Child Development, July/August 2007, Volume 78, Number 4, Pages 1343 – 1359

Cognitive Mechanisms Underlying Achievement Deficits in Children With Mathematical Learning Disability

David C. Geary, Mary K. Hoard, Jennifer Byrd-Craven, Lara Nugent, and Chattavee Numtee

University of Missouri – Columbia

Using strict and lenient mathematics achievement cutoff scores to define a learning disability, respective groups of children who are math disabled (MLD, $n = 15$) and low achieving (LA, $n = 44$) were identified. These groups and a group of typically achieving (TA, $n = 46$) children were administered a battery of mathematical cognition, working memory, and speed of processing measures ($M = 6$ years). The children with MLD showed deficits across all math cognition tasks, many of which were partially or fully mediated by working memory or speed of processing. Compared with the TA group, the LA children were less fluent in processing numerical information and knew fewer addition facts. Implications for defining MLD and identifying underlying cognitive deficits are discussed.

Diagnostic criteria and thus the percentage of children with a learning disability in mathematics (MLD) remain unresolved (Barbaresi, Katusic, Colligan, Weaver, & Jacobsen 2005; Murphy, Mazzocco, Hanich, & Early, in press; Shalev, Manor, & Gross-Tsur, 2005). A common approach is to identify children as MLD if their achievement scores in mathematics fall below a specific cutoff. The cutoffs, however, range from lenient (< 30th percentile) to restrictive (< 5th or 10th percentile). In fact, much of the research in this area has been based on use of lenient criteria and thus may have conflated children with potentially severe to potentially mild forms of MLD. In the only study to separate groups based on potential severity, Murphy et al. (in press) classified children as MLD if they had math achievement scores less than the 11th percentile for 2 or more years and classified children as low achieving (LA) if they had scores between the 11th and 25th percentiles for 2 or more years. The latter group was composed of children who would have been classified as MLD in the majority of previous studies in this area. The MLD and LA groups did not differ on the single math cognition task administered in this study, but the children with MLD had a more general working memory deficit.

We defined MLD and LA groups using a more and less restrictive cutoff criterion, respectively, and compared these groups with each other and with a control group of typically achieving (TA) children. We expanded on the Murphy et al. (in press) study with the inclusion of a more extensive battery of math cognition tasks and working memory measures. Our goals were to determine whether the breadth and severity of the math cognition deficits differed for groups defined with restrictive and lenient diagnostic criteria and the extent to which different components of working memory mediated any group differences in math cognition. In the first section we describe the math cognition tasks used to assess potential deficits, and in the second section we discuss how working memory and speed of processing may contribute to individual and group differences on these tasks. We describe specific goals and predictions in the final section.

Mathematical Development in Children With MLD

Children identified as MLD using a strict criterion are likely to be performing poorly on most of the items that compose math achievement tests, whereas children identified using a lenient criterion are likely to be performing poorly on only some problem types. If so, as compared with TA children, children identified as

The first author acknowledges support from Grant R01 HD38283 from the National Institute of Child Health and Human Development (NICHD), and Grant R37 HD045914 cofunded by NICHD and the Institute of Education Sciences. We thank Linda Coutts, Kendra Andersen, Sara Ensenberger, Nicholas Geary, Larissa Haggard, Rebecca Hale, Mary Lemp, Patrick Maloney, Catherine Ford O'Connor, Amanda Shocklee, Jennifer Smith, and Ashley Stickney for help on various aspects of the project.

Correspondence concerning this article should be addressed to David C. Geary, Department of Psychological Sciences, University of Missouri, Columbia, MO 65211-2500. Electronic mail may be sent to GearyD@Missouri.edu.

MLD with the former criterion (referred to as MLD in this article) should show deficits on most or all math cognition measures, whereas children identified with the latter criterion (referred to as LA in this article) should show more circumscribed deficits. To test these hypotheses, we administered counting and addition tasks that have been shown to discriminate MLD and TA groups in previous studies. We also introduce a new measure, the Number Sets Test, and provide the first study of the ability of children with MLD to use mental representations of number lines to estimate quantity. Because a lenient criterion was used to classify children as MLD in most previous studies, the term MLD in the following reviews refers to groups that where largely composed of children that were classified as MLD or LA by Murphy et al. (in press) and in the current article.

Number sets. Many aspects of mathematics are dependent on a conceptual understanding of sets and the ability to manipulate sets in accordance with mathematical principles. Geary, Hoard, Byrd-Craven, and DeSoto (2004) hypothesized that knowledge of number sets facilitates use of problem-solving strategies involving decomposition of number sets. For instance, one way to solve an addition problem, such as 17 + 6, is to decompose the 6 into two sets of 3 and then to add these in succession (i.e., 17 + 6 = 17 + 3 + 3). Children with MLD use such strategies less frequently and less accurately than their TA peers, and we designed the Number Sets Test to evaluate the corresponding hypotheses: Children with MLD have a poor understanding of number sets or poor fluency of processing these sets (Koontz & Berch, 1996) and this in turn contributes to several of the arithmetical deficits commonly found with these children.

Number estimation. The ability to use a number line is a key component of children's understanding of number (Case et al., 1996; Griffin, Case, & Siegler, 1994), but development of this competency has not been assessed in children with MLD. Potential deficits associated with learning the number line have educational implications and at the same time may be useful in identifying underlying cognitive deficits.

This is because the ability to use a mental number line appears to be dependent on a potentially inherent magnitude representational system (Feigenson, Dehaene, & Spelke, 2004; Gallistel & Gelman, 1992). The inherent system results in estimations that conform to the natural logarithm (ln) of the number. In other words, the representations are compressed for larger magnitudes such that the perceived distance between 2 and 3 is larger than the perceived distance between 89 and 90. If children with MLD have deficits in the number-magnitude system, their number line estimates might not conform to the natural log representation or might show less precision when making estimates based on this representation. Moreover, with schooling, TA children's number line estimates gradually conform to the linear mathematical system (Siegler & Booth, 2004); the difference between two consecutive numbers is identical regardless of position on the number line. If children with MLD do not show evidence of a developing linear representational system, another source of MLD might be ease of modifying the logarithmic system to conform to the school-taught linear system.

Counting knowledge. By the time TA children enter kindergarten, they understand most basic counting concepts and can use counting procedures in many contexts (Briars & Siegler, 1984; Gelman & Gallistel, 1978), although refinement continues for several years (LeFevre et al., 2006). The basic concepts include Gelman and Gallistel's (1978) five implicit principles: (a) one-to-one correspondence (one and only one word tag is assigned to each counted object), (b) stable order (the order of the word tags must be invariant across counted sets), (c) cardinality (the value of the final word tag represents the quantity of items in the set), (d) abstraction (objects of any kind can be collected together and counted), and (e) order irrelevance (items within a given set can be tagged in any sequence). Children's counting behavior suggests they also make inductions about counting rules (Briars & Siegler, 1984). In particular, young children often induce that the unessential features of adjacency (items must be counted contiguously) and start at an end (counting must start on the left) are in fact essential.

Children with MLD understand most of the counting principles proposed by Gelman and Gallistel (1978) but often make errors on items that assess order irrelevance or adjacency (Geary, Bow-Thomas, & Yao, 1992). Performance on these items suggests an inflexible conception of how counting procedures can be executed, and it is correlated with use of immature counting procedures when solving arithmetic problems. These children also fail to detect errors when the first item is double-counted (i.e., the item is tagged *one, two*) but detect these double-counts when they occur with the last item. The pattern suggests children with MLD understand one-to-one correspondence but have difficulty retaining a notation of the counting error in working memory (Hoard, Geary, & Hamson, 1999). Poor skill at detecting counting errors may compromise ability to correct these errors and thus result in more errors in situations in which counting is used to solve arithmetic problems.

Arithmetic. Children use a mix of counting and memory-based processes to solve simple addition problems (Ashcraft, 1982; Siegler & Shrager, 1984). The strategy mix is initially dominated by finger and verbal counting, and the most commonly used procedures are sum and min (Fuson, 1982; Groen & Parkman, 1972). Min involves stating the larger valued addend and then counting a number of times equal to the value of the smaller addend (e.g., counting 3, 4, 5 to solve 3 + 2), and sum involves counting both addends starting from 1. The max procedure is occasionally used and involves stating the smaller addend and then counting the larger addend. With schooling, children use the min procedure more often and eventually rely primarily on decomposition and retrieval.

