Journal of Experimental Child Psychology **77,** 236–263 (2000)
doi:10.1006/jecp.2000.2561, available online at http://www.idealibrary.com on IDEAL®

# Numerical and Arithmetical Cognition: A Longitudinal Study of Process and Concept Deficits in Children with Learning Disability

David C. Geary, Carmen O. Hamson, and Mary K. Hoard

*University of Missouri*

Based on the stability and level of performance on standard achievement tests in first and second grade (mean age in first grade = 82 months), children with IQ scores in the low-average to high-average range were classified as learning disabled (LD) in mathematics (MD), reading (RD), or both (MD/RD). These children ($n$ = 42), a group of children who showed variable achievement test performance across grades ($n$ = 16), and a control group of academically normal peers ($n$ = 35) were administered a series of experimental and psychometric tasks. The tasks assessed number comprehension and production skills, counting knowledge, arithmetic skills, working memory, the ease of activation of phonetic representations of words and numbers, and spatial abilities. The children with variable achievement test performance did not differ from the academically normal children in any cognitive domain, whereas the children in the LD groups showed specific patterns of cognitive deficit, above and beyond the influence of IQ. Discussion focuses on the similarities and differences across the groups of LD children. © 2000 Academic Press

*Key Words:* learning disabilities; mathematical disabilities; reading disabilities; number; counting; arithmetic.

Research on learning disabilities in mathematics (MD) has progressed more slowly than reading disabilities (RD) research. One impediment to research on MD is the complexity of the domain of mathematics and the resulting wide array of cognitive deficits that could contribute to this form of LD. One way to circumvent this impediment is to apply the theoretical models and experimental measures used to study mathematical development in academically normal children to the study of children with poor achievement in mathematics (Geary,

We thank Linda Coutts, Dave Brunda, Bill Williamson, Jean Casteel, Linda Jones, Becky Gregory, Nicole Harris, Aaron Stratman, Sheela Vishwanath, and Natalija Popovic for their assistance with various aspects of the study, and several anonymous reviewers for comments on an earlier draft. The research was supported by Grant HD27931 from the NICHD and a grant from the University of Missouri Research Board.

Address correspondence and reprint requests to David C. Geary, Department of Psychology, 210 McAlester Hall, University of Missouri, Columbia, MO 65211-2500. E-mail: GearyD@Missouri.edu.



0022-0965/00 $35.00

1990; Jordan & Montani, 1997; Siegler, 1988). The current study followed this approach and included a longitudinal component. In a previous longitudinal study, children with low mathematics achievement scores across consecutive grades were found to differ on several cognitive measures from normal children and children with low mathematics achievement scores in one grade but average or better scores in the next (Geary, 1990; Geary, Brown, & Samaranayake, 1991). The latter group did not differ from normal children on measures of numerical working memory or arithmetical cognition. In comparison to these two groups, the children with MD (i.e., consistently low mathematics achievement scores) showed shorter digit spans and an array of deficits and delays in arithmetical competencies.

In the current study, children were classified respectively as MD, RD, or MD and RD (MD/RD) if they showed low achievement scores in mathematics, reading, or both in first grade and in second grade and had low-average or higher IQ scores. These groups were contrasted with groups of children who showed variable achievement levels from one grade to the next or showed average or better achievement levels in both grades. The groups were contrasted on experimental measures derived from models of children's normal development in the number, counting, and arithmetic domains. The groups were also contrasted on measures that assess the basic cognitive systems—such as numerical memory span—that appear to support these early mathematical competencies (Geary, 1994; Siegler, 1996). The use of this approach enabled the study of MD within a well-developed theoretical framework and, at the same time, reduced the frequency of early false positives, that is, children with low mathematics achievement but no real cognitive deficit (Geary, 1990). The first section below provides a brief summary of relevant models in the number, counting, and arithmetic literatures, while the second focuses on several supporting cognitive systems (see Geary, Hoard, & Hamson, 1999, for further discussion).

## NUMERICAL AND ARITHMETICAL COGNITION

### *Number Production and Comprehension*

The comprehension and production of numbers requires the ability to process verbal (e.g., "forty-two") and Arabic number representations (e.g., "42"), as well as an understanding of the meaning of the processed numbers (e.g., that the 4 in 42 represents 4 sets of 10). Number processing is also dependent on the ability to transcode, or translate, numbers from one representation to another (e.g., "forty-two" to "42"; Seron & Fayol, 1994). The few studies that have been conducted suggest—but are not definitive—that the cognitive systems that support number production and comprehension are intact in children with MD, at least for the processing of simple numbers (Badian, 1983; Gross-Tsur, Manor, & Shalev, 1996). The current study is the first longitudinal assessment of these competencies in children with MD/RD or MD, and should help to clarify if these children show number processing deficits.

### *Counting Knowledge*

Counting knowledge appears to emerge from a combination of inherent (i.e., biologically based) and experiential factors (Briars & Siegler, 1984; Geary, 1995; Gelman & Gallistel, 1978). Early inherent constraints can be represented by Gelman and Gallistel's five implicit counting principles. These principles include one–one correspondence (one and only one word tag, such as "one" or "two," is assigned to each counted object), the stable order principle (the order of the word tags must be invariant across counted sets), the cardinality principle (the value of the final word tag represents the quantity of items in the counted set), the abstraction principle (objects of any kind can be collected together and counted), and the order-irrelevance principle (items within a given set can be tagged in any sequence). Children also make inductions about the basic characteristics of counting by observing standard counting behavior (Briars & Siegler, 1984; Fuson, 1988). These inductions result in beliefs that certain unessential features of counting are in fact essential (Briars & Siegler, 1984). These unessential features include adjacency (a consecutive count of contiguous objects) and start at an end (counting proceeds from left to right).

Previous studies have shown that first-grade children with MD/RD understand most of the essential features of counting, such as stable order, but consistently err on tasks that assess adjacency and order irrelevance (Geary, Bow-Thomas, & Yao, 1992). These results suggest that children with MD/RD believe that counting is constrained in such a way that counting procedures can be executed only in the standard way (e.g., from left to right), which, in turn, suggests that they do not fully understand counting concepts. Many of these children also fail to detect counting errors when the first, but not the last, item is counted twice, suggesting difficulties in retaining the "error notation" in working memory while monitoring the counting process (Geary et al., 1999). The current study is the first to assess the counting knowledge of children with MD/RD or MD beyond first grade.

### *Arithmetic*

When first learning to solve simple arithmetic problems (e.g., 5 + 3), children typically rely on their knowledge of counting and the associated procedures (Siegler & Shrager, 1984). These procedures are sometimes executed with the aid of fingers (finger counting) and sometimes without them (verbal counting). The two most commonly used counting procedures are termed min (or counting-on) and sum (or counting-all; Fuson, 1982; Groen & Parkman, 1972). The min procedure involves stating the larger (max) addend and then counting a number of times equal to the value of the smaller (min) addend, such as counting 5, 6, 7, 8 to solve 5 + 3. The sum procedure involves counting both addends starting from 1. Occasionally, children state the value of the smaller addend and then count the larger addend (the max procedure). The development of procedural competencies reflects a gradual shift from frequent use of the sum and max procedures to frequent use of min counting.

The use of procedures also appears to result in the development of memory representations of basic facts (Siegler & Shrager, 1984), which, in turn, support the use of memory-based problem-solving processes. With direct retrieval, children retrieve an arithmetic fact from long-term memory (Ashcraft & Battaglia, 1978). Decomposition involves reconstructing the answer based on the retrieval of a partial sum. For instance, 6 + 7 might be solved by retrieving the answer to 6 + 6 (i.e., 12) and then adding 1 to this partial sum. With the fingers strategy, children uplift a number of fingers corresponding to the addends and then state an answer without counting their fingers. The uplifted fingers appear to prompt retrieval of the answer. The use of retrieval-based processes is moderated by a confidence criterion that represents an internal standard against which the child gauges confidence in the correctness of the retrieved answer. Children with a rigorous criterion state only answers that they are certain are correct, whereas children with a lenient criterion state any retrieved answer, correct or not (Siegler, 1988).

