Child Development, May/June 2003, Volume 74, Number 3, Pages 834–850

A Longitudinal Study of Mathematical Competencies in Children With Specific Mathematics Difficulties Versus Children With Comorbid Mathematics and Reading Difficulties

Nancy C. Jordan, Laurie B. Hanich, and David Kaplan

Mathematical competencies of 180 children were examined at 4 points between 2nd and 3rd grades (age range between 7 and 9 years). Children were initially classified into one of 4 groups: math difficulties but normal reading (MD only), math and reading difficulties (MD–RD), reading difficulties but normal math (RD only), and normal achievement in math and reading (NA). The groups did not differ significantly in rate of development. However, at the end of 3rd grade the MD only group performed better than the MD–RD group in problem solving but not in calculation. The NA and RD only groups performed better than the MD–RD group in most areas. Deficiencies in fact mastery and calculation fluency, in particular, are defining features of MD, with or without RD.

In early elementary school, children with mathematics difficulties (MD) who are good readers progress faster in mathematics achievement than do children with comorbid MD and reading difficulties (RD), independent of their intelligence, income level, ethnicity, and gender (Jordan, Kaplan, & Hanich, 2002). In contrast, children with RD who are good in mathematics and children with comorbid RD and MD progress at about the same rate in reading achievement. Although reading abilities influence growth in mathematics achievement, mathematics abilities do not influence growth in reading achievement. Reading difficulties, regardless of whether they are specific or general in nature, tend to remain stable throughout primary school. MD, on the other hand, seem to be ameliorated by competence in reading.

The subject area of mathematics is complex with multiple domains. Difficulties may result from deficits in one or several cognitive skills (Geary, Hamson, & Hoard, 2000; Jordan & Hanich, 2000). In a recent investigation, Hanich, Jordan, Kaplan, and Dick (2001) found that second-grade children with MD who are good readers (MD only) show a different profile from children with comorbid MD and RD (MD–RD) on cognitive variables related to mathematics competence. In particular, children with MD only showed an advantage over children with MD–RD in areas that can be mediated by language (e.g., story problems and verbal counting) but not in areas that appear to depend on numerical understanding—such as estimation of numerical magnitudes—and on rapid retrieval of number facts. The findings from Hanich et al.'s study also suggested that second-grade children with RD who are good in mathematics (RD only) had weaknesses in rapid retrieval of addition facts, relative to children with normal achievement (NA). It has been posited that there is a relationship between deficits in processing sounds, a hallmark of dyslexia, and accessing arithmetic facts in long-term memory (Geary & Hoard, 2001). Learning number facts is based on counting, which involves number words and the use of the basic phonetic system. However, it is not clear why children with MD only, who presumably have intact phonetic abilities, also show fact-retrieval deficits.

Hanich et al. (2001) examined children's mathematical competencies only at one time point, raising the question of whether the observed patterns of competencies are stable. Faster growth rates of children with MD only compared with children with MD–RD in general mathematics achievement indicate that the former group might outgrow or at least compensate for some of their early weakness in selected domains of mathematical cognition (Jordan et al., 2002). Moreover, studies have suggested that some MD are unstable during primary school (Geary, 1990; Geary, Brown, & Samaranayake, 1991;

Nancy C. Jordan, Laurie B. Hanich, and David Kaplan, School of Education, University of Delaware.

This study was supported by a grant from the National Institute of Child Health and Human Development (R01-HD36672). We are very grateful to the children and teachers who participated in this project.

Correspondence concerning this article should be addressed to Nancy C. Jordan, School of Education, University of Delaware, Newark, Delaware 19716. Electronic mail may be sent to njordan@udel.edu.

 0009-3920/2003/7403-0012

Geary et al., 2000), with children showing low mathematics achievement scores in one grade but average scores or better in another grade. However, weaknesses in specific areas, such as in fact retrieval, appear to be highly stable (Geary et al., 1991; Ostad, 1997, 1998, 1999). It is also possible that the influence of reading problems on mathematical competence becomes greater over time, leading children with RD only in second grade to ''grow into'' mathematical weaknesses by third grade, especially with respect to story problem solving.

The present study directly extended the baseline work of Hanich et al. (2001), tracking the development of mathematical competencies in the same sample of children at three additional time points: spring of second grade, fall of third grade, and spring of third grade. (The original baseline data were collected in the winter of second grade.) Thus, we have a total of four longitudinal data points for children with MD only, MD–RD, RD only, and NA on each of the following tasks: exact calculation of arithmetic combinations, forced retrieval of number facts, approximate arithmetic (estimation), calculation principles, story problems, place value, and written computation. All of the tasks, with the exception of place value, involved the operations of addition and subtraction. The areas of mathematical cognition assessed on our tasks were related to basic calculation, problem solving, and base-ten concepts and are summarized in Table 1. The areas provide a foundation for learning higher level mathematics such as algebra (Miller, 1992) and involve different cognitive processes (Ginsburg, 1997). Our tasks are based on those created for studies of normal mathematical development and have direct relevance to the teaching of mathematics in young children (Hanich et al., 2001).

We examined children's calculation strategies (on exact calculation of arithmetic combinations) at each time point, as it has been shown that children with MD rely on finger counting more (and on retrieval less) than do children without MD and that children with MD only use counting strategies more accurately than do children with MD–RD (Geary et al., 2000; Jordan, Hanich, & Uberti, 2003; Jordan & Montani, 1997). To date, however, little longitudinal data are available on strategy development in children with MD. Our forced-retrieval task, which required children to answer number combinations quickly, allowed us to look at automaticity and fluency in calculation. By third grade, children's knowledge of arithmetic combinations should be strong enough to allow routine retrieval (Lemaire, Barrett, Fayol, & Abdi, 1994).

Because Hanich et al. (2001) did not assess children's IQ in their baseline study, it is possible that performance differences between children with MD only and children with MD–RD was attributable primarily to differences in IQ (Geary, Hoard, & Hamson, 1999). In the present investigation we assessed children's IQ, allowing us to examine achievement-group effects beyond the influence of overall intelligence. We also examined the influences of gender, socioeconomic status (SES), and minority status.

A main question posed in this study relates to rates of growth in mathematics competencies over time. To this end, growth-curve modeling was used to analyze the data, a technique also employed by Jordan et al. (2002) to examine achievement growth. Growth-curve modeling provides an estimate of the average level of mathematical competency at any time point as well as the average rate of growth over time in the mathematical outcomes for the sample (Raudenbush & Bryk, 2002). It also provides an estimate of the average rate of acceleration in growth and is flexible enough to handle nonequidistant measurement occasions. Growth-curve modeling

Table 1
Areas of Mathematical Cognition Assessed in the Present Study

Area	Description
Basic calculation	Accuracy on number combinations
	Use of calculation procedures (e.g., finger counting)
	Automatic fact retrieval and fluency
	Computational estimation
Problem solving	Arithmetic story problems with varied levels of verbal complexity
	Understanding of calculation principles (i.e., relationships within and between arithmetic operations)
Base-ten concepts	Understanding and representation of place value, such as enumeration, number identification, positional knowledge, and digit correspondence
	Written computation with multidigit numbers, including problems involving regrouping

allowed us to predict individual levels and rates of growth by IQ, ethnicity, income, and gender, as well as by the four achievement groups (i.e., MD only, MD–RD, RD only, and NA). Because of its clear focus on growth and its considerable flexibility, growth-curve modeling is seen as a more appropriate analytic strategy than conventional methods for the analysis of longitudinal data, such as repeated-measures ANOVA.

Method

Participants

In the fall of 1999, we screened more than 600 second-grade children (age range between 7 and 9 years) from one school district in New Castle County, Delaware. Using the Reading and Mathematics Composites from the Woodcock-Johnson Psycho-Educational Battery–Revised (WJ), Form A (Woodcock & Johnson, 1990), we classified 210 children into one of four achievement groups: MD only, MD–RD, RD only, and NA. The WJ Mathematics Composite includes Calculation and Applied Problems subtests and the WJ Reading Composite includes Letter-Word Identification and Passage Comprehension subtests. Based on grade-level norms, children with Mathematics Composite Scores at or below the 35th percentile were classified as MD and children with Reading Composite or Letter-Word Identification scores at or below the 35th percentile were classified as RD. We used the Reading Composite or Letter-Word Recognition criteria in reading because on our initial screening data revealed that a number of children had borderline reading composite scores despite low letter-word recognition scores. Because we were interested in children with reading decoding weaknesses, we decided to include these children in the study. It should be noted that all but 2 of the children in the RD only group and 1 in the MD–RD group scored at or below the 35th percentile in Letter-Word Recognition subtest, suggesting that children classified as RD were characterized by decoding problems. A detailed description of our screening procedure is presented in Hanich et al. (2001).

