensuring that no children who are potentially at risk are overlooked, while increasing positive predictive value for capacity tests of BNB as the second stage measurement. Finally, it should be noted that for evaluating the technical adequacy of the mental arithmetic test as a screener for AD, a new GS for this tool (e.g., a standardized math achievement test) would need to be defined. This latter issue is beyond the objectives of this study and should be addressed in further research.

Conclusions

The present study demonstrates a relationship between the basic capacity to represent and process exact numerosities and the individual differences in arithmetical competence in a large range of the developmental time. In-depth analysis revealed differences in the contribution of enumeration and numerical magnitude comparison on arithmetic throughout development until adolescence, but further research focused on longitudinal designs is necessary to confirm this. These findings support the assertion that deficits in basic numerical capacities can serve to identify DD, not only at school entry but also at ages when more sophisticated math skills are acquired. Our results also support that conclusion that capacity tests of BNB are precise and predictive tools for identifying children with DD. Moreover, the contrastive findings of gender ratio data and prevalence for different subgroups strongly suggest that DD due to deficits in basic number abilities could be a distinctive disorder that involves a fraction of the low arithmetical achievers. Finally, this study emphasizes the importance of training strategies on low-level numerical processing and the acquisition of symbolic representations of numerical magnitudes. This is significant not only for the management of atypical development of math skills but also for enhancing math learning potential in typically developing children.

References

Ansari, D., & Karmiloff-Smith, A. (2002). Atypical trajectories of number development: A neuroconstructivist perspective. *Trends in Cognitive Sciences, 6,* 511–516. doi:10.1016/S1364-6613(02)02040-5

Aunola, K., Leskinen, E., Lerkkanen, M.-K., & Nurmi, J.-E. (2004). Developmental dynamics of mathematical performance from preschool to Grade 2. *Journal of Educational Psychology, 96,* 699–713. doi:10.1037/0022-0663.96.4.699

Badian, N. A. (1983). Arithmetic and nonverbal learning. In H. R. Myklebust (Ed.), *Progress in learning disabilities* (Vol. 5, pp. 235–264). New York, NY: Grune and Stratton.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: Incidence in a population-based birth cohort, 1976–1982, Rochester, Minn. *Ambulatory Pediatrics, 5,* 281–289. doi:10.1367/A04-209R.1

Barnes, M. A., Wilkinson, M., Khemani, E., Boudesquie, A., Dennis, M., & Fletcher, J. M. (2006). Arithmetic processing in children with spina bifida: Calculation accuracy, strategy use, and fact retrieval fluency. *Journal of Learning Disabilities, 39,* 174–187. doi:10.1177/00222194060390020601

Bruandet, M., Molko, N., Cohen, L., & Dehaene, S. (2004). A cognitive characterization of dyscalculia in Turner syndrome. *Neuropsychologia, 42,* 288–298. doi:10.1016/j.neuropsychologia.2003.08.007

Butterworth, B. (1999). *The mathematical brain.* London, England: Macmillan.

Butterworth, B. (2003). *Dyscalculia screener.* London, England: NferNelson.

Butterworth, B. (2005). Developmental dyscalculia. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 455–467). Hove, England: Psychology Press.

Butterworth, B. (2010). Foundational numerical capacities and the origins of dyscalculia. *Trends in Cognitive Sciences, 14,* 534–541. doi:10.1016/j.tics.2010.09.007

Butterworth, B., & Reigosa-Crespo, V. (2007). Information processing deficits in dyscalculia. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 65–81). Baltimore, MD: Paul H. Brookes.

Butterworth, B., & Yeo, D. (2004). *Dyscalculia guidance.* London, England: NferNelson.

Butterworth, B., Zorzi, M., Girelli, L., & Jonckheere, A. (2001). Storage and retrieval of addition facts: The role of number comparison. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology, 54*A, 1005–1029. doi:10.1080/02724980143000064

Dehaene, S. (1997). *The number sense.* New York, NY: Oxford University Press.

De Smedt, B., Verschaffel, L., & Ghesquière, P. (2009). The predictive value of numerical magnitude comparison for individual differences in mathematics achievement. *Journal of Experimental Child Psychology, 103,* 469–479. doi:10.1016/j.jecp.2009.01.010

Desoete, A., Roeyers, H., & De Clercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities, 37,* 50–61. doi:10.1177/00222194040370010601

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44,* 43–74. doi:10.1016/0010-0277(92)90050-R

Geary, D. C. (1993). Mathematical disabilities: Cognition, neuropsychological and genetic components. *Psychological Bulletin, 114,* 345–362. doi:10.1037/0033-2909.114.2.345

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist, 50,* 24–37. doi:10.1037/0003-066X.50.1.24

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities, 37,* 4–15. doi:10.1177/00222194040370010201

Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool: The Number Sets Test. *Journal of Psychoeducational Assessment, 27,* 265–279. doi:10.1177/0734282908330592

Geary, D. C., Hamsom, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77,* 236–263. doi:10.1006/jecp.2000.2561

Gelman, R., & Meck, E. (1983). Preschoolers counting: Principles before skill. *Cognition, 13,* 343–359. doi:10.1016/0010-0277(83)90014-8

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities, 38,* 293–304. doi:10.1177/00222194050380040301

Glover, T. A., & Albers, C. A. (2007). Considerations for evaluating universal screening assessment. *Journal of School Psychology, 45,* 117–135. doi:10.1016/j.jsp.2006.05.005

Goldman, S. R., & Pellegrino, J. W. (1987). Information processing and educational microcomputer technology: Where do we go from here? *Journal of Learning Disabilities, 20,* 144–154. doi:10.1177/002221948702000302

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine & Child Neurology, 38,* 25–33. doi:10.1111/j.1469-8749.1996.tb15029.x

Halberda, J., Mazzocco, M. M. M., & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature, 455,* 665–668. doi:10.1038/nature07246

Hasselbring, T. S., Goin, L. I., & Bransford, J. D. (1988). Developing math automaticity in learning handicapped children: The role of computerized drill and practice. *Focus on Exceptional Children, 20,* 1–7.