The use of these models and experimental techniques to study children with MD/RD or MD has yielded a wealth of information on the basic deficits that appear to define these forms of LD (Barrouillet, Fayol, & Lathulière, 1997; Geary, 1990; Geary, Widaman, Little, & Cormier, 1987; Jordan, Levine, & Huttenlocher, 1995; Jordan & Montani, 1997; Ostad, 1997, 1998; Räsänen & Ahonen, 1995; Svenson & Broquist, 1975). As a group, MD/RD, and to a lesser degree MD, children commit more procedural errors and use developmentally immature procedures more frequently than do their normal peers. Many children with MD/RD or MD do not show the shift from procedural-based problem solving to memory-based problem solving that is typically found with normal children (Garnett & Fleischner, 1983; Geary et al., 1991; Jordan & Montani, 1997; Ostad, 1997), suggesting difficulties in storing or accessing arithmetic facts in or from long-term memory. The nature of this retrieval deficit is, however, currently debated.

Geary (1993) suggested that the retrieval deficit is a reflection of a more general deficit in the ability to represent and retrieve information in phonetic and semantic memory. Barrouillet et al. (1997), in contrast, suggest that the retrieval deficit results from difficulties in inhibiting irrelevant associations from entering working memory. These irrelevant associations reduce the functional capacity of working memory and compete with the actual answer for expression and thus can increase error rates and disrupt reactions times (RTs; Conway & Engle, 1994). The current study assessed the predictions of Geary (1993) and Barrouillet et al. (1997). If Geary's prediction is correct, then children with MD/RD or RD should show retrieval deficits when solving arithmetic problems and differ from normal children on the articulation speed measures described below. If Barrouillet et al. are correct, then when children with MD/RD or MD retrieve facts from long-term memory, their errors should reflect frequent intrusions of related associations. For simple addition problems, counting-string associates of the addends might be one form of intrusion for children of this age (Siegler & Robinson, 1982); for instance, the counting-string associates for 5 + 2 would be 6 and 3, respectively, the numbers that follow 5 and 2 in the counting sequence. Of course, some children might show both forms of retrieval deficit.

## SUPPORTING COGNITIVE COMPETENCIES

Working memory, the ease of accessing information from long-term memory, and spatial abilities have been identified as potentially underlying the just-described deficits of children with MD/RD or MD (Geary, 1993; Rourke, 1993).

### *Working Memory*

Siegel and Ryan (1989) showed that children with MD exhibited deficits on span tasks that involved counting but not those involving other forms of language. Hitch and McAuley (1991) confirmed these findings and went on to show that MD was associated with a slow counting speed and difficulties in retaining numbers in working memory during the act of counting. Geary and his colleagues showed that children with MD/RD had shorter numerical digit spans than did their normal peers and had difficulties retaining information in working memory while monitoring the act of counting (Geary et al., 1991, 1999). The current study provided a contrast of the digit spans of children with MD/RD or MD and extended the assessment to children with RD only.

### *Long-Term Memory Retrieval*

If difficulties in representing or accessing information in or from the phonetic and semantic memory systems contribute to the fact retrieval deficits of children with MD/RD, then these children should also show deficits on measures that assess skill at accessing other types of information, such as words, from long-term memory (Geary, 1993). In fact, Geary argued that the comorbidity of MD and RD was related, in part, to difficulties in accessing both words and arithmetic facts from long-term memory and to poor phonetic representations of familiar words and numbers. In this view, a single retrieval deficit might contribute to MD, RD, and their comorbidity (i.e., to MD/RD).

One method that can be used to assess the ease of accessing information from long-term memory is to compare children on tasks that involve articulating familiar words (e.g., numbers) and unfamiliar nonwords (Denckla & Rudel, 1976; Fawcett & Nicolson, 1994; Gathercole & Adams, 1994; Johnston & Anderson, 1998). By definition, familiar words are represented in long-term memory and access to these representations appears to facilitate the encoding of these words into working memory and speed their articulation (Wagner & Torgesen, 1987). Although segments (e.g., familiar sounds) of many nonwords may be represented in long-term memory, the entire nonword will not, and thus there should be less of a long-term memory advantage for encoding nonwords into working memory or for the speed of articulating these words (see Gathercole

& Adams, 1994, for further discussion). As a result, familiar words should be articulated faster than unfamiliar nonwords, and differences in these articulation speeds might provide a useful means of assessing the ease with which information can be retrieved from long-term memory. In particular, this should provide an assessment of the degree to which phonetic representations of familiar words are activated when these words are encoded into working memory.

The prediction is that in comparison to children with MD/RD or RD, normal children will show a speed advantage for articulating familiar words but not for articulating unfamiliar words (Gathercole & Adams, 1994), which, in turn, would suggest poor activation of phonetic representations in children with MD/RD or RD. If the retrieval deficit of children with MD is also related to difficulties in the activation of phonetic representations, then children with MD should also show slow articulation of familiar words. In contrast, if the source of the retrieval deficit of children with MD is related to other mechanisms, such as poor inhibition of irrelevant associations, then children with MD should not show slow articulation of familiar words.

### *Spatial Abilities*

To determine if children with MD/RD or MD who show counting-procedure and retrieval deficits also show disrupted spatial abilities, the Mazes subtest of the WISC-III was administered. This particular measure was included based on Rourke's (1993) finding that performance on the Mazes test is sensitive to the spatial-related deficits of children with LD.

## CURRENT STUDY

The current study defined groups of children as RD, MD, MD/RD, variable (i.e., low academic scores in one grade but average or better scores in the other), and normal. These groups were then compared on tasks that assess basic number, counting, and arithmetical competencies and on tasks that assess several of the more basic cognitive systems, such as working memory, that appear to contribute to skill development in each of these basic mathematical domains. The study thus provided a comprehensive assessment of the basic quantitative competencies of children with LD.

## METHOD

### *Participants*

The 84 participants were selected from a larger group of 114 children (50 boys, 64 girls; mean age = 82 months in first grade) from five elementary schools in Columbia, Missouri. The schools were largely in working-class neighborhoods, but educated children from a variety of socioeconomic backgrounds. The initial selection of participants was based on teacher referral to the Chapter 1 Remedial Education Program and an equal number of children who were not participating in this program but who were in the same regular classrooms as the children participating in Chapter 1. Children participating in the Chapter 1 program received individual phonetics-based instruction in reading several times a week but no individual instruction in mathematics. Information is not available on the instructional approaches used in the regular classrooms.

The 84 participants were chosen so as to exclude children with IQ scores less than 80 or greater than 120 ($n = 23$) and children who did not complete testing in both first and second grade ($n = 7$; most had moved out of district between testing). This exclusionary criterion was used so that the LD groups and the control group included only children in the low-normal to high-normal IQ range, following previous studies (Ackerman & Dykman, 1995; Gross-Tsur et al., 1996). The remaining children were classified based on the pattern and stability of their achievement scores, as described below. Of these children, 29 were Black, 48 were White, and 6 were in other ethnic categories (e.g., Asian; racial information was not available for one child); there were few ethnic differences on the experimental measures, once IQ was covaried. Forty children participated in the Chapter 1 program for at least one academic year, and 29 of these children were classified into one of the three LD groups described below. Thirty-one of the 43 children not receiving Chapter 1 services were classified as academically normal, based on achievement test scores, and the remaining children were classified into one of the three LD groups or the variable group (information on Chapter 1 status was not available for one child).

*Classification scheme.* All children were administered the Vocabulary and Block Design subtests of the Wechsler Intelligence Scale for Children-III (Wechsler, 1991) or their counterparts from the Stanford–Binet Intelligence Scale (Thorndike, Hagen, & Sattler, 1986). The first cohort was administered the WISC-III, but psychologists in the local school requested that we not use this test because it is used in their special education assessments. Thus, the second cohort was administered the Stanford–Binet, based on the high correlation between performance on this IQ test and the WISC-III (Sattler, 1982). An IQ score was estimated based on performance on these two subtests, as prescribed by Sattler. The children were also administered the Mathematics Reasoning subtest of the Wechsler Individual Achievement Test (Wechsler, 1992) and the Word Attack and the Letter–Word Identification subtests of the Woodcock–Johnson Psycho-Educational Battery–Revised (Woodcock & Johnson, 1989, 1990). The Mathematics Reasoning subtest assesses basic arithmetic skills such as counting and subtraction, as well as some more advanced skills such as graph reading and time telling. The Letter–Word Identification subtest assesses the ability to understand that written symbols or groups of symbols (e.g., icons, letters) represent objects, letters, or words, while the Word Attack subtest assesses the ability to apply the rules involved in the pronunciation of written words (e.g., sounding out non-words).

