



Published in final edited form as:

Learn Disabil Res Pract. 2006 May 1; 21(2): 77–88. doi:10.1111/j.1540-5826.2006.00208.x.

Development of the Metacognitive Skills of Prediction and Evaluation in Children With or Without Math Disability

Adia J. Garrett,

University of Maryland, Baltimore County

Michèle M. M. Mazzocco, and

Johns Hopkins University School of Medicine, Johns Hopkins Bloomberg School of Public Health, Kennedy Krieger Institute

Linda Baker

University of Maryland, Baltimore County

Abstract

Metacognition refers to knowledge about one's own cognition. The present study was designed to assess metacognitive skills that either precede or follow task engagement, rather than the processes that occur during a task. Specifically, we examined prediction and evaluation skills among children with ($n = 17$) or without ($n = 179$) mathematics learning disability (MLD), from grades 2 to 4. Children were asked to predict which of several math problems they could solve correctly; later, they were asked to solve those problems. They were asked to evaluate whether their solution to each of another set of problems was correct. Children's ability to evaluate their answers to math problems improved from grade 2 to grade 3, whereas there was no change over time in the children's ability to predict which problems they could solve correctly. Children with MLD were less accurate than children without MLD in evaluating both their correct and incorrect solutions, and they were less accurate at predicting which problems they could solve correctly. However, children with MLD were as accurate as their peers in correctly predicting that they could not solve specific math problems. The findings have implications for the usefulness of children's self-review during mathematics problem solving.

Approximately 6 percent of school-age children are identified as having difficulties in mathematics that cannot be attributed to having low intelligence, a sensory deficit, or lack of economic resources (Badian, 1983; Lyon, 1996). This prevalence rate has been reported by researchers across the globe (Gross-Tsur, Manor, & Shalev, 1996; Ramaa & Gowramma, 2002), as reviewed elsewhere (Shalev, Auerbach, Mannor, & Gross-Tsur, 2000). Moreover, results from the first population-based prevalence study in the United States suggest that the cumulative incidence of math learning disability (MLD) ranges from 6 percent to 14 percent, depending on how MLD is defined (Barbarese, Katusic, Colligan, Weaver, & Jacobson, 2005). Although MLD affects many students, the phenomenon is not well understood (e.g., Butterworth, 2005; Gersten, Jordan, & Flojo, 2005; Mazzocco, 2005). Unlike research in reading disabilities (RD), the field of MLD research has a briefer history, making it more difficult to clearly define what constitutes MLD, and how MLD should be assessed. No core deficit for MLD has been identified; indeed, it may be impossible to do so, given the wide ranging skills related to poor math performance (Geary, 2004). Yet information about MLD is needed to inform the practices of early identification, intervention, or instructional modifications of children with persistent difficulty in mathematics.

Within the last decade, research efforts to characterize cognitive skills or deficits associated with MLD have increased. In this article, we report on a study of the relationship between metacognition and math ability among participants in a longitudinal study. Following a review of studies of MLD, we discuss metacognition and its relation to mathematics. We focus on two metacognitive skills, prediction and evaluation, that have recently been characterized as “offline” (Desoete, Roeyers, & Buysse, 2001), and we address why these metacognitive skills may differ across children with versus without MLD.

Characteristics of Children With MLD

Although several domains of functioning may be impaired among children with MLD, a universally accepted battery of tests designed to assess each domain does not exist. There are several different approaches to identifying children who exhibit overall poor achievement in mathematics. In the past, as described by Sattler (2002), the most common approach was to use a discrepancy between a child’s performance on mathematics achievement and intelligence tests. This approach has been challenged in research and policy settings, and it is no longer the gold standard it once was. An alternative approach is based on low achievement scores in math. However, by focusing on underachievement in isolation, it is possible to overlook the possibility that poor achievement may be unrelated to deficits linked specifically with MLD, and may instead result from factors such as lack of motivation, poor instruction, emotional problems, or school absenteeism. Still, low cutoff scores are often used as indices of MLD in research in the absence of a well-established alternative.

In order to increase the accuracy of assessing MLD, it is necessary to establish the fundamental characteristics of MLD. In one of the earlier studies of MLD, Rourke (1993) reported that children with poor arithmetic achievement have difficulties with spatial organization and with shifting from one mathematical operation to another, make procedural errors, have poor motor skills, and have memory deficits that interfere with math fact retrieval and problem-solving procedures. In addition, Rourke reported that these children make errors in judgment and reasoning (i.e., attempting to solve math problems too difficult for their level of understanding), leading to inaccurate solutions. Such errors in judgment reflect poor metacognitive skills.

In addition to Rourke, several other researchers (e.g., Boudah & Weiss, 2002; Butler, 1998; Geary, 1993; Montague, 1992; Vaidya, 1999) have identified metacognition as an area of difficulty for children with MLD. Children with poor metacognitive skills lack “knowledge concerning [their] own cognitive processes and products or anything related to them” (Flavell, 1976, p. 232). Children with poor metacognition are not able to judge which types of problems they can solve. These children fail to plan which operations are necessary to solve a given problem, have trouble monitoring the procedures they use, and often fail to recognize when they have made an error (Lucangeli, Cornoldi, & Tellarini, 1997). Thus, it is important to understand how these skills change over time for children who struggle in mathematics and for those who do not. Understanding how children’s metacognition develops may be one key to understanding how children become successful problem solvers in mathematics (Lucangeli et al., 1997).

Metacognition and MLDs

Metacognition is a broad concept that refers to any knowledge about one’s own cognition. It includes, but is not limited to, knowledge of the processes of “active monitoring and consequent regulation and orchestration” that occur during task performance, as initially described by Flavell (1976, p. 232). Additionally, metacognition includes knowledge of the processes that either precede or follow active task engagement and which would therefore involve prediction and appraisal of task performance, respectively. Given the breadth of these concepts, it is not

surprising that there is no consensus definition of metacognition (Baker, 1994). Since the term was first used in the 1970s, it has typically been conceptualized as having two primary components—(1) knowledge or awareness and (2) the actual regulation or executive control over tasks (Baker & Cerro, 2000). Some researchers argue that metacognition should be defined on the basis of knowledge about cognition, whereas other researchers believe that the definition should include emphasis on the regulation component. This ongoing debate has led to variation in which components of metacognition have been studied (for a complete review, see Baker & Cerro, 2000).

Metacognition may affect how children learn or perform mathematics. Students must learn how to monitor and regulate the steps and procedures used to meet the goal of solving problems. Academically successful students acquire the self-understanding that supports effective strategies to solve problems. Unfortunately, children with learning disabilities (LD) often lack this self-knowledge (Vaidya, 1999). Children with LD rate themselves as more competent than indicated by ratings from their teachers, despite the fact that the self-ratings are lower than self-ratings of their peers (Meltzer, Roditi, Houser, & Perlman, 1998). Although the study of metacognition in children with LD is not a new area of research, the characteristics of metacognition in children with MLD are not fully understood.