Compared with TA children, children with MLD rely on finger counting for more years, adopt the min procedure at a later age, and commit more counting errors (Geary, 1993; Hanich, Jordan, Kaplan, & Dick, 2001; Jordan & Montani, 1997; Ostad, 1997; Russell & Ginsberg, 1984). The most consistent finding is that children with MLD show a deficit in the ability to use retrieval-based processes (Barrouillet, Fayol, & Lathuliére, 1997; Geary, 1990; Geary, Hamson, & Hoard, 2000; Jordan, Hanich, & Kaplan, 2003). It is not that these children never correctly retrieve answers; rather, they show a persistent difference in the frequency with which they correctly retrieve basic facts, and sometimes in the pattern of retrieval errors.

Working Memory and Speed of Processing

Working memory is the ability to hold a mental representation of information in mind while simultaneously engaging in other mental processes. Working memory is composed of a central executive expressed as attention-driven control of information represented in two core slave systems (Baddeley, 1986). The slave systems are a language-based phonetic buffer and a visuospatial sketch pad. It has been well established that children with MLD do not perform as well as TA children on these working memory tasks (Bull, Johnston, & Roy, 1999; Geary et al., 2004; McLean & Hitch, 1999; Swanson, 1993; Swanson & Sachse-Lee, 2001), but it is not well understood which component or components of working memory contribute to the math cognition deficits of these children. The current study is the first to simultaneously assess the three primary components of working memory as related to performance on a range of math cognition measures and to assess these components as potential mediators of the math cognition deficits of children with MLD.

The potential contributions of working memory to these math cognition deficits are complicated by speed of processing. Children with MLD process information more slowly than TA children (Bull & Johnston, 1997; Murphy et al., in press; Swanson & Sachse-Lee, 2001), which in turn may result in performance deficits in many areas, including on measures of working memory. The relation between speed of processing and working memory, however, is vigorously debated and awaits full resolution. The issues center on whether individual differences in working memory are driven by more fundamental differences in speed of neural processing (Kail, 1991) or whether the attentional focus associated with the central executive speeds information processing (Engle, Tuholski, Laughlin, & Conway, 1999). In any case, a systematic assessment of the potential mechanisms underlying group differences on math cognition tasks requires simultaneous measurement of working memory and speed of processing.

Current Study

The use of multiple math cognition tasks allowed us to test the hypothesis that children with MLD have deficits across a range of math domains, whereas LA children have deficits for only a subset of these domains. For the math cognition tasks, we hypothesized that children with MLD will have a poor understanding of number sets, which will contribute to group differences in use of decomposition and min counting to solve addition problems (Geary et al., 2004). We also hypothesized children with MLD will show a deficit in detection of counting errors, which will contribute to frequent counting-procedure errors when they solve addition problems. The study also allowed us to test hypotheses regarding potential working memory mechanisms underlying math cognition deficits. We assessed whether the central executive, phonological loop, or a combination contribute to the frequency of counting-procedure errors committed by children with MLD, and we tested alternative hypotheses regarding that source of their retrieval deficit: the phonological loop (Geary, 1993) and central executive (Barrouillet et al., 1997).

We also explored potential relations among the visuospatial sketch pad, the central executive, and skill at number line estimation and performance on the Number Sets Test. This is because the processing of numerical information appears to be dependent on visuospatial working memory (Temple & Posner, 1998; Zorzi, Priftis, & Umiltá, 2002) and because the development of school-taught mathematics (e.g., a linear number line) is likely to be dependent in part on the central executive. Based on our hypothesis that LA children may vary in their strengths and weaknesses in math, it is not clear on which measures they will differ from TA children. We consider these comparisons to identify areas of math that are likely to be strengths and weaknesses for these children as a group and to explore variation in math skills within the LA group.

Method

Participants

All kindergarten children from 12 elementary schools were invited to participate in a longitudinal prospective study of MLD. Parental consent and child assent were received for 37% ($n = 311$) of these children. Because the relation between IQ and MLD is not yet known, we first restricted the sample to the 278 children with IQ scores between 80 and 130. We then classified children as MLD if their standard scores for the mathematical achievement test were less than 86 in both kindergarten and first grade. Twenty-two children were below this cutoff in kindergarten and 15 were below it in both grades. Although this cutoff is at the 15th percentile, the requirement that the child score at this level in both grades resulted in a criterion very similar to that used by Murphy et al. (in press). Moreover, this procedure identified 5.4% of our restricted sample and resulted in mean mathematics achievement scores at the 8th and 6th national percentiles in kindergarten and first grade, respectively. The achievement scores and percentage of the sample identified as MLD are consistent with incidence estimates (i.e., 5% to 10% of children) obtained with restrictive criteria used in a large-scale study of MLD (Barbaresi et al., 2005). A sample of 44 (14 male) LA children with percentile rankings between 23 and 39 on the mathematics test in either grade was also identified (38 had rankings < 39 in both grades) and provided a comparison group similar to samples identified as MLD in most previous cognitive studies. The sample of 46 TA children (29 male) had mathematics scores above the 50th percentile in both grades.

At the time of the first assessment, the mean ages of the children in the MLD and LA groups was 73 months ($SD = 4$), which is significantly ($p < .01$) but not practically younger than the mean age of 76 months ($SD = 4$) for the children in the TA group. There were more girls in the LA group and more boys in the TA group, $\chi^2(2) = 8.8$, $p < .01$; for tasks with significant sex differences, repeating all analyses using sex as a covariate did not change the pattern of group differences. The ethnic distribution (67% White, and most of the remaining children Black or Asian) did not differ across the LA and TA groups, $\chi^2(2) = 2.00$, $p > .50$, but there were fewer Asian and more Black children in the MLD group ($p < .01$).

Standardized Measures

Intelligence. In kindergarten, the children were administered the Raven's Coloured Progressive Matrices (Raven, Court, & Raven, 1993), a nontimed test that is considered to be an excellent measure of fluid intelligence. A percentile ranking was obtained for each child and these were converted to IQ scores standardized with a mean of 100 and a standard deviation of 15. In first grade, the children were administered the Vocabulary and Matrix Reasoning subtests of the Wechsler Abbreviated Intelligence Scale (Wechsler, 1999), and these scores were used to estimate IQ based on norms presented in the manual. The Raven and Wechsler scores were significantly correlated, $r(104) = .44$, $p < .0001$, and the final IQ score was the mean of these two measures.

Achievement. The children were administered the Numerical Operations and Word Reading subtests from the Wechsler Individual Achievement Test – II – Abbreviated (Wechsler, 2001). The Numerical Operations subtest assesses number discrimination, rote counting, number production, basic addition and subtraction, multidigit addition and subtraction, and some multiplication and division. The Word Reading subtest includes matching and identifying letters, rhyming, beginning and ending sounds, phoneme blending, letter sounds, and word recognition.

Mathematical Tasks

Number sets. The goal was to develop an easy-to-administer assessment of the speed and accuracy with which children can identify and process number-set information. Two types of number-set stimuli were developed: 0 – 9 small objects (circles, squares, diamonds, and stars) in a 1/2-in. square and one Arabic numeral (18 pt font) in a 1/2-in. square. As shown in Figure 1, these were combined to create domino-like rectangles in the following combinations: objects/same objects, objects/different objects, Arabic numerals/Arabic numerals, and objects/Arabic numerals. These rectangles were presented in lines of five across a page. In addition, there were two lines of three 3-square rectangles for each combination type on a page. The target numbers (5 and 9) were listed in a large font (36 pt) at the top of each page; the items of 5 and 9 where chosen because they represent smaller

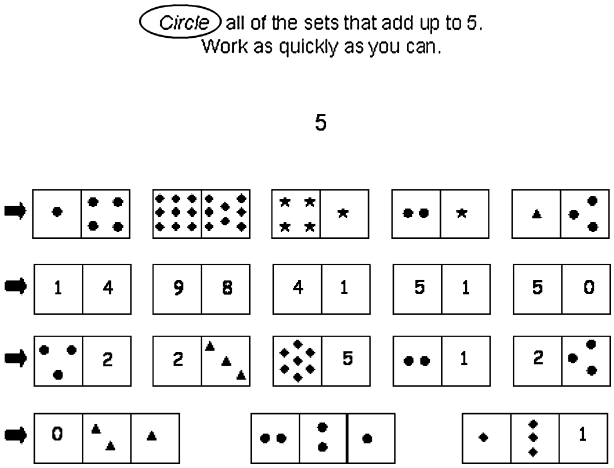

Figure 1. Example items from the Number Sets Test.

and larger values within the range of basic Arabic numerals (i.e., 1 to 9). The first page for each target number contained combination types objects/same objects and objects/different objects. The second page for each target number contained combination types objects/Arabic numerals and Arabic numerals/Arabic numerals. On each page, 18 items matched the target, 12 were larger than the target, 6 were smaller than the target, and 6 contained 0 or an empty square.

