Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Cognition 102 (2007) 361–395
COGNITION
www.elsevier.com/locate/COGNIT

Basic numerical skills in children with mathematics learning disabilities: A comparison of symbolic vs non-symbolic number magnitude processing ☆

Laurence Rousselle *, Marie-Pascale Noël

Unité Cognition et Développement (CODE), Catholic University of Louvain, 10 place C. Mercier, 1348 Louvain-la-Neuve, Belgium

Received 6 May 2005; revised 9 January 2006; accepted 10 January 2006

Abstract

Forty-five children with mathematics learning disabilities, with and without comorbid reading disabilities, were compared to 45 normally achieving peers in tasks assessing basic numerical skills. Children with mathematics disabilities were only impaired when comparing Arabic digits (i.e., symbolic number magnitude) but not when comparing collections (i.e., non-symbolic number magnitude). Moreover, they automatically processed number magnitude when comparing the physical size of Arabic digits in an Stroop paradigm adapted for processing speed differences. Finally, no evidence was found for differential patterns of performance between MD and MD/RD children in these tasks. These findings suggest that children with mathematics learning disabilities have difficulty in accessing number magnitude from symbols rather than in processing numerosity per se.
© 2006 Elsevier B.V. All rights reserved.

Keywords: Mathematical disabilities; Learning disabilities; Dyscalculia; Mathematical development; Number semantic; Number magnitude

☆ This manuscript was accepted under the editorship of Jacques Mehler.

* Corresponding author. Tel.: +32 0 10 47 38 31; fax: +32 0 10 47 37 74.
E-mail address: laurence.rousselle@uclouvain.be (L. Rousselle).

0010-0277/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.cognition.2006.01.005

1. Introduction

Numbers are a pervasive feature of our everyday life. From early childhood, children develop in an environment rich in quantitative information and numerical experiences. They hear adults using numbers to count, measure, when they use money, tell the time and give the date; they see Arabic symbols in shops, streets, and games as well as on cars, on the pages of books, and on television. Very early on, they become able to count on their own, to grasp the quantitative relationship between collections and to compute small calculations. Not being able to count efficiently, to understand the meaning of numbers, or to calculate as other children do, rapidly becomes a handicap, not only at school but also in society in general, in the same way as not being able to read.

Temple (1992) defines mathematical disability (MD) as a "disorder of numerical competence and arithmetical skill which is manifest in children of normal intelligence who do not have acquired neurological injuries". Epidemiological studies showed that this learning deficit is as widespread as reading disorders and affects 3.5–6.5% of the school-age population depending on the country of study (Badian, 1983; Gross-Tsur, Manor, & Shalev, 1996; Kosc, 1974; Lewis, Hitch, & Walker, 1994). However, despite the growing interest observed over these last few years, research on MD is actually much less advanced than on dyslexia. One reason is probably the mosaic nature of mathematics which involves an increasing range of skills as children grow up. As a result, mathematical disabilities can take many different forms and can be generated or accentuated by a variety of related cognitive deficits.

In spite of this diversity, there is general agreement about the main behavioural manifestations of MD. Most MD children exhibit problems in the execution of arithmetical procedures and experience difficulties in learning and/or retrieving arithmetic facts from semantic memory (as manifested by a high frequency of computational/retrieval errors and low computational/retrieval speed) which, in turn, contribute to the persistence of immature problem-solving strategies such as verbal or finger counting (Geary, Bow-Thomas, & Yao, 1992; Geary, 1993; Geary, Brown, & Samaranayake, 1991; Geary, Hamson, & Hoard, 2000; Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Hanich, Jordan, Kaplan, & Dick, 2001; Jordan & Hanich, 2003; Jordan & Montani, 1997; Jordan, Hanich, & Kaplan, 2003; Landerl, Bevan, & Butterworth, 2004; Russell & Ginsburg, 1984).

1.1. The question of etiology: A combination of genetic and environmental influences

Family and twin studies have stressed the importance of genetic factors in the etiology of MD. The examination of concordance rates in family members using a large MD classification criterion (<25th percentile on a standardized test) showed that more than half of the parents and siblings (53% for both groups) of an MD child also presented difficulties in mathematics (Shalev et al., 2001). A twin study using a more stringent MD classification criterion (<7th percentile on a standardized test) showed that 58% of monozygotic and 39% of dizigotic co-twins of an MD child were also classified as MD, that is, a prevalence about tenfold higher than that observed in the general population (Alarcón, DeFries, Light, & Pennington, 1997). Moreover, several genetic abnormalities, such as Turner or Williams syndromes, are known to induce specific patterns of cognitive impairments including a deficit in mathematics, particularly clear in arithmetic (Bellugi, Wang, & Jernigan, 1994; Bruandet, Molko, Cohen, & Dehaene, 2004; Temple & Mariott, 1998). Taken together, these results emphasize the significant contribution of genetic influences to mathematical underachievement but they also highlight the impact of external factors on mathematical development, as heredity was shown to account only partially for the occurrence of MD.

1.2. Mathematical disability as a consequence of deficient non-numerical cognitive processes?

Although extensive studies have been conducted on the factors and cognitive deficits that might contribute to MD, the issue of its origins is still controversial. If normal intelligence is undoubtedly essential to mathematical achievement, the hypothesis of low intellectual efficiency is not able to account for all MD cases as no systematic IQ difference was found between MD children and their normally achieving peers (Gross-Tsur et al., 1996; Landerl et al., 2004; Shalev et al., 2001), which is in accordance with the definition of the pathology.

Other hypotheses have been formulated in the last 15 years but most of them rely on the frequent association of MD with other disorders without highlighting any causal relationship (Landerl et al., 2004). One of these hypotheses assumes that working memory plays a central role in mathematical achievement and that its impairment may lead to a failure to construct a reliable network of arithmetic facts in semantic memory (Geary, 1993, 1994, 2005; Geary et al., 1991). Indeed, to develop long-term memory representations, both the problem and the answer must be simultaneously active in working memory. If information decays too quickly in working memory, the terms of the problem are no longer available when the answer is accessible and long-term memory associations cannot be created. Moreover, as working memory is also responsible for the allocation of attentional resources during problem-solving monitoring, working memory impairment may induce repeated procedural errors leading to incorrect associations in long-term memory. Finally, when retrieval is not possible for solving simple addition problems, children with poor working memory resources may be more inclined to use immature strategies providing external support (i.e., fingers instead of verbal counting) and to apply less resource-demanding algorithms (i.e., counting all instead of counting min) to keep track of the resolution steps and to reduce the load in working memory (Geary et al., 2004; Noël, Seron, & Trovarelli, 2004).

Although the contribution of working memory to problem-solving and arithmetical tasks is not in question (DeStefano & LeFevre, 2004), the numerous studies that have explored the relations between MD and working memory impairment brought contradictory evidence suggesting that these two types of deficit are not steadily associated. While some authors reported data consistent with working memory deficit in MD children, that is, reductions of the forward digit span testing the phonological loop (Geary et al., 1991; Hitch & McAuley, 1991; Koontz & Berch, 1996; McLean & Hitch, 1999), of the visuo-spatial span assessing the visuo-spatial sketch pad (McLean & Hitch, 1999) or of the backward digit span tapping the central executive component of working memory (Geary et al., 1991; Passolunghi & Siegel, 2004), others observed either no difference in any of the span measures (Landerl et al., 2004; Temple & Sherwood, 2002), or only a reduction of backward digit span when MD is associated with reading deficit (Geary et al., 2000; Geary, Hoard, & Hamson, 1999). As MD children do not exhibit systematic impairment in tasks assessing the different components of working memory, working memory and mathematical deficits are unlikely to be causally related in all cases of MD and other underlying factors must be investigated.

Visuo-spatial deficits, reading disorders and finger agnosia are other conditions often associated with MD. However, the reason for these frequent comorbidities remains unclear. According to Rourke and his collaborators (1993, Rourke & Conway, 1997), the presence or absence of reading disabilities (RD) in MD is associated with dissimilar patterns of neuropsychological and arithmetical performances suggestive of different hemispheric dysfunctions. In keeping with this view, they found that children with MD/RD mainly exhibited verbal deficiencies (linked to presumed left-hemispheric dysfunction) while children with isolated MD manifested problems in numerous non-verbal tasks involving visuo-perceptual organisation, psychomotor functioning, perceptivo-tactile abilities and non-verbal reasoning (linked to presumed right-hemispheric dysfunction). Moreover, MD/RD children showed only mild difficulties in arithmetic, essentially attributable to their reading impairment, whereas children with isolated MD experienced deep mathematical disabilities making a wide range of arithmetical errors. This second pattern of deficits, referred to by the authors as the non-verbal learning disability syndrome, bears some resemblance with the Gerstmann syndrome known to occur after dominant hemisphere dysfunction (i.e., usually the left, unlike the non-verbal learning disability syndrome) and associating dyscalculia, dysgraphia without alexia, finger agnosia, left–right disorientation and sometimes visuo-constructional dyspraxia.

Consistent with such patterns of impairments, perceptivo-tactile capacities in 5-year-old children (i.e., simultagnosia, digital gnosia, digital discrimination, and graphisthesia) were found to be powerful predictors of their numerical processing and calculation abilities 6 months later (Fayol, Barrouillet, & Marinthe, 1998; Noël, 2005). In addition, it was shown that visuo-spatial abilities (i.e., visuo-spatial sketch pad of working memory and mental rotation) correlated strongly with high-level mathematical skills (i.e., mental arithmetic, algebra, and trigonometry) in normal development (Reuhkala, 2001). Finally, phonological processing (i.e., phonological memory, rate of access to phonological information, and phonological awareness) were found to contribute significantly to the development of computational skills between second and fifth grade and to account for a large part, if not all, of the association between MD and RD (Hecht, Torgesen, Wagner, & Rashotte, 2001; Robinson, Menchetti, & Torgesen, 2002).

However, in an attempt to examine the validity of Rourke's classification of learning deficit, Shalev and her collaborators (Shalev, Manor, & Gross-Tsur, 1997 Shalev, Manor, Amir, Wertman-Elad, & Gross-Tsur, 1995) failed to find any specific pattern of neuropsychological and arithmetical impairments in children with MD and MD/RD, except that this last group exhibited a deeper deficit in arithmetics than children with isolated MD, even though the mathematical tests did not include verbal problems. These results contradict the predictions of Rourke, who expected children with isolated MD to be more deficient in arithmetics than those with associated RD. Furthermore, a series of recent studies contrasted MD and MD/RD children's performance on various mathematical tasks, that is, exact calculation under timed and untimed conditions, verbal problem-solving, approximate calculation, written computation, and tasks assessing the understanding of place value and calculation principles (Hanich et al., 2001; Jordan & Hanich, 2003; Jordan et al., 2003). In all tasks, their low performances were perfectly comparable except on verbal problem-solving (see also Fuchs & Fuchs, 2002) and untimed exact calculation, where MD children performed still below controls but better than the MD/RD group. Under timed conditions, when back-up counting strategies could no longer be used to compute the answer, the advantage of MD over MD/RD children in the exact calculation task disappeared. Instead of indicating that the presence of RD gives rise to specific patterns of mathematical impairments, these results rather suggest that the absence of RD favours MD children only when they can compensate for their deficit by using their intact verbal abilities such as their reading comprehension to solve story problems or their phonological working memory to count in untimed calculations. Finally, recent investigations identified no difference between MD and MD/RD children in basic numerical processing skills (i.e., number reading, naming and writing, number comparison, reciting number sequence and dot counting) and on neuropsychological tasks (phonological and working memory, vocabulary and visuo-motor planning), providing additional evidence against the hypothesis that there are different causes behind MD with or without RD (Landerl et al., 2004).

