Published in final edited form as:

Ann Dyslexia. 2003 January 1; 53(1): 218–253. doi:10.1007/s11881-003-0011-7.

Complexities in Identifying and Defining Mathematics Learning Disability in the Primary School-Age Years

Michèle M. M. Mazzocco and

Kennedy Krieger Institute and Johns Hopkins School of Medicine, Baltimore, Maryland

Gwen F. Myers

Kennedy Krieger Institute

Abstract

This paper is a descriptive report of findings from a prospective longitudinal study of math disability (MD). The study was designed to address the incidence of MD during primary school, the utility of different MD definitions, and evidence of MD subtypes. The results illustrate the dynamic properties of psychometrically derived definitions of MD. Different groups of children meet criteria for MD depending on which measure(s) are used for identification. Over time, a given individual may not continue to meet MD criteria, even when using the same assessments. Thus, the findings lead to cautions regarding the single-tool/ one-time assessment for a clinical diagnosis of MD. Twenty-two of 209 participants demonstrated “persistent MD” (MD-p), or MD for more than one school grade. Reading disability was relatively more frequent in this MD-p subgroup than in the remaining participants (25 percent vs. 7 percent). Reading-related skills were correlated with math achievement, as were select visual spatial skills. There was minimal overlap between groups who met either a “poor achievement” criteria or an “IQ-achievement discrepancy,” and the latter was far less stable a measure over time than the former. The results highlight the complexities of defining MD and illustrate the need for more research in this area.

To date, research on math disability (MD) is far less extensive than research on reading disability (RD). A search of articles published between 1974 and early 2003 yielded 14 to 33 times as many citations for “dyslexia” versus “dyscalculia” in *PsychInfo* and *Pubmed* searches, respectively. “Reading disability” was listed as a *PsychInfo* key word, whereas no key word category existed for math (or mathematics) disability. Indeed, there were more citations for reading *disability* (2,415) than there were for math *ability* (2,154). Yet, like RD, MD is a significant obstacle to academic achievement for many children. There is a need to better understand its causes and manifestation. In this paper, we address the manifestation of MD as a means by which to help define it.

In view of the widespread research attention devoted to RD, it is not surprising that RD is better understood relative to MD.¹ Phonological decoding deficits have been identified as core symptoms of RD, through replication and extension of seminal studies, including longitudinal research (e.g., Torgesen, Wagner, & Rashotte, 1994). These core deficits are evident across

Address correspondence to: Michèle M.M. Mazzocco, Principal Investigator, Math Skills Development Project, Kennedy Krieger Institute and Associate Professor, Psychiatry, Johns Hopkins School of Medicine, 3825 Greenspring Avenue, Painter Bldg., Top Floor, Baltimore, MD 21211. Telephone: (443) 923-4125; fax: (443) 923-4130. E-mail: mazzocco@kennedykrieger.org.

Publisher's Disclaimer: Copyright of *Annals of Dyslexia* is the property of International Dyslexia Association and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

¹This is not to state that RD is completely understood, for indeed many questions remain and research in this area is ongoing. However, there is a consensus working definition of RD (Consensus Project, 2002), but not for MD.

the various subtypes of RD that have been described (Morris, Stuebing, Fletcher, Shaywitz, Lyon, Shankweiler, Katz, Francis, & Shaywitz, 1998), and that persist over time (Shaywitz, Fletcher, Holahan, Shneider, Marchione, Stuebing, Francis, Pugh, & Shaywitz, 1999). Knowledge of these core deficits leads to an empirically based definition of RD, which, in turn, enhances the ability to identify RD and to provide effective remediation (e.g., Torgesen, Wagner, & Rashotte, 1997). This knowledge also provides a validated framework from which to study the effects of intervention (e.g., Foorman, Francis, Fletcher, Schatschneider, & Mehta, 1998).

Research on mathematics disability is less well developed than RD research. Despite the foundation of research demonstrating cognitive differences in young children with versus without MD (e.g., Geary, Bow-Thomas, & Yao, 1992; Geary, Hoard, & Hamson, 1999; Russell & Ginsburg, 1984; Hanich, Jordan, Kaplan & Dick, 2001), critical gaps in our knowledge of MD remain. So large is this gap that no universally accepted definition of MD exists, in contrast to the consensus definition for RD (Consensus Project, 2002). No core deficit has been identified for MD. It is possible that MD subtypes will not share a unifying core deficit because several different domains of function have been linked to poor math achievement, primarily reading-related, memory, visuospatial skills, and/or executive skills. In the field of MD, work toward establishing a consensus definition is in its early stages.

One promising aspect of early and ongoing MD research is the consistency in reports of cognitive correlates of math difficulties. These consistencies are in both the *domains* of function with which associations are evident and in the *strength* of these associations. For example, specific reading skills, particularly those associated with phonological processing, have been associated with computational math skills in children from grades 2 to 5 (Hecht, Torgesen, Wagner, and Rashotte, 2001), as suggested by earlier studies of children with poor math achievement (Russell & Ginsburg, 1984). Yet not all children with RD also have MD, so additional mitigating factors influence MD outcome. These other factors may be related to the types of math tasks on which an MD definition is based; for example, Hanich and colleagues report language-specific difficulties in children with MD and RD, relative to children with MD only (Hanich, et al., 2001). In their study, children with only MD outperformed their peers with both MD and RD on exact arithmetic tasks, whereas both groups demonstrated comparable difficulty on approximate arithmetic tasks. Bull, Johnston, and Roy (1999) also reported a significant correlation between math and reading abilities, and also found that math performance level was linked to executive function skills, even when statistically controlling for the contributions of reading ability and I.Q. score. Moreover, different components of executive function, as proposed by Miyake and colleagues (Miyake, Friedman, Emerson, Witzki, Howerter, & Wagner 2000), each appears to account for some of the variability in children's math performance level, with particularly strong contributions of poor inhibition and poor working memory (Bull & Scerif, 2001). These findings are consistent with Swanson's (1993) report of working memory deficits in children with learning disability, including children with reading or math difficulties. Thus, there is consistency across reports that both reading-related and executive skills are associated with math achievement level. Still to be explained is the extent to which these cognitive correlates underlie one or more specific MD subtypes.

Unlike the key basic processes that underlie reading achievement, mathematical achievement is cumulative throughout and beyond the elementary school years with quantitative and qualitative changes occurring within and across grade levels. The required changes concern performance demands and the necessary prerequisite skills. Thus, another question that remains concerns whether deficits occur, or at least are manifested, at different points along the trajectory of expected math skills development. In order to define MD, we need to understand its manifestation within developmental levels, within the same children over time, and within

and across MD subtypes. Longitudinal studies can help narrow the current gap in our knowledge, but only a few such studies are currently underway (e.g., Geary, Hamson, & Hoard, 2001; Jordan, Hanich, & Kaplan, 2003; Mazzocco, 2001; Shalev, Manor, Auerbach, & Gross-Tsur, 1998). These longitudinal studies address different questions, depending on how MD is defined in each study, and whether the study is retrospective or prospective.

In this paper, data from a prospective longitudinal study of math learning disability are presented. This study was designed to address the incidence of MD during the primary school age years (Kindergarten to Grade 3), the utility of different definitions of MD, the developmental trajectory of MD or of poor math achievement in primary school years, and evidence for subtypes of MD. The *incidence* of poor math achievement is examined as a function of performance on standardized psychometric instruments and experimental measures. The usefulness of proposed *MD definitions* are examined by applying various MD criteria, all based on widely used standardized measures, to the same group of children within each of four grade levels. Evidence of math LD subtypes is examined, in part, through correlational data, and also through assessment of the stability of MD criteria over time. To examine the *developmental trajectory and stability* of MD, we examine within-subject consistency of poor math achievement across the primary school age years. The information to be derived from the present study may enhance our ability to define and recognize MD and its subtypes. At the very least, this information should provide us with cautions about ignoring the complexities and limitations of current MD definitions when seeking to identify children who need special educational services; it also serves to illustrate the need for more research in this area.

Math Disability: Prevalence and Definitions

The prevalence of MD, reported as approximately 6 percent, parallels the reported frequency of RD (Badian, 1983; Ramaa & Gowramma, 2002; Shalev, Auerbach, Manor, & Gross-Tsur, 2000). Comparable frequencies of this magnitude have been reported across a wide range of studies conducted in different countries, despite variation in investigators' definitions of MD. Thus, one challenge to accurately determining *prevalence* of MD lies in the challenges inherent in *defining* MD. This, in turn, is related to the lack of consensus in defining learning disability (LD) in general, and the recent changes in legal definitions mandated by the Individuals with Disabilities Education Act (IDEA, 1997). Regardless of the distinctions used to define MD or LD, what is clear is that a significant portion of children demonstrates poor achievement in mathematics. In the United States, this observation has been broadly linked to a lack of nationwide core competencies for math knowledge and skills (National Science Foundation Third International Mathematics and Science Study—Repeated in 1999; TIMSS-R study; Mullis, Marting, Gonzalez, Gregory, Garden, O'Connor, Chrostowski, & Smith, 2000), while more marked poor academic achievement overall has also been linked to effects of economic poverty (Brooks-Gunn, Klebanov, & Duncan, 1996). Schooling in general is linked to degree of cognitive function (see Ceci, 1991 for a review), so access to school and the type of school is likely to influence math achievement. Such reports illustrate curriculum-based issues tied to math achievement in a broad sense, but not necessarily the unique developmental differences present in children predisposed to poor mastery of mathematics regardless of educational opportunity. The present report is based on this latter group, children who demonstrate poor math achievement relative to their age-matched and grade-level matched peers within similar educational environments. At issue, then, is how to effectively identify this group of children with developmental MD before their failure in mathematics becomes unequivocally apparent.

