Hammill Institute on Disabilities

Journal of Learning Disabilities
2018, Vol. 51(6) 523–539

DOI: 10.1177/0022219417714773
journaloflearningdisabilities.sagepub.com
SAGE

A Systematic Review of Longitudinal Studies of Mathematics Difficulty

Gena Nelson, MA[1], and Sarah R. Powell, PhD[2]

Abstract

Some students may be diagnosed with a learning disability in mathematics or dyscalculia, whereas other students may demonstrate below-grade-level mathematics performance without a disability diagnosis. In the literature, researchers often identify students in both groups as experiencing *math difficulty*. To understand the performance of students with math difficulty, we examined 35 studies that reported longitudinal results of mathematics achievement (i.e., mathematics performance measured across at least a 12-month span). Our primary goal was to conduct a systematic review of these studies and to understand whether the growth of students with math difficulty was comparable or stagnant when compared with that of students without math difficulty. We also analyzed whether identification of math difficulty was predictive of mathematics achievement in later grades and whether a diagnosis of math difficulty was stable across grade levels. Results indicate that students with math difficulty demonstrate growth on mathematics measures, but this growth still leads to lower performance than that of students without math difficulty. Identification of math difficulty is strongly related to math performance in subsequent grades, and this diagnosis is often stable. Collectively, this literature indicates that students with math difficulty continue to struggle with mathematics in later grades.

Keywords

mathematics, learning disability, longitudinal, predictive, learning trajectories

Infants demonstrate an understanding of number, and this understanding is related to mathematics performance in the toddler years (Starr, Libertus, & Brannon, 2013). Toddler and preschool mathematics experiences influence readiness for kindergarten (Magnuson, Ruhm, & Waldfogel, 2007; Murray & Harrison, 2011). Subsequently, mathematics performance at the elementary grades predicts mathematics performance during middle and high school (Watts, Duncan, Siegler, & Davis-Kean, 2014). High school mathematics performance is related to college readiness, which in turn increases adulthood outcomes (Baum, Ma, & Payea, 2010). It is clear from this cadre of longitudinal research that mathematics knowledge at an earlier age predicts mathematics performance at a later age. (Note that, for the rest of this article, we abbreviate *mathematics* as *math*.)

Wei, Lenz, and Blackorby (2013) recently analyzed the math performance of students across ages 7 to 17 and compared the calculation and problem-solving performance of students in 11 disability categories with that of students without disability. Except at age 7, students without a disability outperformed students with a disability on both types of math assessments at every time point. As students aged, the gap widened, favoring students without a disability. Students in the largest disability category (i.e., 41% of students with a diagnosed disability), specific learning disability, performed better than students in all disability categories besides speech impairment and visual impairment but not on level with students without a disability. Wei and colleagues' (2013) research indicates that math performance follows a longitudinal path for students with disabilities. The disability category of specific learning disability, however, can include students who experience severe difficulty with reading, writing, spelling, or math (Moll, Kunze, Neuhoff, Bruder, & Schulte-Körne, 2014; Scanlon, 2013). For that reason, specific learning disability is a category with tremendous variability, and it is difficult to draw conclusions about the effect of math disability or difficulty from this single category. In this systematic review, we synthesized longitudinal outcomes for students with math difficulty to determine whether growth across years was comparable to students without math difficulty, whether identified math difficulty was predictive of later math achievement, and whether the diagnosis of math difficulty was stable across years. As math is the greatest predictor of later school success (Claessens, Duncan, & Engel, 2009), it

[1]American Institutes for Research, Washington, DC, USA
[2]University of Texas at Austin, USA

Corresponding Author:
Gena Nelson, American Institutes for Research, 1050 Thomas Jefferson St. NW, #3224, Washington, DC 20007
Email: nels8101@umn.edu

is necessary to understand the math trajectories of students who have difficulty with math, to plan for intervention and assessment.

