Journal of Psychoeducational Assessment
Volume 27 Number 3
June 2009 175-196

10.1177/0734282908330578
http://jpa.sagepub.com
hosted at
http://online.sagepub.com

Math Disabilities and Reading Disabilities

Can They Be Separated?

H. Lee Swanson
Olga Jerman
Xinhua Zheng
University of California–Riverside

This article synthesizes some of the published literature that selectively compares the cognitive functioning of children with math disabilities (MDs) with average-achieving children and poor readers (children with reading disabilities [RDs] or comorbid disabilities [RDs + MDs]). All studies in the synthesis report reading, IQ, and math scores for children with MDs and poor readers. A random coefficients model of effect sizes (ESs) show that (a) ESs between MD and normal achievers were moderated by variations in working memory and literacy, (b) ESs between MD- and RD-only children were moderated by working memory and problem solving, and (c) ESs between MD and MD + RD children were moderated by long-term memory and IQ scores. No support was found for the notion that the differentiation between MD children and poor readers (RD and MD + RD) was related to variations in reading across the reviewed studies.

Keywords: *math disabilities; meta-analysis; memory; arithmetic*

Individual differences in mathematical cognition have been studied extensively by cognitive psychologists during recent decades. Furthermore, it is hard not to overemphasize the importance of mathematical ability in a society that requires technical competence among its citizens, a competence that in turn draws on high levels of mathematical literacy. Unfortunately, a significant number of children demonstrate poor achievement in mathematics. Several studies (Badian, 1983; Gross-Tsur, Manor, & Shalev, 1996) estimate that approximately 6% to 7% of the school-age population has math disabilities (MDs). Several studies suggest that the incidence of MDs may be as common as reading disabilities (RDs) (e.g., Geary, 1993).

Numerous studies show that children with MDs do not perform as well as their same-age peers on various cognitive tasks. A recent meta-analysis that synthesized published studies

Authors' Note: This study was partially supported by a U.S Department of Education, Cognition and Student Learning (USDE R305H020055), Institute of Education Sciences grant awarded to the first author. The authors are indebted to Georgia Doukas, Rebecca Gregg, Crystal Howard, and Krista Healy for coding articles, establishing interrater agreements, and/or critical comments on earlier versions of this article. The report does not necessarily reflect the view of the U.S. Department of Education. Please address correspondence to H. Lee Swanson, Graduate School of Education, Area of Educational Psychology, University of California, Riverside, Riverside, CA 92521; e-mail: lee.swanson@ucr.edu.

by comparing children with MDs to average achievers (Swanson & Jerman, 2006) found that the magnitude of differences between the two groups across 194 effect sizes (ESs) was approximately half a standard deviation (M = –.52, SE = .01; negative score was in favor of the average achievers). According to Cohen's (1988) classification of ESs (large = .80, moderate = .50, and small = .20), the comparisons between children with and without MDs were in the moderate range. Of course, this finding cannot be taken at face value because tremendous variability in the magnitude of ESs emerged across different cognitive categories or domains. A comparison of the two groups on ESs across various specific categories isolated major differences on measures related to verbal problem solving (M = –.58), naming speed (M = –.70), verbal working memory (WM) (M = –.70), visual-spatial WM (M = –.63), and long-term memory (LTM) (M = –.72). The magnitude of these ESs was persistent across age and severity of MDs. More important, the magnitude of ESs in overall cognitive functioning between MD children and average achievers was primarily related to verbal WM deficits when the effects of all other variables (e.g., age, IQ, reading level, other domain categories) were partialed out. The results were in line with several study findings suggesting that MDs, when compared to normal achieving children, can be partly attributed to WM deficits (e.g., Geary, 2003; Swanson, 1993).

Unfortunately, in the Swanson and Jerman (2006) synthesis, no clear-cut differences emerged between children with arithmetic and reading difficulties on cognitive measures. The results suggested that ESs were in favor of children with RDs when compared to children with MDs, but the advantages for children with RDs were isolated to measures of naming speed (M = –.23) and visual-spatial WM (M = –.30). The overall findings were problematic, as several studies have suggested that children with RDs can be separated from children with MDs (e.g., Jordan, Hanich, & Kaplan, 2003). The poor differentiation between children with MDs and those with reading difficulties may have occurred because the studies included samples with poor arithmetic skills accompanied by relatively low reading skills. Therefore, it was difficult to determine whether results attributed to MDs were in fact due to arithmetic difficulties or whether they were outcomes related to generally poor academic skills that shared the same process that incorporated both reading and math skills. Another difficulty in the findings was that the variance between studies was not considered in the overall analysis. That is, the overall estimates of ESs did not take into account between-study and within-study variance in the calculation of the observed ESs.

In general, it has been argued that contrasts between children with MDs and RDs are not well understood because reading performance has not been controlled across studies (e.g., see Jordan, 2007, for a review). Thus, there is a question as to whether children with MDs suffer from the same processes associated with RDs. For example, Jordan (2007), in her synthesis of the literature, argued that some authors have incorrectly assumed that MDs are related to language, which in turn suggests some commonality between math and reading. Several studies (e.g., Geary, Hamson, & Hoard, 2000; Landerl, Bevan, & Butterworth, 2004) suggested that all children with MDs, with or without reading problems, showed general deficits in number processing. However, some research suggests that children with RDs only (children with known phonological deficits) do better than those with MDs or MDs and RDs (MDs + RDs) on rapid fact retrieval (e.g., Geary et al., 2000). Other authors also find evidence from other researchers (Fuchs & Fuchs, 2002) that problem solving rather than number and arithmetic skills differentiated children with MDs from children with MDs + RDs. Thus, some authors suggest that what differentiates the two disabilities is the ability to solve complex word problems (e.g., Jordan et al., 2003).

Therefore, based on the previous meta-analysis, the primary cognitive mechanisms that separate MDs from RDs remain unclear. In this study, we reanalyze the meta-analysis originally reported in Swanson and Jerman (2006). However, we placed several restrictions in the sampling of articles selected from this synthesis. First, only studies that yield samples with MDs below the 25th percentile (standard score of 90) were analyzed. The 25th percentile cutoff score is a common measure of risk (e.g., Fletcher, 1985; Koontz & Berch, 1996; Swanson, 1993). The previous synthesis was problematic because the operational criteria for defining MDs (as well as RDs) varied across studies. Cutoff scores for defining MDs in the aforementioned synthesis varied from the 35th percentile to the 8th percentile. By establishing a rigorous standard, we assume that a great deal of the variability across studies would be controlled. Second, only studies that included children with RDs or RDs + MDs compared on cognitive measures were included in the analysis. That is, we attempt to estimate the influence of variations in reading on the cognitive differences between children with MDs and those with reading-only or comorbid difficulties. Specifically, the analysis seeks to determine the cognitive mechanisms that separate the two groups. Finally, the issue of study variance is dealt with more directly. The aforementioned synthesis did not take into consideration the variance that existed across studies. Because study-level variance was assumed to be present as an additional source of random influence in the comparisons, a random effects model was used to analyze ESs to take these sources of variance into account.

In summary, we attempt to answer three questions about the literature that were not addressed in the Swanson and Jerman (2006) synthesis:

1. Are cognitive deficits in children with MDs distinct from those in children with RDs?
2. Are the cognitive deficits a function of variations in age? The majority of studies that have compared children with MDs have focused on the elementary grades. However, we would like to determine if some of the same deficits emerge in studies that include older participants.
3. Do the cognitive deficits that emerge in children with MDs and RDs vary as a function of definitional criteria? We compare studies on cognitive outcomes as a function of severity of the MD and intelligence level.