Two items matching a target number of 4 were first explained for practice. Using 3 as the target number, four lines of two items were then administered as practice; for each line, one set was a match and one was not. Once it was determined the child understood the task, the experimental items were administered. The child was instructed to move across each line of the page from left to right without skipping any and to "circle any groups that can be put together to make the top number, five (nine)" and to "work as fast as you can without making many mistakes." Using a stopwatch the child was given 60 s and 90 s per page for the targets 5 and 9, respectively, and was asked to stop at the time limit. We chose to time the task to avoid ceiling effects and because a timed measure should provide an assessment of fluency in recognizing number combinations. By design, most children (98%) were not able to finish in the allotted time (they were reassured during instruction that they might not be able to finish). The task yielded numbers of hits, misses, correct rejections, and false alarms for each problem type and size.

Number estimation. Stimuli for this task were twenty-four 25-cm number lines printed across the center of a standard 8-1/2 $\times$ 11 in. paper in a landscape orientation. Each number line had a start point of 0 and an endpoint of 100 with a target number printed approximately 5 cm above it in a large font (72 pt). Following Siegler and Booth (2004, Experiment 1), target numbers were 3, 4, 6, 8, 12, 17, 21, 23, 25, 29, 33, 39, 43, 48, 52, 57, 61, 64, 72, 79, 81, 84, 90, and 96. The numbers below 30 were over-sampled to allow for fitting of a logarithmic model of the children's estimates. Experimental stimuli were presented in a random order for each child.

The child was first presented with a number line that included the 0 and 100 endpoints and marked in increments of 10. After discussion about number lines and when it was determined that the child recognized the concept, a blank number line, containing only the endpoints 0 and 100, was presented and the child was asked to determine where the number 50 should go. The child was instructed to mark across the line with a pencil where 50 should fall. A printed number line with the endpoints and the location of 50 marked was then shown to the child. To ensure the child understood the task, the child's response was compared with the printed version and the experimenter discussed with the child how "50 is half of 100, so it goes half way between 0 and 100." The experimental trials were then administered each beginning with "If this is zero (pointing) and this is 100 (pointing), where would N go?" There was no time constraint.

Counting knowledge. The child was first introduced to a puppet that was just learning how to count and needed assistance. During each of the 13 trials, a row of 7, 9, or 11 poker chips of alternating color was aligned behind a screen. The screen was then removed, and the puppet counted the chips. The child was then queried on the correctness of the counting. The experimenter recorded whether the child stated the way the puppet counted was "OK" or "Not OK and wrong." Following previous studies, four types of counting trials were administered (Geary et al., 1992): correct, right–left, pseudo-error, and error. For correct trials, the chips were counted sequentially and correctly, from the child's left to the child's right. Right–left involved counting the chips sequentially and correctly, but from the child's right to the child's left. For pseudo-error trials, the chips were counted correctly from left to right, but first one color was counted and then, beginning again on the left-hand side of the row, the count continued with the other color. For error trials, the chips were counted sequentially from left to right, but the first chip was counted twice; children with MLD commit more errors than other children when the first chip is double-counted (Geary et al., 1992). For all trials, counting rate was about 1 chip per second. Each counting trial type occurred once for each array size (i.e., 7, 9, 11), with one additional pseudo-error count (for 9 chips) as the last trial. If the child stated that this final count was ''not OK,'' the experimenter asked:

> For that last problem, and the others like it where Big Bird counted all of one color first and then all of the other color, you said that he counted the wrong way. Do you think that he counted the wrong way and got the *wrong* answer, or, do you think that he counted the wrong way but still got the *right* answer?

The response was scored as both the answer and counting method were wrong, or the answer was right but the method was wrong. With an adjustment for guessing, children who stated the pseudo-error method was wrong but the count was correct were given credit for pseudo-error trials they initially identified as incorrect.

Addition strategy assessment. Simple and complex addition problems were horizontally presented in a large font (about 2 cm tall), one at a time, at the center of a 5 × 8 in. card. The simple stimuli were 14 single-digit addition problems. The problems consisted of the integers 2 through 9, with the constraint that the same two integers (e.g., 2 + 2) were never used in the same problem. Across stimuli, 1/2 of the problems summed to 10 or less. The complex stimuli were 16 + 7, 3 + 18, 9 + 15, 17 + 4, 6 + 19, and 14 + 8.

Following two practice problems, the simple problems were presented followed immediately by the complex problems. The child was asked to solve each problem (without the use of paper and pencil) as quickly as possible without making too many mistakes. It was emphasized that the child could use whatever strategy was easiest get the answer, and the child was instructed to speak the answer out loud. Based on the child's answer and the experimenter's observations, the trial was classified into one of six strategies: counting fingers, fingers, verbal counting, retrieval, decomposition, or other/mixed strategy (Siegler & Shrager, 1984). A mixed trial was one in which the child started using one strategy but completed the problem using another strategy. Counting trials were further classified as min, sum, max, or other.

During problem solving, the experimenter watched for physical indications of counting, such as finger or mouth movements. For these trials, the experimenter initially classified the strategy as finger counting or verbal counting, respectively. On verbal counting trials, the experimenter probed the child as to how she counted, and the child's response was recorded. If the child held out a number of fingers to represent the addends and then stated an answer without counting them, then the trial was initially classified as fingers. If the child spoke the answer quickly, without hesitation, and without obvious counting-related movements, the trial was initially classified as retrieval or as decomposition if this was the child's predominant retrieval-based strategy on previous trials. After the child had spoken the answer, the experimenter queried the child on how the answer was obtained. If the child's response (e.g., ''just knew it'') differed from the experimenter's observations (e.g., saw the child mouthing counting), a notation indicating disagreement between the child and the experimenter was made. If counting was overt, the experimenter classified it as a counting strategy. If the trial was ambiguous, the child's response was recorded as the strategy. Previous studies indicate this method provides a useful measure of children's trial-by-trial strategy choices (e.g., Siegler, 1987).

Working Memory

The Working Memory Test Battery for Children (WMTB – C; Pickering & Gathercole, 2001) consists of nine subtests that assess the central executive, phonological loop, and visuospatial sketchpad. All of the subtests have six items at span levels ranging from one to six to one to nine. Passing four items at a level moves the child to the next level. At each span level, the number of items (e.g., words) to be remembered is increased by one. Failing three items terminates the subtest. The order of subtests was designed so as not to overtax any one component area of working memory and was generally arranged from easiest to hardest: Digit Recall, Word List Matching, Word List Recall, Nonword List Recall, Block Recall, Mazes Memory, Listening Recall, Counting Recall, and Backward Digit Recall. Each subtest generates a span score and a trials correct score. From these, standard scores and age-based percentile ranks are determined for each subtest and for the three component areas.

Central executive. The central executive is assessed using three dual-task subtests. Listening Recall requires the child to determine if a sentence is true or false and then to recall the last word in a series of sentences. Counting Recall requires the child to count a set of 4, 5, 6, or 7 dots on a card and then to recall the number of counted dots at the end of a series of cards. Backward Digit Recall is a standard format backward digit span.

Phonological loop. Digit Recall, Word List Recall, and Nonword List Recall are standard span tasks with variant stimuli; the child's task is to repeat words spoken by the experimenter in the same order as presented by the experimenter. In the Word List Matching task, a series of words, beginning with two words and adding one word at each successive level, is presented to the child. The same words, but possibly in a different order, are then presented again, and the child's task is to determine if the second list is in the same or different order than the first list.