Math Disability or Difficulty

Approximately, 3% to 8% of school-age students have a diagnosed disability related to math (Desoete, Roeyers, & De Clercq, 2004; Shalev, Auerbach, Manor, & Gross-Tsur, 2000). Math disability may be referred to as *dyscalculia* (e.g., Ashkenazi & Henik, 2010). A greater number of students, however, struggle with low math performance without a disability diagnosis. In the literature, *math difficulty* is the term used to represent students with low math performance as well as students with a diagnosed math disability. This is likely to due to several reasons. First, specific learning disability is often diagnosed in the later elementary grades (O'Connor, Bocian, Beach, Sanchez, & Flynn, 2013), and researchers investigating the effects of early identification and intervention must recruit participants without the help of a school diagnosis (e.g., Dyson, Jordan, Beliakoff, & Hassinger-Das, 2015). Second, it may be difficult to locate the necessary sample of participants with a math disability, so researchers employ methods for identifying students who demonstrate characteristics of math disability (e.g., Fuchs et al., 2014). In this systematic review, we describe 35 studies related to the math performance of students with math disability or difficulty. For consistency, we refer to the students in these studies as experiencing *math difficulty*. To eliminate confusion with the original research synthesized in this manuscript, we do not use any acronym for math difficulty (e.g., MD, MMD, MLD).

How researchers identify math difficulty is quite variable (Watson & Gable, 2012). When researchers use cutoff scores or percentiles, some cutoffs are quite stringent (<10th percentile; Geary, Hoard, Nugent, & Byrd-Craven, 2008), whereas others are more relaxed (≤40th percentile; Jitendra et al., 2012). Other methods of identification include scoring below a standard score or *SD* (e.g., Swanson, 2012) or nonresponse to effective small-group math instruction (e.g., Bryant et al., 2016). Even though one assessment may not adequately portray students with math difficulty (Branum-Martin, Fletcher, & Stuebing, 2012; Mazzocco & Myers, 2003), many research teams rely on this method for identification.

To understand the math profiles of students with math difficulty—in both assessment and intervention research—performance is often compared with that of students with math difficulty comorbid with reading difficulty (e.g., Powell, Fuchs, Fuchs, Cirino, & Fletcher, 2009) and students without math difficulty (e.g., Clarke et al., 2011). Comparing the math performance of students with and without math difficulty allows for an understanding of the deficits associated with math difficulty and the uniqueness of such deficits (Geary, 2000). We review the results from studies where the performance of students with math difficulty was compared with that of students in at least one other group (e.g., math and reading difficulty, typical performance) to (a) understand whether the growth is differential for students with math difficulty and (b) assess the predictive nature and stability of math difficulty identification.

Math Performance of Students With Math Difficulty

Researchers often provide a snapshot of the math performance of students with math difficulty and compare it with the performance of students in other groups (e.g., math and reading difficulty, reading difficulty, typical performance; Andersson, 2008). In the elementary grades, students with math difficulty demonstrate lower performance on counting tasks (Stock, Desoete, & Roeyers, 2010) as well as arithmetic fluency and computation (Tolar, Fuchs, Fletcher, Fuchs, & Hamlett, 2016). Students with math difficulty often exhibit difficulty with comparing quantities, but some of this may be due to symbolic (i.e., numerical) representations of quantities (De Smedt & Gilmore, 2011; Driver & Powell, 2015). Additionally, word problem solving is problematic for students with math difficulty (Fuchs et al., 2008; Kingsdorf & Krawec, 2014). Many of these areas continue to cause problems for students with math difficulty in secondary settings (Calhoon, Emerson, Flores, & Houchins, 2007), which has implications for postsecondary options. Researchers have also determined that cognitive characteristics—such as those related to working memory (Swanson & Beebe-Frankenberger, 2004), processing speed (Cirino, Fuchs, Elias, Powell, & Schumacher, 2015), and phonological processing (Fuchs et al., 2006)—are distinctive for students with math difficulty. Additionally, students with math difficulty often have high math anxiety and low self-efficacy (Rubinsten & Tannock, 2010).

Students with math difficulty often benefit from specific and targeted intervention designed to improve math outcomes (e.g., Bryant et al., 2008; Fuchs et al., 2010; Mancl, Miller, & Kennedy, 2012; Swanson, Moran, Lussier, & Fung, 2014). Some of these interventions can change the learning trajectories of students with math difficulty. Without intervention, however, the outcomes may not be as promising.

Purpose and Research Questions of the Present Study

In this systematic review, we explore the growth in math performance for students with and without math difficulty, to understand the pattern of performance across years when intervention is not applied. We also analyze whether math difficulty can be identified in earlier grades (i.e., prediction) and whether this identification remains constant across grade levels (i.e., stability). Specifically, this systematic review addressed the following research questions:

Research Question 1: What are the characteristics of longitudinal studies related to math difficulty that identified predictors of math achievement?
Research Question 2: What is the growth of students with and without math difficulty on math measures across at least a 12-month period?
Research Question 3: Is math difficulty related to earlier or later math performance?
Research Question 4: Is math difficulty stable over time?