Several components of metacognition have been the focus of studies on children's math skills. For example, Lucangeli and colleagues (1997) focused on regulation in their study of third and fourth graders' mastery of four types of metacognitive skills. Before, during, and after the administration of the math test they completed, participants in that study were asked to predict how they would perform on a given math task, to plan which operations and steps they would use to arrive at a solution, to monitor their strategies used during the task, and to evaluate the effectiveness of the procedure used to complete the task. Children who exhibited more accurate mathematics performance on the test itself had a better understanding of the rules for employing the required problem-solving steps, and they were more accurate in evaluating the correctness of their own solution, relative to children with poor math performance. In fact, students' performance on the metacognitive tasks accurately differentiated good and poor problem solvers, but metacognitive skills were not as critical for automatized arithmetic skills as they were for solving geometry problems.

Desoete and colleagues (2001) carried out two studies to examine the structure of metacognition among third graders. In their first study, participants were classified as having below-average, average, or above-average math achievement. None of these participants was receiving treatment for school-related difficulties, and all were of average intelligence. Each child was assessed on three proposed components of metacognition: beliefs, knowledge, and skills. Principal components analysis revealed that the structure of metacognition did not consist of these three anticipated traditional components. Instead, Desoete and colleagues labeled the three emerging components as global metacognition, offline metacognition (which occurred before or after solving a math problem as opposed to during problem solving), and attributions for success or failure (i.e., ability, effort, luck, or task difficulty). Children who were better at predicting whether they would solve a problem correctly and who were better at evaluating the correctness of their responses had better math problem-solving ability. In other words, these metacognitive skills were associated with better performance, a pattern also shown by Tobias and Everson (2000) in their study of mathematics knowledge monitoring.

In their second study, Desoete and colleagues (2001) examined children's metacognition relative to presence or absence of moderate versus severe MLD. The presence of MLD and its severity were determined on the basis of concurrent math performance on three achievement tests, in combination with teacher ratings. The results of this study, like those from the first study, supported the distinction between a knowledge component (declarative, conditional,

and procedural knowledge), versus the components of prediction and evaluation. Children identified as having a severe MLD showed lower global metacognition than students in the moderate and non-MLD groups. As was true in the first study, each math-ability group significantly differed on the metacognition tasks. Children with severe MLD had significantly lower metacognitive skills than the moderately disabled and the average math problem solvers.

It is possible that poor metacognitive skills in children with LD result from immature, rather than absent, metacognitive skills. According to this developmental lag hypothesis, children with LD should be less accurate on prediction and evaluation than would age- or grade-matched children without LD, but no such group differences would emerge when these groups are matched on problem-solving ability. Desoete and Roeyers (2002) tested this hypothesis in their study of third graders with MLD, whose metacognitive skills were compared with those of third graders without MLD, and with second graders with comparable problem-solving ability. The results partially supported the maturational lag hypothesis: Third graders with MLD had prediction and evaluation skills comparable to those observed in second graders, but much poorer than those observed among third graders without LD. However, on easy math problems (problems designed to be appropriate for grades 1 and 2), third graders with MLD performed more poorly than did the second graders; whereas on difficult items (fourth grade level), second and third graders with MLD had better metacognitive skills than did third graders without MLD. Follow-up interviews revealed that the second and third graders with MLD were better able to judge which types of tasks they were not able to complete, simply because they looked very different from those that they learned in their classroom. The researchers concluded that the maturational lag hypothesis cannot completely explain MLD students' poor metacognitive skills.

Desoete and Roeyers (2002) also found that children with MLD had significantly lower prediction scores on number system knowledge, mental arithmetic, and procedural arithmetic than the younger children. Third graders with MLD with or without comorbid RD performed worse than the younger children on evaluation of number system knowledge and procedural arithmetic tasks. These differences reinforce the notion that MLD students have a different profile of prediction and evaluation skill development than students who do not have MLD, and that these differences were not limited to mathematics problem-solving tasks.

Together, these studies reflect the importance of assessing the metacognitive skills of prediction and evaluation in children who have MLD. Prediction skills allow children to distinguish between which problems are easy or difficult and which problems may require more skill or effort to complete. Children with good prediction skills are able to distinguish between real and apparent difficulties when predicting future performance (Desoete & Roeyers, 2002). Evaluation skills help children reflect on their solutions to problems and determine whether they made errors. If children have poor evaluation skills, it follows that their monitoring skills will also be poor. They will be unable to judge whether a math problem is sensible, whether they selected the correct procedure to solve the problem, or whether their answer is correct (Van Haneghan & Baker, 1989).

The Present Study

The primary goal of the present study was to assess whether prediction and evaluation skills differ in children with versus without MLD. Our emphasis on these processes is not meant to suggest that other aspects of metacognition are unimportant, but only that there are advantages to focusing on "offline" processes. For example, Brown, Bransford, Ferrara, and Campione (1983) emphasized how young children often have difficulty explaining their processes (i.e., the "online" tasks of planning and monitoring) and may provide unreliable reports of those processes. In contrast, assessing metacognition before and after problem solving is quick

(Desoete, Roeyers, & De Clercq, 2003; Lucangeli et al., 1997; Tobias & Everson, 2000) and may more reliably differentiate between very low and high math achievers (Desoete et al., 2001).

A related aim of the study was to improve upon existing developmental studies of how metacognitive and math skills are related. Several metacognitive studies described earlier in this article were carried out with children during only one point in their schooling. Although Desoete and Roeyers (2002) compared second and third graders, they did so using a cross-sectional design rather than following the same group of students over time. Also, although these researchers used several measures and teacher ratings to determine MLD status, these assessments occurred at one point in time. In contrast, the present study was a longitudinal assessment of the development of metacognition and math skills of the same children from grade 2 through grade 4. We were thus able to assess math achievement at multiple time points in order to base MLD status on persistent math difficulty. In addition, by assessing children's metacognition at multiple points over time we were able to address whether there is a reliable correlation between MLD and metacognitive skill.

In summary, our specific research questions were as follows: (1) Do the metacognitive skills of prediction and evaluation improve over time? (2) Are there group differences in these metacognitive skills as a function of whether children are identified as having MLD? (3) Does the pattern of change in metacognition over time differ as a function of whether children are identified as having MLD?

METHOD

Participants

Participants were recruited from seven public elementary schools within one suburban public school district. Schools with a relatively low mobility index were targeted for inclusion in the study in order to reduce attrition. Targeted schools also had a relatively low percentage of students who received free or reduced lunch, to reduce the added influence of low socioeconomic status (SES) on poor math performance (e.g., Leventhal & Brooks-Gunn, 2003). Children who attended the participating schools lived in heterogeneous middle-class suburban neighborhoods with a wide range of SES, excluding very low or very high levels, as described elsewhere in more detail (Mazzocco & Myers, 2002).