Historically, a discrepancy between achievement and IQ test scores served as a primary definition of LD. This practice continues in clinical and educational settings, and the widespread use of such criteria suggests that MD definitions should also follow a discrepancy-

based model. However, there are several reasons to avoid reliance of such a model. In the RD and general LD literature, there exists little if any empirical support for the effectiveness of a discrepancy-based model; moreover, there is strong evidence of the inappropriateness and ineffectiveness of such IQ-achievement discrepancy definitions (e.g., Fletcher, Francis, Shaywitz, Lyon, Foorman, Stuebing, & Shaywitz, 1998; Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1996; Siegel, 1989), and the discrepancy model is being challenged in revision of the Individuals with Disabilities Education Act (IDEA). Although IQ testing remains useful for other purposes, and although a child with LD may demonstrate an IQ achievement score discrepancy, the absence of such a discrepancy does not rule out the presence of LD, or of MD. That is, a child with an IQ math achievement discrepancy may indeed be likely to have MD, indicating that these criteria may have good specificity for identifying MD. However, if many children with MD fail to demonstrate a significant IQ-math achievement discrepancy, use of these criteria will not have sufficient sensitivity for identifying *all* children with MD. Measurements of change over time may be important indices of LD (Francis, Shaywitz, Stuebing, Shaywitz, & Fletcher, 1994; Geary, et al., 1999; Shaywitz & Shaywitz, 1994). The question that remains is, therefore, what measures can be used to enhance the specificity of MD assessment, while maintaining a high rate of overall positive predictive value, the likelihood that a person who “tests” as having MD does indeed have MD? None of these questions can be definitively addressed, however, until there is consensus on how to define MD. Until that time, variation in definitions of MD will persist among and between researchers and educators.

Relative to RD, MD has been defined specifically on the basis of poor math achievement (Strang & Rourke, 1983), or as a component of a generalized LD that includes problems with reading, writing, *and* math (Fleishner, 1994). School-identified children with MD, who may often present with low achievement in math *and* reading (Fleishner, 1994), do not necessarily meet researchers' criteria for MD (Geary, 1990). Different approaches to measuring and defining math achievement appear in the MD research; these include criterion-based (e.g., children in lowest 10th to 45th percentile among their grade level peers) and discrepancy-based (e.g., discrepancy from grade level score, or from IQ score) models. With respect to observing change over time as an indicator of LD (Francis, et al., 1994; Shaywitz & Shaywitz, 1994), Geary reports that poor math achievement in two or more consecutive grade levels is a helpful indicator of MD (Geary, et al., 1999). In view of the lack of consensus in defining and measuring MD, the present report is not based on an *a priori* definition of MD; instead the focus of the report is to assess the outcome of using different criteria to define or classify MD. To analyze MD versus non-MD comparisons, we relied on the persistence of poor achievement over time as a key criterion for determining MD status.

Math Disability Subtypes

When defining MD, it is important to consider the heterogeneity in cognitive profiles observed among children. Geary's (1993) proposed MD subtypes model addresses this heterogeneity. His three proposed subtypes are based on children's performance on specific arithmetic tasks and associated neuropsychological profiles. The *Semantic Memory MD* subtype has received the most consistent support across studies of MD. Semantic Memory MD coexists with RD, and it is characterized by poor math fact retrieval and variable response times to retrieval problems. Support for this subtype is drawn from several studies linking reading and math achievement levels (e.g., Bull & Scerif, 2001; Hanich, et al., 2001; Siegel & Linder, 1984; von Aster, 2000) from evidence of poor short-term memory among children with reading and/or math disability (Siegel & Linder, 1984) and specific evidence of poor math fact retrieval among some children with difficulty in math (e.g., Fletcher, 1985; Geary, et al., 1999; Russell & Ginsburg, 1984; von Aster, 2000). Russell and Ginsburg described poor mastery of math facts as the “most severe difficulties displayed by MD children (p. 241), and hypothesized that this

difficulty might result from “a specific memory discrepancy ... of the type ... in the case of reading” (p. 242).

The two remaining subtypes, as proposed by Geary, include Procedural and Visuospatial MD. The former is characterized by immature strategies, errors in math problem execution, and a delay in acquiring arithmetic concepts. Visuospatial MD involves difficulty with properly aligning numeric information, sign confusion, number omission or rotation, and general misinterpretation of spatially relevant numerical information (e.g., place value). Of the three MD subtypes, the cognitive features associated with the Visuospatial subtype are the least understood, in part because of the lack of empirical studies of visual spatial performance in MD children (Geary, 1993). The relation with a “nonverbal learning disability” (NLD) is also unclear because not all children with NLD reveal poor arithmetic skills, and certainly not all children with poor arithmetic have a profile consistent with NLD (von Aster, 2000). The differentiation of Geary's three subtypes has received empirical support (e.g., Mazzocco, 2001; Shalev, et al., 1998), and these form the basis of our investigation of MD subtypes. Subtype differentiation is essential for guidelines on identification and intervention strategies, and it is important to recognize that these three subtypes may not share one primary, core deficit.

In addition to recognizing qualitative characteristics of MD subtypes, the ability to differentiate a math *delay* versus *disability* is necessary for guidelines on intervention needs and strategies. Normal ranges in performance levels and in rates of change over time must be identified to qualitatively differentiate delay versus deficit (Francis, et al., 1996; Lyon, 1994); this can be achieved most effectively through longitudinal studies. In the present report, we examine the consistency of poor math achievement between Kindergarten through Grade 3 as a means to describe the trajectory of MD and to further differentiate math delay versus disability.

Method

Participants

All participants were recruited from one of seven participating schools within one suburban public school district. Relative to the district-wide mobility index and free/reduced lunch eligibility rate, the participating schools had lower mobility indices and lower percentages of children eligible for free or reduced school lunches. These criteria were used to target participating schools, to enhance long-term retention in the longitudinal study, and to diminish potential influences on poor math achievement that occur with higher mobility or lower socioeconomic status (McLoyd, 1998). Despite these selection criteria, the seven participating schools nevertheless represented a heterogeneous sample of middle class suburban neighborhoods (see table I). Thus, although the sample was not representative of all children in U.S. public schools, it was a sample representative of a wide range of socioeconomic categories in a large, diverse school district.

During the first year of the study, enrollment in the study was open to all English-speaking students attending regular half-day Kindergarten in one of these seven participating schools. Recruitment procedures, described elsewhere in greater detail (Mazzocco & Myers, 2002; Teisl, Mazzocco, & Myers, 2001), included meeting with school principals and faculty, receiving informed consent from parents, and receiving assent from each child participant. During the first year of the study, 249 children were enrolled. A total of 210 continued in the study for four years. One of these 210 children was identified as having a known neurologic condition associated with risk for poor academic achievement, and was thus omitted from the final sample for the present study. The present report is based on the 209 participants who participated in all four years of the study through third grade (including 103 boys and 106 girls). During the course of their study participation, nine of these 209 children subsequently

repeated Kindergarten ($n = 2$), first grade ($n = 6$), or second grade ($n = 1$). Thus, during the fourth year of the longitudinal study, these nine children were actually in second grade. Demographic characteristics of the study participants appear in table II.

Procedures

Each child was tested individually by a female examiner. Children tested at their own school were evaluated during two to three sessions each year; each session lasted approximately one hour. Children left their classroom only if they and their teacher gave approval for the child's absence. Efforts were made to keep the order of test administration constant for all participants, although occasional exceptions occurred due to circumstances beyond the examiner's control (e.g., fire drills). During Year 01 of the study, all assessments occurred at the child's school during the school day. Some children later transferred to nonparticipating schools, so some members of this participant subgroup were assessed at the investigators' offices during Years 02 ($n = 7$), 03 ($n = 10$), and 04 ($n = 13$) of the study. Assessments in our offices took place during one day.

Materials and Variables

During each year of the study, the assessment battery consisted of core and supplemental measures. Only the “core measures” that were administered during three or all four of the four years of the study are included in this report. The core battery administered annually to each child included standardized and experimental measures of basic math, reading-related, and visual spatial skills. Measures across these domains are pertinent to examining the Memory, Visuospatial, and Procedural subtypes of MD proposed by Geary (1993) as discussed in the introduction. Although use of a core measure may ascertain skills across the domains examined, these measures are categorized below by the primary domain of function they are designed to assess.

IQ Measures—In addition to the core battery, a standard “intelligence quotient” (IQ) score was obtained during the first and fourth years of the study. The Stanford Binet fourth edition (SB-IV; Thorndike, Hagen, & Sattler, 1986) was administered one time to provide a full-scale IQ score. For children five to six years of age, the SB-IV is an eight-subtest measure that was nationally standardized in 1985. During Year 04 of the study, the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999) was administered to provide a more current full-scale IQ score. The WASI is a four-subtest measure. Although a full WISC-III would have been desirable, time constraints did not permit that a full WISC-III be administered. The correlation between the WASI-4 subtest and the WISC-III FSIQ scores is .87 (Wechsler, 1999, p. 134). Both the SB-IV and WASI provide a full-scale IQ score, and although the standardization for both is based on a mean of 100, the standard deviation varies ($SD = 16, 15$, respectively). In analyses using SB-IV scores, the full-scale score was converted to conform with a standard score based on a SD of 15.

Despite obvious benefits of using one IQ instrument throughout this study, we used two different measures because the WASI was not available at the onset of the study. Also, the WASI is not normed for children under six years of age, and in the initial year of the study, the majority of our participants were five years old. The SB-IV was initially selected in part because of its standardization for children ages two through to adulthood. Between the ages of five and eight years, the number and types of SB-IV subtests to be administered change, thus leading to inconsistency in subtest administration over these age groups. Time to administer the SB-IV exceeded our allowable time period with the participants, so we used the WASI as an up-to-date norm referenced instrument when it became available, which was when our cohort reached Grade 3. In this study, the SB-IV and WASI provided a standardized IQ score with which to explore discrepancy-based indicators of MD.

Standardized Measures of Math Performance—Measures used to assess math performance and achievement included the age-appropriate subtests of the KeyMath—Revised achievement test (KM-R; Connolly, 1998), the Test of Early Math Ability—second edition (TEMA-2; Ginsburg & Baroody, 1990), and the Woodcock Johnson—Revised (WJ-R; Woodcock & Johnson, 1989) Math Calculations subtest. The KM-R is used as a diagnostic assessment of children's math concepts and skills, and the ability to apply these concepts and skills. It is normed for Grades K through 9, and is based on three areas: Basic Concepts, Operations, and Applications. The TEMA-2, normed for children aged two through eight years, is used to assess formal and informal mastery of mathematics related concepts. The WJ-R Calculations subtest involves paper and pencil math calculations (as does the operations section of the KM-R). Age-referenced standard scores were derived from each of these measures.