Children with Mathematics Reasoning scores less than the 35th percentile in both first and second grade were classified as MD and those children with Word

TABLE 1
Mean Achievement and Estimated Intelligence Scores

| | | Estimated IQ | | Mathematics reasoning | | Word attack | | Letter–word identification | |
|---|---|---|---|---|---|---|---|---|---|
| Group | *N* | *M* | *SD* | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| | | | | Grade 1 | | | | | |
| MD and RD | 16 | 90 | 8 | 17 | 8 | 11 | 7 | 19 | 12 |
| MD | 12 | 94 | 7 | 21 | 9 | 40 | 18 | 41 | 21 |
| RD | 14 | 95 | 9 | 47 | 26 | 15 | 8 | 29 | 20 |
| Variable | 16 | 100 | 11 | 42 | 23 | 37 | 20 | 53 | 22 |
| Normal | 26 | 104 | 10 | 69 | 18 | 67 | 18 | 72 | 23 |
| | | | | Grade 2 | | | | | |
| MD and RD | | | | 15 | 10 | 14 | 8 | 18 | 11 |
| MD | | | | 23 | 10 | 56 | 18 | 56 | 24 |
| RD | | | | 54 | 17 | 21 | 11 | 26 | 14 |
| Variable | | | | 56 | 16 | 56 | 20 | 70 | 23 |
| Normal | | | | 69 | 15 | 71 | 20 | 78 | 20 |

*Note.* Achievement test scores are based on national percentile rankings, based on age.

Attack scores less than the 35th percentile in both first and second grade were classified as RD. A cutoff of the 35th percentile is higher than is normally used in LD research (25th percentile), but was necessary to ensure adequate sample sizes. As noted below, despite the more lenient classification criterion, the children classified as LD showed lower performance on the achievement tests than would be expected based on their IQ and thus do have some form of learning deficit. Word Attack rather than Letter–Word Identification was used because the former test provides a better assessment, at this age, of the basic phonological decoding skills that are often impaired in children with RD (e.g., Morris et al., 1998).

On the basis of these criteria, five groups were identified. As shown in Table 1, the first group was classified as MD and RD (MD/RD) and consisted of 16 children (8 boys, 8 girls; mean age = 81 months in first grade) with mean achievement scores below the 20th percentile for mathematics and reading skills in first and second grade. The second group consisted of 12 children (7 boys, 5 girls; mean age = 83 months) with a mean mathematics achievement score below the 25th percentile in both grades but mean Word Attack and Letter–Word Identification scores at or above the 40th percentile. These children were classified as MD only (MD). The RD only (RD) group consisted of 14 children (4 boys, 10 girls; mean age = 82 months) with mean Word Attack scores below the 22nd percentile in both grades and mean mathematics achievement scores above the 46th percentile. The fourth group consisted of 26 academically normal children (12 boys, 14 girls; mean age = 81 months) with mean achievement scores above the 66th percentile in both first and second grades. The final group included 16 children (7 boys, 9 girls, mean age = 82 months). These children had at least one achievement score below the 35th percentile in one grade and a score on the same test that was above the 34th percentile in the other grade. On the basis of this pattern, these children were termed the variable group. As shown in Table 1, the general pattern was for these children to show improved achievement test performance from Grade 1 to Grade 2, as was found by Geary (1990; Geary et al., 1991).

At the time of testing, the mean ages did not differ across groups in either first or second grade, $F$s $< 1$, nor did the number of boys and girls in each group, $\chi^2(4) = 2.57$, $p > .25$. A one-way analysis of variance (ANOVA) revealed significant group differences in IQ scores, $F(4, 79) = 6.97$, $p < .0001$. Follow-up comparisons using the HSD (honestly significant difference) procedure indicated that the mean IQ score of the normal group was significantly higher than that of the three LD groups but did not differ from the mean of the variable group ($p$s $< .05$). The only other significant difference was for the comparison of the MD/RD and variable groups. On the basis of this effect, IQ was used as a covariate in all analyses.

After partialing IQ scores, group comparisons (HSD) revealed that the children in the MD/RD and RD groups showed significantly lower Word Attack scores than did the children in the MD and normal groups in both grades and lower mean scores than the children in the variable group in second grade ($p$s $< .05$); the same pattern was found for the Letter–Word Identification scores in second grade. The mean Word Attack scores of the children with MD/RD or RD did not differ in either grade ($p$s $> .05$). In both grades, the MD/RD and MD groups showed significantly lower Mathematics Reasoning scores than the RD and normal groups and lower scores than the variable group in second grade ($p$s $<$ .05). The mean Mathematics Reasoning scores of the children with MD/RD or MD did not differ in either grade ($p$s $> .05$). Finally, 2 (test, Mathematics Reasoning and Word Attack) by 2 (grade) repeated measures ANOVAs were conducted for each group and revealed no significant test, grade, or interaction effects for the MD/RD and normal groups ($p$s $> .10$). The children in the MD group showed significantly higher Word Attack than Mathematics Reasoning scores, $F(1, 11) = 25.13$, $p < .001$, whereas the children in the RD group showed the opposite pattern, $F(1, 13) = 43.26$, $p < .001$; no other effects were significant for these groups ($p$s $> .05$). The children in the variable group showed improved performance from first to second grade, $F(1, 15) = 21.01$, $p < .001$, but the test and test by grade effects were not significant ($p$s $> .50$).

The overall pattern indicates that the children in the MD/RD group showed achievement scores in mathematics and basic reading skills that were lower than what would be expected on the basis of their IQ scores, whereas the children in the MD and RD groups showed specific deficits in mathematics and reading, respectively. The children in the variable group showed, on average, improved performance from first to second grade in both mathematics and reading and by second grade did not differ from the normal children.

### *Experimental Tasks and Procedures*

*Number production and comprehension.* The items were developed based on an adaptation of the Johns Hopkins Dyscalculia Battery for children (McCloskey, Aliminosa, & Macaruso, 1991; Shalev, Manor, & Gross-Tsur, 1993) and assess the ability to name and reproduce visually and auditorily presented numbers and to compare the magnitude of visually and auditorily presented numbers.

In the first number recognition/production task, the experimenter showed the child a series of four integers (3, 8, 5,12) arranged vertically on an otherwise blank sheet of paper in a large font. The child was then asked to name each number shown and write the number next to the printed copy. Next, the child was dictated a series of four integers (2, 5, 7, 13), one at a time, and asked to write the number on the same paper used in the visual presentation task. In the first magnitude comparison task, the child was shown a series of four pairs of single-digit integers (1 9, 3 2, 5 7, 9 8) arranged vertically on an otherwise blank sheet of paper in a large font. The child was asked to decide, one pair at a time, "Which is bigger, which is more?". The task was then repeated with a new set of four pairs of single-digit integers (6 3, 2 5, 5 6, 8 1) that were presented by dictation. Both sets of magnitude comparison pairs included numbers that represented relatively small (e.g., 2) and large (e.g., 9) magnitudes, and relatively small (e.g., 5 7) and large (e.g., 1 9) distances between the two magnitudes.

These measures provided four items for each of five basic number processing competencies: number naming (Arabic to verbal transcoding), number writing with visual presentation, number writing with auditory presentation (verbal to Arabic transcoding), magnitude comparison with visual presentation, and magnitude comparison with auditory presentation. The last two of these tasks assess children's understanding of the underlying quantities.

*Counting knowledge.* The child was first introduced to a puppet who was just learning how to count and therefore needed assistance to know if his counting was "OK and right, or not OK and wrong." During each of the 13 trials, a row of 5, 7, or 9 poker chips of alternating color (e.g., red, blue, red, blue, red) were aligned behind a screen. The screen was then removed, and the puppet counted the chips. The child was then queried on the correctness of the counting. The experimenter recorded whether the child stated the puppet's count was "OK, or not OK and wrong."

Following previous studies, four types of counting trials were administered (Briars & Siegler, 1984; Geary et al., 1992; Gelman & Meck, 1983): correct, right–left, pseudo, and error. For correct trials, the chips were counted sequentially and correctly, from the child's left to the child's right. Right–left involved counting the chips sequentially and correctly, but from the child's right to the child's left. For pseudo trials, the chips were counted correctly from left to right, but first one color was counted, and then, returning to the left-hand side of the row, the count continued with the other color. For error trials, the chips were counted sequentially from left to right, but the first chip was counted twice. Each counting trial type occurred once for each array size (i.e., 5, 7, 9), with one additional pseudo trial (for 7 chips).