To sum up, low IQ is not able to explain mathematical underachievement since the intellectual efficiency of most MD children falls in the normal range as specified by the definition of the pathology. Genetic factors play a significant role in mathematical achievement and might be partly responsible for recurrent associations with other deficits such as reading disorder (Light & DeFries, 1995). Nevertheless, heredity is far from accounting for all cases of MD, which points to the involvement of other factors in mathematical development. Finally, it is generally agreed that working memory, visuo-spatial and phonological processing are related and even contribute to mathematical development. However, the precise relations between mathematical achievement, working memory, finger gnosis, visuo-spatial abilities and reading skills need further clarification.

1.3. Mathematical disability as a deficit in numerosity processing

Contrary to the previous theories in which MD emerges as a consequence of basic cognitive deficits, Butterworth and his collaborators (1999, 2005, Landerl et al., 2004) have recently considered a final hypothesis, called the defective number module hypothesis, which assumes that MD directly results from an internal dysfunction of basic numerical cognition. In this view, humans are born with an innate capacity to understand and manipulate numerosities and this capacity would be embodied in specialized neuronal circuits located in the parietal lobe. According to the authors, MD occurs when this basic ability to process numerosity fails to develop normally, resulting in difficulty in understanding number concepts and, consequently, in learning numerical information.

Such a hypothesis had previously been considered to explain the wide range of impairments observed in a case of severe developmental dyscalculia (Ta'ir, Brezner, & Ariel, 1997) but the pattern of performances was inconsistent with this interpretation as Arabic number comparison was perfectly preserved in this child. On the other hand, his recurrent failure in one-to-one correspondence tasks rather suggested that he had difficulty in keeping track of the items already pointed to, which could account for his counting deficit and for much of his abnormally low performance in numerical tasks involving collections.

More convincing evidence for the defective number module hypothesis came from the case of Charles, a Psychology graduate complaining of persistent difficulties with mathematics which constituted a real handicap in his daily life (Butterworth, 1999). Despite having normal intellectual abilities, this 30-year-old adult relied on his fingers to compute even the simplest problems and was unable to solve complex calculations. To compare Arabic digits, his reaction times were on average four times longer than normally expected. Moreover, because he counted on his fingers to achieve the task, he showed an reversed distance effect, taking longer to compare far digits than close ones, unlike normal subjects. Finally, Charles seemed not to process numerical information automatically, which is normally observed in children from the end of first grade (Rubinsten, Henik, Berger, & Shahar-Shalev, 2002). All together, these performances suggested that Charles did not process and understand the meaning of number concepts in the same way as other people and had a basic deficiency of numerical semantic representations.

Consistent with this view, several studies provided evidence that MD affects not only mastery of arithmetic facts and calculation procedures but also tasks requiring manipulation of the quantitative meaning of numbers, such as estimating the approximate result of arithmetic problems or showing the quantities standing for the units and the tens in two-digit numbers (Hanich et al., 2001; Jordan & Hanich, 2003; Jordan et al., 2003). Moreover, MD children were found to be slower at naming digits and quantity from 1 to 4 but not at naming letters or geometric forms (van der Sluis, de Jong, & van der Leij, 2004). Likewise, in a recent study, MD and MD/RD children appeared to be slower than controls in a series of basic numerical processing tasks including number reading, number comparison, reciting number sequences and dot counting (Landerl et al., 2004). Conversely, no difference between groups appeared on working memory and IQ measures, providing additional support to the defective number module hypothesis.

To our knowledge, no study has explored children's brain activation during numerical processing tasks. However, neuroimaging data in adults suggest that the intraparietal sulcus could be the locus of semantic number processing as it is systematically activated in tasks requiring manipulation of the quantitative information conveyed by numbers. In addition, the left angular gyrus has been found to be involved in arithmetic computation, particularly in calculations with significant verbal requirements (see Dehaene, Piazza, Pinel, & Cohen, 2003, for a review).

In the defective number module hypothesis, MD is assumed to result from a failure in the development of the specialized brain systems underlying the processing of number magnitude. Some neuroimaging data support this assumption. Using voxel-based morphometry, a recent study compared the density of grey matter of adolescents with and without arithmetical difficulties who were born pre-term (Isaacs, Edmonds, Lucas, & Gadian, 2001). Results indicated a significant reduction of grey matter in the left intraparietal sulcus in the low performing teenagers. Similarly, the neuroanatomical correlates of arithmetical disabilities in Turner syndrome were studied using functional and structural imaging (Molko et al., 2003). Unlike controls, individuals with Turner syndrome showed no modulation of intraparietal activations as a function of problem size. Additionally, the morphological brain analysis revealed abnormalities in the length, depth, geometry and grey matter density of the right intraparietal sulcus. Finally, a case study using magnetic resonance spectroscopy detected metabolic abnormalities in the left angular gyrus of an 18-year-old adult with lifelong arithmetical disability and superior intellectual abilities (Levy, Reis, & Grafman, 1999).

Taken together, these data strongly support the defective number module hypothesis. Nonetheless, an important issue remains open to question. At present, all work consistent with this hypothesis showed impairments in tasks requiring access to semantic information from symbols (either Arabic or spoken verbal numerals), that is, reading and writing numerals, Arabic number comparison, reciting number sequences, and enumerating collections. However, no studies have investigated numerosity processing with no symbolic processing requirement. Consequently, it remains possible that MD children have no difficulty in processing numerosity per se but rather in accessing numerical meaning from symbols. This alternative account will be called the access deficit hypothesis.

1.4. The present research

The present study addresses three main issues. First, do MD children have difficulty in processing numerosities, as assumed by the defective number module hypothesis, or in accessing number meaning from symbols, as proposed by the access deficit hypothesis? Second, do they process number meaning automatically as normally achieving children do? Third, does the presence of associated RD in MD children give rise to different patterns of performance in basic numerical tasks?

The defective number module hypothesis was contrasted with the access deficit view by comparing performances of MD children in two tasks: an Arabic number comparison task which required relating a symbol to its quantitative meaning, and a collection comparison task involving no symbolic processing. Research on early quantification stressed that collections could be compared on the basis of their numerical properties (i.e., discrete quantities) but also on the basis of the perceptual non-numerical cues that naturally covary with numerosity (i.e., continuous quantities such as area, perimeter, length, etc.; see Mix, Huttenlocher, & Levine, 2002, for a review). As 3-year-old children were found to be able to compare collections when perceptual variables covary with numerosities but not when they are strictly controlled (Rousselle, Palmers, & Noël, 2004), two conditions of perceptual control were used in the collection comparison task. In the low perceptual control condition (i.e., the density condition), most perceptual variables correlated with numerosity and, consequently, children could compare collections on both continuous and discrete quantitative differences. In the high perceptual control condition (i.e., the surface condition), continuous perceptual variables were carefully controlled and children needed to process numerosities to compare collections.

If MD children have difficulty in accessing numerical semantic from symbols, they would be expected to be impaired when comparing Arabic digits but not when comparing collections. By contrast, if MD results from a failure to process numerosity as proposed by the defective number module hypothesis, MD children would be expected to be able to compare continuous quantities in the density condition of the collection comparison task but should be impaired in tasks requiring them to process numerosities, that is, in the surface condition of the collection comparison and in the Arabic number comparison. Finally, MD might result from a defect in a central magnitude system dedicated to the processing of discrete and continuous quantities (Walsh, 2003). In this case, all comparison tasks should be equally affected in MD children.

The second aim of this study was to examine the automaticity of number processing in MD children. As noted earlier, young children, as well as adults, automatically process number magnitude even in non-numerical tasks such as physical size comparison (Girelli, Lucangeli, & Butterworth, 2000; Rubinsten et al., 2002; see Noël, Rousselle, & Mussolin, 2005, for a review). In the defective number module hypothesis, children with MD are not supposed to automatically activate number magnitude as they already have difficulty in processing numerosity under intentional conditions. In accordance with this view, Charles exhibited no sign of automatic number processing (Butterworth, 1999). Using a Stroop paradigm, Landerl et al. (2004) assessed numerical automaticity by asking dyscalculic children to compare the physical size of digits varying in numerical and physical size. The congruency between numerical and physical dimensions was manipulated so that 1/3 of the trials were congruent (i.e., the larger digit in number is also the larger in physical size), 1/3 of the pairs were incongruent (i.e., the larger digit in number is the smaller in physical size), and 1/3 of the trials were neutral (i.e., the digits are the same and only differ in physical size). However, children failed to show any congruity effect whether they were dyscalculic or not. The absence of the congruity effect, even in normally achieving children, might indicate either that they did not process numerical magnitude, or suggest that numerical processing was not fast enough to interfere with physical processing (see Noël et al., 2005, for a discussion).

In the present study, the problem of processing speed difference was overcome by using the procedure finalized by Mussolin (2002). To balance the speed of physical and numerical processing, the physical size difference between digits only appeared after a variable delay depending on participants' processing speed in the numerical and the physical comparison tasks. This delay was individually calculated by computing, for each participant, the difference between the time needed to perform a physical size comparison (same digits, different size) and the time required to compare Arabic digits (different digits, same size). In this way, numerical and physical information are equally likely to influence the decision process.

Thus, in the present study, we examined automatic access to number magnitude in a physical size comparison task using a Stroop paradigm adapted for processing speed difference. The congruency between physical and numerical dimensions was manipulated as well as the numerical size, and the distance between the digits to be compared. If MD children have difficulty in processing numerosity, they would not be expected to show any sign of automatic numerical processing. Conversely, if MD children only have slower access to number magnitude from symbols, they should exhibit automatic activation of semantic properties under conditions adapted for their own symbolic processing speed.

Finally, the performances of MD children with and without RD were contrasted to assess the presence of divergent patterns of numerical achievement. If different cognitive deficits are responsible for MD when associated with RD, MD/RD children are expected to show specific patterns of numerical disabilities compared to children with isolated MD.

2. Method

2.1. Participants

Ninety second-grade children participated in the study: 29 had difficulties in mathematics but not in reading (MD-only group), 16 had difficulties in both mathematics and reading (MD/RD group), and 45 had normal achievement in both mathematics and reading (NA group).

2.1.1. Participant selection procedure

Participants were selected from a screening pool of 499 second graders drawn from twelve state schools in the area of Louvain-la-Neuve, Belgium. All of them attended general education classes and only children whose parents gave their informed consent were included in the final experimental groups. Participant screening took place in the fall of the second grade (mid-November to mid-December) and included mathematics, reading and IQ assessments.

First, a composite test battery devised for mathematics level assessment was set up and collectively administered within the classroom. The test battery comprised two subscales. The number processing subscale involved Arabic number writing, Arabic number comparison and transcoding of tokens into Arabic numerals. The arithmetic subscale included an untimed addition and subtraction test and a timed addition test (see Section 2.2).

Fifteen children were excluded from the initial pool due to their misunderstanding of the instructions. Forty-nine children were also discarded because they failed a year and eight others because they had been academically accelerated. All of the 427 remaining second graders were born in the same year. Standard scores for each of the six subtests (number writing, number comparison, transcoding, untimed addition and subtraction, and timed addition test) were calculated for each child. These scores were then summed within each subscale, giving one composite score for number processing and one for arithmetic processing. As the Pearson correlation indicated that the number processing and the arithmetic subscales were highly correlated ($r = .61$, $p < .001$), the sum of all the standard scores in mathematics was calculated and this composite score was used for the participant selection. To be classified as mathematically disabled (MD), children had to have a mathematics composite score below the 15th percentile (based on the distribution over the 427 second graders) and to have been nominated as having learning difficulties in mathematics by their teacher. The 15th percentile cutoff is more conservative than that typically used in most of the studies on mathematical disabilities (35th percentile: Geary et al., 2000; Hanich et al., 2001; Jordan, Kaplan, & Hanich, 2002, 2003; 30th percentile: Geary et al., 1999, 2004; 25th percentile: Jordan et al., 2003). This cutoff was chosen to reduce the number of false positives. Fifty-four children met these criteria. For each of these MD children, a control child with a global score above the 50th percentile and nominated as having no learning difficulties by the teacher was selected. To reduce the effects of instruction and gender, control and MD children were selected from the same classroom or, when that was not possible, from the same educational establishment and were matched as much as possible for gender.