Visual Spatial/Perceptual Performance—To assess visual perceptual skills, the four motor-reduced subtests of the Developmental Test of Visual Perception—second edition (DTVP-2; Hammill, Pearson, & Voress, 1993) were administered annually. The DTVP-2 subtests involved matching figures on the basis of direction (Position in Space), identifying shapes in embedded designs (Figure Ground), Visual Closure skills, and matching shapes (Form Constancy). Two subtests from the Abstract/Visual Reasoning Area of the SB-IV (Copying and Pattern Analysis), and the Matrix Reasoning subtest of the WASI, were also available for analysis during the first two years (SB-IV) or the third and fourth years (WASI) of the study. Age-referenced standard scores were derived from each of these measures.

Reading-Related Skills—To measure reading-related skills, the Woodcock-Johnson—Revised (WJ-R) (Woodcock & Johnson, 1989) Letter Word Identification (LWID) subtest was administered each year of the study as a measure of single letter or single word recognition. In and after first grade, the WJ-R Word Attack subtest was also administered as a measure of phonological decoding, a skill described throughout the literature on RD as essential to successful reading. Although the WJ-III is now available, it was not available at the onset of the study. We maintained use of the WJ-R during the longitudinal study for consistency in data analysis. Age-referenced standard scores were derived from each of these measures. Nonstandardized reaction time scores were obtained from a measure of rapid automatized naming (RAN; Denckla & Rudel, 1976), which was administered annually. We wished to include measures of both rapid naming and decoding skills since both are essential components and predictors of reading disability (Wolf, 1997).

Variables—From each of the standardized measures, the standardized, age-referenced score was derived. For IQ discrepancy scores, the SB-IV was converted to meet the same standardization as WASI scores, although this conversion led to minimal changes. In addition, qualitative measures were also obtained, including the aforementioned skill-specific items and relative within-group ranks. For example, using data from the total sample ($n = 209$) for each year of the study, RAN percentile scores were calculated to establish which scores fell in or below the bottom 10th percentile, or in or above the 90th percentile. Similarly, relatively restrictive criteria of standard scores < 86 were initially explored as relatively more restrictive criteria for MD, using all measures of math performance that were available; and the less restrictive range included standard scores < 95 .

Results

Incidence of MD

The “incidence” of MD varied greatly depending on the criteria used to define MD. Tables III and IV are summaries of the number of children who met criteria for MD as a function of different criteria. The more restrictive set of criteria included a) a standard score < 86 (or a

scaled score < 7) on any single primary math measure, b) a TEMA-2 standard score in the bottom 10th percentile for the study sample, or c) a >15 point discrepancy between current IQ score and either the TEMA-2 or WJ-R Math Calculations subtest scores. Less restrictive cutoffs were also explored using standard scores < 95. The “incidence” of MD ranges a great deal, even within these two levels of restriction, depending on the specific criteria used. It is important to consider that these incidence figures vary across the same group of children, both *at a given point in time* and *over time*, during the primary school age years. For the analyses and discussion that follow, we chose to rely on the more restrictive criteria presented in table III, in view of the large number of children who met criteria indicated in table IV. (Although the data in table IV were subsequently not used in any of the analyses that follow, table IV is presented because some of these criteria are used in practical or research settings.)

The total percentage of children who met any single criteria for MD ranged from 0 to 45 percent, using the more restrictive criteria in table III. This wide range underrepresents the fact that nearly half of the children (97; 46 percent) met at least one of the MD criteria in Kindergarten, and that over half of all children met a single criterion at some point during the primary school age years (110; 53 percent). This finding may not be surprising in view of the fact that training in psychological assessment includes an emphasis on using multiple tests to arrive at a diagnosis (e.g., Sattler, 2001, p. 9). Although the use of a single test is not advocated as a reliable indicator of a learning disorder, such use may occur in practice. The variability reflected in tables III and IV provide support against such practice, and illustrate the need to consider the appropriateness of individual measures and their combinations to identify MD. It also highlights the importance of being familiar with a wide array of assessment tools, rather than relying on one fixed test battery for all children (as discussed by Kamphaus, Petoskey, & Rowe, 2000). Finally, this variability further illustrates the importance of considering consistency in performance—both at a point in time and over time—as an indicator of MD.

Consistency in Performance on Math Measures

To address the consistency of these defining criteria over time, we examined whether children moved “in” or “out” of MD categories from Grades K through 3. Of particular interest was whether a subset of children showed persistent “deficits” over time, and whether this subset could be retrospectively differentiated from peers in Kindergarten or first grade who had no math difficulties in Grades 2 or 3. To carry out this assessment, it was first necessary to establish an investigator-established definition of MD.

There was no consistent combination of criteria to emerge from our examination of the criteria met by individual children. Children met anywhere from one to all of the criteria listed on table III, as seen in figure 1. The one criterion that was most likely to occur with one or more additional criteria was the TEMA-2 score below the 10th percentile of the peer group. For this reason, we used the TEMA-2 score < 10 percent for many of our subsequent analyses.

Establishing MD Criteria to be Used in the Present Study—Children were assigned to low math performers (Group A), borderline /low average math performers (Group B), and children with average or above average math performance (Group C). We defined these groups on the basis of TEMA-2 performance in part because the scores from this measure had a wider range within each grade level, relative to other math measures used in this study. Also, related to this variability, the TEMA-2 is a comprehensive instrument of formal and informal mathematical ability. Finally, nearly all children (90 percent) who met the criteria of “poor TEMA-2 performance” also met one or more of the remaining criteria for MD listed in table III as described above.

We originally planned to define our TEMA-2-based MD group on the basis of a TEMA-2 score < 86, which is 1 standard deviation from the published age-referenced mean score of 100. Yet

our initial efforts yielded very few children who met this criterion in Grades 2 and 3 because the overall group mean TEMA-2 score increased slightly each year. We examined this change with a 2 (gender) by 4 (grade level) ANOVA, with repeated measures on the second variable, and found significant effects of grade level on TEMA-2 scores, $F(3,207) = 79.66, p < .0001$; there was no main effect of gender, nor was there a significant Gender by Grade level interaction. The TEMA-2 group mean and standard deviations were 101.16 (14.02), 106.59 (13.55), 110.58 (15.54), and 112.54 (15.75) during Grades K, 1, 2, and 3, respectively. Post hoc analyses for all possible pairwise comparisons further revealed significant differences between each pair of grade levels; all $ps < .0001$, except $p < .02$ for Grades 2 versus 3. In view of these grade level differences, we defined “poor math performance” on the basis of a TEMA-2 score that was *below the 10th percentile* for the study sample, $N = 209$.

It was not always possible to delineate the lowest performing group, Group A, with exactly 20 or 21 children because score distributions did not always differ exactly at the 20th or 21st lowest performing child. In Grades 1 and 3, the child in the 20th or 21st “lowest” place received the same score as that obtained by seven or nine other children, respectively. In cases where it was not possible to include only the 20 lowest-scoring children in Group A, we included fewer than 20 children so that we could minimize the absolute value of the difference between “20” and the actual demarcation between the lowest and second lowest groups.

Evaluation of a Discrepancy-based MD Criteria—In view of the frequency with which a discrepancy-based definition of MD is currently used in practice, we compared our investigator-defined classification with outcome using a discrepancy-based definition of MD. Children met this criterion if their TEMA-2 score was > 14 points lower than their full scale IQ score (FSIQ). In figure 2, we illustrate the minimal overlap that exists among children with or without discrepancy-based MD versus our investigator-defined MD (Group A) children at each of the four grade levels.

In a comparison of our low and borderline groups, there was no significant difference in the frequency of discrepancy-based MD among Group A versus B. The frequencies for Groups A and B were as follows: for Kindergarten, 25 versus 20 percent, respectively, $\chi^2(n = 65) = 0.21$, Fisher's Exact test = 0.75; for first grade, 10.5 and 3 percent, respectively, $\chi^2(n = 54) = 1.38$, Fisher's Exact = 0.28; for second grade, 25 and 15 percent, respectively, $\chi^2(n = 53) = 0.79$, Fisher's Exact = 0.48; and for third grade, 11.8 and 10.5 percent, respectively, $\chi^2(n = 54) = 0.02$, Fisher's Exact > 0.99 . Thus, the discrepancy-based definition did not differentiate the more or less “severe” cases of MD (Group A or B), *and the majority of children in both Groups A and B did not meet the discrepancy-based criteria.*

In a comparison of our low versus average/above average groups (Groups A versus C), there were significant differences in the frequency of discrepancy-based MD, but at only some grade levels. The frequencies for Groups A and C were as follows: for Kindergarten, 25 versus 1.4 percent, respectively, $\chi^2(n = 164) = 23.96$, Fisher's Exact < 0.001 ; for first grade, 10.5 and 1.3 percent, respectively, $\chi^2(n = 174) = 6.43$, Fisher's Exact = 0.06; for second grade, 25 and 6.4 percent, respectively, $\chi^2(n = 176) = 7.86$, Fisher's Exact = 0.02; and for third grade, 11.8 and 10.4 percent, respectively, $\chi^2(n = 181) = 0.03$, Fisher's Exact = 0.69. Thus, the discrepancy-based definition was minimally useful for differentiating the lowest performers (Group A) from children who seemed clear of any risk for poor math achievement (Group C).

A comparison of the borderline versus average/above average groups was also minimally informative. At Grades 1, 2, and 3, there was no significant difference in the frequency of discrepancy-based MD, and Fisher's Exact probabilities ranged from 0.14 to 0.99. The frequencies differed significantly only at Kindergarten, during which they were 20 versus 1.4 percent for Groups B and C, respectively, $\chi^2(n = 189) = 21.66$, Fisher's Exact = < 0.0001 . Thus,

the discrepancy-based definition did not differentiate the borderline/at risk group MD (Group B) from those clear of risk for MD (Group C) in first to third grades, and even when the difference was significant, the frequency represented a minority of children at risk for poor math achievement.