During the present study, the participants were in second, third, and fourth grades. During the first year of the study, the average age was 7 years and 9 months, and most students were in the second grade. From the initial group of 249 children enrolled in the study, 202 (81 percent) participated through fourth grade. Over the course of the project, 11 students repeated a grade in school. However, when discussing change over time throughout this article, reference will be made to the grade in which the majority of students were enrolled at the time of testing, although these 11 children were actually in a lower grade than the remaining participants. Initially, there were 25 children with MLD (boys, $n = 17$) and 224 in the non-MLD sample (boys, $n = 103$). Both groups of children were drawn from at least six of the seven participating schools.

Materials

Measure Used to Determine MLD Status—We used the Test of Early Math Ability—Second Edition (TEMA-2; Ginsburg & Baroody, 1990) to determine MLD status, based on findings reported in a previous report (Mazzocco & Myers, 2003). The TEMA-2 is a 65-item measure of formal and informal mastery of math-related concepts and is normed for children aged 2–8 years. Thus this standardized assessment differs from a more limited math

achievement subtest. Test-retest reliability for the TEMA-2 is reported as 0.94 (Ginsburg & Baroody, 1990). The entire TEMA-2 was administered during grades K to 3. Children whose TEMA-2 Composite scores were in the lowest 10th percentile during at least 2 years of primary school were classified as having MLD.

Evaluation Task—Children were asked to evaluate whether they were certain of their response given to each of several math test items. Eight items from the TEMA-2, all concerning place value, were included in the evaluation task (see Appendix). These were items 41, 46, and 57, the latter of which consisted of six questions, for a total of eight test items. These items were selected because the TEMA-2 age norms suggested that we could expect variability in the extent to which students accurately answer these items during the grades 2–4. One item (item 57a) was dropped from analyses because two responses (0 or 1) are acceptable as correct answers to the question it poses, “what is the lowest one-digit number?” With this omission, it was possible for children to make up to seven correct evaluations of their knowledge of place value.

Prediction Task—During each year of the study, the predictions task was given at the beginning of the testing session. Children were shown a set of problems from the Woodcock Johnson—Revised (WJ-R; Woodcock & Johnson, 1989) that constituted our prediction skills task. Without attempting to solve the problems, students were asked to circle “yes” if they thought they could correctly solve the problem, “no” if they thought they could not correctly solve the problem, or “I don’t know” if they were uncertain whether they could successfully solve the problem. (During second grade, children circled a checkmark, an “X,” or a question mark (?), to indicate these same response levels.) Similar to the evaluation items, the prediction items were selected to maximize variability in the students’ ability to solve the problems correctly. That is, there were items that children were expected to pass and some items we expected children would fail at each grade level. When the children were in second grade, there were 10 prediction items; at third grade, there were 31 prediction items. Approximately 30 minutes after making the predictions, students were given the standardized WJ-R calculations subtest, and were asked to solve the problems. This afforded assessment of children’s prediction accuracy.

Procedure

The items included in the present study were administered within the context of a larger, ongoing study, as described earlier. During each year of the study, children worked individually with a trained examiner. Efforts were made to make the testing conditions and test order constant for all children. During each year, at the onset of the testing session, the examiner engaged each child in a warm-up exercise designed to familiarize the child with use of the phrase “to be sure” of something, and to assess whether they understood the meaning of the phrase. The exercise included questions that the child should have been able to answer correctly with certainty, such as, “How old are you now?” The exercise also included questions that the child could not have answered with certainty, such as, “How old do you think I am now?” (with the examiner referring to herself). Following each question, the child was asked, “Are you sure, or are you not sure (of your answer)?” This activity was included to promote children’s awareness that it was acceptable not to be sure of an answer. This was a critical step, because developing a rapport with students and providing them with a supportive environment will increase the likelihood that they will openly disclose their feelings of uncertainty (Bryant & Rivera, 1997).

Following the warm-up trial, the predictions task was administered. After several other cognitive tasks not related to the present study, the complete WJ-R and the TEMA-2 were administered, in that order. (The TEMA-2 was included during only second and third grades,

and the WJ-R was included in all grades.) When given the WJ-R Math calculations sub-test, the child was not informed that some of the problems had appeared during the predictions task earlier in the session. For the TEMA-2 items included in the evaluation task, the examiners read each item aloud according to standardized testing protocols, and the child responded orally to items 41 and 46 and in writing to item 59. These items were administered to all students even if they reached their ceiling (defined as five missed consecutive items). Immediately after each response to an item, the examiner asked the child, “Are you sure, or are you not sure (of your answer)?”

Coding

The evaluation scores obtained from the TEMA-2 were calculated as follows: One point was awarded when students were aware of the accuracy of their response, and no points when students were unaware of their accuracy. These response categories are exemplified in Table 1. Awareness of accuracy level could be demonstrated by giving either a correct response with certainty that it was answered correctly or an incorrect response accompanied by uncertainty. Similarly, unawareness of accuracy occurred when students provided either a correct answer of which they were uncertain or an incorrect answer of which they were certain (Table 1). The total evaluation score equaled the number of points earned, with scores ranging from 0 to 7.

Prediction scores were computed by awarding 1 point for each accurate prediction. That is, students received a point when they predicted that they could solve a calculation and actually provided the correct answer or when they predicted that they would not be able to solve the problem and actually failed the item. Unlike with the evaluation task, the same prediction items were not given from one year to the next. This was because more difficult math problems were added each year, and the easiest items from the previous year were removed in subsequent years. Therefore, in order to examine progress from one year to the next, total prediction scores were converted to reflect the percentage of predictions that were accurate.

In some instances, a student’s response to an item from the evaluation task was, “I don’t know.” When this occurred, the examiner encouraged the child to respond. Nevertheless, some children still refused to at least provide a best guess. In this situation, asking the child, “Are you sure or are you not sure” was not applicable. To address this issue of missing data, participants were included in the analyses only if at least 85 percent of their data were complete (i.e., complete data for six of the seven items). As a result, 28 students were dropped from the evaluation analyses. These 28 included 8 children with MLD and 20 children without MLD. The final sample of participants included 17 children with MLD (12 boys) and 179 children without MLD (85 boys).

RESULTS

Preliminary analyses were conducted to determine whether children with versus without MLD had comparable raw scores on the TEMA-2 and the WJ-R calculations subtest. The group means and standard deviations of these scores are presented in Table 2. By definition, the children with MLD scored significantly lower on the TEMA-2 during both second and third grades, $t(29.802) = 12.027, p < 0.001$, and $t(17.93) = 9.98, p < 0.001$, respectively.¹ Children with MLD also scored significantly lower on the WJ-R calculations sub-test in both second and third grades, $t(20.497) = 10.33, p < 0.001$, and $t(193) = 7.33, p < 0.001$, respectively. Change in raw scores over time was not analyzed, because not all children completed the same number of items.

¹The *t* test for unequal variances was reported for these analyses because the assumption that the two groups had equal variances was not met. The degrees of freedom must be adjusted for such tests, and often they are not whole numbers.