*Addition strategy assessment.* Simple addition problems were presented at the center of the monochrome monitor of an IBM PS/2 Model 30 microcomputer. For each problem, a READY prompt appeared at the center of the screen for 1000 ms, followed by a 1000-ms period during which the screen was blank. Then, an addition problem appeared on the screen and remained until the child responded. Answers were spoken into a microphone that activated a voice-operated relay that was interfaced with the microcomputer. RTs were recorded using a hard card timing mechanism with ±1.0-ms accuracy. The experimenter initiated each problem presentation sequence via a control key.

The stimuli were 14 single-digit addition problems presented horizontally (4 + 5 = ). The problems consisted of the integers 2 through 9, with the constraint that the same two integers (e.g., 2 + 2, 4 + 4) were never used in the same problem, as these appear to be solved differently than other problems (Groen & Parkman, 1972). Across stimuli, each digit was presented two to four times, and half of the problems summed to 10 or less, inclusive. For half of the problems, the larger valued integer was presented in the first position, and for the remaining problems the smaller valued integer was presented in the first position. The order of problem presentation was determined randomly, with the constraint that no integer was presented in the same position across consecutive trials.

Following three practice problems, the 14 experimental problems were presented one at a time. The child was asked to solve each problem as quickly as possible without making too many mistakes. It was emphasized that the child could use whatever strategy was easiest for her to get the answer, and the child was instructed that as soon as she had the answer, she was to speak it into the microphone. Based on the child's answer and the experimenter's observations, the trial was classified into one of five strategies, that is, counting fingers, fingers, verbal counting, retrieval, or decomposition (Siegler & Shrager, 1984). Counting trials were further classified based on where counting began, that is, min, sum, or max.

During problem solving, the experimenter watched for physical indications of counting, such as regular movements (e.g., fingers, mouth). For these trials, the experimenter initially classified the strategy as finger counting or verbal counting, depending on whether or not the child used his or her fingers to count. If the child upraised a number of fingers to represent the addends and then stated an answer without counting them, then the trial was initially classified as fingers. If the child spoke the answer quickly, without hesitation, and without obvious counting-related movements, then the trial was initially classified as retrieval.

After the child had spoken the answer, the experimenter queried the child on how he or she got the answer. If the child's response (e.g., "just knew it") differed from the experimenter's observations (e.g., saw the child mouthing counting), then a notation indicating disagreement between the child and the experimenter was made. If counting was overt, then the experimenter classified it as a counting strategy. If the trial was ambiguous, then the child's response was recorded as the strategy. Previous studies indicate that this method provides a useful measure of children's trial-by-trial strategy choices (e.g., Siegler, 1987). In this study, agreement between the child's description and the experimenter's observation was found for more than 94% of trials.

For the second-grade assessment, 14 additional problems were added to the task. These problems were the reverse of the 14 original problems (e.g., 6 + 3, 3 + 6) and were administered using the same procedures as described above and immediately following the 14 original problems. The only change was that the children were instructed to solve the 14 new problems only by means of retrieval. They were instructed to try to remember the answer and were instructed not to count or use any other type of strategy, other than remembering; if they could not remember, they were told that it was OK to guess. These were thus called retrieval-only problems (Jordan & Montani, 1997; Siegler & Shrager, 1984). These problems were added to increase the number of trials on which children used retrieval for problem solving, which, in turn, enabled a more complete assessment of the arithmetic-fact retrieval characteristics of children with LD.

*Articulation speed.* The stimuli chosen for this task were based on Cowan's (1986) and Edwards's (1978) list of monosyllabic English words and nonwords and included triads of one-digit number names (2, 9, 5), common one-syllable words (school–tree–cake), and one-syllable nonwords (lote, dake, pog). Each word triad was presented by dictation. The child was encouraged to repeat the triad until she could remember all three words. Next, the child was asked to say the triad as quickly as possible, two times in a row. A stopwatch was used to measure articulation speed. This procedure was performed three times for each triad.

*Digit span and mazes.* To assess numerical working memory and basic spatial skills (Rourke, 1993), the Digit Span and Mazes subtests of the WISC-III (Wechsler, 1991) were administered. Both tests were administered and scored according to standard procedures. For Digit Span, both the forward and the backward sections were administered, although performance on the backward measure is of greater theoretical interest. This is because backward digit span requires the retention and manipulation of numbers in working memory, skills that appear to be deficient in many children with MD/RD or MD (Hitch & McAuley, 1991).

## *Procedures*

All children were tested twice during each academic year (a total of four times), once in the fall and once in the spring. Fall testing included the just-described experimental tasks along with Digit Span, whereas the spring testing included the IQ (first grade only) and achievement tests, as well as Mazes in second grade. The achievement and ability tests were administered in the spring, because previous studies suggest that achievement test scores obtained at the end of the academic year are more predictive of achievement in later grades than are scores obtained at the beginning of the year (Geary, 1990). Because of near-ceiling performance in first grade, the number writing with visual and auditory presentation tasks were not administered in second grade.

For fall testing, the experimental tasks were administered in a randomly determined order for each child. For spring testing, the tests were administered in the following order: Letter–Word Identification, Word Attack, and Mathematics Reasoning, followed by the Vocabulary and Block Design subtests of the WISC-III or Stanford–Binet in first grade and the Mazes subtest in second grade. Testing times were about 40 min and 30 min for the fall and spring assessments, respectively. Mean time between testing periods was 125 days ($SD = 31$) for first grade and 160 days ($SD = 25$) for second grade. Time elapsed between testing periods did not differ significantly across groups in either grade, $F$s $< 1$.

## RESULTS

The results are presented in five sections. The first presents results for the number naming, number writing, and magnitude comparison tasks. The second and third sections present results for the counting and strategy assessment tasks, respectively, whereas results for the articulation speed measures are presented in the fourth section. The final section presents results for the Digit Span and Mazes subtests.

### Number Production and Comprehension

The mean percentage of correct responses for the number naming and number writing tasks ranged between 90 and 100% in both grades. The only exception was for the performance of children with MD/RD on the number naming (81% correct) and number writing with auditory presentation (83% correct) tasks in first grade. Because of these ceiling and near-ceiling effects, parametric statistics were not employed. Rather, the first-grade performance of each of the 16 children with MD/RD was examined for these two tasks and revealed that the poor performance was largely the result of difficulties in processing the numbers 12 and 13. Ten of the 16 children with MD/RD could not name the number 12 and 7 of these children could not write the number 13 from dictation.

The mean percentages of correct comparisons for the magnitude comparison tasks were also consistently high, with the mean percent correct ranging from 72 to 100 across groups. The only consistent group difference was a relatively low percentage of correct comparisons for the children with MD/RD for visual (72% correct; 83 to 100% correct for the remaining groups) and auditory (83% correct; 93 to 100% correct for the remaining groups) presentation in first grade. Al-

TABLE 2
Mean Percentage of Correct Identification across Counting Tasks

| Group | Counting task: Correct | Right–left | Pseudo | Error |
|---|---|---|---|---|
| | Grade 1 | | | |
| MD and RD | 98 | 89 | 30 | 67 |
| MD | 100 | 100 | 42 | 64 |
| RD | 97 | 100 | 65 | 95 |
| Variable | 98 | 98 | 63 | 75 |
| Normal | 100 | 91 | 71 | 96 |
| | Grade 2 | | | |
| MD and RD | 100 | 85 | 30 | 88 |
| MD | 100 | 86 | 16 | 86 |
| RD | 100 | 95 | 73 | 95 |
| Variable | 100 | 96 | 52 | 98 |
| Normal | 99 | 85 | 63 | 90 |

though the performance of the children with MD/RD improved to 91% correct for both auditory and visual presentation in second grade, this level of performance was still lower than the 98 to 100% correct for the four other groups. Examination of visual presentation trials in first grade revealed that performance of the children with MD/RD was at chance levels on the near-distance items (i.e., 2 3, 8 9), as 7 students identified the smaller number as being larger and 9 students identified the larger number as being larger. In second grade, all of the children with MD/RD correctly identified 3 as being larger than 2 but 6 of the 16 children identified 8 as being larger than 9. For the auditory presentation task, the errors of the children with MD/RD were distributed across all items in both grades, with no discernable pattern.