Consistency of MD—The consistency with which children remained in Group A across grades varied somewhat with the grade at which a child first met criteria for Group A. As depicted in table V, the number of children who are first assigned to the bottom 10th percentile of their peer group diminishes from Kindergarten to Grade 3, and by Grade 3, only one child who met criteria for MD did not meet criteria during any previous school year. Approximately 63 percent of the 35 children who met Group A criteria at any point in primary school (Grades K to 3) had “persistent” MD (MD-p; $n = 22$); that is, 22 children met Group A criteria for ≥ 2 years and 14 children had met MD criteria during only one of the four years. This latter group was classified as having nonpersistent MD (MD-np), and includes one child whose TEMA-2 score was < 86 in Kindergarten, but who was not in the bottom 10th percentile of TEMA-2 scores during Kindergarten. Thus, only 13 children had MD-np, if we use criteria of TEMA-2 score < 10 percent for all for grades. There was no statistically significant difference in the likelihood that a child would remain in Group A (and be classified as MD-p) as of function of entering Group A in Kindergarten versus first grade, $\chi^2(n = 30) = 2.18$, Fisher's Exact = 0.21. The data presented in table V suggest that children who first entered Group A in Kindergarten or first grade were more likely to be classified as MD in Grade 3, relative to those that entered Group A in second grade. However, most likely because of the small sample, this frequency difference did not reach statistical significance, $\chi^2(n = 35) = 4.59$, Fisher's Exact = 0.052.

MD Subtypes

To explore possible subtypes of MD, we examined whether children with MD also had poor reading ability or poor performance on visual perceptual tasks. Indicators of poor performance in these areas appear in table VI. Patterns of persistence (or the lack thereof) are illustrated in figure 3, which demonstrates the complexity in defining MD. This figure supports the notion of MD subtypes, as evident by the variability with which children with MD also had RD and/or difficulty on the DTVP-2 “position in space” task (Position in Space difficulty, or PSD).

To minimize alpha inflation resulting from multiple statistical analyses, it was important to select a subset of visual spatial and reading scores for MS subtypes analyses. Selection was based in part on outcome of the following correlation analyses.

Correlational Analyses: Reading Measures—We used various measures of math, and of reading, and found the strongest associations occurred when the TEMA-2 score was used as the math performance measure and when the reading measure used was the rapid automatized naming (RAN) response times (rt). The associations may have been strongest in part because of the restricted range in Word Attack scores in Kindergarten. Because of non-normal distribution of RAN rt data, we carried out Spearman's Rank correlations between the TEMA-2 and RAN. RAN-colors was the most appropriate subtest to use for the Kindergarten analyses. It had the strongest association with the TEMA-2 relative to the remaining RAN scores; more importantly, the Kindergarten sample size for RAN-colors was larger than the sample for RAN-numbers and letters since not all kindergartners could automatically and correctly cite numbers or letters. By Grade 1, RAN-numbers was completed by the majority of study participants and yielded the stronger associations. TEMA-2 performance was negatively correlated with speed of rapid naming, in all four grades, as summarized in table VII.

Correlational Analyses: Visual-perceptual Measures—Similarly, the association between visual and math scores was examined through correlations between DTVP-2 and TEMA-2 scores. Particularly noteworthy is the finding that the DTVP-2 “Position in Space” subtest (PS) was more strongly associated with the TEMA-2 score than were any of the remaining DTVP-2 subtest scores, and that it was more strongly associated with the TEMA-2 than it was with the three remaining DTVP-2 subtests (table VII). Moreover, the strength of the association, $\rho = .464, p < .0001$, was comparable to the strength of the TEMA-2 / RAN-colors correlation noted above, $-.458$. Although DTVP-2 subtest scores themselves became significantly correlated with each other with advancing grade levels, the DTVP-2 PS / TEMA-2 correlations remained significant each year, and remained comparable in strength to the TEMA-2 / RAN correlations.

The correlation between the TEMA-2 and either the RAN or DTVP-2 PS scores could be driven by a common underlying construct such as “g.” In order to consider this possible explanation, correlations were examined between RAN and DTVP-2 PS scores during all four years. There was no significant association among these two variables.

Frequency of RD—In addition to correlation analyses, evidence for subtypes can be drawn from the frequency of RD among Groups A, B, and C (seen in figure 3). Children with RD were present among all three of these math-performance based groups. Although the majority of children in Group A *did not* have RD, the frequency of RD was significantly greater in Group A (than in Group B or C). Because expected cell sizes fell below 5 for a 3 (groups) $\times$ 2 (RD versus not RD) Chi Square, we compared (1) Group A versus Groups B and C combined, (2) Group A versus Group B, and (3) Group A versus Group C, in three separate 2 $\times$ 2 analyses for each grade level. Results for the analyses sets were as follows: (1) Significantly more children had RD in *Group A versus Groups B and C combined*, in all four grades; $\chi^2(1, 209) = 13.18, 59.80, 10.55, \text{ and } 10.81$, for Grades K, 1, 2, 3, respectively, Fisher's Exact probabilities ranged from < 0.02 to < 0.0001 ; and (2) in *Group A versus C*, $\chi^2(1, n = 168) = 19.48$, Fisher's Exact $p < 0.001$; $\chi^2(1, n = 174) = 57.31$, Fisher's Exact $p < 0.0001$; $\chi^2(1, n = 176) = 15.71$, Fisher's Exact $p < 0.01$; and $\chi^2(1, n = 171)$, Fisher's Exact = 12.63, $p < 0.01$; for grades K to 3, respectively. (3) Results were inconsistent across grades when comparing *Groups A versus B*. There was no difference in frequency of RD when comparing Groups A and B for Kindergarten, Grade 2, or Grade 3, $ps > 0.19$. During grade 1 only, this difference was statistically significant, $\chi^2(1, n = 54) = 15.68$, Fisher's Exact $p < 0.001$.

Also, not all children with RD had MD, but MD occurred with greater frequency among those with RD; the frequency of MD among those with RD ranged from 33 percent (in grades 2 and 3) to 63 percent (in Grade 1), in contrast with those without RD. Among those without RD, the frequency range of MD was 5 percent, 7 percent, 8 percent, and 10 percent in Grades 1, 3, 2, and Kindergarten, respectively. These differences in frequency among children with versus without RD were significant during all four years of the study, Fisher's Exact probabilities ranged from < 0.02 to < 0.0001 .

Similarly, children with PSD were present among all three math performance-based groups, and although the majority of children in Group A *did not* have PSD, the frequency of PSD was significantly greater in Group A than in Groups B or C. (1) Significantly more children had PSD in *Group A versus Groups B and C combined*, $\chi^2(1; n = 209) = 5.12, 14.22, 11.97, 17.20$, in Grades Kindergarten, 1, 2, and 3, respectively, Fisher's Exact probabilities ranged from < 0.05 to < 0.001 . (2) Differences were also significant for all grades, for comparisons of *Group A versus C*: $\chi^2(1, n = 168) = 8.75$, Fisher's Exact $p < 0.02$; $\chi^2(1, n = 174) = 19.85$, Fisher's Exact $p < 0.01$; $\chi^2(1, n = 176) = 12.47$, Fisher's Exact $p < 0.01$; and $\chi^2(1, n = 177) = 22.54$, Fisher's Exact $p < 0.001$; for Grades Kindergarten, 1, 2, and 3, respectively. There were no

significant differences when comparing frequencies in Groups A versus B, probabilities ranging from 0.06 to 0.71.

There was no difference in frequency of co-occurring RD or PSD between children with MD-p versus MD-np, although our sample size is too small to address this definitively (See table VIII). Finally, although table VIII shows that more boys than girls met our criteria for MD, this gender difference did not reach statistical significance for comparisons using either the MD-p vs. Non-MD subgroups, $p = .07$, or the MD-all versus non-MD subgroups, $p = .20$.

Discussion

Persistent Poor Achievement as a Characteristic of MD

Although the results from this prospective study illustrate the complexities faced when defining MD, these findings lead to important, practical implications for identifying children with MD. The results illustrate the dynamic properties inherent in psychometrically-derived definitions of LD, including MD, and emphasize much caution in relying on a one-time assessment for a definitive diagnosis, a concern expressed by other researchers of MD (Silver, Pennett, Black, Fair, & Balise, 1999). At a given point in time, a very different group of children meets the criteria for MD depending on which measure(s) are used for identification. Over time, a given individual may or may not continue to meet a specific set of criteria for MD, even if the same measurement tools are used. Meeting criteria on one measure is not necessarily predictive of whether a child would meet other criteria at the same point in time, and no single combination of criteria emerged from our study as the most common set of characteristics of children who will continue to demonstrate MD over time (persistent MD, or MD-p).

Regardless of the measure(s) used to define MD, a significant number of children in our study demonstrated poor math achievement, relative to peers in the same school district. A total of 35 of 209 children (17 percent) met our investigator-established definition of MD for at least one of the four years of the study. However, we do not conclude that all of these 35 children had MD. We caution against use of a score from single time point as an indicator of MD because academic strengths and weaknesses can shift over time as a function of children's individual growth processes (Francis, et al., 1994). Thus, not all children classified as MD in a given year can be expected to have a “stable” pattern of MD. Among the 20 children identified as having MD in Kindergarten, 13 (65 percent) continued to be classified as MD (by our standards) in Grades 1, 2, and/or 3. A similar rate of “stability” was seen among the 35 children who had *ever* been classified as MD in Kindergarten through Grade 3; 22 of these 35 children (63 percent) had persistent MD (MD for ≥ 2 years). This rate of stability is comparable to the 50 percent persistence reported by Silver et al. (1999) regardless of whether their participants had received any math intervention. If we omit from the 22 children with MD-p the two children with FSIQ scores < 80 , the 20 remaining children with MD-p make up 9.6 percent of the study sample. This figure is close to other estimates of MD in the general population (Badian, 1983; Ramaa & Gowramma, 2002; Shalev et al., 2000).