The results of the study are presented in order of the three research questions. All dependent variables were continuous scores. For each analysis, 2 (MLD status: MLD/non-MLD) \times 2 (Grade 2 or 3; or Grades 3 or 4) analyses of variance (ANOVAs) were conducted, with grade as the repeated measure. Effect sizes are reported in cases of significant main effects.

Evaluation Accuracy

The means and standard deviations for the number of “sure” responses and for each type of evaluation response are presented in Table 3. A preliminary ANOVA was conducted in order to determine whether the number of “sure” evaluations changed over time, whether children with or without MLD had a comparable number of “sure” evaluations, and whether the number of “sure” responses changed at a different rate for students with MLD versus those in the non-MLD group. Results revealed that the overall number of “sure” evaluations increased significantly from second to third grades, $F(1,194) = 12.68, p < 0.001, \eta^2 = 0.06$. Children with MLD provided fewer “sure” evaluations than children without MLD, $F(1,194) = 12.50, p < 0.001, \eta^2 = 0.06$. There was no significant interaction, $F(1,194) = 0.986$.

The first analysis involving evaluation accuracy examined change in accuracy over time across both groups. Total evaluation scores improved significantly from second to third grade, $F(1,194) = 15.19, p < 0.01, \eta^2 = 0.07$. In other words, there was an increase in the number of items that children were able to accurately determine as to whether they had answered correctly. The children in the non-MLD group evaluated their performance more accurately, $F(1,194) = 54.49, p < 0.001, \eta^2 = 0.22$. The increase in evaluation scores was comparable for each group, showing no evidence of an MLD status \times Grade interaction, $F(1,194) = 0.057$.

Computing a total evaluation score by awarding points for awareness provides little insight into whether children’s accuracy results from children’s awareness of giving correct answers or from their awareness of their errors. To get a deeper understanding of the children’s metacognitive skills, additional analyses were conducted—one for each possible type of response: (1) Hits, or being “sure” when answering correctly, hereafter referred to as *sure-correct* responses; (2) True Negatives, or being “not sure” when answering incorrectly, hereafter referred to as *not sure-incorrect* responses; (3) False Alarms, or being “sure” when answering incorrectly, hereafter referred to as *sure-incorrect* response; and (4) Misses, or being “not sure” when answering correctly, hereafter referred to as *not sure-correct* responses. The means for each response type are presented for MLD and non-MLD students in Table 3. There are two types of accurate evaluations and two types of inaccurate evaluations (as was illustrated in Table 1). Results from the accurate evaluations are provided first.

Sure–Correct—Accurate evaluations in which students were sure of their correct responses occurred most frequently. From second grade to third grade, the number of these responses increased, $F(1,194) = 39.49, p < 0.001, \eta^2 = 0.17$. Children without MLD had more “sure-correct” response than did children with MLD, $F(1,194) = 91.53, p < 0.001, \eta^2 = 0.32$. There was no MLD status \times Grade interaction, $F(1,194) = 0.351$.

Not Sure–Incorrect—Across both groups, there was a significant decrease in the average number of times that students said they were unsure of their answers when they were incorrect, $F(1,194) = 12.69, p < 0.001, \eta^2 = 0.06$. This rate of decrease was constant across the two groups, $F(1,194) = 0.254$, as there was no significant interaction. These responses were more common among children with MLD, $F(1,194) = 27.06, p < 0.001, \eta^2 = 0.12$.

Sure–Incorrect—Children with MLD had more instances of being “sure” of responses when their answer was actually incorrect, relative to children without MLD, $F(1,194) = 69.26, p < 0.001, \eta^2 = 0.26$. From second grade to third, there was a significant decrease in the number

of these sure-incorrect responses, $F(1,194) = 18.57, p < 0.001, \eta^2 = 0.09$; thus children showed improved awareness of their response accuracy. This decrease in responding was constant across the two groups, as there was no significant MLD status \times Grade interaction, $F(1,194) = 0.303$.

Not Sure—Correct—Missed responses were those in which the children indicated that they were not sure of their answer, despite having answered the item correctly. The mean number of such responses increased from second grade to third grade among children with MLD, but decreased over time for children without MLD. However, this change over time was not significant, $F(1,194) = 0.003$. There was no significant group difference, $F(1,194) = 2.45$, nor was there a significant MLD status \times Grade interaction, $F(1,194) = 1.63$.

Prediction Accuracy

The means and standard deviations for the number of “yes” responses and for each type of prediction response appear in Table 4. A preliminary analysis involving the number of “yes” predictions was conducted to examine change over time and group differences. This ANOVA revealed that, from third to fourth grade, there was an increase in the proportion of calculations that students marked “yes” they could solve, $F(1,193) = 24.24, p < 0.001, \eta^2 = 0.11$. Children in the non-MLD group predicted they could correctly solve significantly more calculation problems than did children with MLD, $F(1,193) = 21.05, p < 0.001, \eta^2 = 0.10$. There was no MLD status \times Grade interaction, $F(1,193) = 0.28$.

Prediction accuracy analyses were conducted to examine how well students could predict which calculations they could solve on the WJ-R calculations subtest when they were in third and fourth grade. There was no change in the total proportion of accurate predictions over time, $F(1,193) = 0.22$. Similar to overall evaluation accuracy, the MLD group less accurately predicted the number of items that they could solve correctly, $F(1,193) = 38.26, p < 0.001, \eta^2 = 0.17$. The prediction accuracy for the two groups increased at a consistent rate, $F(1,193) = 0.67$.

Similar to the evaluation task, simply computing a total prediction accuracy score was not sufficient to understand whether children received high prediction scores because they knew which problems they could solve or because they knew which ones were too difficult. Therefore, two separate analyses were conducted. The first analysis provided information about the percentage of correct predictions when the students indicated “yes,” they could solve the problem. The second analysis provided information about the percentage of correct predictions when the students indicated “no,” they could not solve the problem. Students were also given the option to say “I don’t know” when they were unsure that they could correctly solve a problem. Because the percentage of accurate responses cannot be computed for this response type, information about the proportion of items they solved correctly, of those marked “I don’t know,” will be presented separately.

“Yes”—Correct—Of the items marked “yes,” the proportion that children with MLD accurately solved increased slightly from third to fourth grade, from 0.56 to 0.62. For the non-MLD group, the proportion of accurate “yes” predictions remained fairly constant over time, at roughly 0.87. Thus, there was no significant main effect of grade, $F(1,193) = 1.068$. Children with MLD were less accurate in their ability to predict if they could answer a problem correctly, $F(1,193) = 94.167, p < 0.001, \eta^2 = 0.33$. There was no significant interaction, $F(1,193) = 1.626$.

“No”—Incorrect—Overall, children’s accuracy in predicting when they could not solve a problem decreased significantly from third grade to fourth grade, $F(1,104) = 8.96, p = 0.003$,

$\eta^2 = 0.08$. Prediction accuracy did not differ between the MLD and non-MLD groups, $F(1,104) = 0.349$, and there was no interaction, $F(1,104) = 0.24$.