These 108 children were then assessed individually for IQ and reading. IQ was examined using the Similarity and Picture Completion subtests of the Wechsler Intelligence Scale for Children-III (Wechsler, 1996). As these subtests are quite representative of their respective scale, the mean of their standard scores was calculated to estimate IQ. Children with a mean standard score inferior or equal to 7 were discarded ($M = 10$, $SD = 3$). Reading achievement level was examined using two French-speaking reading subtests: the LUM subtest of the LMC-R battery (Khomsi, 1998) and the L3 subtest of the ORLEC battery (Lobrot, 1980 recent normative data in Mousty & Leybaert, 1999). The LUM is a speeded word identification test involving words of increasing spelling complexity. Children were asked to read as many words as possible in 1 min. The L3 is a reading comprehension test in which children have to select the appropriate word to complete a sentence (time limit: 5 min). The mean of the standard scores of each subtest (based on the respective age and grade level norms) was used to determine reading achievement level. Children with a mean standard score in reading below the 15th percentile were classified as reading disabled.

From the 54 MD children, four children were discarded because they were not French native speakers, four children were absent during the individual testing and one child was excluded because of his low IQ score (mean standard scores = 5). The 16 MD children with a mean reading score below the 15th percentile were included in the mathematics and reading disabled group (MD/RD) whereas the 29 MD children with a mean reading score above the 15th percentile were classified as mathematics disabled only (MD-only).

From the 54 controls selected, two children with low reading achievement level (below 15th percentile) were discarded and one was absent during the individual testing. Forty-five control children were then selected in order to match the NA group as closely as possible to the MD-only and MD/RD groups for educational establishment and for gender. Thus, the NA group consisted of two subgroups: 29 controls matched to the 29 MD-only children (NA MD) and 16 controls matched to the 16 MD/RD children (NA MD/RD).

2.2. Selection tasks

2.2.1. Arabic number writing

Children were dictated a series of three single-digit, nine two-digit, and eight three-digit numbers in a random order and were asked to write these numbers on a sheet of paper. All items were drawn from the number writing subtest of the Tedi-Math Battery (Van Nieuwenhoven, Grégoire, & Noël, 2001), a French-speaking standardized diagnostic test battery for basic mathematical competencies. The result of extensive research, this battery was developed for clinical purposes and is widely used to identify mathematical disabilities from kindergarten to third grade. Children were given 1 point for each number correctly written.

2.2.2. Arabic number comparison

A series of 14 pairs of Arabic numbers arranged vertically on a blank sheet of paper was presented in a random order. For each pair, children were asked to circle the larger of the two numbers. The set of comparison pairs included four pairs of single-digit numbers, six pairs of two-digit numbers, and four pairs of three-digit numbers. All pairs were drawn from the number comparison subtest of the Tedi-Math Battery (Van Nieuwenhoven et al., 2001) except for two which were added to enlarge the set of three-digit number pairs (309–370, 815–851). One point was given for each correct response.

2.2.3. Transcoding of tokens into Arabic numerals

This task was devised to assess children's understanding of base-10 in Arabic numeral coding. Children were shown 10 collections of small and large tokens randomly arranged on a sheet of paper. The small tokens were labelled as "1" and the large ones were labelled as "10". Children were told that they had 1 and 10 euro coins in their pocket and were asked to write on the paper how many euros they had in all. Four collections were composed only of 1-tokens (3×1, 5×1, 14×1, 15×1) and six collections included both 1- and 10-tokens (12, 13, 20, 21, 34, and 50). Each correct response was credited with 1 point.

2.2.4. Untimed addition and subtraction tests

Children were asked to solve 12 additions (eight with single-digit numbers only, one with a single-digit and a two-digit number, and three involving two-digit numbers only) and 10 subtractions (five with single-digit numbers only, two with a two-digit and a single-digit number, and three involving two-digit numbers only) presented in a written format. All items were drawn from the Tedi-Math Battery (Van Nieuwenhoven et al., 2001). Children were allowed to take all the time they needed and to use any strategy they wanted to find out the answer. One point was given for each correct response.

2.2.5. Timed addition test

Children were given 36 single-digit additions and asked to solve as many problems as possible in 2 min. This task was carried out twice with two different sets of stimuli. The first set consisted of 36 simple additions of two integers from 2 to 5 and the second set consisted of 36 medium additions of two integers, one between 2 and 5 and the other between 5 and 9. No tie problems (e.g. $2 + 2$, $4 + 4$) were presented. The dependent variable was the number of additions correctly solved through both parts of the test.

2.3. Experimental tasks

All the experimental tasks were carried out on a PC. Responses and latencies were recorded by the computer with a 1 ms precision. As all experimental tasks required choosing between two possible responses, one displayed on the left and one displayed on the right side of the screen, children were always asked to respond by pressing the button on the side of the correct response. Instructions emphasized both speed and accuracy.

2.3.1. Arabic number comparison

Two Arabic digits from 1 to 9 (Arial, 48-point font) were presented on the computer screen and children were asked to select the larger of the two digits. Comparison pairs varied along two numerical size and two distance between digits: small pairs (digits from 1 to 5) were contrasted with large pairs (digits from 5 to 9) and close pairs (distance of 1) were compared with far pairs (distance of 3 or 4).

As shown in Table 1, three different pairs of digits were used in each experimental condition. The side of the correct response was counterbalanced: each pair appeared four times, twice in ascending (e.g., 1–2) and twice in descending (e.g., 2–1) order, for a total of 48 pairs of single-digit number to be compared (3 digit pairs $\times$ 2 sizes $\times$ 2 distances $\times$ 2 sides $\times$ 2 presentations).

Table 1
Single-digit pairings according to the size of the pair and the distance between digits

Size	Distance	
	Close	Far
Small	1–2	1–4
	2–3	1–5
	3–4	2–5
Large	6–7	5–8
	7–8	5–9
	8–9	6–9

Digit pairs appeared in a fixed pseudo-random order (the same item never appeared in two consecutive trials). Each trial started with the presentation of a digit pair until response, followed by an ISI (white screen) of 500 ms.

2.3.2. Collection comparison

Children were simultaneously presented with two collections of sticks and were instructed to select the one that contained more sticks. Stimuli were pairs of white squares (side = 5.5 cm) separated by a red fixation cross and containing a variable number of vertical black rectangles. Both squares were presented on a dark grey background and were separated by a fixation cross (distance between squares = 8 mm).

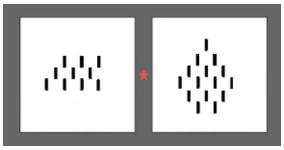
The task was administered twice in two conditions of perceptual control (drawn from Rousselle et al., 2004). In the *density condition*, all sticks were of equal size (2 mm wide and 6 mm high) and the density of the elements was controlled in each pair. In the *surface condition*, the total surface area of the sticks was equated in each pair by reducing the size of the elements in the collection with more sticks. To avoid the larger collection in number being also the one with the smaller elements, sticks of different height were used and the size of the smaller and larger sticks was the same in both arrays to be compared. Furthermore, the external perimeter of both collections was identical within each pair. As all sticks were 2 mm wide, equating the total surface area in both arrays amounted to equating the brightness and total length of the sticks (sum of the individual height). In this way, the differences in the sum of individual perimeters were negligible (i.e., perimeter ratio varies from 0.95 to 0.99).

As small numerosities are supposed to be apprehended by different quantification processes from large ones (Simon, 1997; Trick & Pylyshyn, 1994), only numerosities above the subitizing range, i.e., higher than four, were presented. In each perceptual condition, two sets of comparison pairs were used and intermixed. In the distance-set, comparison arrays involved the same numerosities as those used for large pairs in Arabic number comparison task. Consequently, these comparison pairs varied along two distances (1 vs 3–4). By using the same numerosity for comparing digits and collections, this matching was aimed to directly examine the effect of presentation format, that is, symbolic vs non-symbolic, on children's performance. In the ratio-set, comparison arrays included large numerosities from 6 to 28 and varied along two ratios (1/2 vs 2/3).

Numerosity pairs used in each experimental condition are shown in Table 2. For each pair, the larger numerosity appeared twice on the right and twice on the left, for a total of 48 pairs in each perceptual condition (3 numerosity pairs $\times$ 2 ratios or 2 distances $\times$ 2 sides $\times$ 2 presentations). It should be noted that the 6–9 pair varied along both ratio and distance and, consequently, was included in both set of stimuli. However, this pair did not appear twice as often as the others during the experiment. Thus, only 44 pairs of collections were presented in each perceptual condition during the experiment.

Comparison pairs were presented in a fixed pseudo-random order (the same item never appeared in two consecutive trials). Each trial started with the presentation of a collection pair until response, followed by an ISI of 500 ms during which the fixation cross remained in full view.

Table 2
Numerosity pairings according to their distance or their ratio in each perceptual condition

Condition	Distance-set		Ratio-set	
	Close	Far	1/2	2/3
Density				
	6–7	5–8	6–12	6–9
	7–8	5–9	8–16	10–15
	8–9	6–9	14–28	16–24
Surface				
	6–7	5–8	6–12	6–9
	7–8	5–9	8–16	10–15
	8–9	6–9	14–28	16–24

2.3.3. Physical size comparison

Children were presented with two identical Arabic digits and were asked to select the larger digit in physical size (e.g., 2–2, Arial, 65 pt or 72 pt). The six digit pairs used (2–2, 3–3, 4–4, 6–6, 8–8, and 9–9) were presented four times, twice with the larger digit on the right and twice with the larger digit on the left, for a total of 24 pairs of digits to be compared (6 pairs × 2 sides × 2 presentations).

Digit pairs were presented in a fixed pseudo-random order (the same item never appeared twice in succession). Each trial started with the presentation of a digit pair until response, followed by an ISI (white screen) of 500 ms.

2.3.4. Stroop task

Children were presented with pairs of Arabic digits (1–9) varying along two dimensions: physical size and numerical magnitude. Children were asked to select the larger digit in physical size. Congruency between the physical and the numerical dimensions was manipulated: in congruent pairs, the larger digit in physical size was also the larger in magnitude while in incongruent trials, the larger digit in physical size was the smaller in magnitude.

Each congruity condition included the same digit pairs as those used in Arabic number comparison. So, the digit pairs varied along the numerical size (small vs large pairs), the numerical distance between the digits (close vs far pairs) and the side of the correct response (left vs right). As each pair was presented twice, the final comparison set consisted of 96 pairs of digits to be compared (2 congruity conditions × 3 digit pairs × 2 sizes × 2 distances × 2 sides × 2 presentations).

Comparison pairs were presented in a fixed pseudo-random order (the same item never appeared twice in succession). Each trial started with the appearance of a pair of digits presented in the same intermediate physical size (68 pt, Arial). The physical size difference between digits appeared after a variable delay calculated for each participant by subtracting his/her median reaction time in the physical size comparison task (same digits, different size) from his/her median reaction time in the Arabic number comparison task (different digits, same size; Mussolin, 2002). The difference in milliseconds was rounded to the upper hundred. After the delay, the physical size difference between digits emerged with one digit increasing in size (72 pt) and the other one decreasing (65 pt). Digit pairs were presented until response and were followed by an ISI (white screen) of 500 ms. The task was divided into two blocks of 48 trials and a short break was allowed between blocks.