Method

Literature Review

We reviewed studies published from January 1985 to December 2016 that focused on math difficulty and longitudinal math achievement. The review included a search of the literature via Academic Search Premier, Education Source, Educational Resources Information Center, and PsycINFO databases with combinations of the following search terms: *math*, achievement, longitudinal, growth, predict**, *traject**, *stability, retained, dyscalculia, "math* difficulty"* and *"learning disability*."* This search resulted in the identification of 1760 studies, which we reviewed in two phases. In Phase 1, we reviewed titles, keywords, and abstracts to eliminate studies that were clearly outside the scope of this review. Many studies (n = 1634) were not considered for further review, because the study was on a topic irrelevant to education (39.8%) or was focused on growth in reading or reading disabilities (16.0%), the effectiveness of an intervention (15.2%), or assessment development research (13.6%). Other studies were excluded because they were not longitudinal or they included only unrelated predictors of math achievement (4.8%) and for other reasons (4.0%; e.g., commentaries, focus on teacher perceptions). In Phase 2, we conducted a comprehensive review of 7.2% of articles (n = 126), and 31 studies met inclusion criteria. Most articles were excluded for not meeting more than one inclusion criteria. For example, many studies (70.0%) did not include participants who were at risk for math difficulty (e.g., Jordan, Glutting, & Ramineni, 2010), or participants were considered at risk but the authors did not explicitly determine if the participants had math difficulty (e.g., participants had learning disabilities, but scores on math measures were not available; Wei et al., 2013). Other reasons for exclusion included nonlongitudinal design of at least 12 months (37.6%) and not reporting appropriate results (21.5%; see next section for specific criteria). Finally, in an attempt to conduct an exhaustive search, we also searched relevant researchers' previous works. We identified four more articles that met inclusion criteria and included each of these in this review for a total of 35 studies.

Inclusion Criteria

We used the following criteria to determine study inclusion.

First, the study examined a longitudinal trajectory of math achievement for students with math difficulty without testing the efficacy of a math program or intervention; studies that focused on the success of math interventions were excluded (e.g., Nunes et al., 2009).

Second, participants were school-age (i.e., preschool to 12th grade), and a subset of participants demonstrated math difficulty. Acceptable methods of documentation for math difficulty included an author-specified criterion on a math measure or a disability diagnosis in math. Because one of our research questions focused on identifying differences in author-specified criteria for identifying students with math difficulty and growth in math according to predetermined math difficulty identification, we excluded studies that utilized latent growth curve analysis to retroactively identify groups of students who performed in separate classes in mathematics (e.g., Geary, Bailey, Littlefield et al., 2009). We also excluded studies that focused on special populations who are typically considered at risk, but we did not screen participants for math difficulty, such as students with genetic disorders or learning disabilities (e.g., Murphy & Mazzocco, 2009).

Third, the study included at least one math measure administered at a minimum of two time points at least 12 months apart. We selected 12 months as the minimum time frame to indicate at least one grade-level change from first to last data collection. The math measure did not have to be the same measure administered at different time points.

Fourth, the study included appropriate results (e.g., regression analyses, correlations, stability of math difficulty) to determine the longitudinal predictors of math achievement, growth, and stability of performance. We excluded studies that included a math outcome but included only unfitting predictors for the purpose of this study (e.g., behavior, reading, motion sensitivity; Boets, De Smedt, & Ghesquiere, 2011) or studies that included appropriate predictors of math achievement but included only unsuitable outcomes (e.g., confidence in correct problem solving; e.g., Garrett, Mazzocco, & Baker, 2006).

Finally, the study was published in English in a peer-reviewed journal.

Coding of Studies

We coded the 35 studies that met inclusion criteria for the following information: study characteristics, average age (or grade) of participants at the start of the study, attrition, data collection time points, measures of math achievement administered at each time point, reliability and validity coefficients for measures, and results. Regarding study and sample characteristics, we coded the year of publication, location, and if the study utilized a national data set (e.g., Early Childhood Longitudinal Study–Kindergarten [ECLS-K]). Finally, we recorded the authors' method for determining participants' at-risk status for math difficulty. Many studies included more than one math difficulty group or more than one comparison group; thus, classification methods and sample sizes for all groups were coded separately.