I Don't Know—Forty-two percent of the participating children marked “I don't know” for any WJ-R calculation during third and fourth grade. For these students, the percentage of the items for which they indicated “I don't know” was determined. ANOVA revealed that the number of “I don't know” responses decreased significantly from third grade to fourth grade ($M = 0.16$, $SD = 0.28$; $M = 0.08$, $SD = 0.20$, respectively), $F(1,193) = 11.28$, $p = 0.001$, $\eta^2 = 0.06$. There was no main effect of MLD status with regard to the proportion of “I don't know” responses given ($M = 0.134$, $SD = 0.10$; $M = 0.11$, $SD = 0.09$, respectively), $F(1,193) = 0.855$. Moreover, the percentage of children making “I don't know” responses was the same for children with MLD (44 percent) and children without MLD (42 percent), $\chi^2(1) = 0.021$, $p = 0.89$.

For the “I don't know” responses, accuracy scores could not be computed. Therefore, analyses were conducted to determine if the proportion of correctly solved calculations changed over time for those items marked “I don't know.” ANOVA revealed that the proportion of items answered correctly of those marked “I don't know” did not change from third to fourth grade, $F(1,80) = 3.165$. Additionally, there were no group differences in the proportion of correctly solved calculations that were marked “I don't know,” $F(1,80) = 3.37$, and no interaction effect was present, $F(1,80) = 3.04$.

Secondary Analyses Using a 25th Percentile Criterion

Although the criterion for MLD used in the present study was based on TEMA-2 performance in the lowest 10th percentile, a secondary set of analyses was conducted using performance in the lowest 25th percentile as the criterion. In view of earlier findings that different criteria lead to qualitatively different group profiles (e.g., Murphy, Mazzocco, Hanich, & Early, in press), it is important to learn whether the 25th percentile criterion—a criterion frequently used in research on MLD—leads to different metacognitive profiles of children with MLD, relative to the results reported above.

When the 25th percentile criterion was applied, the number of children identified as having MLD increased from 17 to 56, and the number of children in the non-MLD group decreased from 179 to 140. Consistent with the results from the analyses using the 10th percentile criterion, there were no significant interactions between MLD status and grade. Overall, the results from using the two different criteria differ only in effect size; there were no major differences with regard to patterns of statistical significance. What is interesting, however, is that larger effect sizes did not always occur with the larger sample resulting from using the 25th percentile criterion. Effect sizes for change in metacognitive skills over time were larger when using the 25th percentile criterion relative to the 10th percentile criterion, whereas effect sizes for group differences in metacognitive skills either did not differ as a function of criteria, or in the case of “sure-correct” evaluation responses, were larger when using the 10th percentile criterion. These findings support Murphy et al.'s (2005) conclusion that the two cutoff criteria do not necessarily implicate only quantitative differences in MLD groups.

DISCUSSION

This study was designed to examine metacognitive skills in second, third, and fourth graders, and to assess whether such skills differed as a function of whether children had a persistent MLD. We were interested to learn more about children's thoughts before and after they engage in math problem solving. Our study extends and expands upon the existing literature on metacognitive skills in children with MLD by nature of our longitudinal design and because we defined MLD as persistent low performance at more than one point in time. Following our

discussion regarding the interpretation of the findings, we discuss the study's limitations, its implications, and suggestions for future research.

Evaluation Skills

Items from the TEMA-2 were used to examine total evaluation scores and each possible response pattern of MLD and non-MLD students. The first question addressed was whether evaluation skill improved over time. The number of correct answers to the TEMA-2 items increased from grade 2 to grade 3, which has implications for the main findings. The total number of accurate evaluations improved for both groups from second to third grade, which may have resulted from additional instruction in place value concepts after second grade. The number of "sure"–correct responses also improved from second to third grade. Consistent with the decrease of incorrect answers given, there was a decrease in the number of "not sure"–incorrect responses. Additionally, there was a decrease in the number of instances in which students gave "sure"–incorrect responses. This may be a function of children's improved ability to recognize when they make errors, or may simply result from the significant decrease in the number of items children answered incorrectly.

The next question addressed was whether group differences in evaluation existed for MLD and non-MLD students. Preliminary analyses revealed that children without MLD correctly answered more items than did children with MLD. With respect to total evaluation scores, children with MLD were less accurate in evaluating their answers to the TEMA-2 items. This is consistent with earlier findings that children with higher mathematical competence have superior evaluation skills. Lucangeli and colleagues (1997) attributed better evaluation scores to a better understanding of the steps necessary to perform the tasks at hand and of the rules useful for evaluating answers. It should be noted, however, that participants without MLD in the present study did not have higher scores than participants with MLD for both types of accurate responses. Children in the non-MLD group had more hits (accurate "sure"–correct evaluations) than did children with MLD, whereas children with MLD students had significantly more true negatives ("not sure"–incorrect responses) than did children in the non-MLD group. Additionally, children with MLD had more false alarms (more "sure"–incorrect responses). The two groups were comparable in their number of misses ("not sure"–correct responses), so both groups were comparable in the degree to which they lacked confidence in their correct answers.

Of interest is the finding that the change in evaluation accuracy from second to third grade did not differ for children with MLD versus those without MLD. This finding should be interpreted with caution because of the small number of items (i.e., seven) used to assess evaluation skill (a more detailed discussion of this issue is presented in the limitations section). Future research that uses more test items and different types of math problems may reveal whether an interaction truly exists or whether evaluation accuracy in children without MLD improves at a faster rate than that of children with MLD.

Prediction Skills

The same three questions were explored for the prediction task, concerning prediction accuracy over time, group differences in prediction accuracy for children with versus without MLD, and possible interactions. Over time, there was an increase in the number of items that children believed they could answer correctly. Unlike the finding for the evaluation scores, the total proportion of accurate predictions did not change from third to fourth grade. The difference in performance on the evaluation and prediction task is inconsistent with previous research. Desoete et al. (2001) and Desoete and Roeyers (2002) found a strong correlation between students' prediction accuracy and evaluation accuracy, suggesting that students' performance on both tasks should be comparable. However, the current findings should not be interpreted

as offering evidence that the two skills are indeed different. It is difficult to determine whether the lack of change over time was the result of using two different sets of items in the third and fourth grades, or a result of children's lack of improvement in their ability to accurately predict which items they could solve correctly. Additionally, one must consider that although there was no change in total prediction accuracy, follow-up analyses revealed that the proportion of "no"–*incorrect* predictions significantly decreased from third to fourth grade. That there were fewer instances of this response pattern is consistent with the overall decrease in the proportion of responses that were marked "no."

Among the most important findings was that the total proportion of accurate predictions was higher for children in the non-MLD group compared to those in the MLD group, but that this group difference was related to only a proportion of accurate "yes" predictions. That is, children with MLD were overconfident in the number of problems that they could accurately solve. No group differences were present for "no"–*incorrect* predictions; this is inconsistent with Desoete and Roeyer's (2002) finding that second graders with MLD more accurately predicted when they would not be able to solve a problem compared to children in that grade who did not have MLD.