Method

Identification of Studies

Several approaches were used to locate the relevant studies published in peer-reviewed journals, as discussed in Swanson and Jerman (2006). We will briefly summarize the procedures used for the article selection. The principle method of locating studies comparing children with MDs on psychological variables compared to controls involved a computer search of the PsychINFO, MEDline, and ERIC databases. The search used the following terms: *math disabled*, *math disabilities*, *dyscalculia*, *less skilled math*, *math disabled/reading disabled*, *arithmetic disabled*, *poor problem solvers*, *problem solving in math*, and *problem* *solving and math*. Second, a manual search was conducted of journals in which the majority of articles were published (e.g., *Journal of Learning Disabilities*, *Journal of Experimental Child Psychology*). Finally, a hand search was done on all studies cited in the aforementioned articles. Collectively, these methods identified more than 300 articles published between 1970 and 2003. The pool of literature was then narrowed down to 85 potentially relevant studies based on selecting only comparative studies (e.g., children with MDs compared with an average-achieving non-math-disability [NMD] group) and studies from peer-reviewed journals.

Article Inclusion Criteria

The 85 potential studies were further narrowed to 28. Studies were excluded if (a) they were not published in refereed journals, (b) they failed to provide enough quantitative data to calculate the ESs, (c) they failed to include a nondisabled average-math-achieving comparison group, (d) they failed to provide information of ability group performance on a standardized (norm referenced) math and/or IQ test, and (e) they failed to provide comparison measures independent of the classification measures. From this pool of 28 articles, we eliminated articles that reported an aggregated (mean) standard score of 90 or higher on a norm-referenced standardized math test for their MD sample. We next eliminated articles that failed to provide a comparison group with RDs. Overall, from the 28 studies, 11 studies met our criteria. Of the 11 studies, 8 studies included comorbid subgroups (RD + MD) and 7 included an RD-only subgroup. Four studies included both RD-only and RD + MD samples. The psychometric characteristics of the comparisons are provided in Table 1. One of the studies (McLean & Hitch, 1999) included a group of children with MDs who were compared to a group of younger ability-matched students with significantly lower reading scores. This study was kept in the analysis because chronological age was controlled in the regression analysis. Removal of this study did not change the pattern of results.

Coding Procedure

Each study was coded for the following information: (a) sample characteristics, (b) classification measures, and (c) cognitive measures.

Attributes of the study. Each study provided (a) the year of publication, (b) the name of the first author, (c) the number of coauthors, and (d) the country in which the study was conducted.

Attributes of the participants. There were four identified subgroups: normal-achieving control group, math disabled, reading disabled, and math and reading disabled. According to the inclusion criteria, each study provided at least (a) one MD and one NMD comparison group. Other attributes of the participants that were coded included (b) the number of participants in each subgroup, (c) the number of males in each subgroup, (d) the mean age of the group (converted into months), and (e) the participants' primary language. Studies were also coded for (f) socioeconomic status (SES) and (g) ethnicity.

Table 1
Psychological and Demographic Information on Participants

	Non-MD[a]		MD[b]		RD[c]		MD + RD[d]	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Age	116.76	30.87	117.99	30.79	101.50	24.09	113.25	21.49
IQ	102.67	8.15	98.05	8.23	97.06	8.11	100.77	12.13
Math	106.78	6.46	84.58	4.06	97.58	7.11	91.09	9.29
Reading	107.16	5.42	101.65	4.95	81.09	7.16	90.15	13.57

Note: MD = math disability; RD = reading disability. Age is given in months.
a. $N = 363$.
b. $N = 195$.
c. $N = 183$.
d. $N = 178$.

Comparison measures. All classification measures (IQ, mathematics, and reading) were converted to standard scores. In those cases when only a range was reported, a middle value was assigned. The dependent measures were aggregated into 10 broader domains. Specifically, the domain of literacy included measures related to comprehension, word attack, phonological awareness, writing, and spelling. Measures of rapid naming of objects, letters, and numbers were included under the category of naming speed. The visual-spatial category included measures of both visual-motor and non-visual-motor tasks. Although some of the analyzed categories were closely related (e.g., literacy, problem solving) to the classification variables (i.e., arithmetic calculation for children with MDs and word recognition for children with RDs), only the tasks not used in the classification were used for calculating the ESs. Children with MDs were compared to their counterparts on measures related to the following 10 categories:

1. *Literacy–reading*. The majority of dependent measures in this domain included reading comprehension, writing, vocabulary, and phonological awareness. Tasks presented within this domain required reading, listening and comprehension, vocabulary, phonological processing (e.g., phonemic deletion task), spelling, and recognizing visual forms of words or sounds.
2. *Problem solving–verbal*. This domain included measures of accuracy in solving story problems.
3. *Speed*. This domain included measures of the rapid naming of letters, numbers, and objects and speed measures such as coding.
4. *Problem solving–visual motor*. This domain included measures that required the manual manipulation of objects (e.g., blocks, discs, puzzles) to solve a problem (e.g., Tower of Hanoi).
5. *LTM–general information*. This domain included measures that tapped previous knowledge or memory for general information (e.g., answer questions such as “What’s the capital of California?” or recall a story they heard).
6. *Short-term memory (STM)–words*. This domain included tasks that required the recall of increasingly difficult sets of words and letters. This domain varied from verbal WM (below) in that no distracter question was asked of the participant prior to retrieval.
7. *STM–numbers*. This domain included tasks that required the recall of increasingly difficult sets of digits. This domain varied from WM in that no distracter question was asked of the participant prior to retrieval.
8. *WM–verbal*. This domain included tasks that required the recall of increasingly difficult sets of words and sentences. This domain varied from verbal STM in that process and storage components

were included. An example of a verbal WM test was a semantic association task, in which a child was presented with a set of words reflecting different categories (word sets range from two to nine monosyllabic words). Before recalling the words, however, the participant was asked whether a particular word or word category was included in the set.

9. *WM–visual spatial*. This domain included tasks that required the recall of increasingly difficult sets of dots, designs, and objects. An illustration of spatial WM was the Visual Matrix Task, in which a participant was presented with a series of dots in a matrix and was allowed 5 s to study the matrix. The matrix was then removed, and the participant was asked a process question (e.g., "Are there any dots in the first column?"). After answering the process question, the child was asked to draw the dots in the correct boxes on the blank matrix.
10. *Attention*. This domain included observations that focused mostly on classroom behavior, for which measures of attention or behavior were recorded. For example, in one study, Conners's Continuous Performance Test was used to assess sustained attention; in another study, a developmental questionnaire provided information on participants' activity level, impulsivity, attention, and inattention.

Calculation of ESs

For each measure, an ES was computed (Cohen's [1988] *d*) and was then weighted by the reciprocal in the sampling variance (Hedges & Olkin, 1985). The dependent measure for the estimate of ES was defined as est = d/(1/v), where *d* is the mean the of MD group minus the mean of comparison group divided by the average standard deviation for both groups and *v* is the inverse of the sampling variance, $v = (N_{md} + N_{nmd})/(N_{md} \times N_{nmd}) + d^2/[2(N_{md} + N_{nmd})]$ (Hedges & Olkin, 1985). Means and standard deviations were used in the computation of 98% of the ESs. In the remaining cases, *F* ratios and *t* ratios were converted to ESs. Thus, ESs were computed with each ES weighted by the reciprocal of its variance, a procedure that gives more weight to ESs that are more reliably estimated. The overall results for MD, when compared to NMD, reading disabled, and children with MDs + RDs, are shown in Table 2. As suggested by Hedges and Olkin (1985), outliers were removed from the analysis of main effects. Outliers were defined as ESs lying beyond the first gap of at least 1 standard deviation between adjacent ES values in a positive direction (Bollen, 1989). Cohen's criterion was used for the interpretation of the magnitude of the ESs.