2.4. Experimental procedure

Children were tested individually in a quiet room within their school. Testing took place in the spring of the second grade and was completed in a single, 30–40 min session. The session started with the two IQ subtests (Picture completion and then Similarities) followed by the two reading subtests (LUM and then L3). The experimental tasks were then administered in three steps. First, children were administered the Arabic number comparison and the physical comparison tasks. Second, they were asked to carry out the two perceptual conditions of the collection comparison task. Finally, the Stroop task was administered, once the specific size difference delay had been calculated for the participant. The order of the Arabic number comparison and the physical comparison tasks was counterbalanced across participants as well as the order of the density and the surface conditions of the collection comparison task. Each experimental task was preceded by eight training trials.

2.5. Group differences on descriptive and selection data

Table 3 reports descriptive information, mean scores in the selection tasks, and mean mathematics and reading percentile scores for each achievement group. To compare each learning disabled group with its respective control group, the two NA subgroups were kept separate in the following analyses. The number of boys and girls in each group did not differ, $\chi^2(3) = 2.53$, $p = .20$. However, age difference between groups was statistically significant, $F(3,86) = 4.9$, $p < .005$: MD-only children were on average 3 months younger than their controls NA-MD (Bonferroni adjustment: $p < .005$). Other mean age comparisons did not reach statistical significance (ps $> .10$). As younger children might be expected to be less proficient in mathematics and reading generally, age in months was used as a covariate in the subsequent analyses.

A multivariate analysis of covariance (controlling for age) revealed that there was no significant difference between groups (MD/RD, MD, NA-MD and NA-MD/RD) on verbal and performance IQ subtests nor on mean IQ scores (Picture Completion: $F(3,85) < 1$, $p > .10$; Similarities: $F(3,85) = 1.8$, $p > .10$; IQ mean score: $F(3,85) = 1.62$, $p > .10$). By contrast, differences between groups were significant on all number processing, arithmetic and reading subtests (Fs > 10, ps $< .001$). On mathematics subtests (number writing, number comparison, transcoding, untimed addition and subtraction and timed addition), pairwise comparisons using the Bonferroni adjustment procedure indicated that the mean scores of the MD-only and

Table 3
Descriptive information and mean performance in the selection tasks for each achievement group

	Achievement group			
	MD/RD	MD	NA-MD/RD	NA-MD
Descriptive information				
N	16	29	16	29
Gender (M/ F)	6/10	8/21	7/9	16/13
Age (in month)	88.88 (4.15)	87.86 (3.00)	90.06 (3.00)	90.93 (2.82)
IQ				
Pictures completion[a]	10.69 (3.59)	10.48 (2.97)	11.00 (3.31)	11.00 (3.13)
Similarities[a]	10.81 (2.74)	11.28 (2.30)	12.19 (2.46)	12.17 (3.13)
Subtest mean score[a]	10.75 (2.69)	10.88 (2.06)	11.59 (2.09)	11.59 (2.27)
Mathematics				
Number writing (%)	57.81 (11.69)	55.00 (15.53)	92.81 (10.16)	95.17 (8.29)
Number comparison (%)	76.79 (11.52)	79.06 (9.72)	97.77 (4.30)	97.29 (4.01)
Transcoding (%)	85.63 (17.50)	79.66 (19.73)	98.75 (3.42)	98.97 (3.10)
Untimed additions (%)	69.79 (9.07)	69.25 (10.93)	93.23 (6.95)	93.97 (6.64)
Untimed subtractions (%)	53.13 (11.95)	54.83 (12.71)	84.38 (12.63)	86.21 (9.79)
Timed additions (%)	25.61 (10.36)	31.85 (10.43)	68.58 (10.83)	63.22 (15.47)
Mathematics mean percentile score	6.69 (4.25)	6.86 (4.10)	79.31 (10.49)	77.38 (12.70)
Reading				
LUM[b]	−1.68 (0.50)	−0.29 (0.68)	−0.28 (0.69)	0.21 (0.64)
Lobrot 3[b]	−1.32 (0.45)	−0.07 (0.77)	0.30 (0.68)	0.87 (0.75)
Reading mean percentile score	8.12 (5.16)	42.76 (21.12)	49.31 (20.25)	66.97 (19.83)

Note: Standard deviations are shown in parentheses. MD-only, mathematics difficulties group; MD/RD, mathematics and reading difficulties group; NA-MD, normal achievement group matched to MD children; NA-MD/RD, normal achievement group matched to MD/RD children.
[a] WISC-III: standard note: Mean = 10, $SD = 3$.
[b] Standard score: Mean = 0, $SD = 1$.

MD/RD groups did not differ (*p*s > .50) but were significantly lower than the mean scores of both NA groups (*p*s < .05). As expected, the mean scores of the two NA groups did not differ (*p*s > .50). In keeping with the classification criteria, the MD/RD group had lower scores than the three other achievement groups on both reading subtests (*p*s < .001). However, the NA-MD group had higher reading scores than the MD-only (*p*s < .05) and the NA-MD/RD groups (LUM: $p = .08$, L3: $p < .05$). The reading scores of the two latter groups did not differ (*p*s > .20).

These descriptive analyses suggest that the children had been correctly classified in each achievement group and that the groups were relatively well balanced. The NA-MD group seemed to perform somewhat better in reading than the MD-only and the NA-MD/RD groups but these two last groups had perfectly normal scores considering their age and grade level. On average, MD-only and NA-MD/RD groups, respectively, scored −0.29 and −0.28 standard deviation below the grade level mean on the LUM subtest and −0.07 and 0.30 standard deviation around the grade level mean on the L3 subtest. Conversely, the NA-MD group scored, respectively, .21 and .87 standard deviation above the grade level mean on the LUM and the L3 subtests. The slight discrepancies between the scores of these three groups is sufficient to explain the emergence of statistical differences but in our view, the crucial point is that the MD/RD group scored well under the three other groups as reflected by the percentile scores in reading in Table 3.

2.6. Experimental data analyses

Repeated measures ANOVAs were carried out on both median reaction times (based on correct responses only) and correct response scores. In addition to the mean age difference between groups, the examination of reaction times and accuracy data in the physical comparison task indicated that MD/RD children tended to be generally slower and less accurate than controls (RTs: $F(2,89) = 3.04$, $p = .05$, Accuracy: $F(2,89) = 2.69$, $p = .07$). The other group differences were not significant (RTs and accuracy data: $ps > .39$). Although group differences were unexpected in this perceptual task involving no number processing, this result is in accordance with other studies which consistently reported a general slowdown in learning disabled children (see Censabella & Noël, 2005). Median reaction times were then considered as indicators of children's processing speed and error rate as reflecting children's attentional level. Therefore, group effect and interaction with this factor were further analysed using ANCOVAs controlling for age and general processing speed (i.e., median reaction times in physical comparison) in RT analyses, and for age and general error rate (i.e., error rate in physical comparison) in accuracy analyses (α significance threshold $< .05$). If group effect or group interactions really reflect differential processing, they should be resistant to the introduction of the covariates. For the sake of simplicity, only statistical results of the analyses with covariates will be given when describing group effect or interaction with this factor.

Other precautions were taken before running the analyses. First, subjects with studentized residuals superior to $+2$ or inferior to -2 in all tasks were considered as outliers. Only one MD-only child met this criterion in the reaction time data pool and was removed from all RT analyses. Second, in order not to introduce noise in the RT analyses, subjects whose correct response score was below or equal to chance level in one experimental condition (i.e., $\leqslant 6/12$) were excluded from the RT analyses of that task. In this way, median reaction times included in the analyses were always based on a reasonable number of correct responses in each experimental condition. The exact number of participants included in each RT analysis is mentioned at the beginning of each section. Third, repeated measures and multivariate ANCOVAs are parametric statistics which assume normal data distribution and variance homogeneity when two or more samples are being compared. Shapiro–Wilk's normality test and Levene's variance homogeneity test indicated that the RTs and accuracy data did not completely fit these parametric assumptions.[1] Therefore, logarithmic transformation was used in the RT analyses. In accuracy analyses however, the percentage of correct responses across experi-

[1] See Appendix C for statistics of Shapiro–Wilk's normality test and Levene's variance homogeneity test on RTs and accuracy data for the MD ($n = 45$) and the NA ($n = 45$) groups in each task and in each experimental condition.

mental tasks ranged between 90.6 and 97.1% in the MD/RD group, 91.2 and 98.1% in the MD group and 94.8 and 98.9% in the NA group. As mean scores were at or near ceiling in all tasks, no transformation could improve the fit between accuracy data and the assumptions of the statistical model. The usual solution would have been to use non-parametric statistics to test group effects in the various tasks. However, non-parametric tests do not make it possible to take into account the two covariates. With such confounds, there is no way of determining whether group differences indicate particular processing, or simply reflect age differences between groups or reveal a general group tendency to be more or less error prone. It has long been established that the analysis of variance is a very robust statistical procedure and that violations of parametric assumptions have only a minor effect or no effect at all on substantive conclusions (Howell, 2001, pp. 340). Therefore, accuracy data were analysed using ANCOVAs with age and general error rate covariates. However, only group effects were analysed using this procedure and results should be considered with caution. Due to the near-ceiling performance in all tasks, within-subject effects were not examined on accuracy data.

Finally, performances of the two NA subgroups (RTs and accuracy data) were compared in order to collapse them in the subsequent analyses. Multivariate ANCOVAs were run on both RTs and accuracy data (with their respective covariates, i.e., age and processing speed in RT analysis, and age and general error rate in accuracy analysis) in each experimental condition and in each task (mean of all task conditions). As reported in the Appendix A, the two NA subgroups did not differ on any task, nor on any experimental condition, except for the reaction time data for the large–close pairs in the Arabic comparison task ($p = .04$). As there was no reason to think that this single difference reflected differential processing between our controls, NA-MD and NA-MD/RD subgroups were collapsed into a single group.

3. Results

Results are presented in six subsections. In the first one, the two groups with low mathematics achievement are compared in all tasks and experimental conditions. Sections 3.2 and 3.3 present results for the Arabic number and the collection (ratio-set only) comparison tasks, respectively. Section 3.4 directly assesses the effect of the presentation format by contrasting children's ability to compare the same magnitude in the collection (distance-set) and the Arabic number comparison tasks. Results of the Stroop task are presented in Section 3.5 and the relations between mathematics achievement level and performance in experimental tasks are examined in the final section.

3.1. Comparison of the two mathematics disabled groups

Merging the two NA subgroups led to compare three very unequal-sized samples, i.e., 16 MD/RD children, 29 MD-only children and 45 NA children. These discrepancies between the number of participants in each group automatically reduce the statistical power of the analysis and, consequently, increased the probability of making type II errors (i.e., accepting a false null hypothesis), especially when comparing the MD/RD and the NA groups, which included samples of very different sizes. In keeping with Landerl et al. (2004), who showed that dyscalculic children with and without reading deficit showed exactly the same pattern of numerical disabilities, RTs and accuracy data of the two MD groups were directly compared in multivariate ANCOVAs (with their respective covariates). Results reported in the Appendix B indicated that the groups did not differ on any task, nor in any experimental condition (ps $> .05$). Consequently, both groups with low mathematics achievement were collapsed into a single MD group ($n = 45$) and compared to the 45 NA children in the subsequent analyses.