For the present longitudinal study, only children who completed all phases of the project were included. Thus, our final sample of 180 children included 46 children with MD only, 42 children with MD–RD, 45 children with RD only, and 47 children with NA. The mean mathematics and reading scores for the four achievement groups, along with data related to gender, ethnicity, parental income level, and special education, are presented in Table 2. Within each achievement group, children identified as ethnic minority were primarily African American (>80% for each achievement group). Eligibility for the subsidized lunch program at school was used to determine low-income status. The MD only and MD–RD groups did not differ significantly in mathematics achievement and the RD only and MD–RD groups did not differ significantly in reading achievement. None of the children in NA or MD only groups were retained in second grade. However, 4 children in the MD–RD group and 7 children in the RD only group repeated second grade in Year 2 of our study. The retention issue is addressed in the Results section.

Table 2
Descriptive Information for Participants by Achievement Group

Achievement group	*N*	M/F[a]	Percentage ethnic minority[b]	Percentage low income[c]	Percentage special education[d]	Reading composite percentile scores	Letter-word identification percentile scores	Mathematics composite percentile scores
MD only	46	21/25	61	46	2	71.67 (14.32)	66.34 (18.02)	22.87 (9.74)
MD–RD	42	23/19	48	50	40	24.91 (13.43)	21.01 (11.41)	21.07 (10.80)
RD only	45	29/16	60	56	27	26.96 (10.42)	23.24 (9.65)	60.42 (16.03)
NA	47	23/24	43	40	0	71.96 (13.27)	63.66 (16.46)	68.81 (12.02)

Note. MD = math difficulties but normal reading; MD–RD = math and reading difficulties; RD only = reading difficulties but normal math; NA = normal achievement in math and reading. Standard deviations are in parentheses.
[a]M/F stands for male/female. [b]Within each achievement group, children identified as ethnic minority were primarily African American (>80% for each achievement group). [c]Low income was determined by eligibility for the subsidized lunch program at school. [d]Special education services were provided in either second or third grade.

Materials

Each participant was given seven mathematics tasks, presented in the following order: (a) exact calculation of arithmetic combinations, (b) story problems, (c) approximate arithmetic, (d) place value, (e) calculation principles, (f) forced retrieval of number facts, and (g) written computation. The tasks are exactly the same as those used by Hanich et al. (2001). To emphasize the required operations and to prevent wrong operation errors (e.g., adding instead of subtracting) we separated addition and subtraction problems on exact calculation of arithmetic combinations, approximate arithmetic, forced retrieval of number facts, and written computation.

Exact calculation of arithmetic combinations. Children were given four addition and four subtraction arithmetic combinations (9+8; 3+6; 5+6; 8+7; 9–3; 17–9; 11–5; 15–8), presented individually. A written version of the problem (horizontal format) was shown at the same time it was read aloud by the examiner. The examiner told children to use any method they wanted to figure out the answer. To get a rough estimate of solution time, the experimenter began timing the child with a stopwatch immediately after reading the problem. As soon as the child began to state his or her answer, the examiner stopped timing. If a child answered before the experimenter finished reading the problem a time of 0 was recorded. If a child stated an answer but then wanted to think some more and answer again (in most cases the child immediately said "no" after giving the first answer), the examiner restarted the stopwatch. If the child stated a second answer after the original answer, the examiner made a best estimate of how many additional seconds the child took. Response times were recorded for each arithmetic combination. Although the stopwatch was marked in units of .1 s, times were rounded to the nearest second. The examiners agreed 96% on recording children's response times for a sample set of trials (to the nearest second with rounding).

On each item, the experimenter observed the child's strategy use (e.g., counting verbally or with fingers, no overt strategy, etc.) and wrote down what the child did on the score sheet. Right after the child answered, the experimenter asked the child to give an explanation and recorded the response verbatim. The child and the experimenter agreed on 97% of the trials, which is in keeping with other studies (Geary et al., 2000; Siegler, 1987). If the child and the experimenter disagreed, the experimenter's observation was used when the strategy she observed was readily apparent (e.g., finger counting). In ambiguous cases, the child's explanation was used (Geary et al., 2000).

Story problems. The experimenter read each child 10 story problems, ranging from conceptually simple to conceptually complex (Carpenter & Moser, 1984; Riley & Greeno, 1988; Riley, Greeno, & Heller, 1983). The set included four types of story problems: change problems ($n = 3$; e.g., Nina had 9 pennies. Then she gave 3 pennies to Anthony. How many pennies does Nina have now?), combine problems ($n = 2$; e.g., Emily has 3 pennies. John has 6 pennies. How many pennies do they have altogether?), compare problems ($n = 3$; e.g., Dennis has 7 pennies. Molly has 5 pennies. How many pennies does Dennis have more than Molly?), and equalize problems ($n = 2$; e.g., Claire has 4 pennies. Ben has 9 pennies. How many pennies does Claire need to get to have as many as Ben?). All problems involved simple calculations with sums and minuends of 9 or less.

Children were told to use whatever strategy they wanted to get the correct answer and were given a container of play pennies to use during the activity. The experimenter read each problem aloud with an accompanying written version of the problem in full view of the child. Children were asked to wait until the problem was read in its entirety before giving an answer.

Approximate arithmetic. The approximate arithmetic task was adapted from Dehaene, Spelke, Pinel, Stanescu, and Tsivkin (1999). Children were shown 10 addition and 10 subtraction problems, each having two proposed answers (e.g., 4+5 = 10 or 20; 16–7 = 4 or 8; 40–30 = 11 or 31; 50–9 = 20 or 40). Both of the answers were incorrect but one was within a few units of the correct answer whereas the other was further away. Each problem was shown individually. The child's task was to choose the number that was closest to the correct answer. As the experimenter read each problem, she also displayed a written version in full view of the child. Children were asked to answer right away and not to calculate. To prevent children from calculating, they were allowed only 5 s to answer. If the child did not answer within the time limit, a response of "no answer" was recorded and the problem was scored as incorrect. Children were presented with two practice problems.

Place value. The place value task, adapted from Hiebert and Wearne (1996), Kamii (1989), and Ross (1989), involved three activities: counting and number identification, positional knowledge, and digit correspondence.

First, the child was given 16 colored chips and asked to count the chips. If the child made a mistake, the experimenter counted the chips aloud so that the child understood there were 16 chips in all. Number identification required the child to read two- and three-digit numbers printed on a card (i.e., 16, 37, 415). After the child read the numbers aloud, he or she was asked which number was in the tens place, the ones place, and if relevant, the hundreds place (positional knowledge). The items were scored using a pass–fail criterion in which all of the digit places needed to be identified correctly for the problem to be scored as correct.

The first digit-correspondence activity followed the number-identification task for the number 16. The experimenter circled the 6 in the number 16 and asked the child to use the chips to show what that part stands for in the number 16 (6 chips). The experimenter then circled the 1 and asked the child to use the chips to show what that part stands for (10 chips).

The subsequent digit-correspondence activities directly challenged children's understanding of two-digit numbers. Children were given two conditions: standard place-value partitioning and nonstandard place-value partitioning. In the standard condition, the tens place of the digit was represented by unit squares grouped together in tens and the ones place by individual unit squares. The nonstandard condition was the same except that one of the groups of ten was separated into 10 individual unit squares. The experimenter first showed the child a card showing the number 43, along with a picture of 43 squares grouped with four 10-unit squares and 3 individual unit squares. The experimenter asked the child draw a circle around the squares that the 3 part in the number 43 stands for (i.e., 3 individual unit squares) and then to draw a circle around the squares that the 4 part in the number 43 stands for (i.e., four 10-unit squares). The child was shown a second card with the number 43 printed on it and a corresponding picture of 43 squares. This time the squares were in a nonstandard partitioning arrangement with 3 groups of 10-unit squares and 13 individual unit squares. The same procedure as the one described for the standard partitioning item was used. The standard and nonstandard partitioning activities were repeated with another two-digit number (i.e., 52).

In the last digit-correspondence activity (also nonstandard partitioning), the child was given a card with the number 26 printed on it along with a picture of 26 stars arranged in six groups of four and one group of two. The experimenter asked the child to draw a circle around the number of stars that the 6 stands for in the number 26 and then around the number of stars that the 2 stands for in the number 26. The total number of items on the place-value task was 12.