## Counting Knowledge

The mean percentage of correct identifications across the four types of counting trials is shown in Table 2. A 5 (group) by 4 (counting type) by 2 (grade) mixed ANCOVA, with group as a between-subjects factor, type and grade as within-subjects factors, and IQ as the covariate, revealed significant group, $F(4, 73) = 5.21$, $p < .001$, and group by type, $F(12, 219) = 2.92$, $p < .001$, effects; no other effects were significant ($p$s $> .05$). Follow-up ANCOVAs for each counting type revealed significant group differences for pseudo trials, $F(4, 74) = 3.66$, $p < .01$, and nonsignificant group effects for correct, right–left, and error trials ($p$s $> .25$); the grade by group interaction was nonsignificant for all counting types ($p$s $> .05$). Because this is the first study to examine the counting knowledge of distinct groups of LD children and thus no a priori contrasts could be specified, group differences were examined here and in later analyses by means of post hoc contrasts. Based on the pattern of pseudo-trial means, the performance of the MD/RD and MD groups was contrasted with that of the three remaining groups. The results confirmed that the mean performance of the MD/RD and MD groups was significantly lower than that of these three groups, $F(1, 77) = 13.67$, $p < .001$.

Examination of the performance of first graders revealed that 11 of the 16 children with MD/RD and 7 of the 12 children with MD indicated that pseudo counting was "not OK and wrong" on three ($n = 2$ and 1, respectively) or four ($n = 9$ and 6, respectively) of the pseudo-counting trials. Only four children with MD/RD and four children with MD indicated that pseudo counting was "OK and right" on all four trials. In second grade, 11 of the 16 children with MD/RD and 9 of the 12 children with MD indicated that pseudo counting was "not OK and wrong" on three ($n = 1$ and 1, respectively) or four ($n = 10$ and 8, respectively) trials. Only 4 children with MD/RD and 1 child with MD indicated that pseudo counting was "OK and right" on all four trials.

## Strategy Assessment

The analyses of the strategy choice data are discussed in two sections. The first presents results for the strategy choice trials, those in which the children were encouraged to use whatever strategy made it easiest for them to solve each problem. The second section focuses on the retrieval-only trials presented in second grade. In both sections, the percentage of trials on which each child used each strategy was calculated; the reported group means represent the mean of the individual percentages (strategy data analyses, however, were based on raw frequencies). The degrees of freedom differ across some of the analyses, because not all of the children used all strategies.

### *Strategy Choice Trials*

*Strategy choices.* The pattern shown in Table 3 is consistent with previous studies and indicates that children in all of the groups used similar strategies to solve simple addition problems but differed in their strategy mix (Geary, 1990). Due to the small number of fingers and decomposition trials, the analyses were restricted to the counting fingers, verbal counting, and retrieval strategies. The frequency with which each of these three strategies was used in problem solving was analyzed by means of 5 (group) by 2 (grade) mixed ANCOVAs. The results revealed significant grade by group effects for counting fingers, $F(4, 78) = 2.94$, $p < .05$, and retrieval, $F(4, 78) = 9.31$, $p < .001$, but no other significant effects ($p$s $> .05$).

For counting fingers, a post hoc contrast (MD/RD and MD groups vs the remaining groups)—based on the pattern of group means—confirmed that children in the RD, variable, and normal groups showed a significant decrease in

TABLE 3
Characteristics of Addition Strategies

| Strategy | Mean % of trials[a] | | | | | Mean % of errors | | | | | % of min count | | | | | Mean RTs (s)[b] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | MD/RD | MD | RD | Variable | Normal | MD/RD | MD | RD | Variable | Normal | MD/RD | MD | RD | Variable | Normal | MD/RD | MD | RD | Variable | Normal |
| Grade 1 | | | | | | | | | | | | | | | | | | | | |
| Counting fingers | 46 | 45 | 82 | 58 | 64 | 63 | 55 | 34 | 36 | 22 | 0 | 1 | 34 | 34 | 31 | 12.6 | 10.1 | 12.1 | 12.4 | 12.8 |
| Verbal counting | 4 | 13 | 5 | 22 | 24 | 95 | 35 | 0 | 21 | 29 | 0 | 67 | 33 | 45 | 45 | — | — | — | 6.9 | 10.4 |
| Retrieval | 44 | 34 | 4 | 13 | 4 | 91 | 66 | 100 | 39 | 9 | — | — | — | — | — | — | — | — | 3.8 | 3.8 |
| Grade 2 | | | | | | | | | | | | | | | | | | | | |
| Counting fingers | 47 | 42 | 22 | 26 | 14 | 39 | 4 | 14 | 14 | 11 | 31 | 64 | 89 | 67 | 93 | 8.4 | 9.2 | 5.6 | 5.6 | 5.7 |
| Verbal counting | 10 | 35 | 31 | 25 | 28 | 61 | 32 | 15 | 16 | 5 | 67 | 84 | 98 | 63 | 97 | — | 7.5 | 5.0 | 5.4 | 4.8 |
| Retrieval | 21 | 13 | 44 | 39 | 44 | 77 | 3 | 11 | 3 | 8 | — | — | — | — | — | — | 2.9 | 4.3 | 4.4 | 3.8 |

[a] Trials on which the chld used a mixed approach to problem solving are not included. As an example, the child might have begun a counting procedure and then retrieved an answer. Due to the small percentage of trials, fingers and decomposition are not shown. For these two strategies, all percentages of usage were ≤ 5%, except for the variable group's use of fingers in first grade (6%) and the normal group's use of decomposition in second grade (10%).

[b] All presented means are based on correct trials. Moreover, means are not presented for trials with a low frequency of correct trials. For instance, although the MD/RD children retrieved on 44% of the trials in Grade 1, more than 90% of the associated answers were errors. For Grade 1, the counting fingers means are based on sum trials and on min trials in Grade 2; all verbal counting means are based on min trials.

the use of finger counting from Grade 1 to Grade 2 but the children with MD/RD or MD did not, $F(1, 81) = 10.01$, $p < .005$. For retrieval, the same contrast confirmed that children in the RD, variable, and normal groups showed a significant increase in retrieval from Grade 1 to Grade 2, whereas children in the MD/RD and MD groups showed a significant decrease in retrieval, $F(1, 81) = 37.21$, $p < .001$.

Mixed ANCOVAs—5 (group) by 2 (grade)—were also used to analyze the percentage of errors and revealed significant group differences for counting fingers, $F(4, 26) = 3.08$, $p < .05$, and retrieval, $F(4, 20) = 9.47$, $p < .001$. For counting fingers, a post hoc contrast, based on the pattern of means, confirmed that the children in the MD/RD and MD groups committed more errors than did the children in the three remaining groups in first grade, $F(1, 66) = 7.05$, $p < .01$. For second grade, the only significant effect was for a contrast of the MD/RD group with the four remaining groups, with the former committing more errors, $F(1, 32) = 12.53$, $p < .002$. The analysis of the percentage of retrieval errors in first grade excluded the RD and normal groups, because of the low number of retrieval trials. A post hoc linear contrast (coded −1, 0, and 1, for the MD/RD, MD, and variable groups) confirmed that the percentage of retrieval errors increased linearly across the variable, MD, and MD/RD groups in first grade, $F(1, 21) = 5.39$, $p < .05$. For second grade, a contrast of the MD/RD group with the four remaining groups was significant, $F(1, 57) = 46.80$, $p < .001$.

Again, 5 (group) by 2 (grade) mixed ANCOVAs were used to analyze the percentage of counting trials on which the min procedure was used. The results revealed a significant group effect for counting fingers, $F(4, 26) = 4.85$, $p < .005$, and a significant grade effect for verbal counting, $F(1, 15) = 5.71$, $p < .05$. The latter effect was due to an increase in min counting from Grade 1 to Grade 2. Although the group effect for verbal counting was not statistically significant, the finding that the children with MD/RD never used the min procedure in first grade is noteworthy. For counting fingers, a contrast, based on the pattern of means, of the MD/RD and MD groups with the three remaining groups confirmed that the children in the two former groups used the min procedure much less frequently than the remaining children did in first grade, $F(1, 66) = 6.55$, $p < .02$. In fact, the children with MD/RD never used the min procedure in first grade and the children with MD rarely used this procedure. For second grade, the children with MD/RD used the min procedure less frequently than did the children in the remaining groups, $F(1, 32) = 9.91$, $p < .005$.