3.2. Arabic number comparison task

3.2.1. RTs

Forty-two MD and 45 NA children[2] were included in this RT analysis. A repeated measure ANOVA with size (small, large) and distance (close, far) as within-subject factors and group (MD, NA) as between-subjects factor revealed significant effects of size, $F(1,85) = 81.97$, $p < .001$, $MSE = .239$, $\eta^2 = .49$, power = 1, distance, $F(1,85) = 146.79$, $p < .001$, $MSE = .331$, $\eta^2 = .63$, power = 1, as well as a size by distance interaction, $F(1,85) = 9.21$, $p < .001$, $MSE = .015$, $\eta^2 = .10$, power = .85. Children were faster at comparing small pairs than large ones (mean size effect: 119 ms) and faster at comparing far digits than close ones (mean distance effect: 148 ms). Moreover, when size increases, the distance effect gets smaller, i.e., the distance effect is larger for small pairs (173 ms) than for large ones (122 ms). Such an interaction is predicted by analog numerical models (Gallistel & Gelman, 1992; Dehaene & Changeux, 1993), since the subjective distances between underlying representations become smaller when numerical magnitude increases. Finally, when covariates were introduced into the analysis, the main effect of group was significant, $F(1,83) = 24.55$, $p < .001$, $MSE = .253$, $\eta^2 = .23$, power = 1, as were the size by group, $F(1,83) = 3.92$, $p = .05$, $MSE = .012$, $\eta^2 = .05$, power = .5, and the distance by group interactions $F(1,83) = 4.33$, $p < .05$, $MSE = .01$, $\eta^2 = .05$, power = .54. The size by distance by group interaction was not significant, $F(1,83) = 1$. As shown in Fig. 1, the MD group was on average slower than the NA group (see Table 4) and exhibited a slight reduction of the size ($M = 93$ ms, $SD = 187.5$) and distance effects ($M = 131$, $SD = 136.1$) compared to the NA group (size effect: $M = 146$ ms, $SD = 91.1$, distance effect: $M = 164$ ms, $SD = 126.2$). Further analyses showed that the group effect was significant for each size and distance level, ps $< .001$, but somewhat higher for small and far trials, Fs$(1,83) = 23.70$ and 32.59, respectively, than for large and close pairs, Fs$(1,83) = 14.43$ and 11.55, respectively.

[2] One MD child was removed from all RT analyses because he was extremely slow in all tasks (outlier). The two others were excluded from this analysis because they performed at chance in at least one experimental condition (i.e., correct response score $\leqslant 6/12$).

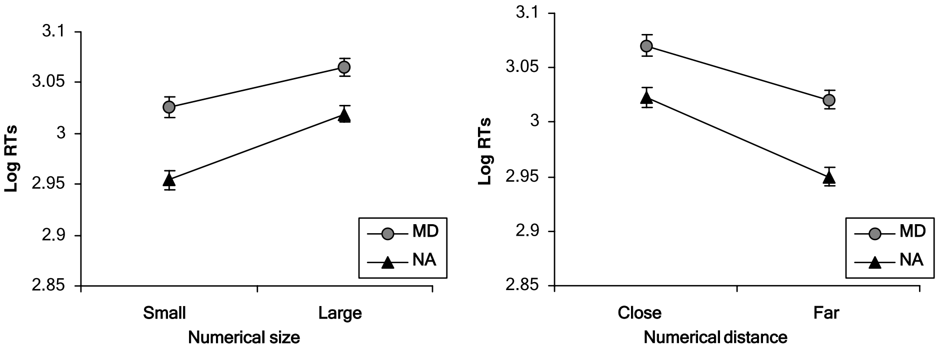


Fig. 1. Numerical size and distance effects in the Arabic digit comparison task: mean reaction times (log) and standard errors for each achievement group (adjustment for covariates).

Table 4
Mean reaction times and accuracy data by task and achievement group

Group	Physical comparison	Arabic number	Collection[a]		Congruity effect[b]
			Density	Surface	
RTs					
MD	771 ms	1156 ms	809 ms	976 ms	88 ms
	(104)	(157)	(104)	(155)	(104)
NA	721 ms	970 ms	759 ms	936 ms	82 ms
	(120)	(151)	(94)	(144)	(60)
Accuracy					
MD	11.73/12	11.01/12	11.62/12	11.10/12	.71
	(.36)	(.56)	(.61)	(.91)	(.89)
NA	11.87/12	11.38/12	11.78/12	11.47/12	.69
	(.27)	(.49)	(.33)	(.86)	(.76)

Note: Standard deviations are shown in parentheses. MD = mathematics difficulties; NA = normal achievement. See Appendix D for an exhaustive presentation of RTs and accuracy data in each task and in each experimental condition.
[a] Only data of the ratio-set are reported.
[b] Difference between data in the congruent and incongruent conditions.

3.2.2. Accuracy

A univariate ANCOVA with group (MD, NA) as between-subjects factor and with age and general error rate as covariates yielded a significant main effect of group, $F(1,86) = 8.25$, $p < .01$, $MSE = 2.248$, $\eta^2 = .09$, power = .81, indicating that the MD group scored on average lower than the NA group (see Table 4).

3.3. Collection comparison task

This section presents the results for the ratio-set. Data on the distance-set were used to assess the effect of presentation format and are presented in the next section.

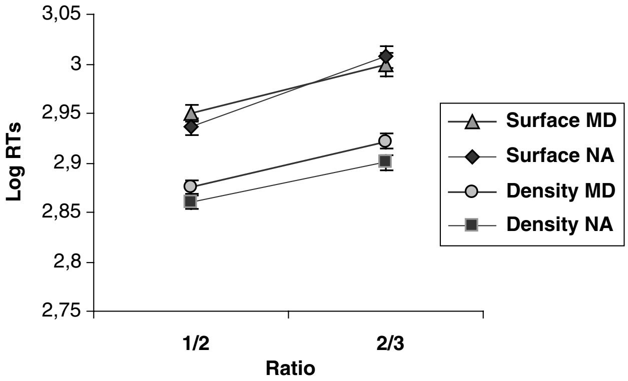


Fig. 2. Ratio effect in the density and surface conditions of the collection comparison task: mean reaction times (log) and standard errors for each achievement group (adjustment for covariates).

3.3.1. RTs

Forty-three MD and 44 NA children[3] were included in this RT analysis. A $2 \times 2 \times 2$ repeated measures ANOVA was run on median reaction times (see Table 4) with condition (density, surface) and ratio (1/2, 2/3) as within-subject factors and group (MD, NA) as between-subjects factor. There was a main effect of perceptual condition, $F(1,85) = 251.43$, $p < .001$, $MSE = .613$, $\eta^2 = .75$, power = 1, a main effect of ratio, $F(1,85) = 189$, $p < .001$, $MSE = .233$, $\eta^2 = .69$, power = 1, and a significant ratio by condition interaction, $F(1,85) = 4.94$, $p < .05$, $MSE = .006$, $\eta^2 = .06$, power = .59. The comparison was on average 152 ms faster in the density than in the surface condition and 110 ms faster for the 1/2 than for the 2/3 ratio. This ratio effect was more pronounced in the surface condition (mean ratio effect: 140 ms) than in the density condition (mean ratio effect: 80 ms). Furthermore, when covariates (age and processing speed) were entered in the analysis, group effect was not significant, $F(1,83) = 1.01$, $p = .31$, $MSE = .009$, nor was any interaction with this factor, condition × group: $F(1,83) = 1.69$, $MSE = .004$, ratio × group: $F(1,83) < 1$, $MSE = .001$, condition × ratio × group: $F(1,83) = 3.24$, $MSE = .004$, all $ps > .05$. Fig. 2 presents mean reaction times by group for each ratio and condition in the collection comparison task.

3.3.2. Accuracy

A multivariate ANCOVA with group (MD, NA) as between-subjects factor and with age and general error rate as covariates was run on accuracy data (see Table 4) in each perceptual condition. Group difference was not significant either in the density, $F(1,86) = 1.34$, $MSE = .33$, or in the surface condition, $F(1,86) = 2.28$, $MSE = 1.769$, both $ps > .10$.

[3] One MD child was extremely slow in all tasks (outlier) and was removed from all RT analyses. Two other children (1 MD and 1 NA) were excluded from this analysis because they performed at chance in at least one experimental condition (i.e., correct response score $\leqslant 6/12$).

3.4. Effect of presentation format

The data presented so far hinted at the presence of group differences in the Arabic digit but not in the collection comparison tasks. However, these divergent results could be due to the use of less salient ratios in the Arabic digit comparison (from .20 to .89) than in the collection comparison task (.5 or .66). In this section, the effect of presentation format was assessed by contrasting children's performance at processing the same magnitude presented in symbolic (large numerosities of the Arabic number comparison task) and non-symbolic stimuli (distance-set of the collection comparison task). However, using the same distance for collections as for digit pairs strongly reduce the discriminability between collections to be compared (mean ratio = .75 for the distance-set vs .58 for the ratio-set of the collection comparison task). As a result, the percentage of errors was very high for non-symbolic stimuli in the distance-set, especially for close pairs where 38% of MD and 31% of NA children performed at chance level.[4] In order not to exclude too many subjects from the RT analysis and not to confuse the analyses in general, only data on far pairs (5–8, 5–9, and 6–9)[5] were examined.

3.4.1. RTs

Forty-three MD and 44 NA children[6] were included in this RT analysis. A 2×2 repeated measures ANOVA with format (non-symbolic vs symbolic) as within-subject factor and group (MD, NA) as between-subjects factor was carried out on RT data. The main effect of format, $F(1,85) = 13.26$, $p < .001$ $MSE = .041$, $\eta^2 = .14$, power = .95, showed that children were faster at comparing non-symbolic magnitudes. Moreover, when covariates were introduced into the analysis, there was a main effect of group, $F(1,83) = 11.44$, $p < .001$, $MSE = .054$, $\eta^2 = .12$, power = .92, and a significant format by group interaction, $F(1,83) = 5$, $p < .05$, $MSE = .016$, $\eta^2 = .06$, power = .60. As illustrated in Fig. 3, group effect was only significant when processing symbols, $F(1,83) = 15.62$, $p < .001$, $MSE = .06$, $\eta^2 = .16$, power = .97, but not when comparing non-symbolic magnitudes, $F(1,83) = 1.52$, $p > .10$, $MSE = .006$.

3.4.2. Accuracy

A multivariate ANCOVA with age and general error rate as covariates and group (MD, NA) as between-subjects factor was run on accuracy data for each presentation format. Group difference failed to reach statistical significance with the symbolic

[4] The number of children performing at chance on close trials in non-symbolic comparison was distributed as follows: 1 MD child in the density condition, 16 MD and 14 NA children in the surface condition and 1 NA child in both conditions.

[5] The symbolic mean was based on 12 observations while the non-symbolic mean was based on 24 observations as a result of the collapsing of the density and surface conditions.

[6] One MD child was extremely slow in all tasks (outlier) and was removed from all RT analyses. The two others (1 MD and 1 NA) were excluded from this analysis because they performed at chance in at least one experimental condition (i.e., correct response score $\leqslant 6/12$).

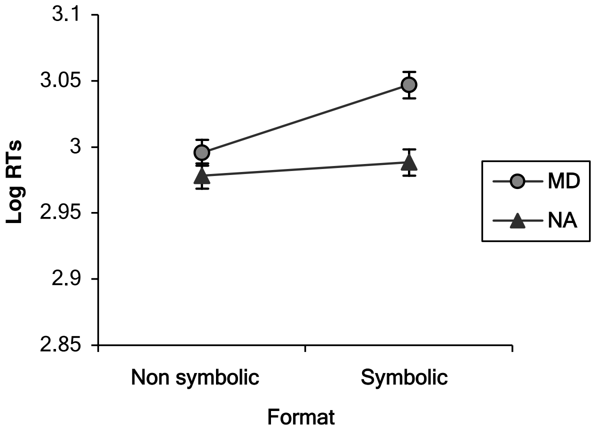


Fig. 3. Mean reaction times (log) and standard errors for each group on far pairs in the collection and the Arabic digits comparison tasks (adjustment for covariates).

presentation, $F(1,86) = 2.82$, p = .097, $MSE = 1.444$, and was completely non-significant for non-symbolic stimuli, $F(1,86) = 1.02$, p = .31 $MSE = 0.799$.