We also determined whether a set of *d*s shared a common ES (i.e., was consistent across the studies) by category. The analysis of each category of measure reported separately is shown in Table 2. For the category of each dependent measure, a homogeneity statistic Q was computed to determine whether separate ESs within each category shared a common ES (Hedges & Olkin, 1985). The statistic Q has a distribution similar to the distribution of chi-square with $k - 1$ degrees of freedom, where k is the number of ESs. A significant chi-square indicated that the study features significantly moderated the magnitude of ESs. If the homogeneity was not achieved, then the influence of outliers was assessed using a 95% confidence interval. Because we expected the absence of homogeneity, the subsequent analyses determined how the characteristics of the sample (e.g., IQ, reading) of the various studies contributed to the variability and the heterogeneity of ESs. To determine the relationship between sample characteristics and the magnitude of ESs, a categorical model was analyzed. Categorical models, analogous to an analysis of variance, show whether the heterogeneity in ESs was isolated to a particular

Table 2
Weighted Effect Sizes, Standard Errors, Confidence Intervals, and Homogeneity of Categories for Comparisons Between MD and NMD, MD and RD, and MD and RD + MD (Corrected for Outliers)

Comparison	*K*	Effect Size	*SE*	Lower	Upper	Homogeneity Q
Total across categories						
MD / NMD	88	–0.58	.03	–0.64	–0.52	395.20***
MD / RD	50	–0.14	.04	–0.23	–0.06	220.46***
MD / MD + RD	68	0.25	.03	0.18	0.32	800.36***
1. Literacy (vocabulary, reading comprehension)						
MD / NMD	10	–1.03	.10	–1.24	–0.82	31.82***
MD / RD	2	0.13	.16	–0.18	0.46	0.29
MD / MD + RD	10	–0.07	.10	–0.26	0.12	64.93**
2. Problem solving–verbal						
MD / NMD	5	–1.18	.12	–1.42	–0.94	27.33**
MD / RD	1	0.59	—	—	—	—
MD / MD + RD	5	–0.31	.16	–0.62	0.002	211.93***
3. Speed-naming						
MD / NMD	11	–0.39	.08	–0.57	–0.22	22.70*
MD / RD	6	–0.43	.15	–0.74	–0.12	2.24
MD / MD + RD	11	0.07	.09	–0.11	0.27	123.22***
4. Visual-spatial problem solving						
MD / NMD	6	–0.46	.12	–0.71	–0.21	44.07**
MD / RD	4	–0.11	.16	–0.42	0.19	29.47***
MD / MD + RD	3	0.63	.15	0.32	0.94	17.36***
5. Long-term memory (e.g., general information)						
MD / NMD	12	–0.87	.12	–1.11	–0.62	41.33***
MD / RD	0					
MD / MD + RD	12	0.58	.09	0.38	0.77	30.10***
6. STM–words						
MD / NMD	1	–0.89	—	—	—	—
MD / RD	0	—	—	—	—	—
MD / MD + RD	1	–0.32	—	—	—	—
7. STM–digits/numbers						
MD / NMD	7	0.07	.10	–0.13	0.28	14.07*
MD / RD	2	0.10	.27	–0.43	0.65	.15
MD / MD + RD	7	–0.55	.13	–0.82	–0.28	73.38***
8. WM–verbal						
MD / NMD	20	–0.53	.06	–0.65	–0.41	69.65***
MD / RD	20	–0.01	.07	–0.15	0.12	127.74***
MD / MD + RD	8	0.88	.08	0.71	1.06	51.99***
9. WM–visual spatial						
MD / NMD	13	–0.63	.07	–0.77	–0.46	28.24***
MD / RD	13	–0.30	.08	–0.46	–0.14	23.43*
MD / MD + RD	8	0.42	.08	0.25	0.59	8.68
10. Attention						
MD / NMD	3	0.09	.13	–0.17	0.35	11.53
MD / RD	2	–0.66	.16	–0.99	–0.33	3.26
MD / MD + RD	3	–0.31	.13	–0.58	–0.05	66.49**

Note: MD = math disability; NMD = non-math-disability–average achiever; RD = reading disability; MD + RD = comorbid group with both low reading and math; *K* = number of measures; lower and upper = 95% level of confidence range; WM = working memory; STM = short-term memory. Negative effect size is in favor of contrast group, and positive effect size is in favor of MD group.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

category. The procedure for calculating categorical models provides a between-class effect. This procedure was considered helpful in determining whether certain characteristics of the sample (e.g., age) made a significant contribution to ES.

Interrater Agreement

Three doctoral students coded studies. The interrater agreement for article inclusion in the analysis was 100%. The overall structure of the coding system yielded a reliable percentage of interrater agreement across all codes (more than 90% agreement). The first author trained the doctoral students in coding the articles.

Results

Study Characteristics

Articles were published most frequently in the *Journal of Experimental Child Psychology*, *Journal of Clinical & Experimental Neuropsychology*, *Journal of Learning Disabilities*, and *Learning Disability Quarterly*. No study separated the math performance as a function of gender, ethnicity, or SES. Therefore, math performance as a function of gender, ethnicity, and/or SES was not compared across the studies. Table 1 provides an overview on the psychometric information (IQ, math, and reading) of participants for the four groups: NMD, MD, RD, and comorbid (RD + MD). The most common intelligence measures were taken from the Wechsler series, and the most common math measures were taken from various revisions of the Wide Range Achievement Test.

The mean ESs for the MD / NMD was –.66 ($K = 88$, $SD = .71$), and the absolute ES (e.g., replacing timed and error responses as positive values) was .77 ($SD = .59$). The mean ESs for the MD / RD was –.23 ($K = 50$, $SD = .70$), and the absolute ES was .54 ($SD = .45$). The mean ESs for the MD / MD + RD was .06 ($K = 68$, $SD = 1.59$), and the absolute ES was 1.08 ($SD = 1.18$). When ESs were averaged within studies, the mean ESs for the MD / NMD was –.23 ($K = 11$, $SD = .62$), and the absolute ES was .79 ($SD = 1.18$). The mean ESs for the MD / RD was –.12 ($K = 7$, $SD = .66$), and the absolute ES was .58 ($SD = .33$). The mean ESs for the MD / MD + RD was .05 ($K = 8$, $SD = .52$), and the absolute ES was .88 ($SD = .53$). Thus, a tremendous amount of variability existed when the unit of analysis shifted from the aggregation of ESs within studies versus an aggregation across all measures. In addition, when the metric included the absolute values (controlling for the direction of the ESs), all the ESs were in the moderate to large range.

Domain Categories

Table 2 provides the weighted means (weighted by the reciprocal in the sampling variance) and standard deviations for ESs for each category and comparison. As shown in Table 2, there were 88 ESs in which we could establish comparisons between MD and NMD children. These dependent measures were averaged and yielded a weighted mean ES of –.58 ($SE = .03$). Negative ESs indicated that the NMD group did better than the MD group.

Table 3
Correlations of Effect Size With Variations in Age, Reading, and Math

	ES NMD vs. MD ($K = 88$)	ES MD vs. RD ($K = 50$)	ES MD vs. MD + RD ($K = 68$)
Math disabled			
Age	–.14	.005	–.26
IQ level	–.24*	–.36**	.18
Math level	.08	–.11	.37**
Reading level	–.17	–.08	.41**
Poor readers			
Low read	–.01	–.15	.17
High read	–.15	–.39**	.48**
Ave read	–.11	–.33*	.47**
Low math	.10	–.28*	.43**
High math	–.17	–.53**	.51***
Ave math	.03	–.52**	.56***
Low IQ	–.18	–.39**	.26*
High IQ	–.17	–.64***	.45***
Ave IQ	–.22*	–.63***	.48**

Note: ES = effect size; NMD = non-math-disability–average achiever; RD = reading disability; MD = math disability; Ave = average of reading scores if both RD only and RD + MD only in the study.
*$p < .05$. **$p < .01$. ***$p < .001$.