Calculation principles. The calculation principles task, based on the work of Baroody (1999) and Russell and Ginsburg (1984), involved solving six pairs of problems in which the given answer to the first of the pair could be used to solve the second. Two items assessed understanding of the commutative principle, that the order of the addends does not affect the sum (i.e., $47+86=133$, so $86+47=?$, and $94+68=162$, so $68+94=?$); two items assessed understanding of the inversion principle, that subtraction is the inverse of addition (i.e., $27+69=96$, so $96–69=?$, and $36+98=134$, so $134–36=?$); and two items assessed understanding of the doubles plus one pattern (i.e., $37+37=74$, so $37+38=?$, and $64+64=128$, so $65+64=?$). Each problem was shown and read to the child at the same time. We used two-digit numbers so children could not retrieve answers by rote. The child was told to give an oral response as quickly as possible. To keep children from calculating, a 5-s time limit was used. If a child did not answer within 5 s, the answer was scored as incorrect.

Forced retrieval of number facts. Four addition and four subtraction combinations were read to each child (i.e., 4+2; 9+4; 7+9; 3+8; 6–4; 13–9; 16–7; 11–8), along with a visual presentation. The experimenter told children to give an answer right away or to tell her that they would need more time. Right after reading the combination, the experimenter started timing. If the child did not answer within 3 s or said he or she needed more time, the item was recorded as "no answer" and scored as incorrect. The forced-retrieval task is based on the method developed by Russell and Ginsburg (1984) and Jordan and Montani (1997). The absence of $n+1$ combinations minimized the possibility of children using a "number-after" rule, rather than retrieval, to get the correct answer (Baroody & Tiilikainen, 2003). Note that a 5-s limit was used on the approximate arithmetic and calculation principles tasks, in contrast to the 3-s limit on the forced-retrieval task. We allowed more time on the former tasks because they required children to evaluate two problems or answers before responding.

Written computation. Children were given 8 two- and three-digit written computation problems, 4 involving addition and 4 involving subtraction. Regrouping was necessary on half of the additions and half of the subtractions.

To assess IQ, each child was given the Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999). The correlation between the WASI and the longer version Wechsler Intelligence Scale for Children–III is .87. The WASI includes both verbal and performance scales.

Data-Collection Procedures

In both years of the project, we had a cadre of four female experimenters who underwent extensive training. Children were tested individually in their schools. They were given the mathematics tasks four times: January of second grade, April of second grade, November of third grade, and May of third grade, constituting a total of 16 months. Each session lasted between 30 and 45 min, with the third-grade sessions typically being shorter than the second-grade sessions. Although the individual items on each task were the same for each test period, we changed their order of presentation during the second and fourth test periods. However, addition items were always presented before subtraction items on exact calculation of arithmetic combinations, approximate arithmetic, forced retrieval of number facts, and written computation. The IQ test was given to children in January of third grade. The experimenters were not given information about children's group membership.

Results

Internal reliability estimates for the mathematics tasks (at Time 4) using coefficient alphas were .66 for exact calculation of number combinations, .72 for story problems, .85 for place value, .80 for calculation principles, .70 for forced retrieval of number facts, .58 for approximate arithmetic, and .67 for written computation.

Growth-curve analyses were performed on each of the mathematics tasks. A brief introduction to growth-curve modeling, including its purpose and

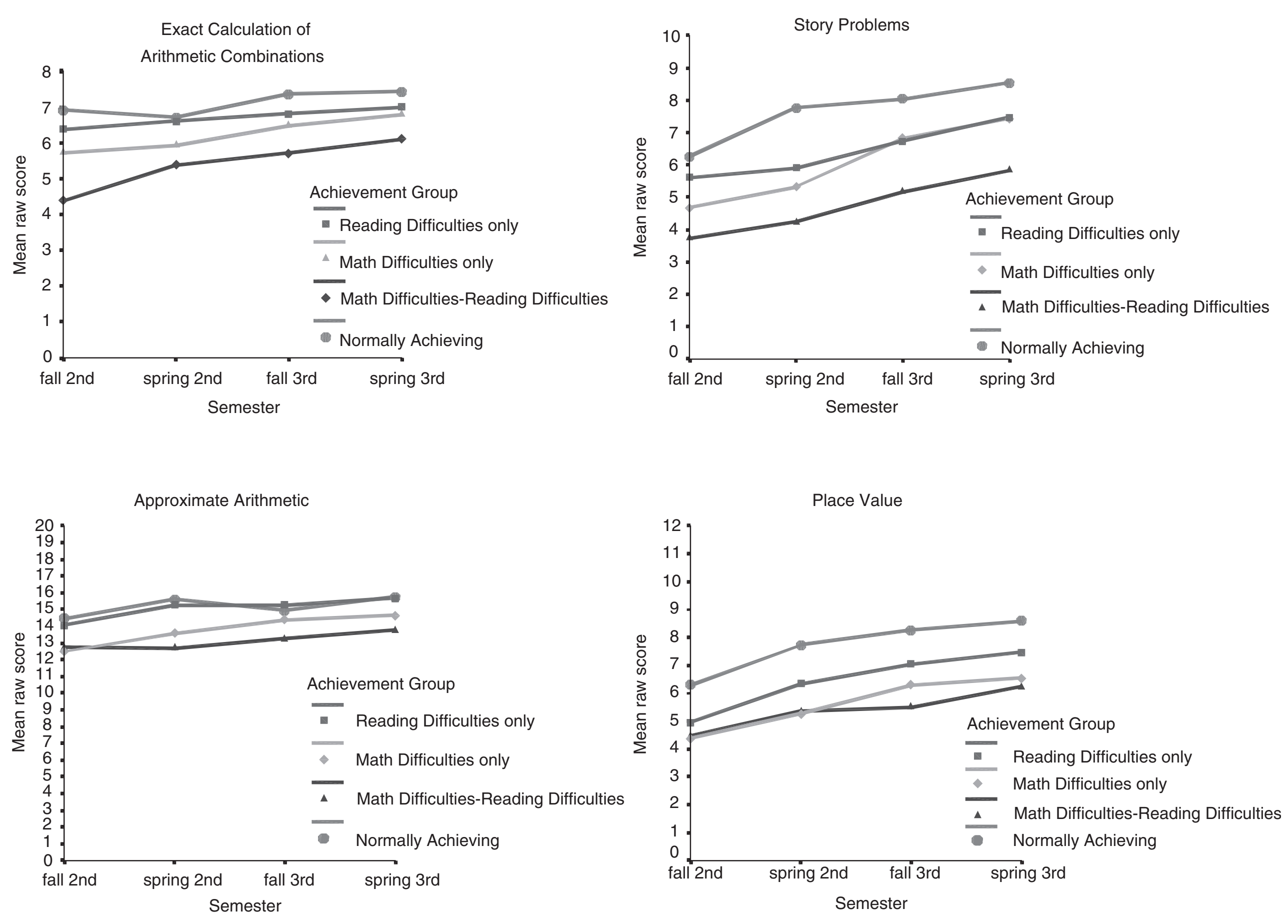

Figure 1. Average empirical growth trajectories on tasks assessing mathematics competencies, by achievement group. Points represent the mean raw scores for each task.