3.5. Stroop task

The following analyses were run to examine the presence of signs indicating automatic processing of the irrelevant numerical dimension in each group, i.e., the congruity, size and distance effects.

3.5.1. RTs

Forty MD and 42 NA children[7] were included in these RT analyses. A $2 \times 2 \times 2 \times 2$ repeated measures ANOVA with condition (congruent, incongruent), size (small, large) and distance (close, far) as within-subject factors and group (MD, NA) as between-subjects factor was run on RT data. The analysis yielded significant effects of condition, $F(1,80) = 113.52$, $p < .001$, $MSE = .44$, $\eta^2 = .59$, power = 1, distance, $F(1,80) = 17.96$, $p < .001$, $MSE = .037$, $\eta^2 = .18$, power = .99, as well as significant distance by condition, $F(1,80) = 24.59$, $p < .001$, $MSE = .055$, $\eta^2 = .24$, power = 1, size by condition, $F(1,80) = 6.16$, $p < .05$, $MSE = .031$, $\eta^2 = .07$, power = .69, and size by distance interactions, $F(1,80) = 55.44$, $p < .001$, $MSE = .167$, $\eta^2 = .41$, power = 1. No other within-subject effects were significant, nor were any group effects and interactions after the covariates had been introduced into the analysis, ps > .05. Physical comparison was on average 85 ms faster for congruent trials than for incongruent ones and this congruity effect was equivalent in both groups (see Fig. 4 and Table 4), as attested by the absence of a condition by

[7] One MD child was extremely slow in all tasks (outlier) and was removed from all RT analyses. The others (four MD and three NA) were excluded from this analysis because they performed at chance in at least one experimental condition (i.e., correct response score $\leqslant 6/12$).

group interaction in the analyses with covariates, $F(1,78) < 1$. Follow-up analyses of the interactions indicated that children were faster at comparing close pairs than far ones (i.e., reversed distance effect), and faster at small trials than large ones in the congruent condition, distance: $F(1,81) = 52.87$, $p < .001$, $MSE = .046$, $\eta^2 = .40$, power = 1, size: $F(1,81) = 7.21$, $p < .01$, $MSE = .014$, $\eta^2 = .08$, power = .76, but none of the distance and size effects were significant in the incongruent condition, both $Fs(1,81) < 1.7$, $ps > .10$. Finally, the size by distance interaction showed that close pairs were processed faster than far ones on small trials (i.e., reversed distance effect), $F(1,81) = 75.67$, $p < .001$, $MSE = .092$, $\eta^2 = .48$, power = 1, but that far pairs were compared faster than close ones on large trials (i.e., classical distance effect), $F(1,81) = 8.58$, $p < .01$, $MSE = .012$, $\eta^2 = .10$, power = .83.

Results regarding size and distance effects are quite confusing as these were expected to have radically different forms in each congruity condition. Logically, all properties that favour access to numerical meaning, that is, small size or far distance, should be an advantage in the congruent condition where physical and numerical information converge to the same response, but should constitute a handicap in the incongruent condition where the two dimensions are conflictual. Consequently, if children automatically processed number meaning, we expected classical size and distance effects in the congruent condition (i.e., better performance for small than for large pairs and for far than for close pairs) but inverted size and distance effects in the incongruent condition (i.e., better performance for large than for small pairs and for close than for far pairs). Our data did not fit these predictions, which makes little sense in our view. Given that none of these effects interacted with group, they were not further investigated.

Finally, a last analysis was conducted on the delay elapsed before the appearance of a size difference between digits which was calculated for each participant. This delay reflects the additional time needed to compare digits when physical encoding is already completed, that is, the time to access to semantic information from sym-

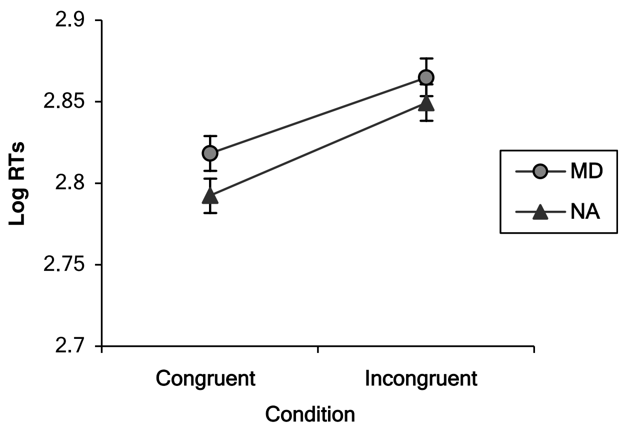


Fig. 4. Congruity effect in the Stroop task: mean reaction times (log) and standard errors for each group (adjustment for covariates).

bols. An univariate ANCOVA controlling for age with group (MD, NA) as between-subjects factor yielded a significant effect of group, $F(1,79) = 15.65$, $p < .001$, $MSE = 324898.71$, $\eta^2 = .17$, power = .97 indicating that MD children ($M = 400$, $SD = 170.97$) needed more time to access to semantic information from symbols than controls ($M = 276.2$, $SD = 112.21$). It is worth noting that NA children needed on average 276 ms more to access to number magnitude after a physical size decision have been reached. To our view, this imbalance between physical and numerical processing speed might account for Landerl et al.'s (2004) failure to find any sign of automatic numerical processing in NA children. In their study, number magnitude might well have been processed by controls, but not fast enough to influence the response as a physical size decision was already made (Noël et al., 2005).

3.5.2. Accuracy

One MD child was excluded from this analysis because he did not follow the instructions and compared digit on the basis of numerical instead of physical information. As a result, his score was near from 0 in the incongruent condition. To examine the modulation of automatic processing by the achievement group in accuracy data, a congruity effect was calculated for each child by subtracting his/her mean score on incongruent trials from his/her mean score on congruent trials (see Table 4). A multivariate ANCOVA with age and general error rate as covariates and group (MD, NA) as between-subjects factor was carried out on this congruity effect and revealed no significant effect of group, $F(1,85) < 1$.

3.6. Experimental measures and mathematics achievement

Previous sections showed that performance in the collection comparison and in the Stroop task could not be used as indicators of mathematical difficulty since group differences were not significant in these tasks. By contrast, results indicated that MD and NA children differed when comparing Arabic digits. In this last section, the relations between performance on experimental measures and mathematics achievement are explored by addressing two issues. First, could performance in Arabic number comparison be considered a powerful indicator of mathematics achievement level? Second, does the absence of any group difference in the collection comparison task persist in children with a very low level in mathematics, namely, in *dyscalculic* children? Due to the near-ceiling performance in all tasks, only RTs were considered in this section.

To address the first issue, standard scores of RTs in the Arabic number comparison task were calculated for each participant on the basis of the mean and the standard deviation of the NA group (NA group: $M = 0$, SD = 1, MD group: $M = 1.24$, SD = 1.04). Considering a reasonable cutoff of 2 standard deviations above the control mean, only 9/42 MD children would have been classified as mathematically disabled. With a more conservative cutoff of 2.5 above the control mean, only 5/42 MD would have been classified as dyscalculic while 2/45 NA children scoring 80 and 85%, respectively in mathematical achievement tests would have been designated as dyscalculic as well. Finally, 13/42 MD children had perfectly normal reaction times in this task since five performed below the control mean reaction time and eight performed between 0 and .5 standard deviation above the control mean. Thus, despite the presence of a group difference in the Arabic number comparison task, the overall pattern suggests that performance in this task could not be used as a reliable and powerful indicator of mathematical difficulties.

The second question concerned the relations between the present results and the inclusion criterion used to screen for MD in this study (i.e., the 15th percentile). This issue was addressed by examining RT data in collection and Arabic digit comparisons using a much more stringent selection criterion comparable to the one used in Landerl et al.'s (2004) study, namely, the 3rd percentile. Ten children had a mathematics composite score inferior or equal to this cutoff in our MD sample and, therefore, could be labelled as dyscalculic. As discussed previously, comparing such a small sample with the 45 NA children using ANOVA would considerably increase the probability of drawing misleading conclusions (i.e., accepting a false null hypothesis). So instead, children's RTs in the collection comparison task[8] were converted into standard scores using the mean and standard deviation of the NA group. On average, these 10 dyscalculic children, respectively, performed 0.57, −0.43 and −0.07 standard deviation around the control mean in the density condition ($SD = 1.14$, $Min = -1.23$, $Max = 2.36$), in the surface condition ($SD = .73$, $Min = -1.73$, $Max = .79$), and finally, in the collection comparison task regardless of the condition ($SD = .86$, $Min = -1.12$, $Max = 1.45$). By contrast, their standard score in the Arabic digit comparison task was on average 1.53 standard deviation above the control mean ($SD = .97$, $Min = -0.30$, $Max = 3.12$). As RTs of dyscalculic children fell in the normal range when comparing collections but not when comparing Arabic numerals, these results confirmed the discrepancy between performances in the collection and in the Arabic digit comparison tasks even in children with a very low level.

4. Discussion

In summary, the present results demonstrated that MD children performed slower and less accurately than NA children when comparing Arabic digits and showed slightly reduced number size and distance effects in this task compared to controls. Conversely, they exhibited a normal ratio effect in the collection comparison task and were able to compare numerosities as well as NA children, regardless of the perceptual condition. In the Stroop task, their performance was influenced by the irrelevant numerical dimension, as attested by the presence of congruity, numerical size and distance effects in each group. Finally, no evidence was found for a dissociation between MD and MD/RD children in tasks assessing basic numerical abilities.

The current study contrasted two alternative views: the defective number module hypothesis assuming that MD results from a core deficit of numerosity processing,

[8] All ratio and distance put together.

and the access deficit hypothesis according to which MD children have difficulty accessing magnitude information from numerical symbols rather than in processing numerosities per se. As MD children were impaired in the Arabic number comparison, our results replicate the findings of Landerl et al. (2004). However, contrary to their predictions, MD children failed to exhibit any deficit in the collection comparison task, even when comparing the same number magnitudes that were underperformed in the symbolic comparison. This failure to find any difference in the collection comparison task was unrelated to the selection criterion used in this study since even severe dyscalculic children showed RTs in the normal range in this task.

Taken together, these findings contradict the defective number module hypothesis which predicts that MD children should be impaired in all tasks requiring them to process number magnitude. As they were able to compare continuous and discrete quantities but were slower and less accurate when processing the number meaning of Arabic symbols, the overall pattern is in accordance with the access deficit view. In line with this hypothesis, the difference between numerical and physical processing speed (i.e., the delay) was found to be greater for MD children, providing strong evidence that they need more time than controls to access semantic properties after encoding the physical characteristics of symbols.

The pattern of performance of MD children bears some resemblance to the profile of younger, academically normal children who are known to be slower and less accurate than older children when comparing Arabic numerals (Duncan & McFarland, 1980; Sekuler & Mierkiewicz, 1977). These similarities are suggestive of immature development in the symbolic numerical processing capacities of MD children. However, unlike children who exhibit a larger distance effect the younger they are, MD children showed a reduction in the distance and the size effects compared to controls. One possibility is that their difficulty in accessing semantic representation from numerical symbols leads some children to use peculiar strategies for comparing digits. Indeed, five MD children showed an reversed distance effect, nine others presented an inverted size effect and one exhibited both inverted size and distance effects. When those children were excluded from the analysis, the distance by group and the size by group interactions disappeared, suggesting that these inverted effects led to a reduction of the overall size and distance effects in the MD group. Although it seems very unlikely given the range of their reaction times ($M = 1146.54$; $SD = 190.89$; $Min = 929$ ms; $Max = 1533$ ms), the reverse distance effect might indicate that these children recited the number sequence to find out the answer just as Charles did (Butterworth, 1999). However, the presence of an inverted size effect is not so easily explainable as it suggests the use of a strategy that would be more efficient for processing large pairs than small ones.