Visuospatial sketch pad. Block Recall is another span task, but the stimuli consist of a board with nine raised blocks in what appears to the child as a "random" arrangement. The blocks have numbers on one side that can only be seen from the experimenter's perspective. The experimenter taps a block (or series of blocks), and the child's task is to duplicate the tapping in the same order as presented by the experimenter. In the Mazes Memory task, the child is presented a maze with more than one solution, and a picture of an identical maze with a path drawn for one solution. The picture is removed and the child's task is to duplicate the path in the response booklet. At each level, the mazes get larger by one wall.

Speed of Processing

Using the same stimuli as Mazzocco and Myers (2003), two Rapid Automatized Naming (RAN; Denckla & Rudel, 1976) tasks were used to assess processing speed. The child is presented with five letters or numbers to determine first if the child can read the stimuli correctly. After these practice items, the child is presented with a 5×10 matrix of 50 incidences of these same letters or numbers and is asked to name them as quickly as possible without making any mistakes. Reaction time (RT) is measured using a stopwatch and errors and reversals for the letters *b* and *d* and *p* and *q* are recorded. For each type of stimulus (letters or numbers), the task generates RT, number correct, and number of reversals for letters. Errors and reversals were too infrequent for meaningful analysis, and thus only RTs were used from this test.

Procedure

All children were tested in the spring of their kindergarten year and in the fall and spring of first grade. The spring assessments included the achievement and intelligence measures, and the fall assessments included the speed of processing and mathematical tasks. Most children were tested in a quiet location at their school site, and occasionally in a testing room on the university campus or in a mobile testing van if the child moved between assessments. Each of these testing sessions required about 40 min. For spring assessments, the Word Reading and Numerical Operations tests were administered, followed by the Progressive Matrices in kindergarten and the Vocabulary and Matrix Reasoning tests in first grade. The order of task administration for fall testing was RAN, number line, counting knowledge, number sets, and addition strategy.

The WMTB–C was added to the study in the summer following kindergarten and we received parental permission to assess 270 children from the original sample; this included 102 of 105 children in the current analysis (1 child was missing from each group). For most children, the battery was administered in the testing van during first grade. The assessment required about 60 min and occurred when the child was not in school (e.g., weekend). The children in the MLD and LA groups were administered the WMTB–C at mean ages of 80 ($SD = 6$) and 82 ($SD = 6$) months, respectively ($p < .05$), which was younger than the mean age for children in the TA group (ps $< .05$), that is, 85 months ($SD = 6$). Age at time of administration was uncorrelated with scores on any of the working memory scales, rs(101) $= -.11$ to $-.07$, ps $> .25$.

Results

Because of the large number of variables, we adopted an alpha of .01, unless otherwise noted, and used two orthogonal contrasts in each analysis rather than overall F tests and comparisons of simple means. The first variable contrasted the TA group with the two other groups, and the second contrasted the LA and MLD groups. For tasks on which the first contrast is significant and the second is not, the implication is the corresponding cognitive deficits are similar for children identified with a strict (MLD group) or lenient (LA group) cutoff criterion. For tasks on which both contrasts are significant, the implication is the corresponding cognitive deficits of children identified with a strict criterion are either more severe or are of a different form than the deficits of children identified with a lenient criterion. To explore these possibilities and on the basis of group means, when the second contrast was significant we performed follow-up analyses. Effect size (d) was calculated as $(M_1 - M_2)/SD$; SD was estimated across all participants, and M_1 was the mean of the lower achieving group.

Standardized Tests

Group differences in IQ (Table 1) emerged for the first contrast, $F(1, 102) = 25.21$, but not the second,

Table 1
Intelligence and Achievement Test Scores

				Kindergarten				First grade			
		IQ		Word Reading		Numerical Operations		Word Reading		Numerical Operations	
	n	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
MLD	15	98	12	39	24	8	5	21	24	6	4
LA	44	101	9	68	19	36	9	57	27	30	7
TA	46	110	9	85	12	81	12	81	19	70	15

Note. MLD = math disabled; LA = low achieving; TA = typically achieving.

$F(1, 102) = 1.05$; ds = −1.1 and −0.8 comparing the TA group with the MLD and LA groups, respectively. Because the first contrast was significant, IQ was used as a covariate. For the two achievement tests, both contrasts were significant in both grades, Fs(1, 101) > 25.00. For Word Reading, the magnitude of the MLD group's deficit relative to the two other groups was large in both grades (ds < −1.3); comparing the LA and TA groups, $d = -0.7$ and −0.8 for kindergarten and first grade, respectively. For Numerical Operations, the difference between the MLD and TA groups was substantial (ds = −2.5) in both grades (ds = −1.0 and −0.9 for the MLD and LA contrasts in kindergarten and first grade, respectively; ds = −1.5 and −1.5 for the corresponding LA and TA contrasts.)

Mathematical Tasks

Number sets. Because preliminary correlations indicated consistency across item content and set size, items were combined to create an overall frequency of hits (Cronbach's $\alpha = .88$), correct rejections ($\alpha = .85$), misses ($\alpha = .70$), and false alarms ($\alpha = .90$), as shown in Table 2. The first contrast was significant for hits, correct rejections, and misses, Fs(1, 100) > 12.00, and the second was significant for misses, $F(1, 100) = 8.75$ ($d = 0.9$ for MLD vs. LA) and showed a trend for hits, $F(1, 100) = 5.53$, $p = .02$ ($d = -0.6$ for MLD vs. LA). The advantage of the TA children was large for hits (ds = −1.7 and −1.1 for contrast with the MLD and LA groups, respectively), misses (ds = 1.4 and 0.6), and correct rejections (ds = −1.4 and −0.5). In all, the TA children were faster and more accurate at processing various forms of number-set information than were LA children, and children with MLD were less likely than their LA peers to identify correct targeted number-set information.

Number estimation. Across trials, the TA children's estimates differed from the correct position by an average of 10 ($SD = 6$) as compared with mean differences of 17 ($SD = 7$) and 25 ($SD = 6$) for the LA ($d = 0.8$) and MLD ($d = 1.9$) groups ($d = 1.0$ for the MLD vs. LA difference). For these differences, both contrasts were significant, Fs(1, 101) > 15.00.

The group differences in placement accuracy could be due to differences in children's underlying mental representation of the number line. Children who make placements consistent with a logarithmic representation would necessarily have larger differences, on average, than would children who make placements consistent with a linear representation. To explore this possibility, we calculated the absolute difference between each child's estimate for each trial and the expected estimate if they were using a linear or log representation. For the linear representations we used the actual magnitude for the trial, and for log representations we used the best fitting log equation

Table 2
Number Sets Scores

	Total attempts		Hits		Correct rejections		Misses		False alarms	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
MLD	72	29	21	12	29	15	15	11	7	14
LA	74	20	28	8	36	11	9	7	2	6
TA	94	22	41	10	48	12	5	3	1	2

Note. MLD = math disabled; LA = low achieving; TA = typically achieving.

found by Siegler and Booth (2004) for first grade children's number line placements; $19\ln(x) - 15$, where $x =$ the value to be estimated. As an example, for "23," children making placements using a linear representation should place estimates near 23, whereas children making placements using a log representation should place estimates near 45.

All trials were classified as linear or logarithmic based on whether the child's estimate was closer to the predicted value for the linear or log model. When the expected value for the linear and log models differed by less than ±5 or the child's estimate was not clearly better fitted by either model, the trial was classified as ambiguous. Using this approach, we were able to make trial-by-trial classifications of whether each child was most likely to have used a linear or logarithmic representation to make the estimate. The percentages of trials classified as linear, log, or ambiguous are shown in Table 3. The first contrast confirmed that children in the TA group had a lower percentage of log trials, $F(1, 101) = 25.89$, and a higher percentage of linear trials, $F(1, 101) = 17.28$, than did children in the LA (*d*s = −1.0 and −0.8 for log and linear trials, respectively) and MLD (*d*s = −1.4 and −1.2) groups. However, the contrast of the LA and MLD groups was not was significant for either strategy, $Fs(1, 101) < 2$ (*d*s = 0.4 and −0.4 for log and linear trails, respectively). The four rightmost columns of Table 3 show the absolute degree of error for linear and log trials. The only significant result was for log trials, $Fs(1, 95) > 12.00$; the MLD group had larger errors than the TA ($d = 1.6$) and LA ($d = 1.0$) groups, and the LA group had larger errors than the TA group ($d = 0.6$).