Across groups and grade level, counting fingers was the only strategy to yield enough correct responses to conduct a meaningful analysis of the associated RTs, although sum trials were analyzed in Grade 1 and min trials in Grade 2. There were no significant group differences in Grade 1 RTs, $F(4, 60) < 1$, but the differences were significant in Grade 2, $F(4, 64) = 4.62$, $p < .005$. A follow-up contrast confirmed that the mean RTs of the MD/RD and MD groups

TABLE 4
Addition Performance for Retrieval-Only Trials

| Group | Retrieval trial % | Retrieval errors % | Counting-string associates: % of retrieval trials | Counting-string associates: % of retrieval errors | Retrieval RTs (s): All trials | Retrieval RTs (s): Correct trials |
|---|---|---|---|---|---|---|
| MD and RD | 88 | 81 | 23 | 29 | 3.5 | 5.7 |
| MD | 79 | 55 | 9 | 17 | 4.6 | 4.7 |
| RD | 90 | 39 | 8 | 21 | 4.2 | 4.0 |
| Variable | 81 | 31 | 2 | 5 | 4.3 | 4.4 |
| Normal | 83 | 23 | 1 | 4 | 5.1 | 5.0 |

were higher than those of the three remaining groups, $F(1, 67) = 18.67$, $p < .001$.

## *Retrieval-Only Trials*

*Strategy characteristics.* Table 4 presents the strategic and RT results for retrieval-only trials. In keeping with the instructions, the majority of problems were solved by means of retrieval, the mean frequency of which did not differ across groups, $F(4, 78) = 1.86$, $p > .15$. There were, however, significant group differences in the percentage of retrieval errors, $F(4, 76) = 10.71$, $p < .001$. A set of post hoc orthogonal contrasts—based on the pattern of group means—indicated that the children in the MD/RD group committed more retrieval errors that did the children in the four remaining groups, $F(1, 78) = 33.78$, $p < .001$, and that the children in the MD and RD groups committed more errors than did the children in the variable and normal groups, $F(1, 78) = 6.25$, $p < .02$.

The most common retrieval error was a counting-string associate of one of the addends (e.g., stating 7 for 3 + 6; Siegler & Shrager, 1984). In fact, 23% of the responses of the children with MD/RD were counting-string associates and 29% of their errors were counting-string associates. *Z* tests confirmed that the children with MD/RD stated more counting-string associates, across all retrieval trials, than did the children in the four remaining groups ($p$s $< .05$), whereas the children in the MD and RD groups stated more counting-string associates than did the children in the variable and normal groups ($p$s $< .05$). In terms of counting-string associates as a percentage of retrieval errors, the MD/RD, MD, and RD groups did not differ from one another ($p$s $> .06$) but the children in these groups committed more counting-string errors than did the children in the variable and normal groups ($p$s $< .05$), who, in turn, did not differ from one another ($p > .05$).

*Reaction times.* Mean reaction times for all retrieval trials (correct and incorrect) and for correct trials are shown in the rightmost columns of Table 4. There

TABLE 5
Mean Articulation Speeds in Seconds

| | Stimulus type | | | | | |
|---|---|---|---|---|---|---|
| | Number | | Word | | Nonword | |
| Group | *M* | *SD* | *M* | *SD* | *M* | *SD* |
| | | | Grade 1 | | | |
| MD and RD | 2.1 | 0.6 | 4.3 | 1.6 | 3.0 | 1.0 |
| MD | 1.8 | 0.4 | 3.8 | 1.4 | 3.0 | 1.0 |
| RD | 2.2 | 0.7 | 4.1 | 1.2 | 3.3 | 1.2 |
| Variable | 1.7 | 0.3 | 3.2 | 0.6 | 2.8 | 1.3 |
| Normal | 1.7 | 0.4 | 3.3 | 1.1 | 2.4 | 0.7 |
| | | | Grade 2 | | | |
| MD and RD | 1.8 | 0.3 | 3.1 | 1.2 | 2.7 | 0.9 |
| MD | 1.7 | 0.2 | 3.0 | 0.7 | 3.1 | 0.8 |
| RD | 1.7 | 0.4 | 2.9 | 0.7 | 2.6 | 0.9 |
| Variable | 1.7 | 0.4 | 2.9 | 0.7 | 2.8 | 1.0 |
| Normal | 1.6 | 0.3 | 2.5 | 0.6 | 2.4 | 0.9 |

were no significant group differences for correct trials, $F(4, 57) = 1.71$, $p > .15$, but there were significant differences for all trials, $F(4, 65) = 3.97$, $p < .01$. The group difference for all trials remained significant after partialing error frequency, $F(4, 64) = 12.85$, $p < .001$, and comparisons of group means (HSD) indicated that the mean for the MD/RD group was lower (i.e., faster) than the means for the four remaining groups ($p$s $< .05$). The pattern across all and correct trials suggests a speed–accuracy trade-off for the children with MD/RD. Although the correlation between error frequency and RTs, $r(12) = -.35$, $p > .20$, was not statistically significant for the MD/RD group, it is consistent with some degree of speed–accuracy trade-off. For the four remaining groups, the correlation between error frequency and RTs was positive and ranged between .42 ($p > .10$) and .87 ($p < .001$). The positive correlations indicate that as problems became more difficult (i.e., error prone), the time required to retrieve an answer, correct or not, increased.

## Articulation Speed

Mean articulation speeds are shown in Table 5. The basic question is whether children with poor reading achievement show slower articulation of familiar words than do other children. Regression techniques rather than ANCOVAs were used to address this issue, because these allow for an assessment of the magnitude of group differences in speed of articulating familiar words while controlling for speed of articulating unfamiliar words. Stated somewhat differently, this approach enabled an assessment of group differences in processes—presumably activation of phonetic representations stored in long-term memory—that are unique to familiar words, once more basic articulatory processes (measured by articulation of unfamiliar words) and IQ are controlled.

First, articulation speeds were standardized for each word type, with a mean of 0 and an *SD* of 1. Based on the earlier noted finding that poor reading achievement is often associated the slow articulation of familiar words (Denckla & Rudel, 1976; Fawcett & Nicolson, 1994; Morris et al., 1998), the standard scores for the word and number articulation speeds were summed to create a familiar-word articulation speed variable. Standard nonword articulation speeds, IQ, and a dummy coded group variable (coded 1 for the two groups with low reading achievement scores, that is, the MD/RD and RD groups, and coded 0 for the three remaining groups) were then regressed on the familiar-word articulation speed variable. The regression weight for the group variable thus provides an estimate for differences in the speed of articulating familiar words comparing the MD/RD and RD groups with the three remaining groups, once IQ and nonword articulation speeds have been statistically controlled.

The results confirmed slower familiar-word articulation speeds for the children with MD/RD or RD relative to the children in the three remaining groups in both first grade, $F(1, 73) = 3.50$, $p = .065$ ($\beta = .62$), and second grade, $F(1, 77) = 5.25$, $p < .05$ ($\beta = .83$). To assess whether children with MD also showed slow articulation of familiar words, the analyses were rerun comparing the MD group to the variable and normal groups. The results revealed nonsignificant group differences in first, $F(1, 48) < 1$, and second, $F(1, 48) = 1.30$, $p > .25$, grade, indicating that children with low achievement in mathematics and average achievement in reading do not show a disadvantage in speed of articulating familiar words.

## Digit Span and Mazes

Mean scores for performance on the forward and backward sections of the Digit Span subtest are shown in Table 6. A 5 (group) by 2 (type) by 2 (grade) mixed ANCOVA revealed significant group, $F(4, 77) = 4.84$, $p < .002$, and type, $F(1, 77) = 11.30$, $p < .002$, effects; except for IQ, $F(1, 77) = 6.06$, $p < .02$, no other effects were significant ($p > .10$). Of course, the significant type effect resulted from consistently higher performance on the forward than on the backward measure. A series of contrasts revealed that the group effect resulted from differences, across type and grade level, in the performance of the MD/RD and normal children, favoring the latter, $F(2, 79) = 9.14$, $p < .001$ (contrast coded 1 for the normal group, −1 for the MD/RD group, and 0 for all other groups).