Regarding data points, we coded each data point (e.g., Time 1, Time 2, Time 3, . . .) and, when available, the time of year for data collection data (e.g., fall Grade 1). At each time point, we recorded which measures were administered, a short description of the measure and skills assessed (e.g., broad math, calculation), and the reported reliability and validity coefficients for the measures. Finally, for each study, we recorded general findings, specific significant results, and significant levels. Specific results included findings such as the percentage of students who retained the same math difficulty classification over time (i.e., stability of math difficulty), correlations, and results from regression and growth curve analyses.

Interrater Agreement

The first author coded all studies, and the second author double-coded 33% of the 35 studies. We discussed all discrepancies and focused on ensuring that the codes were accurate for the classification of students with math difficulty, results, and time points of data collection. Coding discrepancies were resolved to determine the final code used in the analyses. The second author double-checked all information provided in Table 1.

Results

Characteristics of Studies

Table 1 provides a summary of the math difficulty determination within each of the 35 studies included in this systematic review. All studies were published between 2000 and 2016. Data collection occurred in eight countries, with approximately 65% of data collected within the United States. Other countries in which data were collected included Belgium (*n* = 4), Canada (*n* = 3), Germany (*n* = 1), the Netherlands, (*n* = 1), Spain, (*n* = 1), Sweden (*n* = 1), and the United Kingdom (*n* = 1). The mean sample size was 609 participants, but this high mean was due to three studies that analyzed data from the ECLS-K data set (Bodovski & Farkas, 2007; Claessens & Engel, 2013; Morgan et al., 2009). The median sample size was 181, with a range of 28 to 7,892 participants. The majority of studies focused on early elementary participants. Regarding grade level of participants, on average, participants were in first grade at the start of the study and near the end of third grade at the conclusion of the study. Only 20% of studies collected data beyond fifth grade; no studies collected data beyond ninth grade. In the majority of studies (91%), researchers collected data themselves; the other three studies used the aforementioned ECLS-K national database. Only 16 separate research teams collected data for the 35 studies. Of these, two teams produced 2 studies each; two teams produced 3 studies each; two teams produced 5 studies each; and one team produced 6 studies each. Therefore, seven distinct research teams generated almost 75% of the studies.

In terms of classification of math difficulty, researchers in 32 of the studies categorized students into groups using percentiles as a cutoff. The most stringent percentile cutoff for determination of math difficulty was math performance below the 10th percentile; the least stringent cutoff was at or below the 35th percentile. In two studies, performance based on *SD* below the mean was the determinant for math difficulty. Interestingly, these two studies represented research from Spain (Navarro et al., 2012) and Sweden (Andersson, 2010). One of the national database studies identified the math difficulty group by using proficiency levels on the ECLS-K math assessment (Claessens & Engel, 2013).

By default (i.e., study inclusion criteria), all 35 studies had at least one group of students with math difficulty. For 26 studies, determination of math difficulty was based on assessments administered during one data collection wave; in the other 9 studies, researchers identified math difficulty across two or more data collection waves or years. In some instances, math difficulty was predictively identified, whereas in other cases, it was retroactively identified. To provide more detail about math difficulty, researchers in 21 studies identified a second group of students with math difficulty. In 12 of these cases, the second group was identified through less stringent criteria for math difficulty. This group was often described as low achieving in math, whereas the first group of students with math difficulty was described as experiencing a math learning disability. In seven cases, the second group with math difficulty included students demonstrating difficulty with math comorbid with reading. The other two studies with a second math difficulty group focused on a group where a diagnosis of math difficulty was not stable across data collection years. Six studies included a group of students with reading difficulty without math difficulty. In terms of comparison groups, all but two studies (Peng et al., 2016; Toll & Van Luit, 2014) included at least one comparison group without math difficulty. The lowest cutoff percentile for identification of typically achieving students (i.e., not having math difficulty) was performance >10th percentile. In 75% of cases, the cutoff for typical

Table 1. Categorization of Math Difficulty Within Studies.