As the access deficit hypothesis assumes a deficit of the operations relating numerical symbols to the corresponding magnitudes rather than an alteration of the magnitude representation itself, it might simultaneously account for the reduction of size and distance effects in the Arabic digit comparison and for the normal ratio effect in the collection comparison task. In our study, the higher error rate in Arabic digit comparison indicated that MD children's access to number magnitude from symbols is not only slower but also more variable. Logically, this increased variability in accessing semantic representation should result in more diffuse activation of the corresponding magnitudes, making them less distinguishable. In this view, the reduction in discriminability might have more strongly affected the comparison of small and far trials because these pairs are initially easier to distinguish than large and close trials which are more difficult to discriminate even for control children. This might account for the slightly higher group effect on small and far trials and consequently, for the reduction of size and distance effects in the MD group. By contrast, access to number magnitude from non-symbolic stimuli is not affected by such variability, which explains the presence of a normal ratio effect in the collection comparison task.

Ultimately, one final possible explanation for this finding would be to consider that the reduction of size and distance effects reflects the use by controls of various solving procedures for comparing digits while MD children would only apply a single process. According to Tzelgov, Meyer, and Henik (1992), two qualitatively different processes are at work during number comparison in adults: an algorithm-based process in which number magnitudes are compared at the semantic level, and a faster memory-based process which consists of directly retrieving the correct answer from memory. Initially, children would only be able to compare digits using the algorithmic process. But, with age, their growing experience with Arabic digit comparison would lead them to develop long-term memory associations between the digit pair to be compared and the correct response. It should be noted that very similar processes are at work in reading development: at the beginning, children learn to decode words applying letter-to-sound algorithms but progressively, their growing experience leads to the storage of lexical knowledge in memory which allows them to directly recognize words as a whole. Accordingly, in our study, one possibility is that NA children would have accumulated enough experience with digit comparison to retrieve some answers directly from memory, in particular those relative to the more frequent and easily comparable pairs, that is, small and far pairs. As memory retrieval is faster than algorithmic comparison, the use of retrieval in NA children, even for some instances of small and far pairs, might have reduced their global reaction times for these pairs, inducing larger size and distance effects in NA children compared to MD children who processed all pairs algorithmically.

Results in the Stroop task also challenge the defective number module hypothesis, according to which MD children are not supposed to process unintentionally the meaning of numerical symbols due to their lack of understanding of numerosity concepts. However, the presence of comparable congruity effects in both groups demonstrated that they do. Other signs of numerical processing were found, since the children's performance was also influenced by the numerical size and distance of the pairs, independently of the achievement group. These results clearly indicated that they activated number magnitude in a task that did not require such processing. So, despite their slowness when accessing number magnitude from symbols, MD children automatically processed semantic properties under conditions adapted for their own symbolic processing speed, which is in accordance with the access deficit hypothesis.

Finally, the present data add to the large body of studies which attempted to separate mathematical profiles of MD children with and without comorbid RD. In keeping with Landerl et al. (2004) who also contrasted these learning disabled groups in elementary quantification tasks, the present research found no evidence for differential patterns of performance in basic numerical processing skills between MD children with or without RD. In more complex tasks, however, the data collected so far about the difference between MD-only and MD/RD children might seem confusing and sometimes contradictory as both groups can exhibit changing patterns of similarities and divergences in the same tasks from one study to another (i.e., see Jordan & Montani, 1997 vs Hanich et al., 2001; Jordan et al., 2003 & Jordan & Hanich, 2003 for incompatible results in timed exact calculation task). The heterogeneity of participants' grade levels and of inclusion criteria used to constitute learning disabled samples might be partly responsible for these discrepancies. Nevertheless, two areas of the mathematics field seem to be consistently more impaired in MD/RD than in MD children: story problem-solving and exact calculation without time limit (Fuchs & Fuchs, 2002; Hanich et al., 2001; Jordan & Hanich, 2003; Jordan & Montani, 1997; Jordan et al., 2003; Jordan et al., 2002). The advantage of MD-only children over the MD/RD group in these two tasks probably results from their higher verbal abilities which enable them to better understand verbal problems and to execute back up strategies more skillfully under untimed condition than MD/RD children (Jordan & Montani, 1997). In the present study, however, both groups were compared in rudimentary quantification tasks requiring, to our knowledge, no verbal processing, what could account for their equivalent performance. The absence of any difference between learning disabled groups in the most basic numerical processing suggests that these children experience a very specific disorder of numerical cognition and provides additional evidence against Rourke's classification of learning deficits, which assumed different origins for MD according to whether it is associated with RD or not.

To sum up, the convergence of results across the present tasks provides evidence in support of the access deficit view. As MD children exhibit a greater processing time difference between numerical and physical comparison than NA children, the deficit should take place between the visual analysis and the semantic representation. At present however, it is not possible to know exactly which processing step is deficient among those preceding the activation of the semantic representation. For instance, considering the triple-code model (Dehaene, 1992), the deficit could affect the digit visual identification stage as well as the connections between the Arabic visual form and the semantic number representation. In the future, examining the integrity of the visual identification processes (e.g., Arabic digit lexical decision) in MD children should allow determining the precise location of the deficit in cognitive models of number processing. For now, results merely indicate that the locus of the symbolic processing deficit is pre-semantic.

What can we conclude on the basis of the present results? Do they indicate that MD children do not have difficulty in processing numerosities? At least two points must be considered before drawing such a conclusion. First, the psychophysical performance curve for numerosity comparison was not entirely measured in the non-symbolic comparison task, as the ratio between collections varied between .5 and .66. So, it is still possible that the use of less discriminable ratios would have revealed group differences with non-symbolic material as well. Second, with respect to the hypothesis of a genetic etiology of mathematical disabilities, another possibility could be that the numerosity processing deficit was present in the first years of life but was completely compensated by the time children enter to school.

Nevertheless, in the age range tested here, our findings suggest that a core deficit in MD is a difficulty in relating numerical symbols to their underlying meaning. In this respect, Butterworth (2003) recently developed a small test battery wherein basic numerical abilities are assessed through dot counting and Arabic number comparison to screen for dyscalculia. However, in spite of the presence of reliable group differences when comparing Arabic symbols, the present study demonstrated that using performance in the Arabic number comparison to screen for MD could result in numerous misleading classifications (i.e., failure to detect children with low mathematical achievement or conversely, inclusion of normally achieving children). Therefore, this task cannot be considered as a reliable indicator of mathematical disabilities but should rather be viewed as a starting point for further investigations.

These findings indicated that the ability to process the meaning of Arabic digits is undoubtedly a central aspect of mathematical achievement but certainly not the only one. In the course of learning, this ability develops interactively with many other types of conceptual and procedural knowledge. For example, children learn to use counting, to master base-10 numerical coding, to understand basic arithmetical operations and their complementarity, to grasp calculation principles, to develop problem-solving strategies, and so forth. All this knowledge somehow requires connecting numbers and mathematical symbols in general, to their underlying meaning, which in turn favours the acquisition of new procedures and concepts (Gersten & Chard, 1999; Gersten, Jordan, & Flojo, 2005).

Given that mathematical thinking is deeply rooted in this ability to connect numbers to their meaning, children with an access deficit are likely to experience more and more problems in mathematics, which could account for the various behavioural manifestations of MD (Gersten & Chard, 1999; Robinson et al., 2002). As MD children are slower and less accurate when accessing number meaning, they have difficulty acquiring other mathematical concepts and procedures such as decomposing a complex problem in simpler ones or learning arithmetic facts which are less meaningful for them. Moreover, as processing number meaning is effortful, few attentional resources can be allocated to the monitoring of problem-solving steps. Consequently, MD children make recurrent procedural errors which might lead to incorrect associations in long-term memory. In addition, their basic disability should make them less able to recognize gross calculation errors. Finally, as they feel less comfortable with mathematics than their peers, they will prefer to use concrete supports (e.g., fingers) and to resort to overlearned strategies (e.g., counting all) to solve problems, what contributes to the persistence of immature strategies.

These considerations have practical implications for teaching and remediation. Too often, traditional teaching methods tend to dissociate procedures from the underlying concepts (Gersten & Chard, 1999; Gersten et al., 2005; Robinson et al., 2002). They place great emphasis on algorithms and strategies with few opportunities for children to understand the meaning of the arithmetical operations and the relations between them. Instead, instructional programs and mathematical interventions should encourage access to the meaning of numerical and arithmetical concepts (for example, stressing the cardinal and ordinal properties of numbers or the complementarity between addition and subtraction and between multiplication and division). In practice, this involves having a very concrete approach to mathematics, for example, by using visual supports to represent magnitude properties or arithmetic problems.

The present findings open many new perspectives for research on mathematical disabilities. In the future, the dissociation between symbolic and non-symbolic processing in MD children will have to be confirmed with a larger range of numerical ratios and with multi-digit numbers. In addition, ongoing studies will aim to extend these results to other numerical symbols, and in particular, to verbal numerals which are acquired earlier than Arabic digits. Several data support the idea that the access deficit hypothesis generalizes to verbal number processing as MD children were consistently found to be slower in counting and in reciting number sequence (Hitch & McAuley, 1991; Landerl et al., 2004). If the access deficit hypothesis is correct and extends to verbal number processing, performance in verbal numerical tasks could be expected to predict mathematical achievement a few years later. Experiences on verbal number comparison will nevertheless have to face up to methodological problems in controlling the sequential parameters inherent in language. For instance, comparing "six" to "seven" may take more time than comparing "six" to "nine" because of the numerical distance but also because in the last case, the two number words are faster identified as they do not start with the same phoneme.

In conclusion, the overall pattern of performance provides evidence that MD children do not have difficulty in processing numerosity per se but rather in accessing semantic information conveyed by numerical symbols. These difficulties appear to be independent of reading skills. Nonetheless, we do not preclude the possibility that other cognitive processes might contribute to MD. Depending on the child, mathematical disabilities could well result from a single deficiency as well as from a combination of impairments. Currently, the literature only mentions scattered presumptions about the cognitive deficits responsible for MD and there is a growing need to move towards an integrated theory of mathematical development. Further insights into the components implicated in MD and their relationships would help in distinguishing between different MD profiles and setting up specific rehabilitation programs. In the light of the present results, the access deficit to number meaning hypothesis seems to have strong explanatory potential for accounting for MD. Future research should aim to determine precisely how this basic deficit affects the diverse facets of mathematical development.

Acknowledgements

This study was supported by the National Fund for Scientific Research of Belgium (FNRS), by the Concerted Actions Research of the Belgian French

Community (Convention No. 05/10-327, "the development of controlled cognitive processes"), and by a Research Training Network grant from the EEC (MRTN-CT-2003-504927, "Numeracy and brain development"). We thank Julie Engels for her contribution to the participant's screening and to the data collection procedure. We are grateful to the headmasters, teachers and children of the participating schools in the area of Louvain-la-Neuve (Belgium).

Appendix. Supplementary material

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.cognition.2006.01.005.

References

Alarcón, M., DeFries, J. C., Light, J. G., & Pennington, B. F. (1997). A twin study of mathematics disability. *Journal of Learning Disabilities, 30*, 617–623.