Hoge, R. D., & Coladarci, T. (1989). Teacher-based judgments of academic achievement: A review of literature. *Review of Educational Research, 59,* 297–313.

Holloway, I. D., & Ansari, D. (2009). Mapping numerical magnitudes onto symbols: The numerical distance effect and individual differences in children's mathematics achievement. *Journal of Experimental Child Psychology, 103,* 17–29. doi:10.1016/j.jecp.2008.04.001

Jordan, N. C., Hanich, L. B., & Kaplan, D. (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. doi:10.1111/1467-8624.00571

Jordan, N. C., & Montani, T. (1997). Cognitive arithmetic and problem solving: A comparison of children with specific and general mathematics difficulties. *Journal of Learning Disabilities, 30,* 624–634. doi:10.1177/002221949703000606

Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. *Mathematical Cognition, 2,* 1–24. doi:10.1080/135467996387525

Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities, 7,* 164–177. doi:10.1177/002221947400700309

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. doi:10.1016/j.cognition.2003.11.004

Lewis, C., Hitch, G., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- and 10-year old boys and girls. *Journal of Child Psychology and Psychiatry, 35,* 283–292. doi:10.1111/j.1469-7610.1994.tb01162.x

Locuniak, M., & Jordan, N. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities, 41,* 451–459. doi:10.1177/0022219408321126

Mazzocco, M. M. M. (2007). Defining and differentiating mathematical learning disabilities and difficulties. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 29–47). Baltimore, MD: Paul H Brookes.

Nieder, A. (2005). Counting on neurons: The neurobiology of numerical competence. *Nature Reviews Neuroscience, 6,* 177–190. doi:10.1038/nrn1626

Passolunghi, M. C., Vercelloni, B., & Schadee, H. (2007). The precursors of mathematics learning: Working memory, phonological ability and numerical competence. *Cognitive Development, 22,* 165–184. doi:10.1016/j.cogdev.2006.09.001

Paterson, S. J., Girelli, L., Butterworth, B., & Karmiloff-Smith, A. (2006). Are numerical impairments syndrome specific? Evidence from Williams syndrome and Down's syndrome. *Journal of Child Psychology and Psychiatry, 47,* 190–204. doi:10.1111/j.1469-7610.2005.01460.x

Pedhazur, E. J., & Pedhazur-Schmelkin, L. (1991). *Measurement, design and analysis: An integrated approach.* Hillsdale, NJ: Erlbaum.

Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., . . . Zorzi, M. (2010). Developmental trajectories of number acuity in normal and dyscalculic children. *Cognition, 116,* 33–41. doi:10.1016/j.cognition.2010.03.012

Rousselle, L., & Noël, M.-P. (2007). Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs. nonsymbolic number magnitude processing. *Cognition, 102,* 361–395. doi:10.1016/j.cognition.2006.01.005

Royer, J. M., & Walles, R. (2007). Influences of gender, ethnicity, and motivation on mathematical performance. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 349–367). Baltimore, MD: Paul H Brookes.

Shalev, R. S. (2007). Prevalence of developmental dyscalculia. In D. B. Berch & M. M. M. Mazzocco (Eds.), *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (pp. 49–60). Baltimore, MD: Paul H Brookes.

Shalev, R. S., Manor, O., & Gross-Tsur, V. (2005). Developmental dyscalculia: A prospective six-year follow up. *Developmental Medicine & Child Psychology, 47,* 121–125. doi:10.1017/S0012162205000216

Shrout, P. E., Skodol, A. E., & Dohrenwend, P. B. (1986). A two-stage approach for case identification and diagnosis: First-stage instruments. In J. E. Barret & J. M. Rose (Eds.), *Mental disorders in the community* (pp. 286–303). New York, NY: Guilford Press.

Simon, T. J., Bearden, C. E., Mc-Ginn, D. M., & Zackai, E. (2005). Visuospatial and numerical cognitive deficits in children with chromosome 22q11.2 deletion syndrome. *Cortex, 41,* 145–155. doi:10.1016/S0010-9452(08)70889-X

Spelke, E., & Kinzler, K. (2007). Core knowledge. *Developmental Science, 10,* 89–96. doi:10.1111/j.1467-7687.2007.00569.x

Torbeyns, J., Verschaffel, L., & Ghesquiere, P. (2004). Strategic aspects of simple addition and subtraction: The influence of mathematical ability. *Learning and Instruction, 14,* 177–195. doi:10.1016/j.learninstruc.2004.01.003

Ulrich, R., & Miller, J. (1994). Effects of outlier exclusion on reaction time analysis. *Journal of Experimental Psychology: General, 123,* 34–80. doi:10.1037/0096-3445.123.1.34

Wilson, A. J., Dehaene, S., Pinel, P., Revkin, S. K., Cohen, L., & Cohen, D. (2006). Principles underlying the design of "The Number Race," an adaptive computer game for remediation of dyscalculia. *Behavioral and Brain Functions, 2,* 19. doi:10.1186/1744-9081-2-19

Wynn, K. (1992). Addition and subtraction by human infants. *Nature, 358,* 749–750. doi:10.1038/358749a0

Received June 11, 2010
Revision received June 6, 2011
Accepted June 9, 2011 ■