Defining MD

It is evident that for the children with MD-p, particularly the 20 who also had IQ scores > 79 , MD is an appropriate group assignment. What is not evident is whether *additional children* in our sample also have MD. For example, there was some overlap between children in our groups A and B, in terms of greater frequency of RD and/or PSD, and with respect to “movement” from Groups A and B over time (see figure 3). It is important to determine whether some children in this “borderline” group have significant obstacles to math achievement, as do the children from Group A. It is also apparent that some children who do not meet an IQ achievement discrepancy-based criterion for MD demonstrate poor math achievement, and that

this poor math achievement persists over time. *This lack of IQ math discrepancy occurred for the majority of children in Groups A and B.* Our findings support the notion that an IQ achievement discrepancy is not necessary for defining MD, but we should consider whether it is ever an appropriate indicator of MD.

The characteristics of children who met the discrepancy-based criteria were quite variable, and this (in part) led us to exclude the criterion from our definition of MD. For example, when our participants were in third grade, the 13 children in the discrepancy-based subgroup had significantly higher FSIQ scores than the remaining 196 children (mean = 124 versus 108, respectively; $p < 0.0001$), and significantly lower TEMA-2 scores (101 versus 113, respectively; $p < 0.01$), as would be expected. Yet over half (7) of these 13 children had TEMA-2 scores above 106, including children with TEMA-2 scores as high as 122. Although some children in this subgroup may have MD, overall, this group seems less all-inclusive of children with math difficulty than our group defined by TEMA-2 < 10th percentile. The discrepancy-based group is certainly less inclusive of children with poor math achievement than is our investigator-defined MD subgroup. Also, the discrepancy-based criteria were less persistent than those for low TEMA-2 performance, as depicted in table IX (versus table V). Of the 35 who *ever* met our low TEMA-2 score criteria, 22 (63 percent) had MD-*p*; whereas only eight (18 percent) of the 44 children who *ever* met the discrepancy-based criteria did so *persistently* (for $\geq$ two years). We chose not to rely on the discrepancy-based MD criteria in our study in view of this lack of persistence over time. The question remains whether any child with a discrepancy has MD, in the absence of poor math achievement, and how to define MD in such cases.

Core Deficits and MD Subtypes

Evidence in support of MD subtypes came from this assessment of MD stability over time and from correlational analyses. The strength of the correlations between TEMA-2 performance and either rapid naming or DTPV-2 performance was comparable to the correlation reported by Bull and Scerif (2001) between math performance and executive function skills. Among their six- to eight-year-old participants, math performance was negatively correlated with perseverative responses on the Wisconsin Card Sort Task (WCST), -0.43 , and on an interference task (-0.46); and positively correlated with performance on a counting span measure, 0.44 . In the present study, TEMA-2 performance in Kindergarten was negatively correlated with speed on rapid automatized naming (RAN), $\rho = -0.458$, and was positively correlated with accuracy on the DTVP-2 “position in space” (PS) subtest, $\rho = 0.449$. It is unlikely that all of these correlation coefficients represent a single underlying cognitive correlate because the RAN and PS scores in the present study were not correlated during first to third grade, ρ s = 0.001 to 0.027 .

The TEMA-2/RAN data support the notion of an MD subtype associated with RD. Also, RD occurred more frequently in “Group A,” the group with poorest math performance, than in Groups B and C. However, there are some discrepancies between our findings and those in the literature, primarily with regard to the lower frequency of RD in our MD sample. Whereas RD is reported to occur among 50 percent of children with MD, in our study, about 25 percent of children with MD also had RD. Our lower figure may relate to our criteria for RD or to how we defined MD. It may be that the TEMA-2 captures primarily subgroups of MD children who do not have RD. However, we feel that this latter explanation is not likely because of the inclusion of fact retrieval items on the TEMA-2, items on which children with RD are reported to perform poorly (Geary, et al., 1999).

The TEMA-2 / DTVP-2 PS association offers partial support for a visual-spatial subtype of MD, as does the finding that poor PS performance occurred with greater frequency in the MD group than in the non-MD groups. Alternatively, it is possible that the PS subtest measures an

executive skill rather than a basic visual perceptual skill of orientation, and that it implicates a procedural subtype of MD. This PS subtest requires that the child compare an initial target figure (always presented on the left side of the page) with five forced choice alternatives, and to select which of the five choices is identical to the target figure in both shape and orientation. The correct responses vary across items in terms of the number of attributes that can be (or that must be) compared across the choices to select a correct response. We are currently exploring the contributions of visual perceptual and executive skills to success on this task, through studies of PS subtest correlates with other measures of visual perception versus measures of executive function skill, and through manipulations of the PS subtest stimuli. We hope that with further examination we may provide evidence addressing the notion of the Visuospatial versus Procedural MD subgroups. Despite not being able to assign either category definitively, we do feel that other MD subtypes are exemplified in addition to the Semantic Memory subtype discussed above.

There is empirical support for the notion that neuropsychological influences affect math performance, and this support can guide future studies of MD subtypes. Neuropsychological variables have been found to predict later math achievement at least as well, if not more accurately, than IQ scores alone; and when added to IQ scores, they contribute to the accuracy of such predictions (Clarren, Martin, & Townes, 1993). Neuropsychological features associated with poor math performance include (but are not limited to) poor reading (Badian, 1983; Donlan & Hutt, 1991; Geary & Hoard, 2001) and attentional or executive function skills (Geary, 1990). Similarities between RD and MD have been described (Kulak, 1993), and the two do co-occur (Badian, 1983; Light & DeFries, 1995; Geary & Hoard, 2001) in elementary school children. Exact arithmetic appears to have a stronger language component than the more visuospatial related approximate skills (Dehaene, Spelke, Pinel, Stanescu, & Tsivkin, 1999), and a comparison of these skills may differentiate models of Memory MD versus Visuospatial MD. Executive function and attentional components of math are also relevant to math performance; these include organizational skills, and success with understanding and applying strategies. The frequent errors in math execution seen in the Procedural subtype of MD may be associated with attention difficulties.

The evidence for MD subtypes receives further empirical support from the notable similarities between acquired MD and developmental MD (McCloskey, 1992). Theoretical links have been postulated between math achievement and visual spatial skill (Battistia, 1980; Luria, 1966; Rourke, 1993; Semrud-Clikeman & Hynd, 1990) or lexical skill (Benson & Denckla, 1969; Geary & Hoard, 2001). Both visual and linguistic components of math are implicated by functional imaging studies (Deheane et al., 1999), and MD does not appear indicative of any single, specific lateralized dysfunction (Shalev, Manor, Amir, & Gross-Tsur, 1993) in view of the multiple components that may underlie it (as reviewed by Geary & Hoard, 2001). These component skills represent several neuropsychological domains. Future studies of specific or very basic skills will enhance our ability to identify MD subtypes as we move toward a consensus definition of MD. Such assessments are ongoing in our research program as well as by others (e.g., Jordan, et al., 2003, Geary, Hamson, & Hoard, 2001).

Limitations of the Present Study

The limitations of the present study pertain primarily to our sample size or to our methods of measuring and defining MD. Although our MD sample is indeed very small, it reflects the frequency of MD reported across many other studies. Our definition of MD was far more restrictive than criteria used in other studies of MD; such studies often include children with standard scores in the bottom 25th or 35th percentile in the MD group. Our restrictive range may have led to some false negatives (excluding children who do have MD). However, we chose to err on the side of decreasing our false positives in order to increase the likelihood that

all children in our MD subgroup did indeed have MD. It is appropriate for other research questions to err in the opposite direction, thereby increasing a cutoff of criteria to ensure inclusion of *all* children with MD at the expense of including children without MD in the target sample. Both approaches are scientifically sound, but address different questions. This may explain some differences noted in our findings and those of other researchers, such as our relatively lower rate of MD/RD co-occurrence.

Another potential source of inconsistent findings concerns the measures we used to assess math skills. Researchers use different available instruments to identify MD subgroups in studies, such as the WRAT and PIAT (Alarcon, DeFries, Light, & Pennington, 1997; Siegel & Linder, 1984; Silver, et al., 1999), the Group Mathematics Test (Bull & Scerif, 2001), the TEMA-2 (e.g., Mazzocco, 2001), investigator-devised testing batteries (e.g., Gross-Tsur, Manor, & Shalev, 1996; Russell & Ginsburg, 1984), school performance measures and teacher ratings (e.g., von Aster, 2000). Our tables III and IV illustrate how dramatically group assignment can vary as a function of which measure is used.

Finally, a potential interacting variable is gender, and we did not systematically examine this in the present study. The small sample size prohibited any systematic analysis of possible gender effects, across MD overall or children with both MD and RD. Share, Moffitt, and Silva (1988) reported that girls with MD did not show visual/spatial deficits and von Aster (2000) did not find gender differences among MD children who were not from a clinical sample. In table VIII, there is a suggestion of gender differences, but it lacks the support of statistical significance. To fully address possible gender effects, a larger sample size is required, and this was not possible in our study.

Conclusion

The aim of this report was to address the incidence of MD, the usefulness of different criteria for defining or identifying MD, the trajectory of MD over the primary school age years, and evidence for MD subtypes. It is not possible to provide a single frequency figure of MD because our findings yield a wide range of frequency counts dependent on which MD criteria are used. The variability seen in incidence rates, as a function of measures used to define MD, illustrates the importance of understanding the sensitivity, specificity, and predictive value of a given instrument at different developmental levels, and the developmental appropriateness (or inappropriateness) of using any single tool to define MD. What is evident is that a subgroup of children showed persistent poor math achievement over time, relative to their peers from the same schools. Approximately 22 children from the initial study sample of 209 (11 percent) met MD criteria for $\geq$ two years of the study; of these, 20 had FSIQ scores $\geq$ 80, and represent 9.6 percent of the initial study sample, a figure consistent with frequencies reported by researchers of MD. Of course, our figures are based on the measure we determined to be the primary single most useful measure, and yet we are not advocating that this (or any other measure) be used alone to identify children with MD. Our results illustrate how highly inconsistent different measures of MD are, among the same group of children within a grade level, and among the same children across grade levels. Finally, the results support the notion of MD subtypes, with approximately 25 percent of our MD sample also meeting our criteria for RD. Measures of rapid naming, and of visual perception, were correlated with math performance. The latter correlation may reflect a role of executive function skills, so it is unclear whether this other “subtype” implicates a visuospatial or procedural subtype, or both. Future studies will help clarify the nature of MD subtypes, which will, in turn, contribute to efforts to define MD as a function of the underlying deficits. Until that time, the present study indicates the importance of including persistence over time in the definition of MD.