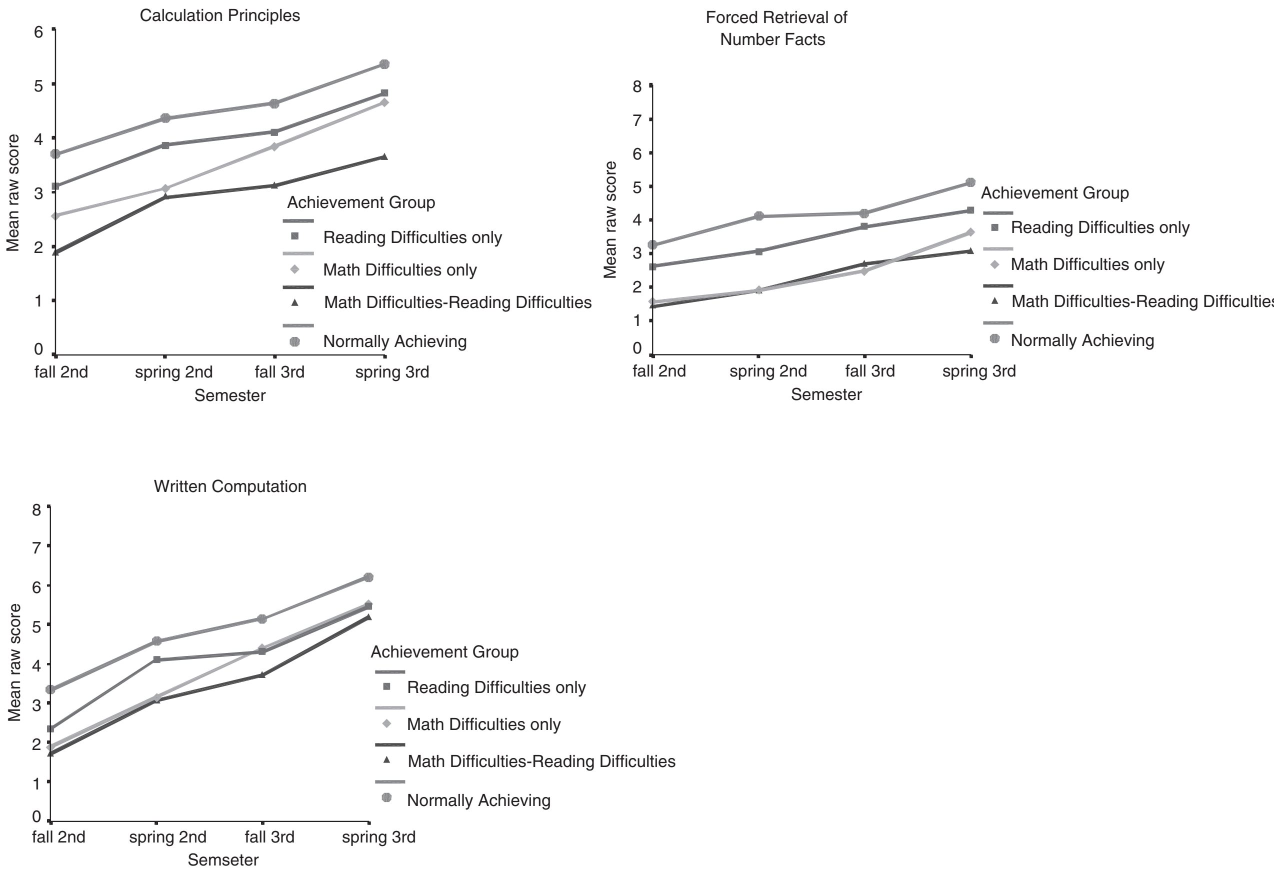

Figure 1. Continued.

Table 3
Baseline Model: Growth Curve Results for Mathematics Tasks

Estimate	Exact calculation of arithmetic combinations	Story problems	Approximate arithmetic	Place value	Calculation principles	Forced retrieval of number facts	Written computation
Intercept	6.84*	7.34*	14.96*	7.22*	4.60*	4.02*	5.57*
Slope	0.02	0.07*	0.03	– 0.03	0.08*	0.10*	0.14*
Var(intercept)	1.39*	3.60*	5.37*	7.42*	1.60*	2.70*	1.37*
Var(slope)	0.01*	0.01*	0.01	0.01*	0.00	0.00*	0.01*
R(int. slope)	0.20	0.43*	0.64*	0.71*	0.16	0.54*	0.34
Acceleration variable	– 0.00	– 0.00*	– 0.00	– 0.01*	– 0.00	– 0.00	– 0.00

Note. On exact calculation of arithmetic combinations, the dependent number reflects accuracy rather than latency or strategies. Var( ) stands for the variance of the parameters in parentheses.
$^*p < .05$.

the interpretation of statistical values, is presented in the Appendix. The average empirical growth trajectories, by achievement group, for each mathematics task are shown in Figure 1. Models included the effects of achievement-group membership and time-invariant predictors (i.e., gender, ethnicity, income, and IQ) on average performance level at the end of third grade (intercept at Time 4) and rate of growth

Table 4
Model 1: Growth Curve Results for Mathematics Tasks with Effects of Achievement Group Membership

Estimate	Exact calculation of arithmetic combinations	Story problems	Approximate arithmetic	Place value	Calculation principles	Forced retrieval of number Facts	Written computation
Intercept	6.05*	5.83*	13.73*	6.22*	3.58*	3.07*	5.15*
Slope	– 0.00	0.10	0.06	0.06	0.02	0.04	0.21*
Var(intercept)	1.19*	2.60*	4.68*	6.47*	1.31*	2.19*	1.17*
Var(slope)	0.01*	0.01*	0.01	0.01*	0.00*	0.00*	0.01*
R(int. slope)	0.35	0.52*	0.70*	0.74*	0.23	0.65*	0.42*
Acceleration variable	– 0.01*	– 0.00	– 0.00	– 0.00	– 0.01	– 0.00	0.00
Intercept on NA	1.41*	2.63*	1.90*	2.36*	1.74*	1.95*	1.03*
Slope on NA	0.03	– 0.13	0.06	– 0.13	0.08	0.04	– 0.08
Intercept on RD only	0.94*	1.64*	1.98*	1.23*	1.19*	1.21*	0.26
Slope on RD only	0.03	0.03	0.08	– 0.11	0.06	0.02	– 0.10
Intercept on MD only	0.73*	1.62*	1.07	0.33	1.06*	0.53	0.35
Slope on MD only	0.05	– 0.03	0.10	– 0.09	0.11	0.16	– 0.08
Acceleration on NA	0.01	– 0.01	0.01	– 0.01	0.01	0.00	– 0.00
Acceleration on RD only	0.01	0.00	0.01	– 0.01	0.00	0.00	– 0.00
Acceleration on MD only	0.01	– 0.00	0.01	– 0.01	0.01	0.01	– 0.01

Note. On exact calculation of arithmetic combinations, the dependent number reflects accuracy rather than latency or strategies. Var( ) stands for the variance of the parameters in parentheses.
$^*p < .05$.

Table 5
Model 2: Growth Curve Results for Mathematics Tasks with Effects of Achievement Group Membership and Time-Invariant Predictors

Estimate	Exact calculation of arithmetic combinations	Story problems	Approximate arithmetic	Place value	Calculation principles	Forced retrieval of number facts	Written computation
Intercept	6.15*	6.39*	13.30*	6.83*	3.68*	2.92*	5.32*
Slope	– 0.01	0.09	0.07	0.08	0.02	0.04	0.18*
Var(intercept)	1.06*	1.70*	3.91*	4.64*	1.04*	1.95*	0.98*
Var(slope)	0.01*	0.01*	0.01	0.01*	0.00*	0.00*	0.01*
R(int. slope)	0.41*	0.59*	0.73*	0.73*	0.31	0.63*	0.43
Acceleration variable	– 0.01*	– 0.00	– 0.00	– 0.00	– 0.01	– 0.00	0.00
Intercept on NA	0.93*	1.44*	1.41*	1.04	1.12*	1.48*	0.60
Slope on NA	0.04	– 0.14	0.06	– 0.16	0.08	0.03	– 0.07
Intercept on RD only	0.76*	1.29*	1.69*	0.86	0.98*	1.00*	0.11
Slope on RD only	0.03	0.03	0.08	– 0.11	0.06	0.02	– 0.10
Intercept on MD only	0.51	1.18*	1.04	0.04	0.88*	0.40	0.14
Slope on MD only	0.05	– 0.03	0.11	– 0.10	0.10	0.16	– 0.08
Acceleration on NA	0.01	– 0.01	0.01	– 0.01	0.01	0.00	– 0.00
Acceleration on RD only	0.01	0.00	0.01	– 0.01	0.00	0.00	– 0.00
Acceleration on MD only	0.01	– 0.00	0.01	– 0.01	0.01	0.01	– 0.01
Intercept on gender	– 0.25	– 0.36	1.48*	0.90*	0.16	0.41	– 0.19
Slope on gender	– 0.00	– 0.02	0.03	0.01	– 0.02	0.02	0.04*
Intercept on ethnicity	0.33	0.10	– 0.22	– 0.56	– 0.07	0.04	0.40
Slope on ethnicty	– 0.01	0.02	– 0.02	– 0.01	– 0.00	– 0.01	0.02
Intercept on income	0.21	0.23	– 0.04	– 0.56	0.26	0.28	– 0.14
Slope on income	0.02	0.03	– 0.04	– 0.04	0.02	0.01	– 0.01
Intercept on IQ	0.04*	0.10*	0.05*	0.10*	0.05*	0.04*	0.04*
Slope on IQ	0.00	0.00	0.00	0.00	– 0.00	0.00	– 0.00

Note. On exact calculation of arithmetic combinations, the dependent number reflects accuracy rather than latency or strategies. Var( ) stands for the variance of the parameters in parentheses.
$^*p < .05$.