In the present study, children with versus without MLD had comparable proportions of math problems for which they marked "I don't know" when asked to predict whether they could solve specific math problems correctly. There was no change in the number of items that they solved correctly, nor were there any group differences in the number of items solved correctly. These findings should be interpreted with caution because of the small number of children who actually gave this type of response (42 percent). However, although more than half of the students never gave an "I don't know" response, this was true for both children with or without MLD (56 percent and 58 percent, respectively).

As was seen with the evaluation task, the pattern of change in prediction accuracy was comparable for children with versus without MLD, as there was no interaction between MLD status and grade. That children with MLD differed from children without MLD on both total evaluation and total prediction accuracy is consistent with earlier findings that these metacognitive skills are associated with students' problem-solving ability (Desoete et al., 2001; Desoete & Roeyers, 2002; Tobias & Everson, 2000). That is, students with low math achievement or those with MLD exhibit lower metacognitive skills than do students with average or above-average achievement. Also, the results presented here are consistent with the finding of Desoete and colleagues (2001) that requesting metacognitive judgments before or after mathematical problem solving can be sensitive enough to differentiate between ability groups.

Implications for Assessment and Instruction

The finding that children with MLD have poor metacognitive skills has implications for the identification of MLD and for classroom instruction. Until recently, a common method of MLD identification was to assess for a discrepancy between a student's math achievement and intelligence (Sattler, 2002). It is now well established that this method is inappropriate and empirically unsound (e.g., Francis et al., 2005), but the alternative for identifying MLD is not so well established. Until an alternative is clearly determined, it is useful to be maximally aware of the cognitive characteristics of children with MLD. While we do not suggest that metacognitive skills are a primary determiner of MLD, nor that only children with MLD have poor metacognition, the results of the present study suggest that assessing children's evaluation and prediction skills may be a useful contribution to assessments of children's mathematics abilities. Such a recommendation is consistent with the notion that using flexible interviews during students' problem solving enhances assessment of children's mathematical knowledge (e.g., Ginsburg, Jacobs, & Lopez, 1998).

Direct classroom instruction in metacognitive skills may be beneficial for children who have difficulty determining which problems they will be able to solve and who cannot identify whether they solved a problem correctly. If children are taught how to predict and evaluate their performance, they may become aware of how much effort they will need to exert in order to answer a problem, and they can review and correct answers that are wrong (Desoete & Roeyers, 2002). It should not be assumed that children will develop these skills naturally through mere repetitive exposure to various math problems. In a sample of children without MLD, Desoete and colleagues (2003) found that students who received direct instruction in metacognitive skills improved their accuracy on a test of problem-solving ability in addition to their prediction and evaluation scores.

Of course, if a child lacks the ability to recognize an incorrect response, then instruction to review and self-correct will not be an effective intervention strategy unless strategies for self-correction itself are included. In the present study, children with MLD made more true negative (“not sure”–incorrect) responses than children without MLD, but also made more false alarms (“sure”–incorrect). Both of these response types pertain to incorrectly solved math problems and suggest that children with MLD will be inconsistent in recognizing when their incorrect solutions to math problems are, in fact, incorrect. Moreover, children with MLD made fewer hits (“sure”–correct), suggesting difficulty recognizing when their solutions to math problems are indeed correct. Thus, for children with MLD, it is insufficient to intervene only at the level of teaching strategies for arriving at a solution. It is also necessary to provide metacognitive support to enhance reviewing skills and to help students learn to recognize both correct and incorrect solutions to their math problems.

Limitations and Directions for Future Research

A primary limitation of the present study is the small number of children with MLD, which resulted from the prospective nature of the longitudinal study from which they were drawn. However, our findings were comparable when we reanalyzed data using a larger sample resulting from a higher cutoff criterion to determine MLD status, as reported earlier. Also, although our sample of children with MLD was quite small, this group of children was fairly heterogeneous; they were students enrolled in one of six of our seven participating schools, and thus came from a range of lower- to upper-middle class neighborhoods and had different schoolteachers. As a group, their reading ability scores were quite variable, with decoding skills ranging from below average to well above average (standard scores in third grade ranged from 70 to 118). Fewer than half (6) of the 17 children with MLD scored sufficiently below average on decoding skills to fall in the range for RD. Given the small sample size, we did not analyze the effects of comorbid RD, although this question has been addressed with this sample in another report (Mazzocco & Myers, 2003).

The primary contributions of the present study were its longitudinal design, enabling the examination of metacognitive skills over time, and its use of multiple testing times to establish presence of a math disability. With a longitudinal design comes the challenge of developing metacognitive tasks that can be used for multiple years. This challenge was seen with the evaluation task, which consisted of only seven items. Although the children were asked to evaluate more than seven items on the TEMA-2, the items selected for analysis were limited to those that were likely to result in some variability in the extent to which students would be able to successfully answer over each relevant year of the study. This limited number of items may have resulted in a ceiling that was too low. This low ceiling was also due to the emphasis in the present study on examining MLD, and thus the need to avoid including problems that the majority of children with MLD would be incapable of completing. Similarly, the items included as being sufficiently challenging for children with MLD could not all be items that children without MLD would correctly solve with great ease. For these reasons, the number

of items included in the present study was small, and the small number of items may have prevented an interaction effect from being detected, if one truly exists. Future studies with additional items, and additional types of math items, may result in a much larger increase in the number of correct evaluations from one year to the next.

A limitation of the prediction task was that a different set of prediction items was used from one year to the next, making comparisons difficult. This limitation is a result of the current study's being a part of a larger research project that was focused on MLD, with a secondary aim of measuring metacognition. In future studies, the same items could be used both years by selecting items that represent various levels of ability required in order to solve them correctly. Additionally, the same set of items should be used to assess both evaluation and prediction skill. An improved design would use a single set of problems to assess both prediction and evaluation. For each item, students could predict which problems they believe they will accurately solve. Then, after solving each problem, they would identify those that they accurately answered. This procedure could be repeated for various domains of mathematics such as arithmetic, geometry, and problem solving, as was done in the work of Lucangeli and colleagues (1997).

Summary

Regardless of its limitations, the present study demonstrates changes over time, from second to fourth grade, in children's prediction and evaluation skills, and important differences in these skills as a function of MLD status. These findings are consistent with some of the earlier work on metacognitive skills, and they add to a growing body of research on characteristics of children with MLD—as well as to the broader body of research on metacognitive skills in children with LD. The findings demonstrate that children with MLD are no more likely than their peers to report that they do not know whether they can correctly solve a problem. That is, children with MLD are as likely as their peers to have an opinion as to whether they can or cannot solve a problem. Yet children with MLD are less accurate than their peers at *predicting* that they can solve a problem correctly, but as accurate as their peers at being correct when predicting that they cannot solve a problem correctly. It is interesting that *evaluation* assessments are not aligned with these results, because children with MLD are less accurate at evaluating both correct and incorrect solutions to math problems, relative to their peers. This misalignment of findings suggests that a given child's accuracy at predicting which math problems will be too difficult to solve may not be a valid indicator that the child can accurately review whether completed math problems were solved correctly.