Acknowledgments

This research was supported by grant R01 HD 34061 from the National Institute of Child Health and Human Development, awarded to Dr. Michele M. M. Mazzocco. The authors thank the students whose dedicated participation contributed to this research; all of the children's parents; the faculty and staff members of the Baltimore County Public School District who welcomed us into their elementary schools to carry out this research; and research assistants Jennifer Lachance, Megan M. Kelley, Laurie Thompson, and Gwyn Gerner.

References

- Alarcon M, DeFries JC, Light JG, Pennington BC. A twin study of mathematics disability. *Journal of Learning Disabilities* 1997;30:617–623. [PubMed: 9364899]
- Badian, N. Dyscalculia and nonverbal disorders of learning. In: Myklebust, H., editor. *Progress in learning disabilities*. New York: Stratton; 1983. p. 235–264.
- Battistia M. Interrelationships between problem solving ability, right hemisphere processing facility and mathematical learning: Focus on learning problems in mathematics. *Psychological Bulletin* 1980;2:53–60.
- Benson DF, Denckla MB. Verbal paraphasia as a source of calculation disturbance. *Archives of Neurology* 1969;21:96–102. [PubMed: 5772552]
- Brooks-Gunn J, Klebanov PK, Duncan GJ. Ethnic differences in children's intelligence test scores: Role of economic deprivation, home environment, and maternal characteristics. *Child Development* 1996;67:396–408. [PubMed: 8625720]
- Bull R, Johnston RS, Roy JA. Exploring the roles of the visual-spatial sketch pad and central executive in children's arithmetic skills: Views from cognition and developmental neuropsychology. *Developmental Neuropsychology* 1999;15:421–442.
- Bull R, Scerif G. Executive functioning as a predictor of children's mathematics ability: Inhibition, switching, and working memory. *Developmental Neuropsychology* 2001;19:273–293. [PubMed: 11758669]
- Ceci SJ. How much does schooling influence general intelligence and its cognitive components? A reassessment of the evidence. *Developmental Psychology* 1991;27:703–722.
- Clarren SB, Martin DC, Townes BD. Academic achievement over a decade: A neuropsychological prediction study. *Developmental Neuropsychology* 1993;9:161–176.
- Connolly, AJ. *The key math revised: A diagnostic inventory of essential mathematics manual*. Circle Pines, MN: American Guidance Service; 1998.
- Consensus Project. Definition Consensus Project sponsored by the International Dyslexia Association and the National Institutes of Child Health and Human Development. 2002 August;
- Dehaene S, Spelke E, Pinel P, Stanescu R, Tsivkin S. Sources of mathematical thinking: Behavioral and brain imaging evidence. *Science* 1999;284:970–973. [PubMed: 10320379]
- Denckla MB, Rudel RG. Rapid automatized naming (R.A.N.): Dyslexia differentiated from other learning disabilities. *Neuropsychologia* 1976;14:471–479. [PubMed: 995240]
- Donlan, C.; Hutt, E. Teaching maths to young children with language disorders. In: Durkin, K.; Shire, B., editors. *Language in mathematical education*. Philadelphia: Open University Press; 1991. p. 198–207.
- Fleishner, JE. Diagnosis and assessment of mathematics learning disabilities. In: Lyon, GR., editor. *Frames of reference for the assessment of learning disabilities*. Baltimore: Paul H. Brookes Publishing Company; 1994. p. 441–458.
- Fletcher JM. Memory for verbal and nonverbal stimuli in learning disability subgroups: Analysis by selective reminding. *Journal of Experimental Child Psychology* 1985;40:244–259. [PubMed: 4045379]
- Fletcher JM, Francis DJ, Shaywitz SE, Lyon GR, Foorman BR, Stuebing KK, Shaywitz BA. Intelligence testing and the discrepancy model for children with learning disabilities. *Learning Disabilities Research & Practice* 1998;13:186–203.
- Foorman BR, Francis DJ, Fletcher JM, Schatschneider C, Mehta P. The role of instruction in learning to read: Preventing reading failure in at-risk children. *Journal of Educational Psychology* 1998;90:37–55.

- Francis, DJ.; Shaywitz, SE.; Stuebing, KK.; Shaywitz, BA.; Fletcher, JM. Measurement of change: Assessing behavior over time and within a developmental context. In: Lyon, GR., editor. *Frames of reference for the assessment of learning disabilities*. Baltimore: Paul H. Brookes Publishing Company; 1994. p. 29-58.
- Francis DJ, Shaywitz SE, Stuebing KK, Shaywitz BA, Fletcher JM. Developmental lag versus deficit models of reading disability: A longitudinal, individual growth curves analysis. *Journal of Educational Psychology* 1996;88:3-17.
- Geary DC. A componential analysis of an early learning deficit in mathematics. *Journal of Experimental Child Psychology* 1990;49:363-383. [PubMed: 2348157]
- Geary DC. Mathematical disabilities: Cognitive, neuropsychological, and genetic components. *Psychological Bulletin* 1993;114:345-362. [PubMed: 8416036]
- Geary DC, Bow-Thomas CC, Yao Y. Counting knowledge and skill in cognitive addition: A comparison of normal and mathematically disabled children. *Journal of Experimental Child Psychology* 1992;54:372-391. [PubMed: 1453139]
- Geary D, Hamson C, Hoard M. Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology* 2001;77:236-263. [PubMed: 11023658]
- Geary D, Hoard M. Numerical and arithmetical deficits in learning disabled children: Relation to dyscalculia and dyslexia. *Aphasiology* 2001;15:635-647.
- Geary DC, Hoard MK, Hamson CO. Numerical and arithmetical cognition: Patterns of functions and deficits in children at risk for a mathematical disability. *Journal of Experimental Child Psychology* 1999;74:213-239. [PubMed: 10527555]
- Ginsburg, H.; Baroody, A. *Test of Early Mathematics Ability*. Vol. 2nd. Austin, TX: PRO-ED; 1990.
- Gross-Tsur V, Manor O, Shalev RS. Developmental dyscalculia: Prevalence and demographic features. *Developmental Medicine and Child Neurology* 1996;38:25-33. [PubMed: 8606013]
- Hammill, D.; Pearson, N.; Voress, J. *Developmental test of visual perception*. Vol. 2nd. Austin, TX: PRO-ED; 1993.
- Hanich LB, Jordan NC, Kaplan D, Dick J. Performance across different areas of mathematical cognition in children with learning disabilities. *Journal of Educational Psychology* 2001;93:615-626.
- Hecht SA, Torgesen JK, Wagner RK, Rashotte CA. The relations between phonological processing abilities and emerging individual differences in mathematical computation skills: A longitudinal study from second to fifth grades. *Journal of Experimental Child Psychology* 2001;79:192-227. [PubMed: 11343408]
- Individuals with Disabilities Education Act (IDEA). Public Law. 1997 June 4;:105-117.1997
- Jordan NC, Hanich LB, Kaplan D. A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with co-morbid mathematics and reading difficulties. *Child Development* 2003;74:834-850. [PubMed: 12795393]
- Kamphaus RW, Petoskey MD, Rowe EW. Current trends in psychological testing of children. *Professional Psychology: Research and Practice* 2000;31:155-164.
- Kulak AG. Parallels between math and reading disability: Common issues and approaches. *Journal of Learning Disabilities* 1993;26:666-673. [PubMed: 8151206]
- Light GJ, DeFries JC. Comorbidity for reading and mathematics disabilities: Genetic and environmental etiologies. *Journal of Learning Disabilities* 1995;28:96-106. [PubMed: 7884303]
- Luria, AR. *Higher cortical functions in man*. New York: Basic Books; 1966.
- Lyon, GR. Critical issues in the measurement of learning disabilities. In: Lyon, GR., editor. *Frames of reference for the assessment of learning disabilities*. Baltimore: Paul H. Brookes Publishing Company; 1994. p. 3-13.
- Mazzocco M. Mathematics disability and MD subtypes: Evidence from studies of Turner syndrome, fragile X, and neurofibromatosis type 1. *Journal of Learning Disabilities* 2001;34:520-533. [PubMed: 15503567]
- Mazzocco MMM, Myers GF. Maximizing efficiency of enrollment for school-based educational research. *Journal of Applied Social Psychology* 2002;32:1577-1587. [PubMed: 19750148]