Table 4
HLM Regression Predicting Effect Sizes for All Cognitive Measures Comparing MD With Various Groups: Unconditional Model

	NMD vs. MD		MD vs. RD		MD vs. MD + RD	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Fixed effects						
Intercept	–.67***	.11	–.09	.19	.10	.20
Random effects						
Study[a]	.0		.10	.42	.12	.13
Domain (study)[a]	.19*	.11	.20	.35	.34*	.19
Residual[b]	.36***	.11	.16***	.03	.26***	.05
Deviance	184.6		75.4		121.6	
AIC	190.6		83.4		129.6	
BIC	191.8		85.0		131.2	

Note: NMD = non-math-disability–average achiever; MD = math disability; RD = reading disability; HLM = hierarchical linear model; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion.
a. Variance between studies.
b. Variance within studies.
*$p < .05$. ***$p < .001$.

Table 5
HLM Regression Predicting Effect Sizes for All Cognitive Measures Comparing MD With Various Groups: Conditional Model

	NMD vs. MD		MD vs. RD		MD vs. MD + RD	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Fixed effects						
Intercept	–.67***	.11	–.12	.28	.04	
1.24						
Age	–.001	.006	.006	.01	–.03**	.01
IQ	.02	.05	–.02	.04	–.30*	.11
Read	.007	.04	–.07	.10	.19	.32
Math	–.01	.04	.03	.04	.05	.36
Across sample						
Ave read	–.02	.03	–.02	.03	.006	.03
Ave math	.07	.05	–.06	.04	–.21	.16
Ave IQ	–.06	.05	–.005	.04	.44*	.15
Category						
WM–verbal	.27*	.12	.27*	.13	–.39	.90
WM–visual	.28*	.14	.29*	.13	–.58	.90
STM–words	.16*	.09	—	—	—	—
STM–digits	.10	.06	—	—	–.004	.05
LTM	.16*	.08	—	—	–.23*	.08
Speed	.14*	.06	.09	.08	–.05	.06
Literacy	.23*	.07	—	—	–.04	.06
Problem solving	.17*	.06	–.25**	.09	–.12	.07
Random effects						
Study[a]	0	—	0	—	0	—
Domain (study)[a]	0	—	0	—	0	—
Residual[b]	.34**	.05	.13*	.02	.24**	.03
Deviance	156.8		41.5		85.5	
AIC	190.8		65.5		117.5	
BIC	197.6		70.3		123.9	

Note: MD = math disability only; NMD = non-math-disability–average achiever; RD = reading disability only; MD + RD = comorbid group with both low reading and math; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; WM = working memory; LTM = long-term memory; STM = short-term memory; Ave = average of reading scores if both RD only and RD + MD only in the study; HLM = hierarchical linear model.
a. Variance between studies.
b. Variance within studies.
$^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

A chi-square was computed and weighted for the sample size. When comparing the weighted ESs, a significant effect was found for domain or category of the outcome measure, $\chi^2(9, N = 87) = 53.77$, $p < .001$. A Scheffé test indicated that ESs were significantly higher (in favor of children without MDs) for literacy and problem solving relative to the other categorical domains (literacy = problem solving–verbal > problem solving–visual = naming speed = problem solving–visual = LTM–WNM–verbal = WM–visual = STM–word > STM

Table 6
HLM Regression Predicting Effect Sizes for All Cognitive Measures Comparing MD With Various Groups: Model Reduction

	NMD vs. MD		MD vs. RD		MD vs. MD + RD	
	Estimate	*SE*	Estimate	*SE*	Estimate	*SE*
Fixed effects						
Intercept	–.78**	.11	–.27	.16	–.13	.10
Age					–.01**	.002
IQ					–.06**	.02
Read						
Math						
Across sample						
Ave read						
Ave math	.05**	.01	–.09**	.01		
Ave IQ	–.04**	.01			.09**	.01
Category						
WM–verbal	.15*	.06	.27**	.12		
WM–visual	.16*	.07	.28 **	.13		
STM–word						
STM–digits						
LTM	.04	.04			.16**	.04
Speed	.05	.03				
Literacy	.12**	.04				
Problem solving	.07	.03	–.27**	.09		
Random effects						
Study[a]	.01	.02	.03	.03	0	
Domain (study)[a]	0		0		.04	.05
Residual[b]	.37**	.05	.16**	.03	.28**	.06
Deviance	165.6		57.4		101.1	
AIC	187.6		71.4		117.0	
BIC	192.0		74.2		120.1	

Note: MD = math disability only; NMD = non-math-disability–average achiever; RD = reading disability only; MD + RD = comorbid group with both low reading and math; AIC = Akaike's Information Criterion; BIC = Bayesian Information Criterion; WM = working memory; LTM = long-term memory; STM = short-term memory; Ave = average of reading scores if both RD only and RD + MD only in the study; HLM = hierarchical linear model.
a. Variance between studies.
b. Variance within studies.
$*p < .05$. $**p < .01$.

digits = attention). However, these results should be interpreted with caution because of the infrequent number of ESs (e.g., STM–word) and unaccounted variance between and within studies.

As shown in Table 2, the weighted ES comparing MD- and RD-only children on cognitive measures was –.14 (*SE* = .04). Although ESs were generally in the low range, disadvantages were found for MD children when compared with RD children on measures of naming speed (*M* = –.43). Although an advantage was found for children with MDs for verbal problem solving (e.g., word problems), only one ES was reported. When comparing the weighted ESs, a significant effect was found for domain, $\chi^2(7, N = 49) = 33.85$, $p < .01$. A Scheffé test indicated that the negative ESs in favor of the RD group emerged for behavioral measures of attention and naming speed.

As also shown in Table 2, a comparison was also made between children with MDs and children with MDs + RDs. Sixty-eight dependent measures were averaged and yielded a mean ES in favor of the children with MDs of .25 (SE = .03). Positive ESs in Table 2 indicated that the children with MDs did better than the comorbid group. As shown in Table 2, MD children did better than the comorbid group (moderate ES range) on measures of problem solving–visual spatial, LTM–general information, and WM–visual spatial and had a large ES on measure of WM–verbal. MD children were inferior to the comorbid group on measures STM (digits and words), problem solving (verbal), and literacy. For the weighted ES, a significant ($p < .05$) effect was found for domain, $\chi^2(9, N = 67) = 152.19$, $p < .001$. A Scheffé test indicated that ESs were significantly larger ($p < .05$) for the LTM and measures of WM.

We further explored whether variations in reading, math, and IQ were related to the overall magnitude of the ES. Table 3 shows the correlation between the ESs of children with MDs and average achievers, children with MDs vs. RDs only, and MDs vs. RDs + MDs. These ESs were correlated with variations in MD children's age, math, and reading scores across studies. Because all studies had a poor reading sample (e.g., RDs only, RDs + MDs), ESs were correlated with variations in reading, math, and IQ scores in the poor reading samples. Across the studies, three values were computed within each study: highest, lowest, and average reported score. The highest value was the highest reading score reported for either the RD or the MD + RD group. The lowest level was the lowest mean reading score reported for either the RD or the MD + RD group. The average score was the mean value averaged across the RD or the MD + RD group. This comparison was done so we would not need to establish artificial cutoff scores and would allow for variations across the comparative samples. Similar scores (high, low, and average) were computed for math and intelligence.