In all, the estimates of TA children were more frequently consistent with a linear representation of the number line than were the estimates of their lower achieving peers. However, lower achieving children, including children with MLD, are just as accurate as TA children when they appeared to use a linear-based representation. Children with MLD and their LA peers made placements consistent with a log representation for more than half of their number line trials, but the LA children were more accurate on these trials.

Table 3
Number Line Strategy and Fit Accuracy

	Strategy			Fit difference			
	Linear	Log	Amb	Linear		Log	
	M	*M*	*M*	*M*	*SD*	*M*	*SD*
MLD	19	63	18	10	7	18	6
LA	27	54	20	7	4	11	6
TA	45	30	25	5	3	7	6

Note. Strategy percentage may not equal 100% because of rounding. Amb = ambiguous trial; MLD = math disabled; LA = low achieving; TA = typically achieving.

Counting knowledge. The percentages of correct identifications for correct trials was near ceiling for all groups (96% to 100%) and was nearly as high for right–left trials (>88%). The percentage of correct identifications for pseudo-error trials (i.e., stating the count was correct) was 93, 74, and 79, for the MLD, LA, and TA groups, respectively, and 67, 92, and 98 for double-count error trials, respectively. The only significant effects to emerge were for double-count error trials; both contrasts were significant, $Fs(1, 101) > 11.00$. The pattern of means indicates that children with MLD found double-count error trials to be more difficult than did children in the LA ($d = -1.1$) and TA ($d = -1.4$) groups, who did not differ ($p > .05$).

Addition strategy choices. The majority of both simple and complex problems were solved by means of finger counting, verbal counting, retrieval, or decomposition (Table 4). To reduce the number of variables and to capture the dynamics of children's strategy choices (Siegler, 1996; Siegler & Shrager, 1984), we used a coding procedure developed by Geary and Burlingham-Dubree (1989). Independently for simple and complex problems, the first strategy choice variable was termed *memory accuracy* and was coded as the sum of the number of correct retrieval and correct decomposition trials. A high score indicated that when a memory-based process was used, it produced the correct answer. The second variable was termed *backup* and was coded to represent what the children did to solve the problem when memory-based processes were not used accurately. If the number of retrieval errors was greater than the number of problems correctly solved by finger or verbal counting, backup was coded (0 − number of retrieval errors). If the number of problems solved correctly by finger or verbal counting was greater than number of retrieval errors, backup was coded [(2 × correct min counts) + (correct sum and max counts) − total counting errors]. A high score indicated frequent and accurate use of the most mature counting procedure and a low score indicated frequent guessing. The strategy variables were uncorrelated (*p*s > .15), indicating they captured different sources of variation in addition skills.

The first contrast was significant for the memory accuracy variable for simple (*d*s = −1.2 and −0.8 for TA contrast with the MLD and LA groups, respectively) and complex (*d*s = −0.8 and −0.1) problems, $Fs(1, 101) > 9.80$. This contrast was also significant for the backup variable for simple (*d*s = −1.4 and −0.2 for TA contrast with the MLD and LA groups,

Table 4
Addition Strategy Choices

	Percent usage				Percent errors				Min usage	
	CF	VC	R	D	CF	VC	R	D	CF	VC
Simple problems										
MLD	48	9	31	0	45	44	77	–	33	50
LA	56	23	18	1	19	26	39	17	67	82
TA	37	27	22	12	10	8	7	4	86	91
Complex problems										
MLD	41	6	42	0	83	25	100	–	29	100
LA	71	18	10	0	34	22	90	–	80	98
TA	57	22	6	13	13	19	17	3	94	100

Note. Strategy percentage may not equal 100% because of rounding. Min usage = use of the min counting procedure; CF = counting fingers; VC = verbal counting; R = retrieval; D = decomposition; MLD = math disabled; LA = low achieving; TA = typically achieving.

respectively) and complex (ds = −1.5 and −0.4) problems, Fs(1, 101) > 6.80. The second contrast was significant for the backup variable for both simple (d = −1.2), F(1, 101) = 16.31, and complex (d = −1.1), F(1, 101) = 17.49, problems, but neither effect was significant for memory accuracy, Fs(1, 101) < 2. Examination of the MLD–LA differences on the variables that defined the backup variable revealed children with MLD committed more finger counting errors (d = 0.6), F(1, 86) = 13.96; used the min procedure less often (d = −1.1), F(1, 86) = 11.02; and committed more retrieval errors (d = 0.9), F(1, 101) = 120.58, when solving simple addition problems. When solving complex problems, the children with MLD used finger counting less often (d = −.8), F(1, 96) = 7.83, and committed more finger counting errors (d = 1.3), F(1, 78) = 11.90, than did the LA children. The children with MLD also used the min procedure less often when finger counting (d = −1.5), F(1, 78) = 17.56, and committed more retrieval errors (d = 1.0), F(1, 101) = 12.58.

Compared with their lower achieving peers, children in the TA group knew more addition facts and were more skilled in the use of decomposition and in the execution of counting procedures. Neither the children with MLD nor their LA peers appeared to know many addition facts, and these children almost never used decomposition. Compared with LA children, children with MLD used less mature counting procedures, were more error prone in the execution of these procedures, and were more likely to commit retrieval errors.

Working Memory and Speed of Processing

For the age range assessed in this study, our overall sample of 270 children was substantially larger than the WMTB–C standardization subsample of 86 children (Pickering & Gathercole, 2001). We thus standardized (M = 100, SD = 15) raw scores for the Phonological Loop, Visuospatial Sketch Pad, and Central Executive scales for our overall sample and used these for subsequent analyses. The associated mean scores are shown in Table 5. The first contrast was not significant for the phonological loop, F(1, 98) = 4.07, or visuospatial sketch pad, F(1, 98) = 2.42, but it was significant for the central executive, F(1, 98) = 9.90. The second contrast was significant for the phonological loop, F(1, 98) = 12.50, and central executive, F(1, 98) = 16.90, and a trend was evident for the visuospatial sketch pad, F(1, 98) = 6.22, p = .014. Although the first contrast was significant for the central executive, a post hoc t test indicated no significant difference comparing the TA and LA groups (p > .05); the significance was due to the low scores of the MLD group. The children with MLD had substantial deficits relative to the TA and

Table 5
Working Battery Test Scores and Speed of Processing Reaction Times

	Working memory system						Speed of processing	
	Phonological loop		Visuospatial sketch pad		Central executive		RAN RTs	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
MLD	85	18	85	16	81	14	58	16
LA	101	13	99	13	99	14	43	11
TA	104	11	105	18	106	14	36	7

Note. RAN RTs = are the mean reaction times in seconds for the letter-naming and number-naming measures; MLD = math disabled; LA = low achieving; TA = typically achieving.

LA children for the phonological loop (*d*s = −1.3 and −1.1, respectively), visuospatial sketch pad (*d*s = −1.3 and −0.9), and central executive (*d*s = −1.7 and −1.2). The speed of processing variable was the mean of the letter and number naming RTs (in seconds); both contrasts were significant, *F*s(1, 101) > 21. The children in the TA group had clear advantages over children in the LA (d = 0.5) and MLD (d = 1.7) groups (d = 1.2 for the LA and MLD comparison).

Cognitive Deficits for MLD Children

We assessed whether working memory and speed of processing might mediate the differences comparing the LA and MLD groups on the math cognition variables. Inclusion of the TA group in these analyses was less informative because the first contrast was not significant for two of the four potential mediators, whereas the LA/MLD contrast differed for all potential mediators. If the math cognition deficits of children with MLD are due to working memory or speed of processing, one or several of these variables will emerge as significant mediators of the group contrast. In keeping with Baron and Kenny (1986), we focused on the math cognition variables for which the contrast of the LA and MLD groups was significant. The variables were the number of hits and misses from the Number Sets Test and (hits − misses), overall accuracy from the number line task, error detection from the counting knowledge task, and the frequency of retrieval errors, finger counting errors, and min usage from simple and complex addition. The latter variables contributed to the group difference in use of backup strategies. The external validity of these variables was confirmed with significant correlations with Numerical Operations scores at the end of first grade, independently, $|r_{(57)}|$ = .28 to .58, *p*s < .05, and in combination, $R^2 = .45$, $p < .02$.