An ANCOVA revealed no significant group differences on the Mazes subtest, ($p > .25$); mean standard scores were 8.3 ($SD = 2.4$), 8.9 ($SD = 2.5$), 9.1

TABLE 6
Mean Digit Span Scores

| Group | Forward | | Backward | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| | | Grade 1 | | |
| MD and RD | 4.0 | 0.9 | 1.6 | 1.2 |
| MD | 4.5 | 1.2 | 2.4 | 0.9 |
| RD | 4.3 | 0.6 | 2.4 | 1.2 |
| Variable | 4.4 | 0.8 | 2.9 | 0.6 |
| Normal | 4.8 | 0.8 | 3.1 | 0.5 |
| | | Grade 2 | | |
| MD and RD | 4.3 | 1.1 | 2.6 | 0.5 |
| MD | 5.0 | 1.0 | 2.9 | 1.0 |
| RD | 4.5 | 1.0 | 2.9 | 1.0 |
| Variable | 5.0 | 1.0 | 3.4 | 0.7 |
| Normal | 5.5 | 1.2 | 3.4 | 0.8 |

($SD$ = 3.2), 10.0 ($SD$ = 3.9), and 12.0 ($SD$ = 4.2) for the MD/RD, MD, RD, variable, and normal groups, respectively.

## DISCUSSION

The current study provided insights as to how the early mathematical development, and the underlying cognitive systems, of LD children differ from those of their normal peers and how the associated competencies may differ across LD groups. The first three sections below provide discussion of the number production and comprehension, counting knowledge, and arithmetic competencies of LD children, whereas the final section focuses on group differences in the supporting cognitive systems. However, before this discussion it should be noted that the failure to find differences between the variable and normal groups on the experimental measures is in keeping with previous findings (Geary, 1990; Geary et al., 1991). Specifically, it appears that children who show low achievement levels in one grade but average or better achievement levels in another grade (i.e., the variable group) do not have the cognitive deficits associated with MD/RD and MD. Rather, the intermittent academic difficulties of these children appear to be related to other factors (e.g., emotional difficulties).

### *Number Production and Comprehension*

In first grade, the children with MD/RD showed relatively low performance on the number naming and writing tasks, and relatively low but improving performance across grades on the magnitude comparison tasks. The results for the number naming and writing tasks suggest that early in first grade many children with MD/RD are not familiar with Arabic representations greater than 10, but this does not appear to reflect a cognitive deficit. Many of the children with MD/RD also showed difficulties in discriminating the value of adjacent numbers (e.g., 7 8). As a group, they showed some improvement in this area from first to second grade, specifically in the ability to discriminate smaller valued adjacent numbers (e.g., 2 3), but many of these children still had difficulties with larger valued sets (e.g., 8 9).

The results are not definitive but suggest that the processes involved in either the accessing of magnitude representations with visually presented Arabic numbers or the formation of boundaries between adjacent representations develops slowly in many children with MD/RD. If it is assumed that the boundaries between smaller valued numbers (e.g., 2 and 3) matures sooner than the boundaries between larger valued numbers (e.g., 8 and 9; Geary, 1994), then a developmental delay in the formation of these boundaries might explain the across-grade performance of the MD/RD group. Specifically, such a delay would result in difficulties in determining the larger of smaller valued (2 3) and larger valued (8 9) sets in first grade but only larger valued sets in second grade, as was found for the MD/RD group. This pattern, however, was not evident for the auditory presentation of number pairs, perhaps because the counting sequence is more accessible with the auditory presentation of numbers and thus can be used to make inferences about relative magnitude.

Either way, the number comprehension difficulties of the children with MD/RD were not evident in the children with MD, suggesting that any such difficulties are not a feature of MD per se. Rather, these apparent delays in basic number production and comprehension skills are found primarily in children with a more general learning disability, even after IQ has been statistically controlled.

### *Counting Knowledge*

Children with MD/RD or MD scored below chance levels on pseudo-counting trials but performed above chance levels and as well as children in the other groups on the remaining counting tasks. The overall pattern suggests that children with consistently low mathematics achievement scores understand most basic counting principles (e.g., one–one correspondence) but, at the same time, do not appear to understand the order-irrelevance principle or incorrectly view adjacency as an essential feature of counting (Briars & Siegler, 1984; Gelman & Gallistel, 1978). These results confirm earlier findings (Geary et al., 1992) and indicate that this pattern extends beyond first grade and includes children with MD as well as children with MD/RD; Geary et al. found this pattern for first-grade children with MD/RD.

Although the children with MD/RD or MD detected actual counting errors—trials where the first item was counted twice—less frequently than did the children in the three remaining groups, the group differences on these trials were not statistically significant. The lack of significant differences is inconsistent with earlier results (Geary et al., 1992), where children with MD/RD, as a group, failed to detect double-counting errors when the first item was counted twice. One possible reason for the discrepant results is because error trials were administered across sets of 5, 7, or 9 objects in the current study, as compared to 8, 12, or 16 objects in the Geary et al. (1992) study. One result of these procedural differences would be that an error notation would need to be held in working memory for a shorter duration in the current than in the previous study and thus the current results do not necessarily indicate normal working memory skills for these children. In fact, the pattern across studies suggests that many children with MD/RD or MD do have difficulties retaining the notation of a counting error in working memory when it occurs early in the counting process (Geary et al., 1999; Hitch & McAuley, 1991). However, these difficulties may be evident only with relatively long counts, as in our previous study (Geary et al., 1992; Hoard, Geary, & Hamson, 1999), or on tasks that require children to maintain information in working memory across a series of counting tasks, as in Hitch and McAuley's study.

### *Arithmetic*

As in previous studies, first-grade children with MD/RD or MD showed more counting-procedure and retrieval errors and used the min counting procedure less frequently than did children with average or better mathematics achievement scores (Barrouillet et al., 1997; Geary, 1990, 1993; Geary et al., 1991; Gross-Tsur et al., 1996; Jordan & Montani, 1997; Ostad, 1997, 1998; Svenson, & Broquist, 1975). One unique finding is that the children with MD/RD never used the min procedure in first grade and the children with MD used it only rarely.

Another unique finding is that the strategy choices of the MD/RD and MD groups began to diverge across grades, although some common features were still evident. Unlike the children in the RD, variable, and normal groups, the children in the MD/RD and MD groups did not show a decrease in finger counting from Grade 1 to Grade 2. However, across grades, the children with MD showed a substantive increase in the use of min counting, and significant reductions in the frequency of procedural and retrieval errors, with their performance approaching that of the children in the RD, variable, and normal groups in second grade. The children with MD/RD, in contrast, showed reductions in the frequency of procedural and retrieval errors and a substantial increase in min counting, but still differed from the children in the four remaining groups on these dimensions. The finding of more skilled use of counting procedures in second-grade children with MD relative to their MD/RD peers mirrors the findings of Jordan and Montani (1997).

The use of the retrieval-only procedure yielded a number of important findings (see also Jordan & Montani, 1997). First, the children with MD/RD committed more retrieval errors than did the children in the four remaining groups, whereas the children with MD or RD committed more errors than did the children in the variable and normal groups. Of particular interest is the high proportion of retrieval errors that were counting-string associates of one of the addends (Siegler & Shrager, 1984) in each of the LD groups. The pattern is consistent with the results of Barrouillet et al. (1997). In this study, seventh-grade LD children had difficulties inhibiting the retrieval of irrelevant associations to simple multiplication problems (e.g., retrieving 18 to $6 \times 4$). The current results and those of Barrouillet et al. are in keeping with the hypothesis that the retrieval difficulties of LD children are related, in part, to the inefficient inhibition of irrelevant associations (see also Conway & Engle, 1994).

Whereas inefficient inhibition of irrelevant associations might explain the retrieval deficits of children with MD, the pattern appears to be more complicated for the children with MD/RD or RD. As a group, the children with MD/RD used retrieval less frequently (21%) on the strategy choice problems than on the retrieval-only problems (88%), suggesting a moderately rigorous confidence criterion (Siegler, 1988). However, unlike the pattern of substantively lower retrieval errors on strategy choice than on retrieval-only problems for the MD, RD, variable, and normal groups, the MD/RD group showed essentially the same percentage of retrieval errors on the retrieval-only (81%) and strategy choice problems (77%). The pattern across problem type suggests that the children with MD/RD had a lower confidence criterion than did the children in the four remaining groups (Siegler, 1988). The slow speed of articulating familiar words (discussed below) also suggests complicated retrieval deficits for the children with MD/RD or RD.