As shown in Table 3, the ESs for the NMD / MD comparison group were significantly related to IQ level. The ESs for RD vs. MD were significantly correlated with IQ level of the MD group, as well as variations in reading, math, and IQ scores in the poor reading comparison group. ESs that included a group of MD + RD children were significantly related to variations in IQ, math, and reading. Thus, the results clearly show that the ESs for RD, RD + MD, and MD comparison were moderated by variations in reading, math, and IQ.

Multilevel Mixed Modeling

A random effects model was used to analyze the data set.[1] The purpose of the analysis was to assess the average MD effects and to gauge the amount of variability among the studies. We studied whether ESs varied across age, IQ, math level, and reading level. Also entered in the model was the type of domain or category for the dependent measures. These cognitive category variables were dummy coded (1 represented the category and 0 was the comparison to all other categories that did not include this particular domain). As shown in Table 4, we used an HLM where Level 1 equations represented the ESs for comparisons of MD children versus average achievers, MD children versus RD-only children, and MD children vs. MD + RD–only children. Level 2 reflected between-domain study effects nested within studies, and Level 3 reflected study effects. The results of the NMD versus MD analysis are shown on the left-hand side of the table. The middle of the table shows the analysis when presented ESs related to the MD- and RD-only sample. The right hand of the table shows the comparison of the MD and the comorbid group.

NMD Versus MD

As shown on the left side of Table 4, the unconditional model yielded parameter estimates for the fixed effects (the intercept) for the average ES in the sample of studies that compared MD children with average achievers. For an unconditional model, there was only one fixed effect that provided an estimate. The estimated average ES across studies was –.67. As shown in Table 4, both the random effects for domains nested within studies and the residual were significantly different from zero. No significant variance emerged for the intercepts between studies. The nested effects indicated that the studies differed significantly in their ESs. Furthermore, there was also substantial variation (according to the size of the estimate of the residual) within the studies. For the unconditional model, we computed an interclass correlation by taking the ratio of the variance component between studies (0, .19) to the sum of the variance between and within ESs (0 + .19 + .36 = .55). The intraclass correlation tells us the total proportion of variance across each individual study. The intraclass correlation was .35 (.19 / .55). Thus, 35% of the variance in ESs was at the study level, whereas 65% of the variance was at the within-study level.

The unconditional model provided a baseline to compare our first conditional model that included main effects for age, IQ, math, and reading level of the MD group and comparisons among categories. The question of interest related to this conditional modeling was whether any of the classification measures and various categories would predict ESs. Prior to determining the significant moderators, all moderator variables were centered on the mean performance across studies. We coded the cognitive measures of WM, STM, LTM, speed, and problem solving as dichotomous variables (present as 1 vs. absent as 0). These measures reflected a point biserial correlation with the overall ESs. That is, the cognitive variables represented the presence of the measures (coded as 1) when compared to all other measures (coded as 0). Table 6 shows a conditional model that contains fixed effects for age, IQ, reading scores, and math level for the MD participants and other disabled children, as well as dummy variables for the various domain categories. The estimates for each variable shown in Table 6 have been partialed for the influence of all other variables. As shown in Table 6, categories related to WM, literacy, and problem solving contributed significant variance to ES. When comparing Table 5 and Table 6, the variance component representing the difference between the studies in the conditional model accounted for 100% of the explainable variance relative to the unconditional model (.19 – .0 / .19 = 1.0). However, the within variance was reduced by only 5% ([.36 – .34] / .36).

Given that between-study variance as a function of domains nested within studies was significant, it was appropriate to examine whether a conditional model that introduced the between-study variables provided a good fit to the data. As can be calculated from the deviance values in Table 4 (184.6) and Table 5 (156.8), the difference score was significant, $\Delta\chi^2(17) = 27.80$, $p < .05$. The significant chi-square indicated that the conditional model showed a better fit to the data than the unconditional model.

We next determined if a reduced model was a better fit to the data. The reduced model entered only those variables that were significant (in this case, $p < .10$) in the full model. Because the reduced model was a submodel of the full model (the covariance stayed the same), a likelihood ratio test (deviance test) was again computed. A nonsignificant difference was the preferred outcome, suggesting that the reduced model was an adequate representation of the saturated model. As shown in Table 6, the reduced model (165.6) and full model (156.8) were not significantly different, $\Delta\chi^2(17 - 8 = 9) = 8.80$, $p > .05$, suggesting that the reduced model was a good fit to the data. Both the AIC and BIC estimates were lower in the reduced model when compared to the unconditional model, suggesting a good fit to the data.

The important finding from the reduced model was that the significant moderators were math and IQ (in the reading comparison groups) and the cognitive categories of WM (verbal and visual-spatial) and literacy. The positive estimates for the cognitive categories suggest that these variables, partialed for the influence of remaining variables, moderated overall ESs.

MD Versus RD

Because there were only 50 ESs for the MD- versus RD-only comparisons, the number of moderator variables that could be analyzed was limited. To reduce the number of independent variables, categories that were with infrequent ESs (three or less; e.g., literacy [$K = 2$] or problem solving–verbal [$K = 2$]) were not entered in the conditional models (the same rule was applied to comparisons of MD vs. MD + RD). In addition, conduct or attention behavior ratings were not considered in the analysis because our focus was on cognitive measures. As shown in the middle of Table 4, the unconditional model yielded a parameter estimate for the fixed effect (the intercept) of –.09, which was no better than chance. Also shown in Table 4 was that the random effects between studies were no better than chance. However, there was substantial variation (according to the size of the estimate of the residual) within the studies. The intraclass correlation (ratio of the variance component between studies [.10 and .20] to the sum of the variance between and within ESs [.10 + .20 + .16 = .46]) was .65 (.30 / .46), suggesting that 65% of the variance was attributed to the differences between studies.

The middle of Table 5 shows a conditional model that entered the fixed effects for age, IQ, reading scores, math level, and dummy-coded variables for the various domain categories for children with MDs and RDs only. As shown, categories related to WM and problem solving significantly moderated ESs. When comparing Table 4 and Table 5, the variance component representing the difference between the studies in the conditional model accounted for 100% of the explainable variance relative to the unconditional model. To evaluate the compatibility of the data with the full conditional model, we tested the significance of the model change. The significant chi-square indicated that the conditional model showed a better fit to the data than the unconditional model, $\Delta\chi^2(11) = 33.90$, $p < .05$.

The reduced model entered only those variables that were significant (in this case $p < .10$) in the full model. As shown in Table 6, the reduced model (57.4) and full model (41.1) were not significantly different, $\Delta\chi^2(7) = 16.30$, $p > .05$, suggesting the reduced model was a good fit to the data. This was confirmed by the lower estimates for the BIC and AIC parameters. The important finding from the reduced model was the categories of verbal WM, visual-spatial WM, and visual-spatial problem solving significantly moderated the ESs.

MD Versus MR + RD

The estimated average ES across studies was .10, and it was no better than chance (Table 4). As shown on the right side of the table, significant variance in the intercepts and the residual existed between studies. For the unconditional model, intraclass correlation between studies (.12 and .34) to the sum of the variance between and within ESs (.12 + .34 + .26 = .72) was .63 (.46 / .72), suggesting that 63% of the variance related to ESs was between studies.

The right side of Table 5 shows a conditional model that entered the fixed effects for age, IQ, reading scores, math level for the MD and MD + RD participants, as well as dummy variables for the various domain categories. As shown, IQ, variation in age, and the category related to LTM contributed significant variance. When comparing Table 4 and Table 5, the variance component representing the difference between the studies in the conditional model accounted for 100% of the explainable between-study variance relative to the unconditional model. A significant chi-square indicated that the conditional model showed a better fit to the data than the unconditional model, $\Delta\chi^2(14) = 36.10$, $p < .05$.