We first used a stepwise regression procedure to determine which of the three working memory or speed of processing variables predicted variation on each of the math cognition tasks. On the basis of predictions detailed in the Introduction, we also conducted several specific analyses: For finger counting errors we included the error detection variable from the counting knowledge task and for min usage we included the number sets variables as potential predictors. Based on the predicted relation between visuospatial working memory and skill at identifying number sets and making number line estimates, we performed separate analyses using only the visuospatial sketch pad variable as a potential mediator. Mediational effects were assessed following Baron and Kenny (1986) and using Sobel's (1988) test. If the mediational effect was significant and reduced the group contrast to nonsignificance ($p > .05$), full mediation is implied; partial mediation is implied if the group contrast remained significant.

The stepwise equations revealed higher central executive scores were associated with better performance on the number sets task (i.e., hits − misses; β = .31), smaller errors on the number line task (β = −.46), better skill at detecting errors on the counting knowledge task (β = .57), and fewer retrieval errors when solving simple (β = −.59) and complex (β = −.40) addition problems (*p*s < .05); no other predictors were significant for these variables. The central executive emerged as a significant and full mediator of group differences in error detection on the counting knowledge task ($Z = 2.14$, $p < .05$) and number of retrieval errors for simple addition ($Z = -2.38$, $p < .02$), and a partial mediator of the group difference in overall accuracy on the number line task ($Z = 2.26$, $p < .05$). However, the central executive did not emerge as a significant mediator of the group differences on the number sets task ($Z = 1.25$, $p > .20$) or for retrieval errors for complex addition ($Z = -1.26$, $p > .20$). When evaluated as a potential mediator, higher visuospatial sketch pad scores were found to be associated with more hits (β = .30, $p < .02$), and a weak trend emerged for a partial mediational effect ($Z = 1.56$, $p < .12$). Higher visuospatial sketch pad scores were also associated with less error for number line estimates (β = −.41, $p < .01$), and a trend emerged for a partial mediational effect ($Z = 1.60$, $p < .11$)

During the solving of simple addition problems, frequent finger counting errors were associated with lower phonological loop scores (β = −.37) and fewer detections of counting errors on the counting knowledge task (β = −.28, *p*s < .05). The mediation analyses revealed a trend for the phonological loop ($Z = 1.87$, $p < .07$) and a nonsignificant effect for error detection ($Z < 1.5$), but inclusion of both variables dropped the group contrast to nonsignificance ($p > .05$), suggesting full mediation. During the solving of complex addition problems, frequent finger counting errors were associated with slower speed of processing (β = −.48, $p < .002$) and partial mediation of the group difference ($Z = 2.11$, $p < .05$). No significant predictors emerged for min usage while solving simple problems, but frequent use of min counting while solving complex problems was associated with higher visuospatial sketch pad scores (β = .49) and higher scores on the number sets (hits − misses) variable (β = .47, *p*s < .05). A trend for a full mediation effect emerged for the visuospatial sketch pad ($Z = 1.79$, $p < .08$), but the effect for number sets was not significant ($Z < 1.5$).

A summary of the regression and mediational results is shown in Table 6. For the children with MLD, the central executive was implicated as a potent source of their deficits across math cognition tasks that involve counting, number representation, and aspects of addition. Phonological and visuospatial working memory contributed to more specific math cognition deficits, as did speed of processing. Individual differences in error detection on the counting knowledge task and performance on the number sets task contributed to counting errors while solving simple addition problems and min usage while solving complex addition problems, as predicted, but did not emerge as mediators of the differences comparing the MLD and LA groups. Evidence for a contribution of visuospatial working memory to skill at detecting number sets and making number line estimates was found, but the contribution of the central executive was more important in explaining the MLD–LA contrast.

Cognitive Deficits for LA Children

One way to assess the pattern and severity of the LA children's deficits in relation to the TA children is to examine effect sizes. Across the math cognition variables, the d values ranged from 0.6 to -1.9 comparing the MLD and TA groups, indicating moderate to severe deficits. The comparison of the LA and TA groups revealed a more varied pattern, with differences ranging from small ($d = 0.1$) to large ($d = -1.1$). Controlling for the group difference in IQ yielded eight d values greater than 0.5. These were degree of error ($d = 0.6$) and percentage of linear ($d = -0.7$) and log ($d = 0.8$) trials for the number line task; hits ($d = -0.8$), misses ($d = 0.6$), and correct rejections ($d = -0.5$) for the Number Sets Test; and memory accuracy for simple ($d = -0.6$) and complex ($d = -0.6$) addition. Examination of the distributions of scores for these variables revealed between 36% and 98% of the LA children were more than 1 *SD* below the mean of the TA group, and more than 45% were more than 0.5 *SD* below the mean for all variables. More than 50% of the LA children were more than 1 SD below the mean of the TA group for percentage of log trials on the number line task (66%), number of hits (61%) and correct rejections (59%) for the Number Sets Test, and memory accuracy for simple (57%) and complex (98%) addition. In comparison to the TA children, most LA children are much more highly dependent on the log representation for making number line estimates, lacked fluency in the processing of number set information, and have not memorized or cannot retrieve many addition facts.

Table 6
Summary of Regression and Mediational Results

	Potential mediators						
Math variable	Phonological loop	Visuospatial sketch pad	Central executive	Reaction time	Count error	Number set	Mediation effect
Number and counting							
Number set			.31				No
Hits		.30					Partial
Number line		(−.41)	−.46				Partial
Count error			.57				Full
Simple addition							
CF: Error	−.37				−.28		Full
Min							No
Retrieval error			−.59				Full
Complex addition							
CF: Error				−.48			Full
Min		.49				.47	Full
Retrieval error			−.40				No

Note. Reported values are significant beta estimates from the predictors identified in the stepwise regression analyses. The visuospatial sketch pad estimates for number line (in parentheses) and hits are from separate a priori regression analyses. Number set = hits − misses; number line = overall accuracy of the estimate; count error = detection of double-counting errors on the counting knowledge task; CF = counting fingers; min = use of the min counting procedure.

Discussion

We addressed issues regarding the criteria used to classify children as MLD, the range and nature of associated deficits on math cognition tasks, and potential working memory and speed of processing mediators of these deficits. The definition of MLD concerned whether the use of a stringent versus a lenient cutoff criterion on mathematics achievement tests will affect the range and nature of deficits that emerge on math cognition tasks. As we discuss in the first section, use of a stringent cutoff is associated with a broader range of math cognition deficits, but more subtle deficits may still emerge with use of a lenient cutoff. In the second section, we discuss the specifics of the deficits in the domains of number, counting, and arithmetic. In the final section, we focus on the implications of the mediational analyses for understanding the potential core cognitive deficits that may underlie more severe forms of MLD.

Defining Mathematical Disability

Our strict cutoff criterion was mathematics achievement scores below the 15th percentile in both kindergarten and first grade. The use of this cutoff for two consecutive grades identified 5% of our sample as MLD. Their mean mathematics achievement scores at the 8th and 6th national percentile in kindergarten and first grade, respectively, is consistent with the incidence of MLD (i.e., 5% to 10%) found in a large-scale birth cohort study (Barbaresi et al., 2005). One very practical implication is that children scoring below the 15th percentile on a mathematics achievement test in kindergarten are at risk for MLD, but confirmation of MLD status requires continually low performance in later grades (Murphy et al., in press). In our study, 22 children (8% of the sample) met this criterion in kindergarten and 15 remained below it in first grade. The breadth of math cognition deficits of these 15 children, combined with an average IQ, suggests a severe form of MLD. These children also scored below average on the reading achievement test but not at a level consistent with a diagnosis of reading disability. Nonetheless, their slow speed of processing on the RAN suggests risk for slow development of reading fluency and below average reading development (Denckla & Rudel, 1976).