Authors (year)	Location	N	Math difficulty 1 (category;[a] n)	Math difficulty 2 (category; n)	Other difficulty (category; n)	Comparison 1 (category; n)	Comparison 2 (category; n)	Measures for determining math difficulty
Andersson (2010)	Sweden	249	Special education services in math and ≥1.5 *SD* below mean on math (MD-only; *n* = 39)	Special education services in math and reading and ≥1.5 *SD* below mean on math and reading (MD-RD; *n* = 80)	Special education services in reading and ≥1.5 *SD* below mean on reading (RD-only; *n* = 36)	No special education services (control; *n* = 94)		Math screener (addition, subtraction, multiplication), reading screener (cloze passage comprehension)
Bodovski and Farkas (2007)	US	3,206	<25th (LA; *n* = 267)			≥25th (*n* = 2,939)[b]		ECLS-K *Mathematics Test*
Chong and Siegel (2008)	Canada	214	≤10th (MLD; *n* = not reported)	11th–25th (LA; *n* = not reported)		≥35th (TA; *n* = not reported)		WJ-III Calculation; WJ-III Math Fluency
Claessens and Engel (2013)	US	7,154	Not mastered Proficiency Level 2 (low-achieving; *n* = 2,329)			Mastered Proficiency Level 2 (main analytic; *n* = 7,154)[c]		ECLS-K *Mathematics Test*
Cowan and Powell (2014)	UK	258	≤10th (MLD single time point; *n* = 29; MLD persistent subset; *n* = 11)	10th–25th (LA single time point; *n* = 42; LA persistent subset; *n* = 14)		>25th (TA single time point; *n* = 187; TA persistent subset; *n* = 166)		WIAT-II (UK)
Desoete et al. (2012)	Belgium	395	≤10th (MD; *n* = 16)	10th–25th (LA; *n* = 64)		>25th (TA; *n* = 315)		*Arithmetic Number Facts Test* and/or *Kortrijk Arithmetic Test*; confirmation by school psychologist
Desoete and Grégoire (2006); Study 1	Belgium	66	≤15th (delay; *n* = 22)			≥50th (*n* = 44)		*Arithmetic Number Facts Test*
Geary, Bailey, and Hoard (2009)	US	158	<15th (MLD; *n* = 45)	15th–30th (LA; *n* = 17)		>30th (TA; *n* = 96)		WIAT-II Numerical Operations
Geary et al. (2000)	US	84	<35th in math (MD; *n* = 12)	<35th in math and reading (MDRD; *n* = 16)	<35th in reading (RD; *n* = 14)	>35th in math and reading (normal; *n* = 26)	<35th in math at only 1 grade level (variable; *n* = 16)	WIAT Mathematics Reasoning, WJ-R Letter-Word Identification
Geary et al. (2007)	US	278	<15th, 2 years (MLD; *n* = 15)	23rd–39th, at least 1 year (LA; *n* = 44)		>50th, 2 years (TA; *n* = 46)		WIAT-II Numerical Operations
Geary et al. (2012)	US	177	<25th and "slow growth" across 5 years (MLD; *n* = 16)	>25th but "slow growth" across 5 years (LA; *n* = 29)		>25th and "average growth" across 5 years (TA; *n* = 132)		WIAT-II Numerical Operations
Geary et al. (2008)	US	261	<11th, 2 years (MLD; *n* = 19)	11th–25th, 2 years (LA; *n* = 43)		26th–74th, 2 years (TA; *n* = 50)		WIAT-II Numerical Operations
Hecht and Vagi (2010)	US	181	≤25th (MD; *n* = 55)			>40th (TA; *n* = 126)		WJ-III Calculation composite (Calculation and Math Fluency)

(continued)

Table 1. (continued)