### *Supporting Cognitive Competencies*

Consistent with many previous studies, children in the three LD groups showed lower mean forward and backward digit span scores than did the normal children (e.g., Geary et al., 1991). However, once IQ was partialed the only statistically significant comparison was for the MD/RD and normal groups on backward digit span. This does not mean that LD children do not have working memory difficulties. Rather, the current results suggest that the apparent working memory deficits of many LD children might be better understood in terms of IQ and not working memory per se. Nonetheless, this does not appear to be the whole story. In the earlier described study of Hitch and McAuley (1991), children with MD and normal children were matched on IQ and the children with MD still evidenced a deficit in counting span and in the speed of implicit counting. In all, it appears that the relation between LD and working memory deficits is more complex than can be captured by simple digit span measures and future studies will need to take this into consideration (Bull & Johnston, 1997; Bull, Johnston, & Roy, 1999).

The finding that children with MD/RD or RD had slower familiar-word articulation speeds than did the other children, even when IQ and nonword articulation speeds were statistically controlled, supports the view that poor reading achievement is associated with phonetic processing deficits (Denckla & Rudel, 1976; Fawcett & Nicolson, 1994; Morris et al., 1998). In particular, these findings suggest that children with RD, independent of their mathematics achievement levels, might experience relatively low activation levels of phonetic representations of familiar words when these words are encoded into working memory (Gathercole & Adams, 1994). Although this pattern is consistent with the position that the comorbidity of MD and RD might result from a common activation/retrieval deficit (Geary, 1993), the arithmetic retrieval patterns of the children with MD/RD or RD suggest that an additional cognitive deficit—poor inhibition of irrelevant associations—might contribute to the comorbidity of MD and RD.

Finally, the finding of no group differences in spatial ability, once IQ was partialed, is consistent with the findings of Morris et al. (1998). In the Morris et al. study, most subtypes of children with RD (most of whom showed a core phonological processing deficit), children with MD, and many children with MD/RD did not show deficits in spatial cognition. These findings do not rule out the possibility of spatial deficits in LD children (Rourke, 1993), but they do suggest that any such deficits are rare in comparison to the other types of cognitive deficits (e.g., in phonological processing) found in studies of children with LD.

## GENERAL CONCLUSION

The current study indicated that the use of theoretical models of normal mathematical development, and the associated measures, is a fruitful approach to the study of the cognitive deficits of children with LD. The use of this method demonstrated a wide array of numerical, counting, and arithmetical deficits in children with MD/RD, as well as deficits in digit span and low activation of familiar words when these words are encoded into working memory. Many of the deficits showed different developmental patterns—for instance number comprehension and production deficits appear to be most common in first grade but retrieval deficits are more persistent—and are thus likely to be the result of disruptions of different cognitive systems (Geary, 1993). In contrast, children with MD showed a more circumscribed pattern of cognitive deficit. The most salient of these was a poor understanding of the order-irrelevance or adjacency counting principle, arithmetic-fact retrieval difficulties, and a high frequency of counting-procedure errors, although the last was more common in first than in second grade (Jordan & Montani, 1997). The children with RD did not evidence any obvious deficits in number, counting, and basic arithmetic competencies, although their performance on the retrieval-only task suggests some difficulties in inhibiting irrelevant associations when retrieving arithmetic facts from long-term memory. The most salient of their deficits was a slow speed of articulating familiar words, suggesting low activation of the phonetic representations of familiar words when these words are encoded into working memory.

## REFERENCES

Ackerman, P. T., & Dykman, R. A. (1995). Reading-disabled students with and without comorbid arithmetic disability. *Developmental Neuropsychology,* **11,** 351–371.

Ashcraft, M. H., & Battaglia, J. (1978). Cognitive arithmetic: Evidence for retrieval and decision processes in mental addition. *Journal of Experimental Psychology: Human Learning and Memory,* **4,** 527–538.

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

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

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

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

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

Conway, A. R. A., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General,* **123,** 354–373.

Cowan, N. (1986). A matrix of consonant-cluster-free monosyllabic words in English. *Behavior Research Methods, Instruments, & Computers,* **18,** 434–446.

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

Edwards, P. (1978). *Reading problems: Identification and treatment.* London: Heinemann.

Fawcett, A. J., & Nicolson, R. I. (1994). Naming speed in children with dyslexia. *Journal of Learning Disabilities,* **27,** 641–646.

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

Fuson, K. C. (1988). *Children's counting and concepts of number.* New York: Springer-Verlag.

Garnett, K., & Fleischner, J. E. (1983). Automatization and basic fact performance of normal and learning disabled children. *Learning Disabilities Quarterly,* **6,** 223–230.

Gathercole, S. E., & Adams, A.-M. (1994). Children's phonological working memory: Contributions of long-term knowledge and rehearsal. *Journal of Memory and Language,* **33,** 672–688.

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

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

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

Geary, D. C. (1995). Reflections of evolution and culture in children's cognition: Implications for mathematical development and instruction. *American Psychologist,* **50,** 24–37.

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

Geary, D. C., & Brown, S. C (1991). Cognitive addition: Strategy choice and speed-of-processing differences in gifted, normal, and mathematically disabled children. *Developmental Psychology,* **27,** 398–406.

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

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

Geary, D. C., Widaman, K. F., Little, T. D., & Cormier, P. (1987). Cognitive addition: Comparison of learning disabled and academically normal elementary school children. *Cognitive Development,* **2,** 249–269.

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

Gelman, R., & Meck, E. (1983). Preschooler's counting: Principles before skill. *Cognition,* **13,** 343–359.

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

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

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

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

Johnston, R. S., & Anderson, M. (1998). Memory span, naming speed, and memory strategies in poor and normal readers. *Memory,* **6,** 143–163.

Jordan, N. C., Levine, S. C., & Huttenlocher, J. (1995). Calculation abilities in young children with different patterns of cognitive functioning. *Journal of Learning Disabilities,* **28,** 53–64.

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

McCloskey, M., Aliminosa, D., & Macaruso, P. (1991). Theory-based assessment of acquired dyscalculia. *Brain and Cognition,* **17,** 285–308.

Morris, R. D., Stuebing, K. K., Fletcher, J. M., Shaywitz, S. E., Lyon, G. R., Shankweiler, D. P., Katz, L., Francis, D. J., & Shaywitz, B. A. (1998). Subtypes of reading disability: Variability around a phonological core. *Journal of Educational Psychology,* **90,** 347–373.

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

Ostad, S. A. (1998). Developmental differences in solving simple arithmetic word problems and simple number-fact problems: A comparison of mathematically normal and mathematically disabled children. *Mathematical Cognition,* **4,** 1–19.

Räsänen, P., & Ahonen, T. (1995). Arithmetic disabilities with and without reading difficulties: A comparison of arithmetic errors. *Developmental Neuropsychology,* **11,** 275–295.

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

Sattler, J. M. (1982). *Assessment of children's intelligence and special abilities* (2nd ed.). Boston: Allyn & Bacon.

Seron, X., & Fayol, M. (1994). Number transcoding in children: A functional analysis. *British Journal of Developmental Psychology,* **12,** 281–300.

Shalev, R. S., Manor, O., & Gross-Tsur, V. (1993). The acquisition of arithmetic in normal children: Assessment by a cognitive model of dyscalculia. *Developmental Medicine and Child Neurology,* **35,** 593–601.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development,* **60,** 973–980.

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

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

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

Siegler, R. S., & Robinson, M. (1982). The development of numerical understandings. In H. Reese & L. P. Lipsitt (Eds.), *Advances in child development and behavior* (Vol. 16, pp. 241–312). New York: Academic Press.

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

Svenson, O., & Broquist, S. (1975). Strategies for solving simple addition problems: A comparison of normal and subnormal children. *Scandinavian Journal of Psychology,* **16,** 143–151.

Thorndike, R. L., Hagen, E. P., & Sattler, J. M. (1986). *The Stanford–Binet Intelligence Scale: Fourth edition.* Chicago, IL: Riverside Publishing.

Wagner, R. K., & Torgesen, J. K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin,* **101,** 192–212.

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children–Third edition.* San Antonio, TX: Psychological Corporation.

Wechsler, D. (1992). *Wechsler Individual Achievement Test.* San Antonio, TX: Psychological Corporation.

Wheeler, L. R. (1939). A comparative study of the difficulty of the 100 addition combinations. *Journal of Genetic Psychology,* **54,** 295–312.

Woodcock, R. W., & Johnson, M. B. (1989, 1990). *Woodcock–Johnson Psycho-Educational Battery–Revised.* Allen, TX: DLM Teaching Resources.

Received July 12, 1999; revised October 29, 1999