(slope) in the mathematics tasks. For each task, three growth-curve models were computed: (a) baseline model, which provided estimates of the slope and the intercept (Table 3); (b) Model 1, which adds the effects of achievement-group membership (Table 4); and (c) Model 2, which adds the effects of the predictor variables (i.e., gender, income, ethnicity, and IQ; Table 5). The MD–RD group was used as the reference group for achievement-group comparisons. In Model 2, males, minority students, and students participating in the subsidized lunch program were dummy coded 1. In addition, for interpretability, IQ was centered on the sample mean. Thus, as noted in the Appendix, the average Time 4 measurement and average slope are with reference to White female MD–RD students with IQ scores at the mean of the sample.

We included both linear and quadratic models in the growth-curve analyses. Although the acceleration variable was significant in some models, the small size of the coefficients indicates minimal effects. Moreover, none of the analyses revealed significant achievement-group effects with regard to acceleration of growth. Therefore, we limit our discussion to linear models.

It should be noted that growth-curve modeling is flexible enough to handle unequal spacing of measurement occasions. For this study, the measurement occasions were parameterized in such a way as to reflect rates of growth in terms of monthly increments over the 16 months of the study.

Exact Calculation of Arithmetic Combinations

We guide the reader through Tables 3 to 5 on exact calculation of arithmetic combinations. This description will help the reader interpret results for subsequent tasks, which are discussed in less detail.

The intercept and slope on exact calculation of arithmetic combinations for the total sample are displayed in the baseline model in Table 3. The average raw score at the end of third grade was 6.84 (of 8) and the average growth was 0.02 points over 2 years.

The effects of adding achievement-group membership to the model are summarized under Model 1 in Table 4. The average score at the end of third grade on exact calculation of arithmetic combinations for the MD–RD group was 6.05, and the slope was –0.00. The methodology of growth-curve modeling allows for a large-sample test of differences in the intercept and growth rate as a function of grouping variables. This is discussed more fully in the Appendix. In the context of this study, an inspection of the effects of achievement-group membership shows that the NA, RD only, and MD only groups had significantly higher scores at the end of third grade on exact calculation of arithmetic combinations than did the MD–RD group.

Model 2 in Table 5 shows the effects of adding time-invariant predictors of gender, ethnicity, income, and IQ to the model. Because of the choice for coding dummy variables, the slope and intercept refer to White females with average IQs in the MD–RD group from middle-income families. The intercept for this group was 6.15 and the slope was –0.01. Holding predictor variables constant, the NA and RD only groups had significantly higher scores at Time 4 than did the MD–RD group. Although the MD only group had significantly higher scores at Time 4 than did the MD–RD group in Model 1 (without the predictor variables), the MD only group did not differ from the MD–RD group on Time 4 scores when the predictors were taken into account in Model 2. There were no significant differences in slope values among the achievement groups. IQ was a significant predictor of the intercept at Time 4; that is, children with higher IQs scored better on exact calculation of number combinations than did children with lower IQs. The effect of IQ was not significant on the slope. Gender, ethnicity, and income did not predict the intercept or the slope.

We performed post hoc mean comparisons using ANOVAs with Tukey tests ($p < .05$) to examine mean differences among the achievement groups on exact calculation of number combinations at Time 4 (i.e., differences among NA, MD only, and RD only groups). The MD only, RD only, and NA groups did not differ significantly from each other.

Story Problems

Holding predictor variables constant (Model 2 in Table 5), the NA, RD only, and MD only groups ended third grade (Time 4) with significantly higher story problems scores than did the MD–RD group. There were no achievement-group differences in growth rate. There was a significant effect of IQ on the intercept for story problems, favoring children with higher IQ scores. There were no effects of gender, ethnicity, or income on the intercept or on the slope.

Post hoc mean comparisons at Time 4 showed that the NA group performed significantly better than the MD only and RD only groups. The MD only and RD only groups did not differ from each other.

Approximate Arithmetic

The NA and RD only groups had significantly higher scores at the end of third grade on approximate arithmetic than did the MD–RD group, independent of the predictor variables (Model 2 in Table 5). The MD only group did not differ from the MD–RD group on Time 4 approximate arithmetic scores. Achievement-group membership did not predict growth rate. There was relatively little growth in approximate arithmetic across all achievement groups. IQ and gender were significant predictors of Time 4 scores, favoring boys and children with higher IQs. There were no significant predictors of the slope.

Post hoc mean comparisons at Time 4 showed that both the NA and RD only groups performed better than did the MD only group. The NA and RD only groups did not differ from each other.

Place Value

In Model 1 (Table 4), the MD–RD group performed significantly worse than the NA and RD only groups at Time 4. However, when predictor variables were added in Model 2 (Table 5), these achievement-group differences disappeared. There was no achievement-group effect for growth rate. There was a significant effect of IQ and gender on the intercept but not on the slope. Boys ended third grade with significantly higher place value scores than girls, and children with higher IQs had higher scores than did children with lower IQs. Ethnicity and income did not predict either the intercept at Time 4 or the slope.

Post hoc mean comparisons at Time 4 showed the NA group performed better than the MD only and RD only groups. The MD only and RD only groups did not differ from each other.

Calculation Principles

Holding the predictor variables constant, children in the NA, RD only, and MD only groups had higher scores at the end of third grade than did children in the MD–RD group (Table 5). There were no significant achievement-group differences on the slope. IQ was a significant predictor for Time 4, favoring children with higher IQs. Gender, ethnicity, and income did not predict the slope or Time 4 scores.

Post hoc comparisons at Time 4 revealed that the NA group performed better than the MD only group but not the RD only group. The MD only and RD only groups did not differ from each other.

Forced Retrieval of Number Facts

Controlling for predictor variables, the NA and RD-only groups had significantly higher scores at the end of third grade than did the MD–RD group, while the MD only group did not differ from the MD–RD group (Table 5). There was a significant effect of IQ on Time 4 scores, favoring children with higher IQs, but not on the slope.

Post hoc comparisons at Time 4 showed that the NA group performed significantly better than the MD only group but not the RD-only group. The RD only group performed significantly better than the MD only group.

Written Computation

When we held predictor variables constant (Model 2 in Table 5), there were no significant

Table 6

Mean Number of Times Finger Counting Was Used on Exact Calculation of Arithmetic Combinations and Mean Percentage of Times It Yielded a Correct Answer Across Four Time Points (Number of Items = 8)

Achievement group	Time 1	Percentage correct	Time 2	Percentage correct	Time 3	Percentage correct	Time 4	Percentage correct
MD only	4.74	77	4.87	81	4.33	81	3.96	81
	(2.67)	(0.27)	(2.76)	(0.20)	(2.95)	(0.21)	(2.82)	(0.26)
MD–RD	5.60	55	5.21	68	5.07	79	5.14	71
	(2.64)	(0.28)	(2.93)	(0.29)	(2.78)	(0.23)	(2.75)	(0.31)
RD only	4.71	76	4.44	84	4.02	82	3.11	85
	(2.60)	(0.27)	(2.62)	(0.20)	(2.45)	(0.23)	(2.51)	(0.33)
NA	4.06	84	3.55	79	3.28	88	2.79	85
	(2.65)	(0.20)	(2.54)	(0.29)	(2.46)	(0.25)	(2.56)	(0.27)

Note. MD = math difficulties but normal reading; MD–RD = math and reading difficulties; RD only = reading difficulties but normal math; NA = normal achievement in math and reading. Standard deviations are in parentheses. Time 1 = winter 2000; Time 2 = spring 2000; Time 3 = fall 2000; Time 4 = spring 2001.

Table 7
Growth Curve Results for Mean Number of Times Finger Counting Was Used on Exact Calculation of Arithmetic Combinations

Estimate	Baseline model	Model 1 with effects of achievement group membership	Model 2 with effects of achievement group membership and time-invariant predictors
Intercept	3.77*	5.16*	5.26*
Slope	0.01	0.05	– 0.02
Var(intercept)	5.38*	5.50*	4.09*
Var(slope)	0.00	0.01*	0.00*
R(int. slope)	0.31	0.43	0.25
Acceleration variable	0.01*	0.00	0.00
Intercept on NA		– 2.35*	– 1.36*
Slope on NA		– 0.10	0.10
Intercept on RD only		– 2.03*	– 1.26*
Slope on RD only		– 0.20	0.08
Intercept on MD only		– 1.22*	– 0.65
Slope on MD only		– 0.13	0.01
Acceleration on NA		– 0.00	0.01
Acceleration on RD only		– 0.01	0.01
Acceleration on MD only		– 0.01	0.00
Intercept on gender			– 1.17*
Slope on gender			0.02
Intercept on ethnicity			– 0.07
Slope on ethnicity			0.02
Intercept on income			0.07
Slope on income			0.01
Intercept on IQ			– 0.07*
Slope on IQ			– 0.01*

Note. Var( ) stands for the variance of the parameters in parentheses.
*$p < .05$.

achievement-group differences on the slope or intercept. The NA group, however, had a significantly higher score at the end of third grade than did the MD–RD group when the predictors were not considered (Model 1 in Table 4). There were no significant achievement-group differences in growth rates. There was a significant effect of IQ on Time 4 scores, favoring children with higher IQs. Gender was a significant predictor of growth rate, with boys having a 0.04-point advantage in growth over girls.