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

Bellugi, U., Wang, P. P., & Jernigan, T. L. (1994). Williams syndrome: An unusual neuropsychological profile. In S. H. Broman & J. Grafman (Eds.), *Atypical cognitive deficits in developmental disorders: Implication for brain function*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Bruandet, M., Molko, N., Cohen, L., & Dehaene, S. (2004). A cognitive characterisation of dyscalculia in Turner syndrome. *Neuropsychologia, 42*, 288–298.

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

Butterworth, B. (2003). *Dyscalculia screener*. London: nferNelson Publishing Company.

Butterworth, B. (2005). Developmental dyscalculia. In J. I. D. Campbell (Ed.), *Handbook of Mathematical Cognition*. New York: Psychology Press.

Censabella, S., & Noël, M.-P. (2005). The inhibition of exogenous distracting information in children with learning disabilities. *Journal of Learning Disabilities, 38*, 400–410.

Dehaene, S. (1992). Varieties of numerical abilities. *Cognition, 44*, 1–42.

Dehaene, S., & Changeux, J.-P. (1993). Development of elementary numerical abilities: A neuronal model. *Journal of Cognitive Neuroscience, 5*, 390–407.

Dehaene, S., Piazza, M., Pinel, P., & Cohen, L. (2003). Three parietal circuits for number processing. *Cognitive Neuropsychology, 20*, 487–506.

DeStefano, D., & LeFevre, J.-A. (2004). The role of working memory in mental arithmetic. *European Journal of Cognitive Psychology, 16*, 353–386.

Duncan, E. M., & McFarland, C. E. (1980). Isolating the effects of symbolic distance and semantic congruity in comparative judgments: An additive-factors analysis. *Memory & Cognition, 8*, 612–622.

Fayol, M., Barrouillet, P., & Marinthe, C. (1998). Predicting arithmetical achievement from neuropsychological performance: A longitudinal study. *Cognition, 68*, B63–B70.

Fuchs, L. S., & Fuchs, D. (2002). Mathematical problem-solving profiles of students with mathematics disabilities with and without comorbid reading disabilities. *Journal of Learning Disabilities, 35*, 563–574.

Gallistel, C. R., & Gelman, R. (1992). Preverbal and verbal counting and computation. *Cognition, 44*, 43–74.

Geary, D. C., Bow-Thomas, 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. (1993). Mathematical disabilities: Cognitive, neuropsychological and genetic components. *Psychological Bulletin, 114*, 345–362.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications*. Washington, DC: American Psychological Association.

Geary, D. C. (2005). Les troubles d'apprentissage en arithmétique: rôle de la mémoire de travail et des connaissances conceptuelles. In M.-P. Noël (Ed.), *La dyscalculie : trouble du développement numérique de l'enfant*. Marseille: Solal.

Geary, D. C., Brown, S. C., & Samaranayake, V. A. (1991). Cognitive addition: A short longitudinal study of strategy choice and speed-of-processing differences in normal and mathematically disabled children. *Developmental Psychology, 27*, 787–797.

Geary, D. C., Hamson, 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.

Geary, D. C., Hoard, M. K., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk of mathematical disability. *Journal of Experimental Child Psychology, 74*, 213–239.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, C. M. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology, 88*, 121–151.

Gersten, R., & Chard, D. (1999). Number sense: Rethinking arithmetic instruction for students with mathematical disabilities. *The Journal of Special Education, 33*, 18–28.

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.

Girelli, L., Lucangeli, D., & Butterworth, B. (2000). The development of automaticity in accessing number magnitude. *Journal of Experimental Child Psychology, 76*, 104–122.

Gross-Tsur, V., Manor, O., & Shalev, R. S. (1996). Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology, 38*, 25–33.

Hanich, L. B., Jordan, N. C., Kaplan, D., & Dick, J. (2001). Performance across different areas of mathematical cognition in children with learning difficulties. *Journal of Educational Psychology, 93*, 615–626.

Hecht, S. A., Torgesen, J. K., Wagner, R. K., & Rashotte, C. A. (2001). The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology, 79*, 192–227.

Hitch, G. J., & McAuley, E. (1991). Working memory in children with specific arithmetical learning difficulties. *British Journal of Psychology, 82*, 375–386.

Howell, D. C. (2001). *Statistical methods for psychology* (5th ed.). Belmont, CA: Duxbury Press.

Isaacs, E. B., Edmonds, C. J., Lucas, A., & Gadian, D. G. (2001). Calculation difficulties in children of very low birthweight. *Brain, 124*, 1701–1707.

Jordan, N. C., & Hanich, L. B. (2003). Characteristics of children with moderate mathematics deficiencies: A longitudinal perspective. *Learning Disabilities Research and Practice, 18*, 213–221.

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

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.

Jordan, N. C., Kaplan, D., & Hanich, L. B. (2002). Achievement growth in children with learning difficulties in mathematics: findings of a two-year longitudinal study. *Journal of Educational Psychology, 94*, 586–597.

Khomsi, A. (1998). Epreuve d'évaluation de la compétence en lecture. Lecture de mots et compréhension-révisée (L.M.C-R. Paris: Editions du Centre de Psychologie Appliquée).

Koontz, K. L., & Berch, D. B. (1996). Identifying simple numerical stimuli: Processing inefficiencies exhibited by arithmetic learning disabled children. *Mathematical Cognition, 2*, 1–23.

Kosc, L. (1974). Developmental dyscalculia. *Journal of Learning Disabilities, 7*, 164–177.

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8- to 9-year-old students. *Cognition, 93*, 99–125.

Levy, L. M., Reis, I. L., & Grafman, J. (1999). Metabolic abnormalities detected by 1H-MRS in dyscalculia and dysgraphia. *Neurology, 53*, 639–641.

Lewis, C., Hitch, G. J., & Walker, P. (1994). The prevalence of specific arithmetic difficulties and specific reading difficulties in 9- to 10-year-old boys and girls. *Journal of Child Psychology and Psychiatry, 35*, 283–292.

Light, J. G., & DeFries, J. C. (1995). Comorbidity of reading and mathematics disabilities: Genetic and environmental etiologies. *Journal of Learning Disabilities, 28*, 96–106.

Lobrot, M. (1980). *Lire avec épreuves pour évaluer la capacité de lecture (D-OR-LEC)*. Paris: Editions ESF.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetic learning difficulties. *Journal of Experimental Child Psychology, 74*, 240–260.

Mix, k. S., Huttenlocher, J., & Levine, S. C. (2002). Multiple cues for quantification in infancy: Is number one of them?. *Psychological Bulletin 128*, 278–294.

Molko, N., Cachia, A., Rivière, D., Mangin, J. F., Bruandet, M., Le Bihan, D., et al. (2003). Functional and structural alterations of the intraparietal sulcus in a developmental dyscalculia of genetic origin. *Neuron, 40*, 847–858.

Mousty, P., & Leybaert, J. (1999). Evaluation des habilités de lecture et d'orthographe au moyen de BELEC: Données longitudinales auprès d'enfants francophones testés en 2ème et 4ème années. *Revue Européenne de Psychologie Appliquée, 4*, 325–342.

Mussolin, C. (2002). *L'automaticité de l'accès à la représentation numérique chez l'enfant*. Unpublished Master's thesis under the supervision of MP. Noël. Université Catholique de Louvain, Belgium.

Noël, M.-P. (2005). Finger gnosia: a predictor of numerical abilities in children?. *Child Neuropsychology, 11*, 413–430.

Noël, M-P., Rousselle, L., & Mussolin, C. (2005). Magnitude representation in children: Its development and dysfunction. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition*. New York: Psychology Press.

Noël, M.-P., Seron, X., & Trovarelli, F. (2004). Working memory as a predictor of addition skills and addition strategies in children. *Current Psychology of Cognition, 22*, 3–25.

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.

Reuhkala, M. (2001). Mathematical skills in ninth-graders: Relationship with visuo-spatial abilities and working memory. *Educational Psychology, 21*, 387–399.

Robinson, C. S., Menchetti, B. M., & Torgesen, J. K. (2002). Toward a two-factor theory of one type of mathematics disabilities. *Learning Disabilities Research and Practice, 17*, 81–89.

Rourke, B. P. (1993). Arithmetic disabilities, specific and otherwise: A neuropsychological perspective. *Journal of Learning Disabilities, 26*, 214–226.

Rourke, B. P., & Conway, J. A. (1997). Disabilities of arithmetic and mathematical reasoning: Perspective from neurology and neuropsychology. *Journal of Learning Disabilities, 30*, 34–46.

Rousselle, L., Palmers, E., & Noël, M-P. (2004). Magnitude comparison in preschoolers: What counts. Influence of perceptual variables. *Journal of Experimental Child Psychology, 87*, 57–84.

Rubinsten, O., Henik, A., Berger, A., & Shahar-Shalev, S. (2002). The development of internal representations of magnitude and their association with Arabic numerals. *Journal of Experimental Child Psychology, 81*, 74–92.

Russell, R. L., & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematics difficulties. *Cognition and Instruction, 1*, 217–244.

Sekuler, R., & Mierkiewicz, D. (1977). Children's judgments of numerical inequalities. *Child Development, 48*, 630–633.

Shalev, R. S., Manor, O., Kerem, B., Ayali, M., Badichi, N., Friedlander, Y., & Gross-Tsur, V. (2001). Developmental dyscalculia is a familial learning disability. *Journal of Learning Disabilities, 34*, 59–65.

Shalev, R. S., Manor, O., & Gross-Tsur, V. (1997). Neuropsychological aspects of developmental dyscalculia. *Mathematical Cognition, 3*, 105–120.

Shalev, R. S., Manor, O., Amir, N., Wertman-Elad, R., & Gross-Tsur, V. (1995). Developmental dyscalculia and brain laterality. *Cortex, 31*, 357–365.

Simon, T. J. (1997). Reconceptualizing the origins of number knowledge: A "non-numerical" account. *Cognitive Development, 12*, 349–372.

Ta'ir, J., Brezner, A., & Ariel, R. (1997). Profound developmental dyscalculia: Evidence for a cardinal/ordinal skills acquisition device. *Brain and Cognition, 35*, 184–206.

Temple, C. M. (1992). Developmental dyscalculia. In S. J. Segalowitz & I. Rapin (Eds.), *Handbook of neuropsychology: Vol. 7, Child Neuropsychology* (pp. 211–222). New York: Elsevier.

Temple, C. M., & Mariott, A. J. (1998). Arithmetical ability and disabilitiy in Turner's syndrome: A cognitive neuropsychological analysis. *Developmental Neuropsychology, 14*, 47–67.

Temple, C. M., & Sherwood, S. (2002). Representation and retrieval of arithmetical facts: Developmental difficulties. *The Quarterly Journal of Experimental Psychology, 55A*, 733–752.

Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently. A limited-capacity preattentive stage in vision. *Psychological Review, 101*, 80–102.

Tzelgov, J., Meyer, J., & Henik, A. (1992). Automatic and intentional processing of numerical information. *Journal of Experimental Psychology: Learning, Memory and Cognition, 18*, 166–179.

van der Sluis, S., de Jong, P. F., & van der Leij, A. (2004). Inhibition and shifting in children with learning deficits in arithmetic and reading. *Journal of Experimental Child Psychology, 87*, 239–266.

Van Nieuwenhoven, C., Grégoire, J., Noël, M.-P. (2001). Test diagnostique des compétences de base en mathématiques (TEDI-MATH. Paris: Les Editions du Centre de Psychologie Appliquée). (Available in Dutch and Spanish; translation in English in preparation).

Walsh, V. (2003). A theory of magnitude: Common cortical metrics of time, space and quantity. *Trends in Cognitive Sciences, 7*, 483–488.

Wechsler, D. (1996). *Echelle d'intelligence de Wechsler pour enfants, troisième édition (WISC-III)*. Paris: Les Editions du Centre de Psychologie Appliquée.