The children in our LA group had average IQ and reading achievement scores but scored between the 23rd and 39th percentiles on the mathematics achievement test in kindergarten or first grade. Their mathematics achievement was thus consistently below expectations and in a range that would have resulted in their inclusion with the MLD group in many other studies (Murphy et al., in press). The finding of consistent math cognition differences across our LA and MLD groups indicates children identified with a stringent versus a lenient cutoff criterion differ in important ways and should not be conflated. Comparisons with the TA group revealed the LA children had near normal math skills in some areas (e.g., use of backup strategies to solve simple addition problems) but deficits that clustered in other areas, specifically, in fluency of processing number sets, skill at making number line estimates, and speed in retrieving addition facts. It cannot be determined from this study whether the math cognition deficits of the LA children are simply due to insufficient exposure to the associated topics or whether these children have a circumscribed learning disability in fluency of number processing and addition-fact retrieval. Response to instruction studies are needed to differentiate these two alternatives (Fuchs, 2002).

Mathematical Cognition Deficits

Number. With use of the number sets and number line tasks (Siegler & Booth, 2004; Siegler & Opfer, 2003), we expanded the assessment of the math cognition competencies of children with MLD and their LA peers in novel ways. The Number Sets Test was constructed to assess children's understanding that numbers can be combined and decomposed to form larger and smaller sets, respectively, and to assess fluency of processing number-set information. The measure was also developed to test the hypotheses that poor knowledge of number sets contributes to the tendency of children with MLD to use decomposition and min counting less frequently than their TA peers during the solving of addition problems (Geary et al., 2004). Utility of the test was confirmed by significant group differences across the TA, LA, and MLD groups, and by its emergence as a predictor of individual differences in use of min counting to solve complex addition problems. The LA and MLD groups did not use decomposition and thus the predicted relation between number sets performance and use of this strategy could not be tested. Moreover, the finding that individual differences in number of hits was related to visuospatial working memory is consistent with the finding that access to number information engages visuospatial systems (Temple & Posner, 1998).

The TA children were faster and more accurate at identifying targeted number sets and rejecting nonmatching sets than were children in the other groups. The children with MLD had nearly as many hits as misses (ratio of 1.4:1), whereas the ratio was 3:1 for the LA children and 8:1 for the TA group. The pattern suggests the processes underlying number set performance may be similar for the TA and LA groups, but the latter group processes this information more slowly or more frequently engages in effortful counting to determine set size. The poor identification of sets by the children with MLD and the finding that the LA – MLD contrast for number of hits was partially mediated by visuospatial working memory suggest potential deficits in the number representational system of children with MLD (Koontz & Berch, 1996).

The number line estimation task developed by Siegler and colleagues (Siegler & Booth, 2004; Siegler & Opfer, 2003) also proved to be useful for the study of MLD. Analyses of trial-by-trial variation indicated more frequent reliance on the formal linear representation by the TA children and heavy reliance on the log representation for children in the LA and MLD groups. Even though the LA and MLD groups were similar in their reliance on the log representation, some differences were evident. The ratio of log to linear trials was 3:1 for the MLD children and 2:1 for the LA children, suggesting the latter group is progressing more rapidly in the learning of the formal linear representation of the number line. The finding that the central executive was a partial mediator of the LA – MLD contrast for overall accuracy on this task suggests that attentional control or other subcomponents of this system are involved in the construction of linear representations and contribute to the slow learning of children with MLD.

This is not the whole picture, however, as the children with MLD were less accurate than the LA children even when both groups used the log representation. Because we used the log curve that Siegler and Booth (2004) fitted to first graders' number line estimates, our results indicate the log representation accessed by our LA children was similar to that used by the children studied by Siegler and Booth. The large difference in the accuracy of log trials ($d = 1.0$) comparing the LA and MLD groups suggests there may be differences in the nature of the underlying log representation of the number line. The inaccuracy of the MLD group would most likely result from the compression of the log representation at lower values. This is because the nature of the log representation will result in less precision in the placement of larger valued numbers for all groups, and a compressed representation means less discrimination of numbers toward the beginning of the number line and thus higher error rates. Our follow-up longitudinal assessments will reveal whether the number line estimates for children with MLD are normal but delayed, or whether a more fundamental deficit exists in this competency and possibly the underlying visuospatial system (Zorzi et al., 2002).

Counting knowledge. The prediction of group differences on pseudo-error trials was not confirmed. In previous studies, children with MLD have tended to state this form of correct but irregular counting is "wrong" and this tendency predicted low use of the min counting procedure (Geary et al., 1992; Geary et al., 2004). This pattern did not emerge for the current sample and in fact the children with MLD scored somewhat higher on these trials. Using a similar procedure, LeFevre et al. (2006) found a curvilinear performance pattern from kindergarten to second grade for children's responses to irregular counts, including pseudo-error counts. Children with low numerical test scores tended to say these counts were correct in kindergarten and first grade, whereas their high-ability peers tended to say these counts were incorrect. The pattern reversed in second grade. LeFevre et al. suggested children have an initial bias to state all counts are correct, except for those with obvious errors, as in double-counts. In this view, the development of counting knowledge initially involves an awareness and use of unessential counting features when judging the correctness of the count. An early awareness leads to rejection of irregular but correct counts; thus it is an initial disadvantage for high-ability children. With experience, children eventually understand irregular ways of counting do not violate core principles (e.g., cardinality) and at this point they accept these counts as correct. If LeFevre et al.'s hypothesis is correct, in our longitudinal follow-up the performance of children with MLD will deteriorate and that of the LA and TA children will improve.

In any case, our prediction regarding group differences for double-count error trials was confirmed. The LA and TA children almost always detected these errors, whereas the children with MLD failed to detect them in 1 of 3 trials. The mediational analysis indicated that the failure of children with MLD to detect these errors was fully mediated by the central executive (Hoard et al., 1999). Overall, it appears that children with MLD understand basic counting concepts as well as TA and LA children, and any difficulties that children with MLD have on counting tasks may be due to working memory failures (Klein & Bisanz, 2000).

Arithmetic. The use of the strategy choice variables to examine group differences for simple and complex addition placed our analyses in the context of Siegler's (1996; Siegler & Shrager, 1984) developmental model and yielded results consistent with previous studies of MLD (Geary et al., 2000; Jordan & Montani, 1997; Ostad, 1997). The memory access variable indexed retrieval of correct addition facts from long-tem memory and accurate use of decomposition. The children in the TA group knew more facts and used decomposition more effectively than did children in the LA and MLD groups, but the latter two groups did not differ. As described earlier, one of the most consistent findings in this literature is that children with MLD have difficulty correctly retrieving arithmetic facts from long-term memory (Geary, 1993; Jordan et al., 2003). The finding that our LA children did not differ from our children with MLD and were substantially below the mean of the TA group on this variable suggests the finding of retrieval deficits in earlier studies may not have been strongly affected by use of different cutoff criteria.

The MLD and LA groups did, nonetheless, differ in skilled use of backup strategies. The children with MLD committed more finger counting errors, used the min procedure less frequently, and committed more retrieval errors than children in the LA group. The LA children in turn were as skilled as the children in the TA group when using backup strategies to solve simple problems, but they showed a moderate deficit when solving complex problems. There were no consistent mediators of the LA – MLD group differences in skilled use of counting, as the phonological loop, visuospatial sketch pad, and RAN RT all emerged as full mediators of the group difference in one of the component counting skills. It is possible that skilled use of counting requires simultaneous use of multiple components of working memory and is influenced by speed of processing, but multicollinearity among these measures resulted in only one predictor emerging for each skill.

The final key finding was that the MLD and LA groups differed in frequency of retrieval errors, but not correct retrieval, as noted previously. The pattern suggests the children with MLD have a lower confidence criterion for stating retrieved facts than the LA children (Geary et al., 2000; Siegler, 1988). This leaves open the possibility that the low number of correctly retrieved facts for the latter group is not a memory deficit per se, but rather results from a strict confidence criterion. Compared with the LA children, the children with MLD may also have difficulty inhibiting irrelevant associations from entering working memory, as proposed by Barrouillet et al. (1997) and in keeping with the finding that the central executive was a partial mediator of the LA – MLD difference in frequency of retrieval errors.