Post hoc comparisons at Time 4 showed that the NA group performed significantly better than the MD only and RD only groups, who did not differ from each other.

Summary of Growth-Curve Analyses for the Mathematics Tasks

We did not find any significant differences in growth rate by achievement group. The NA group had significantly higher ending scores than did the MD–RD group on all of the mathematics tasks. However, the group difference on written computation and place value disappeared when we added predictor variables (i.e., gender, income ethnicity, and IQ) in Model 2.

The RD only group also had higher ending scores than did the MD–RD group on most tasks, with the exception of place value (only after predictors were added) and written computation (before predictors were added).

The MD only group showed significantly higher ending scores than did the MD–RD group on exact calculation of arithmetic combinations, story problems, and calculation principles. However, the group effect for exact calculation of number combinations was not significant after predictor variables were added in Model 2. The two MD groups did not differ on approximate arithmetic, forced retrieval of number facts, place value, or written computation.

IQ predicted performance at the end of third grade on all of the mathematics tasks, whereas gender (favoring boys) predicted performance only on approximate arithmetic and place value. Gender (favoring boys) predicted growth for written computation. Ethnicity and income level, by themselves, did not predict performance or growth.

Strategy Analyses

Across all achievement groups at each time of testing, finger counting was the most commonly used calculation strategy. Table 6 presents the mean number of trials on which finger counting was used as well as the mean percentage of time finger counting yielded a correct answer, by achievement group and time of testing.

To determine whether there were achievement-group differences in the frequency and growth of finger counting, we performed a growth-curve analysis on the mean number of times children used finger counting. (We did not perform analyses on the other strategies because they were used so infrequently.) The model included the effects of achievement-group membership and time-invariant predictors (i.e., gender, ethnicity, income level, and IQ) on performance level at Time 4 (intercept) and rate of growth (slope) in finger counting. As in the previous analyses, the MD–RD group was used as the reference group for achievement-group comparisons.

The growth-curve model is shown in Table 7. Holding predictors constant, the MD–RD group used finger counting significantly more often than did the NA and RD only groups but not the MD only group at the end of third grade. No achievement-group differences in growth rate were observed. There was a significant effect of gender and IQ on finger counting at the end of third grade. Girls used their fingers more often than did boys, and children with lower IQs used their fingers more often than did children with higher IQs. IQ was also a significant predictor of the slope, suggesting that children with higher IQs decline in finger use more rapidly than do children with lower IQs. Ethnicity and income did not predict finger use at Time 4 or the slope.

Although children with MD only and MD–RD did not differ significantly in frequency in finger use, it is interesting to note that with the exception of the third time point (fall of third grade), children with MD only were more accurate with their fingers (see Table 6). The accuracy gap was especially wide (55% accuracy for MD–RD vs. 77% accuracy for MD only) at the first time point (winter of second grade).

Retention

We examined children who were retained in second grade ($n = 11$; 4 MD–RD children and 7 RD only children) in Year 2 of the study versus children who were promoted to third grade. Although retained children performed below promoted children on the mathematics tasks overall, ANOVAs revealed no Time $\times$ Group interactions. That is, retained children grew at about the same rate on the mathematics tasks as promoted children in Year 2 of our study.

Discussion

Over a 16-month period, we investigated the development of mathematical competencies in children with different patterns of mathematics and reading achievement (i.e., MD only, MD–RD, RD only, and NA). We examined children's ending level of performance as well as their growth rates, using growth-curve modeling procedures. Of particular interest was the comparison between two subgroups of children with MD, namely children with MD only versus children with MD–RD.

Hanich et al.'s (2001) previous baseline work revealed that in mid second grade, children with MD only had an advantage over their MD–RD counterparts in two areas of mathematical cognition: exact calculation of arithmetic combinations and story problems. In the present study, which examined performance in the same children on the same tasks at the end of third grade, the MD only group continued to show an advantage over the MD–RD group on these two tasks. However, the end-of-third-grade performance difference between the two MD groups on exact calculation of number combinations disappeared when we considered predictor variables of IQ, gender, ethnicity, and income level. To determine whether the early difference between the MD only and MD–RD groups on exact calculation of number combinations at Time 1 reported by Hanich et al. was a consequence of IQ differences in particular, we performed the analysis again at Time 1 with IQ (as measured in third grade) as a covariate. Hanich et al's original findings held, even when we considered IQ. In the present study, the MD only group also performed better than the MD–RD group at the end of third grade on calculation principles, irrespective of predictor variables.

The tasks that did not differentiate the MD only group from the MD–RD group (either at the end of third grade in the present study or in second grade in Hanich et al., 2001) were forced retrieval of number facts, approximate arithmetic, place value, and written computation. The two MD groups did not differ in their growth rates (from mid second to the end of third grade) on any of the mathematics tasks.

What do the present findings tell us about the characteristics and growth of children with learning difficulties in mathematics? First, the difficulties of children identified as MD–RD in early second grade are pervasive and stable over second and third grades, even when we hold predictors, such as IQ, constant. Despite their weaknesses, however, children with MD–RD achieved on the mathematics tasks in this study at a rate similar to other children.

Jordan et al. (2002) reported that second graders with MD who are good readers achieved faster over a 2-year period on a standardized diagnostic mathematics achievement test than did children with MD who are poor readers. The findings suggested that MD only children grew out of or at least compensated for some of their earlier achievement weaknesses. One reason for this may be the MD-only group's strengths in problem solving, as evidenced by their advantage over children with MD–RD on calculation principles as well as on story problems in the present investigation. Children's understanding of calculation principles, such as the commutative and inversion principles, reflects a grasp of relationships within and between arithmetic operations (Hanich et al., 2001). The ability to solve story problems involves comprehending the words of the problem and translating the verbal information to mathematical representations (Geary, 2000; Jordan & Montani, 1997). Story-problem-solving skill is associated with verbal comprehension more generally, which reflects or is even determined by reading skill (Jordan et al., 2002; Stanovich, 1991).

It is interesting that children with MD only and RD only performed at about the same levels in problem solving throughout the present study. (The level was slightly below that of NA children who had no learning difficulties, however.) It is possible that children in the two achievement groups use different pathways to solving problems; that is, children with RD only may exploit their mathematical strengths to compensate for reading and verbal weaknesses, whereas children with MD only may take advantage of reading and verbal strengths to compensate for mathematical weaknesses. Children's performance may be differentiated further on a wider range of story-problem tasks, for example, tasks assessing the ability to translate and integrate several sentence problems and the ability to recognize superficial or irrelevant structures in story problems (Fuchs & Fuchs, 2002).

In the present study children with MD only appeared to have consistent difficulties in rapid fact retrieval and, by extension, calculation fluency. They performed as poorly as children with MD–RD when required to respond to number combinations quickly (on a forced-retrieval task) and relied on their fingers as much as children with MD–RD (on the exact calculation task), even at the end of third grade. It should be noted, however, that children with MD only used finger-counting strategies more accurately than did children with MD–RD, suggesting better facility with counting procedures in the MD only group than in the MD–RD group (Bull & Johnston, 1997; Geary et al., 1999; Jordan & Hanich, 2000; Jordan & Montani, 1997). Children with MD only seem to have trouble making the transition from procedure-based to memory-based problem calculation (Geary, 2003).

Our data do not support the suggestion that difficulties in reading and fact retrieval share a core-underlying deficit related to phonological processing (e.g., Geary, 1993; Hanich et al., 2001; Miles, 1993). When we isolated reading difficulties from mathematics difficulties, we found that children with RD only (who were characterized by decoding weaknesses) performed better than children with MD–RD on the forced-retrieval task but children with MD only, who are proficient readers, did not. The RD only group also performed as well as the NA group in forced retrieval. Fact-retrieval deficits appear to be a central characteristic of MD (i.e., both MD only and MD–RD) but not of RD (i.e., RD only), at least with respect to addition and subtraction operations.