- McCloskey M. Cognitive mechanisms in numerical processing: Evidence from acquired dyscalculia. *Cognition* 1992;44:107–157. [PubMed: 1511584]
- McLoyd VC. Socioeconomic disadvantage and child development. *American Psychologist* 1998;53:185–204. [PubMed: 9491747]
- Miyake A, Friedman NP, Emerson MJ, Witzki AH, Howerter A, Wagner TD. The unity and diversity of executive functions and their contributions to complex frontal lobe tasks: A latent variable analysis. *Cognitive Psychology* 2000;41:49–100. [PubMed: 10945922]
- Morris RD, Stuebing KK, Fletcher JM, Shaywitz SE, Lyon GR, Shankweiler DP, Katz L, Francis DJ, Shaywitz BA. Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology* 1998;90:347–373.
- Mullis, IVS.; Marring, MO.; Gonzalez, EJ.; Gregory, KD.; Garden, RA.; O'Connor, KM.; Chrostowski, SJ.; Smith, TA. TIMSS 1999: International mathematics report. Chestnut Hill, MA: TIMSS International Study Center; 2000.
- 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]
- Russell RL, Ginsburg HP. Cognitive analysis of children's mathematics difficulties. *Cognition and Instruction* 1984;1:217–244.
- Sattler, JM. Assessment of children: Cognitive applications. Vol. 4th. La Mesa, CA: Jerome Sattler Publisher, Inc.; 2001.
- Semrud-Clikeman M, Hynd GW. Right hemispheric dysfunction in nonverbal learning disabilities: Social, academic, and adaptive functioning in adults and children. *Psychological Bulletin* 1990;107:196–209. [PubMed: 2181523]
- Shalev RS, Auerbach J, Manor O, Gross-Tsur V. Developmental dyscalculia: Prevalence and prognosis. *European Child and Adolescent Psychiatry* 2000;9:58–64.
- Shalev R, Manor O, Amir N, Gross-Tsur V. The acquisition of arithmetic in normal children: Assessment by a cognitive model of dyscalculia. *Developmental Medicine and Child Neurology* 1993;35:593–601. [PubMed: 9435775]
- Shalev R, Manor O, Auerbach J, Gross-Tsur V. Persistence of developmental dyscalculia: What counts? Results from a 3-year prospective follow-up study. *The journal of Pediatrics* 1998;133:358–362. [PubMed: 9738716]
- Share DL, Moffit TE, Silva PA. Factors associated with arithmetic-and-reading disability and specific arithmetic disability. *Journal of Learning Disabilities* 1988;21:313–320. [PubMed: 3404067]
- Shaywitz SE, Fletcher JM, Holahan JM, Shneider AE, Marchione KE, Stuebing KK, Francis DJ, Pugh KR, Shaywitz BA. Persistence of dyslexia: The Connecticut longitudinal study at adolescence. *Pediatrics* 1999;104:1351–1359. [PubMed: 10585988]
- Shaywitz, BA.; Shaywitz, SE. Measuring and analyzing change. In: Lyon, GR., editor. *Frames of reference for the assessment of learning disabilities*. Baltimore: Paul H. Brookes Publishing Company; 1994. p. 59–68.
- Siegel LS. IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities* 1989;22:469–479. [PubMed: 2794763]
- Siegel LS, Linder BA. Short-term memory processes in children with reading and arithmetic learning disabilities. *Developmental Psychology* 1984;20:200–207.
- Silver CH, Pennett D, Black JL, Fair GW, Balise RR. Stability of arithmetic disability subtypes. *Journal of Learning Disabilities* 1999;32:108–119. [PubMed: 15499712]
- Strang, J.; Rourke, B. Arithmetic disability subtypes: The neuropsychological significance of specific arithmetical impairment in childhood. In: Rourke, B., editor. *Neuropsychology of learning disabilities: Essential subtype analysis*. New York: Guilford Press; 1983. p. 167–183.
- Swanson HL. Working memory in learning disability subgroups. *Journal of Experimental Child Psychology* 1993;56:87–114. [PubMed: 8366327]
- Teisl JT, Mazzocco MMM, Myers G. The utility of kindergarten teacher ratings for predicting low academic achievement in first grade. *Journal of Learning Disabilities* 2001;34:286–293. [PubMed: 15499882]

- Thorndike, RL.; Hagen, EP.; Sattler, JM. The Stanford Binet intelligence scale: Guide for administering and scoring. Vol. 4th. Chicago: The Riverside Publishing; 1986.
- Torgesen JK, Wagner RK, Rashotte CA. Longitudinal studies of phonological processing and reading. *Journal of Learning Disabilities* 1994;27:276–286. [PubMed: 8006506]
- Torgesen JK, Wagner RK, Rashotte CA. The prevention and remediation of severe reading disabilities: Keeping the end in mind. *Scientific Studies of Reading* 1997;1:217–234.
- von Aster M. Developmental cognitive neuropsychology of number processing and calculation: Varieties of developmental dyscalculia. *European Child & Adolescent Psychiatry* 2000;9:41–57.
- Wechsler, D. Wechsler abbreviated scale of intelligence. San Antonio, TX: The Psychological Corporation; 1999.
- Wolf, M. A provisional, integrative account of phonological naming speed deficits in dyslexia: Implications for diagnosis and intervention. In: Blachman, B., editor. *Foundations of reading acquisition and dyslexia*. Mahwah, NJ: Lawrence Erlbaum Associates; 1997. p. 67-93.
- Woodcock, R.; Johnson, M. Woodcock-Johnson psycho-educational battery - Revised. Allen, TX: DLM Teaching Resources; 1989.

Figure 1. Number of children meeting one or more criteria for Math Disability, per grade. Maximum number of criteria varied across grades; there were six possible for Kindergarten and Grade 1, five for Grade 2, and two for Grade 3

Figure 2. Number of children meeting one or two criteria for Math Disability

Figure 3. Number of children, per grade, meeting criteria for Group A (< 10th %ile on TEMA-2), Group B TEMA-2 score > 10th %ile and < 25th %ile), or Group C (TEMA-2 > 25th %ile). Also noted is the number of children per group who met criteria for Reading Disability (RD) or Position in Space Difficulty (PSD). Line thickness between grades corresponds to the number of children moving from one group to another, over time. Also noted is the number of children who “entered” Group A in Kindergarten (K) or Grade 1 or 2. For example, of the 13 children who moved from Group A in grade 1 (Group A-1) to Group A in grade 2 (Group A-2), nine first entered Group A in Kindergarten (9 A-K) and four first entered Group A in grade 1 (4 A-1)

TABLE I
Demographic Data for the Seven Participating Schools

Participating schools	Mobility ^a index	% Free ^a or reduced lunch	Distance to downtown (in miles)	Parents level ^b of education <i>Mean</i>	<i>SD</i>	Average home sale price ^c
1	19.5	27.69	8.35	5.3	1.2	\$163,064.08 ^d
2	6.8	1.58	13.22	5.8	0.7	\$240,865.00
3	18.9	29.04	7.98	4.7	1.0	\$118,740.42
4	16.2	14.71	16.33	5.4	1.1	\$231,556.00
5	20.1	11.11	9.25	5.5	1.1	\$146,669.34
6	18.0	12.67	7.55	5.0	0.9	\$123,123.98
7	16.2	18.70	8.63	5.0	0.9	\$118,740.42

^aMobility index and percentage of free/reduced lunches based on 1997 data. District-wide figures ranged from 4.8 to 51.0 ($X = 20.79$) and 1.58–79.70 ($X = 30.76\%$), respectively.

^bLevel of education was coded as follows: 4 = completed grade 12; 5 = completed 1 to 3 years of college; 6 = completed 4 year college; 7 = University /graduate school. Data were available for parents of 202 of the 209 participants; most recently available information was used. Mother's education level was used except for one household headed by a single-parent father.

^cBased on Metropolitan Regional Information Systems, Inc. from homes sold in 2001.

^dSale price figure may be inflated compared to income because of a significant number of rental apartments in this area.

Table II

Demographic Data on Study Sample

	Boys <i>n</i> = 103		Girls <i>n</i> = 106		Total group <i>n</i> = 209	
	Mean	(<i>SD</i>)	Mean	(<i>SD</i>)	Mean	(<i>SD</i>)
Parent education level ^a	5.24	(0.98)	5.19	(1.04)	5.21	(1.01)
Full Scale IQ. ^b	108.02	(14.83)	110.41	(12.82)	109.23	(13.87)
Age (years) at time of testing:						
Kindergarten	5.86	(0.34)	5.72	(0.33)	5.79	(0.34)
Grade 1	6.84	(0.34)	6.69	(0.31)	6.77	(0.33)
Grade 2	7.85	(0.34)	7.70	(0.31)	7.77	(0.34)
Grade 3	8.72	(0.30)	8.60	(0.31)	8.66	(0.31)

^aParent education level was coded as follows: 4 = completed grade 12; 5 = completed 1 to 3 years of college; 6 = completed 4 year college; 7 = University /graduate school. Data were available for parents of 202 of the 209 participants; using the most recent information available. Mother's education level was used except for one household headed by a single-parent father.

^bBased on all four subtests of the Wechsler Abbreviated Scale of Intelligence (WASI), administered one time (Year 04), except for one child whose score was based on 2 subtests of the WASI administered in Year 03.

Table III
Number (and Percent)^a of Children Meeting “More Restrictive” Criteria for Math Disability (MD)

MD Criteria	Grade in school (<i>N</i> = 209) ^a			
	Kindergarten	First ^b	Second ^c	Third ^d
I.				
TEMA-2 scores				
TEMA-2 score < 86	24 (11%)	9 (4%)	6 (3%)	11 (5%)
TEMA-2 score < 10th %ile	20 (10%)	19 (9%)	20 (10%)	17 (8%)
II.				
WJ-R Calc score < 86	89/199 (45%)	17 (8%)	n/a	26 (12%)
III.				
KeyMath-Revised subtests < 7;				
Numeration	9/208 (4%)	3 (1%)	0 (0%)	n/a
Geometry	5/207 (2%)	2 (1%)	1 (<1%)	n/a
Addition	5/169 (3%)	4 (2%)	2 (1%)	n/a
Measurement	4/207 (2%)	3 (1%)	0 (0%)	n/a
IV.				
Discrepancy based scores				
SB-IV minus TEMA>14	16 (8%)	5 (2%)	n/a	n/a
SB-IV minus WJ-R>14	51/199 (26%)	17 (8%)	n/a	n/a
WASI minus TEMA>14	n/a	n/a	20 (10%)	13 (6%)
WASI minus WJ-R>14	n/a	n/a	n/a	44 (21%)

^a*N* = 209, unless otherwise indicated. *N* < 209 in cases of missing data.

^bIncludes 2 children who repeated kindergarten.

^cIncludes 8 first graders, 6 who repeated first grade.

^dIncludes 9 second graders, 1 who repeated second grade.