Acknowledgments

This research was supported by NIH grant R01 HD 034061 awarded to Dr. Mazzocco. We wish to thank the study participants, their parents and teachers, and research coordinator Gwen Myers, who collected much of the data over all 3 years of the study. Ms. Myers made a significant contribution to this research.

References

- Badian, N. Dyscalculia and nonverbal disorders of learning. In: Myklebust, H., editor. Progress in learning disabilities. New York: Stratton; 1983. p. 235-264.
- Baker, L. Fostering metacognitive development. In: Reese, H., editor. Advances in child development and behavior. Vol. 25. San Diego, CA: Academic Press; 1994. p. 201-239.
- Baker, L.; Cerro, LC. Assessing metacognition in children and adults. In: Schraw, G.; Impara, J., editors. Issues in the measurement of metacognition. Lincoln, NE: Buros; 2000. p. 99-145.
- Barbarese WJ, Katusic SK, Colligan RC, Weaver AL, Jacobsen SJ. Math learning disorder: Incidence in population-based birth cohort, 1976–82, Rochester, MN. Ambulatory Pediatrics 2005;5:281–289. [PubMed: 16167851]

- Boudah DJ, Weiss MP. Learning disabilities overview: Update 2002. ERIC Digest 2002:ED4628080.
- Brown, AL.; Bransford, JD.; Ferrara, RA.; Campione, JC. Learning, remembering, and understanding. In: Flavell, JH.; Markman, EM., editors. Handbook of child psychology: Cognitive development. Vol. 3. New York: Wiley; 1983. p. 77-166.
- Bryant BR, Rivera DP. Educational assessment of mathematics skills and abilities. *Journal of Learning Disabilities* 1997;30:57–68. [PubMed: 9009875]
- Butler, DL. Metacognition and learning disabilities. In: Wong, BYL., editor. Learning about learning disabilities. 2. San Diego, CA: Academic Press; 1998. p. 277-307.
- Butterworth, B. Developmental dyscalculia. In: Campbell, J., editor. Handbook of mathematical cognition. New York: Psychology Press; 2005. p. 455-467.
- Desoete A, Roeyers H. Off-line metacognition—A domain specific retardation in young children with learning disabilities? *Learning Disabilities Quarterly* 2002;25:123–138.
- Desoete A, Roeyers H, Buysse A. Metacognition and mathematical problem solving in grade 3. *Journal of Learning Disabilities* 2001;34:435–449. [PubMed: 15503592]
- Desoete A, Roeyers H, De Clercq A. Can offline metacognition enhance mathematical problem solving? *Journal of Educational Psychology* 2003;95:188–200.
- Flavell, JH. Metacognitive aspects of problem solving. In: Resnick, LB., editor. The nature of intelligence. Hillsdale, NJ: Erlbaum; 1976. p. 231-235.
- Francis DJ, Fletcher JM, Stuebing KK, Lyon GR, Shaywitz BA, Shaywitz SE. Psychometric approaches to the identification of learning disabilities: IQ and achievement scores are not sufficient. *Journal of Learning Disabilities* 2005;38:98–108. [PubMed: 15813593]
- Geary DC. Mathematical disabilities: Cognitive, neuropsychological and genetic components. *Psychological Bulletin* 1993;114:345–362. [PubMed: 8416036]
- Geary DC. Mathematics and learning disabilities. *Journal of Learning Disabilities* 2004;37:4–15. [PubMed: 15493463]
- Gersten R, Jordan N, Flojo JR. Early identification and interventions for students with mathematics difficulties. *Journal of Learning Disabilities* 2005;38:293–304. [PubMed: 16122059]
- Ginsburg, H.; Baroody, A. Test of early mathematics ability. 2. Austin, TX: Pro-Ed; 1990.
- Ginsburg, HP.; Jacobs, SF.; Lopez, LS. The teacher's guide to flexible interviewing in the classroom: Learning what children know about math. Needham Heights, MA: Allyn & Bacon; 1998.
- Gross-Tsur V, Manor O, Shalev RS. Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology* 1996;38:25–33. [PubMed: 8606013]
- Leventhal T, Brooks-Gunn J. Children and youth in neighborhood contexts. *Current Directions in Psychological Science* 2003;12:27–31.
- Lucangeli D, Cornoldi C, Tellarini M. Mathematics and metacognition: What is the nature of the relationship? *Mathematical Cognition* 1997;3:121–139.
- Lyon GR. Learning disabilities. *Future of Children* 1996;6:54–76. [PubMed: 8689262]
- Mazzocco MMM. Challenges in identifying target skills for math disability screening and intervention. *Journal of Learning Disabilities* 2005;38:318–323. [PubMed: 16122063]
- Mazzocco MMM, Myers GF. Maximizing efficiency of enrollment for school-based educational research. *Journal of Applied Social Psychology* 2002;32:1577–1587. [PubMed: 19750148]
- Mazzocco MMM, Myers GF. Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia* 2003;53:218–253. [PubMed: 19750132]
- Meltzer L, Roditi B, Houser RF, Perlman M. Perceptions of academic strategies and competence in students with learning disabilities. *Journal of Learning Disabilities* 1998;31:437–451. [PubMed: 9763774]
- Montague M. The effects of cognitive and metacognitive strategy instruction on the mathematical problem solving of middle school students with learning disabilities. *Journal of Learning Disabilities* 1992;25:230–248. [PubMed: 1573333]
- Murphy MM, Mazzocco MMM, Hanich LB, Early M. Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cut-off criterion used to define MLD. *Journal of Learning Disabilities*. in press.

- Ramaa S, Gowramma IP. A systematic procedure for identifying and classifying children with dyscalculia among primary school children in India. *Dyslexia* 2002;8:67–85. [PubMed: 12067188]
- Rourke BP. Arithmetic disabilities, specific and otherwise: A neuropsychological perspective. *Journal of Learning Disabilities* 1993;26:214–226. [PubMed: 8515186]
- Sattler, J. *Assessment of children: Behavioral and clinical implications*. San Diego, CA: Jerome Sattler Publishers; 2002.
- Shalev RS, Auerbach J, Manor O, Gross-Tsur V. Developmental dyscalculia: Prevalence and prognosis. *European Child and Adolescent Psychiatry* 2000;9:58–64.
- Tobias, S.; Everson, H. Assessing metacognitive knowledge monitoring. In: Schraw, G.; Impara, J., editors. *Issues in the measurement of metacognition*. Lincoln, NE: Buros; 2000. p. 147-222.
- Vaidya SR. Metacognitive learning strategies for students with learning disabilities. *Education* 1999;12:186–189.
- Van Haneghan, JP.; Baker, L. Cognitive monitoring in mathematics. In: McCormick, CB.; Miller, GE.; Pressley, M., editors. *Cognitive strategy research*. New York: Springer; 1989. p. 215-238.
- Woodcock, RW.; Johnson, MB. *Woodcock-Johnson Psycho-Educational Battery-Revised*. Tests of Achievement. Itasca, IL: Riverside Publishing; 1989.