The reduced model entered only those variables significant (in this case $p < .10$) in the full model. As shown in Table 6, the reduced model (101.1) and full model (85.5) were not significantly different, $\Delta\chi^2(10) = 15.50$, $p > .05$, suggesting the reduced model was a good fit to the data. The important finding from the reduced model was that significant moderators were variations in age, IQ, and LTM.

Discussion

This synthesis reviewed studies that compared children with MDs, average achievers, and poor readers on various cognitive measures. The common features of these studies were that samples with MD children had math scores at or below the 25th percentile and IQs within the average range. Of particular interest was determining whether cognitive differences emerged between the groups when reading scores and variations between studies were taken into consideration. Three important findings emerged. First, the full regression model showed that the overall cognitive functioning between MD children and average achievers was significantly moderated by dependent measures related to WM, LTM, literacy, and problem solving when the effects of all other variables (e.g., age, IQ, reading level, other cognitive domain categories) were partialed out. That is, when comparing children with MDs to normal achievers, reading level, IQ, and severity of math differences played little role in moderating the ESs of the cognitive variables. Second, the magnitude of the ESs between MD- and RD-only children was moderated by only three variables: verbal and visual-spatial WM and visual-spatial problem solving. Variations in IQ, math, and reading level did not play a significant role in moderating differences between these two ability groups. Finally, variables that significantly moderated differences between MD and MD + RD children were variations in the age and IQ of the MD group and variations of IQ in the comorbid group. The overall ESs were also moderated by performance on measures of LTM. Thus, the present study found no support in the regression modeling for the assumption that variations in reading level moderated comparisons of MD children to their counterparts.

We now address three questions that directed this synthesis. First, we were interested in whether cognitive deficits in children with MDs were distinct from their average-achieving counterparts, children with RDs, and children with comorbid disorders (MDs + RDs). The results clearly indicate that moderate (.50 to high) weighted ESs in favor of age-matched average-achieving children emerged on measures of literacy (M = –1.03), verbal problem solving (M = –1.18), verbal WM (M = –0.53), visual-spatial WM (M = –0.63), and LTM (M = –0.87). Children with MD were also differentiated from children with combined RDs and MDs. Specifically, the ESs in favor of the MD group when compared to the comorbid group was found on measures of visual-spatial problem solving (M = .63), verbal WM (M = .88), LTM (M = .58), and verbal WM (M = .88). An advantage was found for the comorbid group on measures of STM for digits (M = –.55). In contrast to comparisons with the comorbid group, children with MDs could not be clearly differentiated from children with RDs only across cognitive measures (M ES = –.14). However, we did find weak to moderate (between .20 and .49) ESs in favor of children with RDs on measures of naming speed (M = –0.43). Yet these results do not account for variations across studies or the variables that moderate the magnitude of these outcomes. When the reduced models were analyzed, we found that cognitive variables of WM and literacy were significantly related to NMD and MD comparisons. In contrast, significant moderators of the ESs between MD and RD children were measures of WM and problem solving. For the comorbid group, we found in the reduced model that age and IQ of the MD group and IQ of the poor reading comparison group moderated ESs. In the reduced model, the only significant cognitive moderator between children with MDs and children with MDs + RDs was LTM.

Second, we examined whether cognitive deficits in children with MDs relative to their comparison group varied as a function of age. The results of the HLM analysis for the reduced (parsimonious) models clearly indicated that age was unrelated to the magnitude of ESs (see Table 6). This finding emerged even when the type of domain assessed, IQ, math level, and reading level were partialed out of the analysis. Thus, the results support the notion that MDs are persistent across age.

Finally, we tested whether ESs were a function of severity in MDs and intellectual level. In the regression modeling, we did not find that IQ, math scores, or reading scores for the MD group in the sample moderated the magnitude of the ESs. We found that variation in math for the RD comparison and IQ in the comorbid group moderated the magnitude of the outcomes. It is interesting that variations in reading scores for the MD or the comparison groups did not significantly moderate the magnitude of the cognitive outcomes.

In general, our results are consistent with previous syntheses of the literature that have attributed MDs when compared to average achievers to WM deficits. No significant moderators of ESs were found between RD and MD children, except for measures of WM and visual-spatial problem solving. No significant moderators were found to underlie ESs between MD and MD + RD on cognitive measures, except for the measure of LTM. More important, we did not find support in the full regression model that reading level in the poor reading group or the MD group moderated the magnitude of the ESs. Rather, at least for the comorbid group, the differences between MD children and poor readers were moderated by IQ level.

A primary interest in this analysis was whether unique differences would emerge between MD children and poor readers when the random effects related to study were taken into consideration. As shown in the unconditional model, the selection of articles for analysis yielded no significant between-study variance for studies at Level 3 in the analysis. Thus, the selection criteria for inclusion in the analysis were satisfactory. However, the between-study variance related to domains nested within studies was significant across all comparisons in the unconditional model. Moreover, that variance was effectively eliminated in the full and reduced conditional models. That is, the between-study variance related to domains was equal to zero, which suggests that the fixed effects in the model were appropriate for analysis. This does not imply that there was no variance between the studies, but rather, the estimated value of the random effect was set to zero because the residual on the between-study level was very small relative to the residual on the within-study level.

Thus, this question emerges: How can the cognitive deficits in children with MD be explained in relation to average-achieving and poor readers? We think a key process that differentiates the groups is memory. A contrasting position is provided by Landerl et al. (2004), who have challenged the underlying assumptions that children with MDs primarily suffer memory deficits. Landerl et al. argued that memory deficits have been confounded with numerical processing. They indicate that there is little evidence of nonnumerical semantic deficits in children with MDs. We found in our conditional model, when variables related to various classification measures, naming speed, and problem solving were partialed from the analysis, WM was related to the magnitude of the ESs. No significant parameters were found on STM measures for numbers.

Focusing on variables independent of the classification variables, Landerl et al. (2004) compared children of different subtypes and found that children with MDs were normal on several tasks involving phonological STM, tests accessing nonverbal information, language abilities, and psychomotor abilities. They concluded that children with MDs were best defined in terms of deficits in processing numerical information. They also found that children with RDs performed slightly similar to controls on numerical processing tasks. MD and RD children were slower than controls in reciting number sequences, although unlike children with MDs, the number naming trend in children with RDs disappeared once general ability was controlled. Although several studies (e.g., see Shalev, Manor, & Gross-Tsur, 1997) along with Landerl et al. have found that children with MDs differ more on measures that include numerical information than other measures, we found in our meta-analysis that these tasks were comparable between the groups (e.g., ES for STM–number was .10). No doubt, our findings did not tap all the basics of numerical concepts (especially numerosity; i.e., dot counting, number comparison, and subsidizing).

Some studies have documented that children with MDs also perform poorly on very complex math tasks, such as word problems, and that this is not necessarily due to just a numerical deficit but to both phonological and executive processing deficits (Swanson & Sachse-Lee, 2001). Thus, one could argue that differences between math ability groups, such as children with RDs and the comorbid group, become much more reliable with greater manipulations of phonological information. Phonological STM is certainly believed to comprise rehearsal components and phonological skills that are deficient in children with MDs and RDs. As shown in Table 2, the two groups could not be differentiated on measures attributed to phonological memory. That is, the ES between these two groups was .10 for STM–digits. One difficulty with the phonological explanation, however, was that we found an advantage for RD children in terms of naming speed, a measure assumed to tap phonological processing. Thus, although we do not completely discount the fact that RD and MD children share similar deficits in phonological processing, other areas of memory components (executive processing) are also in need of exploration.