Number (and Percent)^a of Children Meeting “Less Restrictive” Criteria for Math Disability (MD)

MD Criteria	Grade in school (<i>N</i> = 209) ^a			
	Kindergarten	First ^b	Second ^c	Third ^d
I. TEMA-2 score < 95	65 (31%)	31 (15%)	37 (18%)	30 (14%)
II. WJ-R Calc score < 95	119 / 199 (60%)	35 (17%)	n/a	50 (24%)
III. KeyMath-Revised subtests < 8(=)				
Numeration	17/208 (8%)	6 (3%)	4 (2%)	n/a
Geometry	7/207 (3%)	3 (1%)	6 (3%)	n/a
Addition	15/169 (9%)	8 (4%)	4 (2%)	n/a
Measurement	6/207 (3%)	6 (3%)	2 (1%)	n/a

^a *N* = 209, unless otherwise indicated. *N* < 209 in cases of missing data, including tests not scored because of lack of age referenced norms.

^b Includes 2 children who repeated kindergarten.

^c Includes 8 first graders, 6 who repeated first grade.

^d Includes 9 second graders, 1 who repeated second grade.

Table V
Number of Children who Met Restrictive Criteria for MD^a between Kindergarten and Grade 3

Number of children who first met restrictive MD criteria ^a in Kindergarten or Grade 1, 2, or 3				
All grade(s) during which criteria were met	Starting in Kindergarten	Starting in Grade 1	Starting in Grade 2	Starting in Grade 3
During only one grade	7	1	4	1
During only two grades				
K,1	2	^b	—	—
K,2	2	—	—	—
K,3	1	—	—	—
1,2	—	1	—	—
1,3	—	1	—	—
2,3	—	—	1	—
Total	5	2	1	—
During three grades			—	—
K,1,2	1	—	—	—
K,1,3	2	—	—	—
K,2,3	0	—	—	—
1,2,3	—	6	—	—
Total	3	6	—	—
During all four grades				
K,1,2,3	5 (2 ^c)	—	—	—
Grade level totals				
Total to first emerge in given grade [<i>N</i> = 35]	20	9	5	1
Total for one grade only [<i>N</i> = 13]	7	1	4	1
Number “persistent” for (2 years [<i>N</i> = 22])	13/20 (65%)	8/9 (89%)	1/5 (20%)	—
Number “persistent” for (3 years [<i>N</i> = 14])	8/20 (40%)	6/9 (67%)	—	—
	8/13 (62%)	6/8 (75%)	—	—

^aMD criteria = TEMA-2 score < 10%ile for study sample.

^bIndicates a cell where a count is not applicable.

^cNumber in parentheses indicates the number of children with FSIQ < 80. This occurred only in the Kindergarten cell, for children meeting criteria for all four school years. One of these children repeated preschool. The other child attends a non-graded school for children with learning disabilities.

Table VI
Number (and Percent)^a of Children Meeting Specific Criteria for Poor Reading or Visual Spatial Performance

Criteria	Grade in school (<i>N</i> = 209) ^a			
	Kindergarten	First ^b	Second ^c	Third ^d
More Restrictive				
WJ-R LWID < 86	18/204 (9%)	5 (2%)	2 (1%)	4 (2%)
WJ-R WA < 86	9/197 (5%)	10 (5%)	3 (1%)	6 (3%)
RAN ^e RT > X+1.5 <i>SD</i>	11/202 (5%)	16/208 (8%)	15 (7%)	12 (6%)
DTVP-2 subtests < 7:				
Position in Space	13 (6%)	9 (4%)	36 (17%)	23 (11%)
Figure Ground	38 (18%)	32 (15%)	36 (17%)	24 (11%)
Visual Closure	74/208 (36%)	81 (39%)	65 (31%)	34 (16%)
Form Constancy	1/208 (<1%)	5 (2%)	2 (1%)	2 (1%)
Less Restrictive				
WJ-R LWID < 95	50/204 (25%)	17 (8%)	13 (6%)	26 (12%)
WJ-R WA < 95	66/197 (34%)	25 (12%)	13 (6%)	28 (13%)
RAN ^e RT > X+1 <i>SD</i>	19/202 (9%)	24/208 (12%)	30 (14%)	23 (11%)
DTVP-2 subtests < 8:				
Position in Space	29 (14%)	46 (22%)	62 (30%)	33 (16%)
Figure Ground	63 (30%)	52 (25%)	41 (20%)	28 (13%)
Visual Closure	118/208 (57%)	103 (49%)	71 (34%)	40 (19%)
Form Constancy	3/208 (1%)	15 (7%)	7 (3%)	5 (2%)

^a *N* = 209, unless otherwise indicated. *N* < 209 in cases of missing data.

^b Includes 2 children who repeated kindergarten.

^c Includes 8 first graders, 6 who repeated first grade.

^d Includes 9 second graders, 1 who repeated second grade.

^e RAN numbers subtest was used for all grades with the exception of Kindergarten where RAN colors subtest was used.

Table VII
Spearman Rho Correlation Values for TEMA-2, RAN, and DTVP-2 Subtest Scores

Variable Pairs	Grade Level			
	Kindergarten	Grade 1	Grade 2	Grade 3
TEMA-2, RAN colors	-.458 ^d	-.254 ^c	-.258 ^c	-.188 ^b
TEMA-2, RAN numbers	n/a	-.348 ^d	-.288 ^d	-.256 ^c
TEMA-2, DTVP-2 PS	.464 ^d	.397 ^d	.332 ^d	.313 ^d
TEMA-2, DTVP-2 FG	.176 ^a	.002 ^e	.213 ^b	.130 ^e
TEMA-2, DTVP-2 VC	.168 ^a	.271 ^d	.442 ^d	.316 ^d
TEMA-2, DTVP-2 FC	.130 ^e	.276 ^c	.327 ^d	.325 ^d
DTVP-2 PS, RAN	-.142 ^a	-.027 ^e	-.025 ^e	.001 ^e

^a $p < .05$;

^b $p < .01$;

^c $p < .001$;

^d $p < .0001$;

^e rho not statistically significant.

Table VIII
Frequency of MD, RD, and PSD as a Function of Grade and Gender

	MD-P ^a <i>n</i> = 22		MD all ^b <i>n</i> = 35		Non-MD ^c <i>n</i> = 174	
	Girls	Boys	Girls	Boys	Girls	Boys
Total Sample Number	7	15	14	21	92	82
Mean Age (<i>SD</i>)						
Kindergarten	5.47(0.41)	5.92 (0.46)	5.67 (0.42)	5.90 (0.42)	5.72 (0.31)	5.85 (0.32)
FSIQ Grade 3						
Mean	93.57	89.00	97.64	93.52	112.35	111.73
(<i>SD</i>)	(10.55)	(9.99)	(8.76)	(11.48)	(12.23)	(13.27)
Number (%) with RD ^d						
Kindergarten	1 (17%)	3 (23%)	1 (9%)	5 (26%)	3 (3%)	2 (3%)
Grade 1	1 (17%)	8 (53%)	3 (23%)	8 (38%)	2 (2%)	3 (4%)
Grade 2	1 (14%)	6 (40%)	2 (14%)	7 (33%)	3 (3%)	3 (4%)
Grade 3	0 (0%)	4 (27%)	1 (7%)	4 (19%)	1 (1%)	6 (7%)
Number (%) with PSD ^e						
Kindergarten	0 (0%)	3 (20%)	2 (14%)	4 (19%)	3 (3%)	4 (5%)
Grade 1	0 (0%)	4 (27%)	1 (7%)	5 (24%)	3 (3%)	0 (0%)
Grade 2	1 (14%)	7 (47%)	2 (14%)	8 (38%)	11 (12%)	15 (18%)
Grade 3	4 (57%)	5 (33%)	5 (36%)	5 (24%)	8 (9%)	5 (6%)
Number (%) with RD+PSD						
Kindergarten	0 (0%)	1 (7%)	0 (0%)	1 (5%)	0 (0%)	0 (0%)
Grade 1	0 (0%)	2 (13%)	1 (7%)	2 (10%)	0 (0%)	0 (0%)
Grade 2	0 (0%)	2 (13%)	0 (0%)	2 (10%)	0 (0%)	2 (2%)
Grade 3	0 (0%)	1 (7%)	0 (0%)	1 (5%)	0 (0%)	0 (0%)
Number (%) with FSIQ<80	0 (0%)	2 (13%)	0 (0%)	2 (10%)	0 (0%)	1 (1%)

^aMD-P = TEMA-2 < 10% for 2 or more years during grades K-3.

^bMD all = TEMA-2 < 10% at least one year. Includes MD-P children.

^cNon MD = MD TEMA-2 criteria were not met during any year (grades K-3).

^dRD = RAN response time > 1.5 SD from overall group mean (*N* = 209).

^ePSD = DTVP-2 position in space standard score < 7.

Table IX
Number of Children who Met Discrepancy Criteria between Kindergarten and Grade 3

All grade(s) during which criteria were met	Number of children who first met discrepancy criteria ^a in Kindergarten or Grade 1, 2, or 3			
	Starting in Kindergarten	Starting in Grade 1	Starting in Grade 2	Starting in Grade 3
During only one grade	10	2	15	9
During only two grades				
K,1	1	^b	–	–
K,2	1	–	–	–
K,3	2	–	–	–
2,3 ^c	–	–	2	–
Total	4	–	2	–
During three grades				
K,1,2	2	–	–	–
Total	2	–	–	–
Grade level totals				
Total to first emerge in given grade [<i>N</i> = 44]	16	2	17	9
Total for one grade only [<i>N</i> = 36]	10	2	15	9
Number “persistent” for 2 ≥ years [<i>N</i> = 8]	6/16 (38%)	–	2/17 (12%)	–
Number “persistent” for ≥ 3 years [<i>N</i> = 2]	2/16 (13%)	–	–	–
	2/6 (33%)	0	–	–

^aDiscrepancy criteria = SBIV minus TEMA-2 score > 14 for Kindergarten and Grade 1, WAS1 minus TEMA-2 score > 14 for Grade 2 and Grade 3.

^bIndicates a cell where a count is not applicable.

^cNo discrepancy criteria was met for the following grade combinations: 1,2 - 1,3 - K,1,3 - K,2,3 - 1,2,3 - K,1,2,3.