Biographies

Adia J. Garrett is a doctoral student in the Applied Developmental Psychology Ph.D program at the University of Maryland, Baltimore County. Her research interests include strategies children use to be successful test takers, early interventions that promote both academic and social competence, and educational policy. The research presented in this article was completed as a partial fulfillment for her research requirements.

Michèle M. M. Mazzocco is a developmental psychologist with research interests in cognitive development. She is Associate Professor of Psychiatry at the Johns Hopkins School of Medicine, and Director of the Math Skills Development Project at the Kennedy Krieger Institute. In 1997, she initiated an NICHD-supported longitudinal research program focused on math abilities in typically and atypically developing children. Through this program she conducts studies of elementary school children, children with math learning disability, and children with fragile X, Turner, or Barth syndrome.

Linda Baker is Professor of Psychology and Director of the Applied Developmental Psychology Ph.D. program, University of Maryland, Baltimore County. Her research focuses on early literacy development and motivation, parental influences on educational achievement, metacognition and comprehension monitoring, and instructional interventions to improve achievement.

TABLE 1

Coding for Evaluation and Prediction Tasks

Question Posed to Student and Examples of Each Response Category	Samples of Student Responses	Score
Evaluation Task		
“How many nickels make up one quarter?”		
“Are you sure or not sure?”		
Response Categories		
Sure—correct	“5, and I am sure”	1
Not sure—incorrect	“25, and I am not sure”	1
Sure—incorrect	“25, and I am sure”	0
Not sure—correct	“5, and I am not sure”	0
Prediction Task		
“Can you solve this problem correctly?”		
(Sample item) $17 \times 3 =$		
Response Categories		
Yes—correct	“Yes, I can” $17 \times 3 = 51$	1
No—incorrect	“No, I cannot” $17 \times 3 = 14$	1
Yes—incorrect	“Yes, I can” $17 \times 3 = 121$	0
No—correct	“No, I cannot” $17 \times 3 = 51$	0

TABLE 2

Mean Raw Scores on Math Tests for Children With or Without MLD

Math Test	Group	N	Grade 2		Grade 3		Grade 4	
			M	SD	M	SD	M	SD
TEMA-2	MLD	17	35.71	4.12	41.88	6.79	—	—
	Non-MLD	179	49.79	8.17	58.79	5.34	—	—
	Total	196	48.57	8.85	56.64	7.25	—	—
WJ-R calculations	MLD	16	—	—	9.00	2.58	16.00	4.68
	Non-MLD	179	—	—	16.21	3.57	22.43	3.23
	Total	195	—	—	15.62	4.01	21.90	3.79

Note. MLD = math disabled group; non-MLD = non-math disabled group.

TABLE 3
 Mean, Standard Deviation, and Range of Evaluation Responses for Children With or Without MLD

Responses	Group	Grade 2			Grade 3		
		M	SD	Range	M	SD	Range
Total "sure" responses	MLD ^c	5.06	2.08	0-7	5.47	1.87	0-7
	Non-MLD ^d	5.79	1.27	0-7	6.52	0.83	0-7
Evaluation score ^e	Total ^e	5.73	1.37	0-7	6.43	1.00	1-7
	MLD	3.88	1.96	0-7	4.53	1.97	0-7
"Sure"—correct ^{a,b}	Non-MLD	5.63	1.23	1-7	6.36	0.91	2-7
	Total	5.47	1.39	0-7	6.20	1.16	0-7
"Sure"—correct ^{a,b}	MLD	2.24	1.39	0-4	3.29	1.93	0-6
	Non-MLD	4.90	1.51	0-7	6.18	1.10	0-7
"Not sure"—incorrect ^{a,b}	Total	4.67	1.68	0-7	5.93	1.44	0-7
	MLD	1.65	1.97	0-7	1.24	1.79	0-7
"Sure"—incorrect ^{a,b}	Non-MLD	0.73	0.96	0-4	0.18	0.52	0-4
	Total	0.81	1.11	0-7	0.27	0.77	0-7
"Not sure"—correct	MLD	3.00	2.24	0-6	2.18	1.91	0-7
	Non-MLD	0.98	1.24	0-5	0.34	0.69	0-4
Total	Total	1.15	1.46	0-6	0.50	1.00	0-7
	MLD	0.06	0.24	0-1	0.18	0.53	0-2
Non-MLD	Non-MLD	0.37	0.72	0-5	0.24	0.50	0-3
	Total	0.34	0.69	0-5	0.23	0.50	0-3

Note. MLD = math disabled group; non-MLD = non-math disabled group.

^a Significant main effect of MLD status.

^b Significant main effect of grade.

^c N = 17.

^d N = 179.

^e N = 196 (sample size was consistent for all analyses).

TABLE 4

Mean Proportion, Standard Deviation, and Range for Prediction Responses

Responses	Group	N	M	SD	Range	M	SD	Range
"Yes" (total) ^{a,b}	MLD	16	0.52	0.20	0.10–1	0.68	0.14	0.35–0.97
	Non-MLD	179	0.70	0.20	0.20–1	0.83	0.15	0.32–1
	Total	195	0.61	0.38	0.10–1	0.75	0.27	0.32–1
"Yes"—correct ^a	MLD	16	0.56	0.32	0–1	0.59	0.22	0–0.83
	Non-MLD	179	0.87	0.16	0–1	0.87	0.11	0.35–1
"No"—incorrect ^b	Total	195	0.85	0.19	0–1	0.85	0.13	0–1
	MLD	12	0.97	0.01	0.67–1	0.78	0.25	0.33–1
	Non-MLD	94	0.91	0.25	0–1	0.77	0.31	0–1
"I don't know"—correct ^b	Total	106	0.91	0.24	0–1	0.77	0.31	0–1
	MLD	7	0.00	0.00	0–0	0.33	0.36	0–1
	Non-MLD	75	0.39	0.42	0–1	0.39	0.38	0–1
Total	82	0.35	0.41	0–1	0.39	0.37	0–1	

Note. MLD = math disabled group; non-MLD = non-math disabled group.

^a Significant main effect of MLD status.

^b Significant main effect of grade.

APPENDIX

Evaluation Items from TEMA-2

Item 41	An applied question concerning how many 10s are in 100
Item 46	An applied question concerning how many 100s are in 1,000.
Item 57	(a) “What is the smallest 1-digit number?”
	(b) “What is the largest 1-digit number?”
	(c) “What is the smallest 2-digit number?”
	(d) “What is the largest 2-digit number?”
	(e) “What is the smallest 3-digit number?”
	(f) “What is the largest 3-digit number?”

Note. From Ginsburg, H., and Baroody, A. (1990). *Test of Early Mathematics Ability* (2nd ed.). Austin, TX: PRO-ED.