Thus, we think that WM deficits may underlie MDs. Because verbal and visual-spatial WM tasks were deficient in children with MDs compared to average achievers, it appears that their memory deficits may operate outside a verbal system. This finding differs from other studies suggesting that WM deficits in MD children are domain specific. For example, Siegel and Ryan (1989) found that children with MDs perform poorer on WM tests related to counting and remembering digits. They did not have difficulties on nonnumerical WM tasks. A study by McClean and Hitch (1999) also suggested that children with MDs do not have general WM deficits but have specific problems with the numerical information. In contrast, Koontz and Berch (1996) tested children with and without MDs on digit and letter span tasks. They found that the children with MDs performed below average on both types of tasks, indicating a general WM difficulty (also see Swanson, 1993, for a similar finding). In contrast, Temple and Sherwood (2002) found no difference between groups on any of the measures for forward and backward digit span, and no correlation was found between memory and arithmetic ability. Landerl et al. (2004) suggests that there is no convincing evidence that WM is a causal feature of MDs. Our results showed, however, support for a WM deficit when the influences of age, IQ, reading ability, and related domain categories (e.g., STM–number information, naming speed) are partialed out. We would argue that because variables related to STM, LTM, and visual-spatial WM were partialed from the analysis, the residual variance related to the WM measures may reflect measures of controlled processes and therefore tap a general system. No doubt, this speculation will have to be tested in subsequent studies.

Our results also suggest that children with MDs have deficits related to cognitive processes that support reading. Across all categories, we found that the ES between MD and RD was only –.14. Thus, it is possible that an important correlate of MD is RD. Our findings are similar to that of Shalev et al. (1997), who found no quantitative differences between children with RDs and MDs. We did find, however, from our regression analysis that differences between children with RDs and MDs were moderated by measures of WM and problem solving. These findings suggest that variations between the two groups must be placed in the context of demands placed on WM and problem solving. Overall, however, we found weak support for the notion that distinct processes separate children with MDs from children with RDs.

In summary, the analysis of the experimental research identified cognitive differences between MD children and average math achievers. The most important conclusion is that

MD children as a group are distinctively disadvantaged when compared to their peers who are average in math performance across a broad range of tasks. However, the results showed that differences between MD children and normal achievers were significantly moderated by variations in WM and literacy. In contrast, difficulties between MD and RD children were moderated by variations in WM and problem solving. Few studies were available to document differences on measures of phonological processing, and therefore, it could not be determined if the two groups share a common problem in phonological processing. A different picture appeared when comparing children with MDs to the comorbid group. The cognitive differences between MD and MD + RD children were moderated by variations in IQ as well as measures of LTM.

Note

1. Statistical Analysis

The data reflected ESs nested within domains nested within studies. Thus, a hierarchical linear model (HLM; Bryk & Raudenbush, 1992; Singer, 2002) was developed to analyze ESs nested between domains and studies. To examine ESs, we used a random effects model (Singer, 2002). The unconditional means model is expressed as follows:

$$y_{ij} = \beta_{01} + U_{01j} + U_{02j} + R_{ij},$$

where y_{ij} is the dependent variable (e.g., ES), β_{01} is the grand mean, U_{01j} is the random intercept for study *j* in the sample representing variation between studies, U_{02j} is the random intercept representing variation of ESs for domains nested within studies, and R_{ij} is the residual. The between-study variance components, τ^2_0 = Var (U_{0j}) and τ^2_{01} = Var (U_{01j}), reflected individual studies in ESs as a function of categories of the measures embedded within studies and ESs across studies. An important extension of the multilevel regression model for meta-analysis is to allow for more than two levels (see Hox, 2002, p. 152, for a review). That is, there are several outcome measures for each study. The typical approach for analysis is to combine ESs into a single outcome per study or to carry out separate analyses for each different outcome. However, multivariate modeling allows for an analysis of all the different outcomes and provides an estimate for missing data for studies that do not provide data for all available outcome measures. ML procedures were used to determine parameter estimates because the ML estimation procedure has several advantages over other missing data techniques (Peugh & Enders, 2004). A simple conditional model can be expressed as follows:

$$y_{ij} = \beta_0 + \beta_{01} \text{ (IQ level)} + \beta_{02} \text{ (reading level)} + \beta_{03} \text{ (math level)} + \beta_{04} \text{ (RD only)}$$
$$+ \beta_{05} \text{ (IQ level of RD only)} + \beta_{06} \text{ (domain)} + U_{0j} + U_{01j} + R_{ij},$$

where y_{ij} is the dependent variable (e.g., ES), β_0 is the grand mean, β_{01} to β_{05} are the classification measures, and β_{06} is a binary variable related to domain comparison (e.g., ES related to the domain of literacy). The domain variables were entered as binary variables (e.g., literacy + 1, other domains 0). Thus, the conditional model also included the same two random effects and the residual as included in the unconditional model. The fixed and random effect parameter estimates were obtained using PROC MIXED in SAS 9.1 (SAS Institute, Inc., 2003). Presented are both unconditional (Table 4) and conditional models (Tables 5 and 6). The first conditional model (Table 5) tested whether the classification variables and the type of domains contributed significantly to the magnitude of ES. The second conditional model shown in Table 6 tested whether a more parsimonious model was a better fit for the ESs.

As shown at the bottom of Table 6, we tested whether adding one or more predictors to the model reduced the magnitude of the various random components related to study effects. The random effects of the unconditional model (see Table 4) represented the proportion of variance in those effects that were parameter specific rather than related to error variance. We also determined if the model provided a good fit to the data. This was done by using the differences between the deviance value (i.e., lack of correspondence between model and data) from the unconditional and conditional model as chi-square values and the number of parameters that were added for the conditional model as degrees of freedom. A significant chi-square indicated that the conditional model showed a better fit to the data than the unconditional model. Models were compared using several methods (deviance statistic, Akaike's Information Criterion [AIC], and Schwarz's Bayesian Information Criterion [BIC]). In all three methods, the smaller the value of the criterion relative to the other models, the better the fit of the model. Different models are compared by subtracting their deviances. The degrees of freedom for the $\Delta\chi^2$ equals the number of independent constraints imposed. Differences are then compared to critical values in a χ^2 distribution. Both AIC and BIC attempt to find the model that can best explain the data with the minimum number of parameters. Both measures penalize for added parameters, but BIC also penalizes for sample size. When deviance goes down, indicating a better fit, both AIC and BIC also tend to go down. The BIC places a larger penalty on sample size and therefore leads to a preference of more parsimonious models (fewer parameters). When comparing variations in the fixed models, a maximum likelihood (ML) function should be used (Hox, 2002, p. 46).

Snijders and Bosker (1999) argued that the power to detect significant parameters in multilevel research is frequently low because of reductions in parameter reliability. For this reason, we maintained all multiple comparisons at $p < .05$. We tested the models using both restricted maximum likelihood and ML estimation to compute the parameters in the various models. However, because we compared variations in both fixed and random effects, the results of the ML estimation are shown in Tables 4 through 6. Prior to the analysis, we computed the intraclass correlations for ESs related to comparisons between MD and NMD, MD versus RD, and MD versus MD + RD. In all three cases, the intraclass correlations exceeded .10, indicating that ESs within studies were not necessarily independent of one another. Thus, it was necessary to portion the total outcome variance into between-study variance (random intercepts, $\tau_{01}{}^2$), between-study variance within domains (e.g., literacy, memory $\tau_{02}{}^2$), and within-study variance (residual error, σ^2).

References

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

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.

Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical Linear Models: Applications and data analysis methods*. London: SAGE.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). New York: Academic Press.

Fletcher, J. M. (1985). Memory for verbal and nonverbal stimuli in learning disability subgroups: Analysis by selective reminding. *Journal of Experimental Child Psychology*, *40*, 244-259.

Fuchs, L., & Fuchs, D. (2002). Mathematical problem-solving profiles of students with mathematics disabilities with and without reading disabilities. *Journal of Learning Disabilities*, *35*, 564-574.