It is possible to argue, however, that even with the stringent 3 s criterion on the forced-retrieval task some children used a strategy other than retrieval to get a correct answer rapidly (Baroody, 1999). Although we did not use "$n+1$" combinations, which can be solved quickly with a number-after rule principle, children may have used other calculation shortcuts related to relational knowledge (such as subtraction being the inverse principle: $4+2=6$ so $6–4=2$). On inversion problems, such as 28+36–36, Siegler and Stern (1998) found that second-grade children employed calculation shortcuts unconsciously, with solution times averaging less than 3 s. Regardless, however, of whether children used retrieval or shortcut strategies—conscious or unconscious—when solving timed arithmetic combinations, competence clearly depends on fast mental processing.

Cognitive mechanisms underlying retrieval deficits in children may include problems in representing and retrieving information in long-term memory (Ashcraft, 1992; Geary, 1993; Rasanen & Ahonen, 1995) and difficulties inhibiting the retrieval of irrelevant associations (Barrouillet, Fayol & Lathulière, 1997). Robinson, Menchetti, and Torgesen

(2002) have maintained that weakly consolidated number sense contributes to fact-retrieval, or fluency, deficits in children with MD. However, number sense is a general construct that encompasses several cognitive areas (e.g., Dowker, 1998). And the relatively strong performance of children with MD only in problem solving and in using counting procedures suggests that at least some elements of number sense are intact in this population. Children with MD only and MD–RD, however, did not differ significantly on approximate arithmetic in the present study, which required them to estimate answers to addition and subtraction problems. Weaknesses in spatial representations related to numerical magnitudes (rather than weaknesses in verbal representations) may underlie rapid fact-retrieval deficits. When solving approximate arithmetic (or estimation) problems the child must form a "mental number line" to manipulate and estimate quantities (Dehaene & Cohen, 1991). Neuropsychological studies indicate that estimation of quantities involves visual-spatial abilities that are independent of language (Dehaene & Cohen, 1991; Dehaene et al., 1999). Before learning conventional arithmetic combinations, normally developing children are proficient in estimating set sizes and using numerical reference points (Baroody & Gatzke, 1991). Whether very early deficits in numerical estimation are a marker for later mathematics difficulties should be investigated.

It is interesting that growth on approximate calculation was relatively flat across achievement groups, indicating that this task was hard for all children (although all groups performed above the chance level of 50%). The relatively low reliability estimate for approximate arithmetic also reflects its difficulty level. Estimation skills may receive little emphasis in school. In fact, research with college students reveals they are better at performing exact calculations than at estimating answers (Hanson & Hogan, 2000).

In terms of base-ten concepts (i.e., written computation and place value), there were no achievement-group differences at the end of third grade when we accounted for predictors in our model. Before we added predictors to the model, however, the MD–RD group performed worse than both the RD only and the NA groups on place value and worse than the NA group on written computation. The findings are in keeping with what Hanich et al. (2001) observed in second grade. The present study shows that knowing a child's IQ and gender add more to our understanding of group differences in these tasks than knowing only the group differences.

It is not surprising that knowing a child's IQ level is important for predicting performance on all of the mathematics tasks, beyond the influences of minority status and income level. (However, on five of the seven tasks, IQ added little information beyond our initial achievement-group classification.) Additionally, boys had a small but significant advantage over girls in tasks assessing place value and estimation, and they used their fingers less than girls on exact calculation of arithmetic combinations. Our gender findings are consistent with previous work showing few differences between third-grade boys and girls in mathematical competencies, except that girls use more concrete strategies than do boys when solving number combinations (Fennema, Carpenter, Jacobs, Franke & Levi, 1998). Mental computational strategies are closely related to computational estimation ability and are needed to develop a full understanding of place value (Sowder, 1998).

In conclusion, the results of the present study, together with earlier findings, suggest that deficiencies in fact retrieval, and by extension calculation fluency, are a defining feature of mathematics difficulties, specific or otherwise. However, we do not fully understand the impact of instructional intervention on the development of mathematical competencies in children with MD. In primary school most interventions are targeted at mathematics difficulties rather than at reading difficulties (Jordan et al., 2002). Although it is tempting to suggest that children with MD only—who have a greater range of competencies than do children with MD–RD—bypass their relatively circumscribed deficiencies in number fact mastery with calculators or other aids, it may be wiser to provide instruction aimed at fostering calculation fluency, in addition to methods that emphasize problem solving. Studies on the maintenance of mathematical competencies in adults indicate that the degree of extended rehearsal and practice provided during the school years is the best predictor of performance levels in adulthood (Bahrick & Hall, 1991). The extent to which the skill and growth trajectories of kindergarten and first-grade children in math and reading precursors predict later mathematics difficulties in second grade and beyond remains an open yet important question.

References

Arbuckle, J. L. (1999). *AMOS 4.0 Users' Guide*. Chicago: Smallwaters Inc.

Ashcraft, M. H. (1992). Cognitive arithmetic: A review of data and theory. *Cognition*, *44*, 75–106.

Bahrick, H. P., & Hall, L. K. (1991). Lifetime maintenance of high school mathematics content. *Journal of Experimental Psychology: General*, *120*, 22–33.

Baroody, A. (1999). Children's relational knowledge of addition and subtraction. *Cognition and Instruction, 17,* 137–175.

Baroody, A. J., & Gatzke, M. R. (1991). The estimation of set size by potentially gifted kindergarten-age children. *Journal for Research in Mathematics Education*, *22*(1), 59–68.

Baroody, A. J., & Tiilikainen, S. H. (2003). Two perspectives on addition development. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: The construction of adaptive expertise* (pp. 75–125). Mahwah, NJ: Erlbaum.

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

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods.* Newbury Park, CA: Sage.

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.

Carpenter, T. P., & Moser, J. M. (1984). The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education*, *15,* 179–202.

Dehaene, S., & Cohen, L. (1991). Two mental calculation systems: A case study of severe acalculia with preserved approximation. *Neuropsychologia, 29,* 1045–1054.

Dehaene, S., Spelke, E., Pinel, P., Stanescu, R., & Tsivkin, S. (1999). Sources, of mathematical thinking: Behavioral and brain-imaging evidence. *Science, 284,* 970–974.

Dowker, A. D. (1998). Individual differences in normal arithmetical development. In C. Donlan (Ed.), *The development of mathematical skills* (pp. 275–302). Hove, England: Taylor & Francis.

Fennema, E., Carpenter, T. P., Jacobs, V. R., Franke, M., & Levi, L.W. (1998). A longitudinal study of gender differences in young children's mathematical thinking. *Educational Researcher, 27*(5), 6–11.

Fuchs, L. S., & Fuchs, D. (2002). Mathematical problem-solving profiles of students with mathematics disabilities with and without co-morbid reading disabilities. *Journal of Learning Disabilities*, *35,* 563–573.

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. (2000). From infancy to adulthood: The development of numerical abilities. *European Child and Adolescent Psychiatry, 9,* II11–II16.

Geary, D. C. (2003). Learning disabilities in arithmetic: Problem solving differences and cognitive deficits. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 199–212). New York: Guilford.

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

Geary, D. C., Hamson, C. O., & Hoard, M. K. (2000). Numerical and arithmetical cognition: A longitudinal study of process and concept deficits in children with learning disability. *Journal of Experimental Child Psychology, 77,* 236–263.

Geary, D. C., & Hoard, M. K. (2001). Numerical and arithmetical deficits in learning-disabled children: Relation to dyscalculia and dyslexia. *Aphasiology, 15*(7), 635–647.

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.

Ginsburg, H. P. (1997). Mathematics learning disabilities: A view from developmental psychology. *Journal of Learning Disabilities*, *30*, 20–33.

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

Hanson, S. A., & Hogan, T. P. (2000). Computational estimation skill of college students. *Journal for Research in Mathematics Education*, *31*, 483–499.

Hiebert, J., & Wearne, D. (1996). Instruction, understanding, and skill in multidigit addition and subtraction. *Cognition and Instruction*, *14,* 251–283.

Jordan, N. C., & Hanich, L. B. (2000). Mathematical thinking in second-grade children with different types of learning difficulties. *Journal of Learning Disabilities*, *33*, 567–578.

Jordan, N. C., Hanich, L. B., & Uberti, H. (2003). Mathematical thinking and learning disabilities. In A. J. Baroody & A. Dowker (Eds.), *The development of arithmetic concepts and skills: The construction of adaptive expertise* (pp. 359–383). Mahwah, NJ: Erlbaum.

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

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.