Working Memory and Speed of Processing

The children in our MLD group scored 1 *SD* below their LA peers on measures of each of the working memory systems and showed a deficit of about the same magnitude on the speed of processing measure, consistent with Swanson and colleagues' findings of pervasive working memory deficits in children with MLD (Swanson, 1993; Swanson & Sachse-Lee, 2001). Comparison of the TA and LA groups revealed moderate differences for the speed of processing measure, favoring the TA children, but no differences across the three working memory measures. The implication is that varying the stringency of the cutoff criterion for defining MLD can result in variation in the severity of working memory and speed of processing deficits, but even children identified with a lenient criterion may differ from their IQ-equated TA peers in speed of processing.

Although there is wide agreement that some components of working memory and speed of processing are contributing factors to MLD, there are disagreements regarding the relative importance of these factors (Bull & Johnston, 1997; Bull et al., 1999; Geary et al., 2004; McLean & Hitch, 1999; Swanson, 1993; Swanson & Sachse-Lee, 2001). The mediational analyses revealed the central executive as a core component in the deficits of children with MLD across math cognition tasks, but at the same time the phonological loop, visuospatial sketch pad, and speed of processing may contribute to more specific math cognition deficits. In other words, different components of working memory may affect different aspects of mathematical performance and learning, which may explain previous inconsistent findings.

Summary

Compared with the LA and TA groups, the children with MLD had deficits across all of the math cognition tasks and on measures of working memory and speed of processing. Many but not all of the math cognition deficits of these children were partially or fully mediated by the central executive, although the phonological loop, visuospatial sketch pad, and speed of processing contributed to several more specific math cognition deficits. Compared with the TA group, the math deficits of the LA children centered on fluency of processing numbers, making number line estimates, and retrieving addition facts. Overall, use of a restrictive cutoff criterion identifies children with pervasive and often severe math cognition deficits and underlying deficits in working memory and speed of processing, whereas use of a lenient criterion identifies children that may have more subtle deficits in a few math domains.

References

Ashcraft, M. H. (1982). The development of mental arithmetic: A chronometric approach. *Developmental Review*, *2*, 213–236.

Baddeley, A. D. (1986). *Working memory.* Oxford, UK: Oxford University Press.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: Incidence in a population-based birth cohort, 1976–82, Rochester, Minn. *Ambulatory Pediatrics*, *5*, 281–289.

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*, 1173–1182.

Barrouillet, P., Fayol, M., & Lathuliére, E. (1997). Selecting between competitors in multiplication tasks: An explanation of the errors produced by adolescents with learning disabilities. *International Journal of Behavioral Development*, *21*, 253–275.

Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, *20*, 607–618.

Bull, R., & Johnston, R. S. (1997). Children's arithmetical difficulties: Contributions from processing speed, item identification, and short-term memory. *Journal of Experimental Child Psychology*, *65*, 1–24.

Bull, R., Johnston, R. S., & Roy, J. A. (1999). Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetical skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology 15*, 421–442.

Case, R., Okamoto, Y., Griffin, S., McKeough, M., Bleiker, C., Henderson, B., et al. (1996). The role of central conceptual structures in the development of children's thought. *Monographs of the Society for Research in Child Development*, *66* (1-2, Serial No. 246).

Denckla, M. B., & Rudel, R. (1976). Rapid automatized naming (RAN): Dyslexia differentiated from other learning disabilities. *Neuropsychologia*, *14*, 471–479.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*, 309–331.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Science*, *8*, 307–314.

Fuchs, L. S. (2002). Three conceptualizations of "treatment" in a responsiveness-to-treatment framework for LD identification. In R. Bradley, L. Danielson, & D. P. Hallahan (Eds.), *Identification of learning disabilities: Research to practice* (pp. 521–529). Mahwah, NJ: Erlbaum.

Fuson, K. C. (1982). An analysis of the counting-on solution procedure in addition. In T. P. Carpenter, J. M. Moser, & T. A. Romberg (Eds.), *Addition and subtraction: A cognitive perspective* (pp. 67–81). Hillsdale, NJ: Erlbaum.

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

Geary, D. C. (1990). A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology*, *49*, 363–383.

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

Geary, D. C., Bow-Thomas, C. C., & Yao, Y. (1992). Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology*, *54*, 372–391.

Geary, D. C., & Burlingham-Dubree, M. (1989). External validation of the strategy choice model for addition. *Journal of Experimental Child Psychology*, *47*, 175–192.

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., Byrd-Craven, J., & Desoto, 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.

Gelman, R., & Gallistel, C. R. (1978). *The child's understanding of number.* Cambridge, MA: Harvard University Press.

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 25–49). Cambridge, MA: Bradford Books/MIT Press.

Groen, G. J., & Parkman, J. M. (1972). A chronometric analysis of simple addition. *Psychological Review*, *79*, 329–343.

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.

Hoard, M. K., Geary, D. C., & Hamson, C. O. (1999). Numerical and arithmetical cognition: Performance of low- and average-IQ children. *Mathematical Cognition*, *5*, 65–91.

Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003). Arithmetic fact mastery in young children: A longitudinal investigation. *Journal of Experimental Child Psychology*, *85*, 103–119.

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.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin*, *109*, 490–501.

Klein, J. S., & Bisanz, J. (2000). Preschoolers doing arithmetic: The concepts are willing by the working memory is weak. *Canadian Journal of Experimental Psychology*, *54*, 105–115.

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

LeFevre, J.-A., Smith-Chant, B. L., Fast, L., Skwarchuk, S.-L., Sargla, E., Arnup, J. S., et al. 2006). What counts as knowing? The development of conceptual and procedural knowledge of counting from kindergarten through Grade 2. *Journal of Experimental Child Psychology*, *93*, 285 – 303.

Mazzocco, M. M., & Myers, G. (2003) Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, *53*, 218 – 253.

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.

Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B., & Early, M. C. (in press). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*.

Ostad, S. A. (1997). Developmental differences in addition strategies: A comparison of mathematically disabled and mathematically normal children. *British Journal of Educational Psychology*, *67*, 345 – 357.

Pickering, S., & Gathercole, S. (2001). *Working Memory Test Battery for Children (WMTB – C) manual*. London: Psychological Corporation.

Raven, J. C., Court, J. H., & Raven, J. (1993). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. London: Lewis.

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

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

Siegler, R. S. (1987). The perils of averaging data over strategies: An example from children's addition. *Journal of Experimental Psychology: General*, *116*, 250 – 264.

Siegler, R. S. (1988). Individual differences in strategy choices: Good students, not-so-good students, and perfectionists. *Child Development*, *59*, 833 – 851.

Siegler, R. S. (1996). *Emerging minds*: The process of change in children's thinking. New York: Oxford University Press.

Siegler, R. S., & Booth, J. L. (2004). Development of numerical estimation in young children. *Child Development*, *75*, 428 – 444.

Siegler, R. S., & Opfer, J. (2003). The development of numerical estimation: Evidence for multiple representations of numerical quantity. *Psychological Science*, *14*, 237 – 243.

Siegler, R. S., & Shrager, J. (1984). Strategy choice in addition and subtraction: How do children know what to do? In C. Sophian (Ed.), *Origins of cognitive skills* (pp. 229 – 293). Hillsdale, NJ: Erlbaum.

Sobel, M. E. (1988). Direct and indirect effect in linear structural equation models. In J. S. Long (Ed.), *Common problems/proper solutions: Avoiding error in quantitative research* (pp. 46 – 64). Beverly Hills, CA: Sage.

Swanson, H. L. (1993). Working memory in learning disability subgroups. *Journal of Experimental Child Psychology*, *56*, 87 – 114.

Swanson, H. L., & Sachse-Lee, C. (2001). Mathematical problem solving and working memory in children with learning disabilities: Both executive and phonological processes are important. *Journal of Experimental Child Psychology*, *79*, 294 – 321.

Temple, E., & Posner, M. I. (1998). Brain mechanisms of quantity are similar in 5-year-old children and adults. *Proceedings of the National Academy of Sciences USA*, *95*, 7836 – 7841.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation, Harcourt Brace.

Wechsler, D. (2001). *Wechsler Individual Achievement Test – II – Abbreviated*. San Antonio, TX: Psychological Corporation, Harcourt Brace.

Zorzi, M., Priftis, K., & Umiltá, C. (2002, May 9). Neglect disrupts the mental number line. *Nature*, *417*, 138.