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

Geary, D. C. (2003). Math disabilities. In H. L. Swanson, K. Harris, & S. Graham (Eds.), *Handbook of learning disabilities* (pp. 199-212). New York: Guildford.

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.

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

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.

Hox, J. (2002). *Multilevel analysis: Techniques and applications*. Mahwah, NJ: Lawrence Erlbaum.

Jordan, N. (2007). Do words count? Connections between mathematics and reading difficulties. In D. Berch & M. M. Mazzocco (Eds.), *Why is math so hard for some children: The nature and origins of mathematical learning difficulties and disabilities* (pp. 107-120). Baltimore: Brookes.

Jordan, N., Hanich, L. B., & Kaplan, D. (2003). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with co-morbid mathematics and reading difficulties. *Child Development*, *74*, 834-850.

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

Landerl, K., Bevan, A., & Butterworth, B. (2004). Developmental dyscalculia and basic numerical capacities: A study of 8–9 year old students. *Cognition*, *93*, 99-125.

McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetical difficulties. *Journal of Experimental Child Psychology*, *74*, 240-260.

Peugh, J. L., & Enders, C. K. (2004). Missing data in educational research: A review of reporting practices and suggestions for improvement. *Review of Educational Research*, *74*, 525-556.

SAS Institute, Inc. (2003). *SAS/STAT user's guide*. Cary, NC: Author.

Shalev, R. S., Manor, O., & Gross-Tsur, V. (1997). Neuropsychological aspects of developmental dyscalculia. *Mathematical Cognition*, *3*(2), 105-120.

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.

Singer, J. D. (2002). Fitting individual growth models using SAS PROC MIXED. In D. S. Moskowitz & S. Hersberger (Eds.), *Modeling intraindividual variability with repeated measures data: Methods and applications* (pp. 135-170). Mahwah, NJ: Lawrence Erlbaum.

Snijders, T. A., & Bosker, R. J. (1999). *Multilevel analysis*. Thousand Oaks, CA: SAGE.

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

Swanson, H. L., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research*, *76*, 249-274.

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

Temple, C., & Sherwood, S. (2002). Representation and retrieval of arithmetical facts: Developmental difficulties. *Quarterly Journal of Experimental Psychology*, *55A*(3), 733-752.

Meta-Analysis

**Denotes studies included in present meta-analysis. *Denotes additional studies in the original analysis.

**Badian, N. A. (1999). Persistent arithmetic, reading, or arithmetic and reading disability. *Annals of Dyslexia*, 49, 45-70.

*Brookshire, B. L., Butler, I. J., Ewing-Cobbs, L., & Fletcher, J. M. (1994). Neuropsychological characteristics of children with Tourette syndrome: Evidence for a nonverbal learning disability? *Journal of Clinical and Experimental Neuropsychology*, *16*, 289-302.

**Fletcher, J. M. (1985). Memory for verbal and nonverbal stimuli in learning disability subgroups: Analysis by selective reminding. *Journal of Experimental Child Psychology*, *40*, 244-259.

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

**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., & 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.

*Gonzalez, J. E. J., & Espinel, A. I. G. (1999). Is IQ-achievement discrepancy relevant in the definition of arithmetic learning disabilities? *Learning Disability Quarterly*, *22*, 291-301.

*Gonzalez, J. E. J., & Espinel, A. I. G. (2002). Strategy choice in solving arithmetic word problems: Are there differences between students with learning disabilities, G-V poor performance and typical achievement students? *Learning Disability Quarterly*, *25*, 113-122.

*Klorman, R., Thatcher, J. E., Shaywitz, S. E., Fletcher, J. M., Marchione, K. E., Holahan, J. M., et al. (2002). Effects of event probability and sequence on children with attention-deficit/hyperactivity, reading, and math disorder. *Biological Psychiatry*, *52*, 795-804.

*Lennox, C., & Siegel, L. S. (1993). Visual and phonological spelling errors in subtypes of children with learning disabilities. *Applied Psycholinguistics*, *14*, 473-488.

*Lindsay, R. L., Tomazic, T., Levine, M. D., & Accardo, P. J. (2001). Attentional function as measured by a continuous performance task in children with dyscalculia. *Journal of Developmental & Behavioral Pediatrics*, *22*(5), 287-293.

**Loveland, K. A., Fletcher, J. M., & Bailey, V. (1990). Verbal and nonverbal communication of events in learning-disability subtypes. *Journal of Clinical and Experimental Neuropsychology*, *12*(4), 433-447.

*Lucangeli, D., Coi, G., & Bosco, P. (1997). Metacognitive awareness in good and poor math problem solvers. *Learning Disabilities Research & Practice*, *12*(4), 209-212.

*Lund, A. M., Hall, J. W., Wilson, K. P., & Humphreys, M. S. (1983). Frequency judgment accuracy as a function of age and school achievement (learning disabled versus non-learning-disabled) patterns. *Journal of Experimental Child Psychology*, *35*, 236-247.

*Mattson, A. J., Sheer, D. E., & Fletcher, J. M. (1992). Electrophysiological evidence of lateralized disturbances in children with learning disabilities. *Journal of Clinical and Experimental Neuropsychology*, *14*(5), 707-716.

*Mazzocco, M. M. (2001). Math learning disability and math LD subtypes: Evidence from studies of Turner syndrome, Fragile X syndrome, and Neurofibromatosis type 1. *Journal of Learning Disabilities*, *34*(6), 520-533.

**McLean, J. F., & Hitch, G. J. (1999). Working memory impairments in children with specific arithmetical difficulties. *Journal of Experimental Child Psychology*, *74*, 240-260.

**Miles, J., & Stelmack, R. M. (1994). Learning disability subtypes and the effects of auditory and visual priming on visual event-related potentials to words. *Journal of Clinical and Experimental Neuropsychology*, *16*(1), 43-64.

*Montague, M., & Applegate, B. (1993). Mathematical problem-solving characteristics of middle school students with learning disabilities. *The Journal of Special Education*, *27*(2), 175-201.

*Nolan, D. R., Hammeke, T. A., & Barkley, R. A. (1983). A comparison of the patterns of the neuropsychological performance in two groups of learning disabled children. *Journal of Clinical Child Psychology*, *12*(1), 22-27.

*Passolunghi, M. C., Cornoldi, C., & De Liberto, S. (1999). Working memory and intrusions of irrelevant information in a group of specific poor problem solvers. *Memory & Cognition*, *27*(5), 779-790.

*Passolunghi, M. C., & Siegel, L. S. (2001). Short-term memory, working memory, and inhibitory control in children with difficulties in arithmetic problem solving. *Journal of Experimental Child Psychology*, *80*, 44-57.

*Shafrir, U., & Siegel, L. S. (1994). Subtypes of learning disabilities in adolescents and adults. *Journal of Learning Disabilities*, *27*, 123-134.

*Share, D. L., Moffitt, T. E., & Silva, P. A. (1988). Factors associated with arithmetic-and-reading disability and specific arithmetic disability. *Journal of Learning Disabilities*, *21*, 313-320.

*Siegel, L. S., & Ryan, E. B. (1988). Development of grammatical-sensitivity, phonological, and short-term memory skills in normally achieving and learning disabled children. *Developmental Psychology*, *24*(1), 28-37.

**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.

**Sikora, D. M., Haley, P., Edwards, J., & Butler, R. W. (2002). Tower of London test performance in children with poor arithmetic skills. *Developmental Neuropsychology*, *21*(3), 243-254.

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

**Swanson, H. L. (1994). The role of working memory and dynamic assessment in the classification of children with learning disabilities. *Learning Disabilities Research & Practice*, *9*(4), 190-202.