Kamii, C. (1989). *Young children continue to reinvent arithmetic*. New York: Teacher College Press.
Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Newbury Park, CA: Sage.
Lemaire, P., Barrett, S. E., Fayol, M., & Abdi, H. (1994). Automatic activation of addition and multiplication facts in elementary school children. *Journal of Experimental Child Psychology, 57*(2), 224–258.
Miles, T. R. (1993). *Dyslexia: The pattern of difficulties* (2nd ed.). London: Whurr.
Miller, K. (1992). What a number is: Mathematical foundations and developing number concepts. In J. I. D. Campbell (Ed.), *The nature and origin of mathematical skills* (pp. 3–38). New York: Elsevier Science.
Muthén, B. (1991). Analysis of longitudinal data using latent variable models with varying parameters. In L. Collins & J. Horn (Eds.), *Best methods for the analysis of change: Recent advances, unanswered questions, future directions* (pp. 1–17). Washington, DC: American Psychological Association.
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.
Ostad, S. A. (1999). Developmental progression of subtraction strategies: A comparison of mathematically normal and mathematically disabled children. *European Journal of Special Needs Education, 14*, 21–36.
Rasanen, P., & Ahonen, T. (1995). Arithmetic disabilities with and without reading difficulties: A comparison of arithmetic errors. *Developmental Neuropsychology, 11*, 275–295.
Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousands Oaks, CA: Sage.
Riley, M. S., & Greeno, J. G. (1988). Developmental analysis of understanding language about quantities and of solving problems. *Cognition and Instruction, 5*, 49–101.
Riley, M. S., Greeno, J. G., & Heller, J. I. (1983). Development of childrens' problem-solving ability in arithmetic. In H. P. Ginsburg (Ed.), *The development of mathematical thinking* (pp. 153–196). New York: Academic Press.
Robinson, C. S., Menchetti, B. M., & Torgesen, J. K. (2002). Toward a two-factor theory of one type of mathematics disabilities. *Learning Disabilities Research & Practice, 17*(2), 81–89.
Ross, S. (1989). Parts, whole, and place value: A developmental view. *Arithmetic Teacher, 36*, 47–51.
Russell, R. L., & Ginsburg, H. P. (1984). Cognitive analysis of children's mathematics difficulties. *Cognition and Instruction, 1*, 217–244.
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., & Stern, E. (1998). Conscious and unconscious strategy discoveries: A microgenetic analysis. *Journal of Experimental Psychology: General, 127*, 377–397.
Sowder, J. T. (1998). Perspectives from mathematics education. *Educational Researcher, June-July*, 12–13.
Stanovich, K. E. (1991). Discrepancy definitions of reading disability: Has intelligence led us astray? *Reading Research Quarterly, 26*, 7–29.
Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: Psychological Corporation.
Willett, J. B., & Sayer, A. G. (1994). Using covariance structure analysis to detect correlates and predictors of individual change over time. *Psychological Bulletin, 116*, 363–381.
Woodcock, R. W., & Johnson, M. B. (1990). *Woodcock–Johnson Psycho-Educational Battery–Revised*. Allen, TX: DLM Teaching Resources.

Appendix

In this appendix, we provide a brief didactic discussion of the method of growth-curve modeling. For a more detailed overview of growth-curve modeling, see Raudenbush and Bryk (2002), and for an application to research in MD, see Jordan et al. (2002). Briefly, growth-curve models can be viewed as falling within the class of multilevel linear models (i.e., multilevel linear models; Raudenbush & Bryk, 2002) where Level 1 represents intra-individual differences in status and growth rate and Level 2 represents interindividual differences in status and growth rate. Consider a simple growth model for any one of the mathematics tasks for person p at time i, which we denote here as $MATH_{ip.}$ We write the Level 1 equation expressing the math task scores over time within an individual as

$$MATH_{ip} = \pi_{0p} + \pi_{1p}t_i + \varepsilon_{ip}, \tag{A1}$$

where π_{0p} represents the math task score for person p at time $t = 0$. For this study, the reference point at $t = 0$ (so-called status point) will be the last wave of testing. Continuing, t_i represents a temporal dimension that is assumed to be the same for all individuals. In our study, the temporal dimension is months, with students measured in January, April, November, and May. Note that although these testing dates represent unequal time intervals, this presents no difficulty for the estimation of growth-curve models. The parameter π_{1p} represents the linear growth rate over the 16 months of testing. Finally, ε_{ip} is the error term in the Level 1 equation representing the influence of omitted variables for student p at time i.

The simple linear growth-curve model in Equation A1 can be extended to incorporate a quadratic trend (acceleration) in the growth trajectory. To do so, Equation A1 is rewritten as

$$MATH_{ip} = \pi_{0p} + \pi_{1p}t_i + \pi_{2p}t_i^2 + \varepsilon_{ip}, \tag{A2}$$

where π_{2p} is the acceleration rate.

A major benefit of growth-curve modeling is that it can be extended to handle predictors of individual differences in the growth parameters (ending status, linear growth rate, and acceleration rate). In the case of Equation A2, three models are specified: one for the end of testing status parameter, one for the linear growth rate parameter, and one for the acceleration parameter. Predictors of the growth parameters are referred to as time-invariant predictors. In our study, achievement-group membership, gender (male = 1), ethnicity (minority = 1), income level (low income = 1), and IQ centered on the sample mean were used as time-invariant predictors.

In our study, the time-invariant predictors of major interest were group membership and the remaining predictors are used as controls. Because achievement-group membership was dummy coded with MD–RD serving as the reference group, three dummy-coded vectors were specified. Thus, the Level 2 model for the exit status, growth rate, and acceleration, can be written as

$$\begin{aligned}\pi_{0p} = \mu_{ES} &+ \gamma_{1,ES}(NA)_p + \gamma_{2,ES}(RD\ only)_p \\ &+ \gamma_{3,ES}(MD\ only)_p + \gamma_{4,ES}(MALE)_p \\ &+ \gamma_{5,ES}(ETHNICITY)_p \\ &+ \gamma_{6,ES}(INCOME)_p + \gamma_{7,ES}(IQ)_p + \zeta_{0p},\end{aligned} \tag{A3}$$

$$\begin{aligned}\pi_{1p} = \mu_{GR} &+ \gamma_{1,GR}(NA)_p + \gamma_{2,GR}(RD\ only)_p \\ &+ \gamma_{3,GR}(MD\ only)_p + \gamma_{4,GR}(MALE)_p \\ &+ \gamma_{5,GR}(ETHNICITY)_p \\ &+ \gamma_{6,GR}(INCOME)_p + \gamma_{7,GR}(IQ)_p + \zeta_{1p},\end{aligned} \tag{A4}$$

and

$$\begin{aligned}\pi_{2p} = \mu_{AC} &+ \gamma_{1,AC}(NA)_p + \gamma_{2,AC}(RD\ only)_p \\ &+ \gamma_{3,AC}(MD\ only)_p\, \gamma_{4,AC}(MALE)_p \\ &+ \gamma_{5,AC}(ETHNICITY)_p \\ &+ \gamma_{6,AC}(INCOME)_p + \gamma_{7,AC}(IQ)_p + \zeta_{2p},\end{aligned} \tag{A5}$$

where μ_{ES}, μ_{GR}, and μ_{AC} are intercept parameters representing mean growth parameters for White female MD–RD students with IQ scores at the mean of the sample. (Note that this interpretation of the intercept is arbitrary, depending only on the coding of the variables.) The parameters γ_1, γ_2, and γ_3 give the mean differences in the growth parameters between the specific achievement groups and the MD–RD group, holding constant gender, ethnicity, income, and IQ; the coefficient γ_4 gives the mean difference between boys and girls on the growth parameters, holding constant achievement-group membership, ethnicity, income, and IQ; the coefficient γ_5 gives the mean difference between minority and nonminority students on the growth parameters, holding constant achievement-group membership, gender, income, and IQ; the coefficient γ_6 gives the mean difference between low-income and middle-income students on the growth parameters, holding constant achievement-group membership, gender, ethnicity, and IQ; and the coefficient γ_7 gves the effect of IQ on the growth parameters, holding constant achievement-group membership, gender, ethnicity, and income. The ζs are error terms containing the effects of omitted variables in their respective equations. The derivations of these parameters are discussed in detail elsewhere (e.g., Bryk & Raudenbush, 1992).

Research by Muthén (1991) and Willett and Sayer (1994) has shown how the models in Equations A1 through A5 can be incorporated into a structural equation modeling framework (see also Kaplan, 2000). For this paper, we used the structural equation modeling framework, employing the software program AMOS (Arbuckle, 1999).