Authors (year)	Location	*N*	Math difficulty 1 (category;[a] *n*)	Math difficulty 2 (category; *n*)	Other difficulty (category; *n*)	Comparison 1 (category; *n*)	Comparison 2 (category; *n*)	Measures for determining math difficulty
Jordan and Hanich (2003)	US	74	≤15th on math (MMD-only; *n* = 11)	≤15th on math and reading (MMD/MRD; *n* = 8)	≤15th on reading (MRD-only; *n* = 8)	>35th (TA; *n* = 47)		WJ-R Broad Math (Calculation and Applied Problems), WJ-R Broad Reading (Letter-Word Identification and Passage Comprehension)
Jordan et al. (2003a)	US	105	≤25th (poor fact mastery; *n* = 45)			51st–75th (good fact mastery; *n* = 60)		Forced retrieval of number facts test
Jordan et al. (2003b)	US	180	≤35th on math (MD only; *n* = 46)	≤35th on math and reading (MD-RD; *n* = 42)	≤35th on reading (RD only; *n* = 45)	>35th (TA; *n* = 47)		WJ-R Mathematics Composite (Calculation and Applied Problems), WJ-R Reading Composite (Letter-Word Identification and Passage Comprehension)
Jordan et al. (2002)	US	180	≤35th in math (MD only; *n* = 46)	≤35th in math and reading (MD-RD; *n* = 42)	≤35th in reading (RD only; *n* = 45)	≥40th in math and reading (*n* = 47)		WJ-R Broad Math (Calculation and Applied Problems), WJ-R Broad Reading (Letter-Word Identification and Passage Comprehension)
Locuniak and Jordan (2008)	US	198	≤25th (at risk; *n* = 48)			>25th (not-at risk; *n* = 158)		*Number Knowledge Test, Number Combinations Test*
Martin et al. (2013)	US	144	<32nd on math; >40th on reading (MD; *n* = 83)			>40th on math and reading (no LD; *n* = 61)		WRAT-3 Arithmetic or WJ-III Math Fluency; WRAT-3 Reading
Mazzocco and Devlin (2008)	US	106	≤10th (MLD; *n* = 12)	11th–25th (LA; *n* = 18)		>25th (TA; *n* = 76)		WJ-R Calculation
Mazzocco et al. (2011)	US	71	≤10th (MLD; *n* = 10)	11th–25th (LA; *n* = 9)		25th–95th (TA; *n* = 37)	>95th (high achieving; *n* = 15)	TEMA-2, WJ-R Calculation
Mazzocco and Myers (2003)	US	209	<10th, 2 or more years (MD-persistent; *n* = 22)	<10th, at least 1 year (MD all; *n* = 35)[d]		≥10th (non-MD; *n* = 174)		TEMA-2
Mazzocco et al. (2013)	US	122	<10th (MLD; *n* = 11)	11th–25th (LA; *n* = 18)		>25th (TA; *n* = 93)		WJ-R Calculation
Mazzocco and Thompson (2005)	US	209	≤10th (MLD; *n* = 23)			>10th (non-MLD; *n* = 186)		TEMA-2, WJ-R Calculation
Morgan et al. (2009)	US	7,892	≤10th (MD; *n* = 988)			>10th (no MD; *n* = 6,904)		ECLS-K *Mathematics Test*
Murphy et al. (2007)	US	249	≤10th, 2 or more years (MLD; *n* = 22)	11th–25th, 2 or more years (MLD; *n* = 42)		Did not meet MLD criteria for at least 2 years (non-MLD; *n* = 146)		TEMA-2

(continued)

Table 1. (continued)

Authors (year)	Location	*N*	Math difficulty 1 (category;[a] *n*)	Math difficulty 2 (category; *n*)	Other difficulty (category; *n*)	Comparison 1 (category; *n*)	Comparison 2 (category; *n*)	Measures for determining math difficulty
Navarro et al. (2012)	Spain	127	≤1 *SD* below the mean (LA; *n* = 46)			1 *SD* below the mean to 1 *SD* above the mean (middle achiever; *n* = 76)	≥1 *SD* above the mean (high achiever; *n* = 26)	*Early Numeracy Test* (Spanish version)
Peng et al. (2016)	US	176	≤25th math and below average factor score on reading measures (MDRD; *n* = 176)					Teacher nominations, followed by WRAT-4, Rapid Sound Naming, Phonemic Decoding Fluency (TOWRE)
Schwenck et al. (2015)[e]	Germany	166	≤25th math (MD; *n* = 22)	≤25th math and reading (COM; *n* = 35)	≤25th reading (RD; *n* = 24)	> 26th math and reading (TD; *n* = 28)		*Wurzburg Silent Reading Test*; spelling test (DERET 1–2+); *Heidelberg arithmetic test*
Swanson et al. (2008)	US	353	≤25th mean on 2 measures (SMD; *n* = 134)			>26th mean on 2 measures (not at risk; *n* = 219)		WISC-III Arithmetic; CTOPP Digit Naming Speed
Stock et al. (2010)	Belgium	362[f]	≤10th (AD; *n* = 16)	11th–25th (LA; *n* = 27)		>25th (TA; *n* = 319)		*Arithmetic Number Facts Test* and/or *Kortrijk Arithmetic Test*
Toll and Van Luit (2014)	Netherlands	199	<15th (very low numeracy; *n* = 199)					*Early Numeracy Test–Revised*
Vanbinst et al. (2014)	Belgium	28	<25th, all 3 time points (MLD-persistent; *n* = 14)			>35th, all 3 time points (TD; *n* = 14)		Standardized mathematics test from Flemish Student Monitoring System
Vukovic (2012)	Canada	203	<25th on math at least twice in 4 years;[g] ≥25th on reading (MD-only; *n* = 19)	<25th on math at least twice in 4 years[g] and reading at least once in Grade 1, 2, or 3 (MD+RD; *n* = 19)		≥25th math and reading (TD; *n* = 165)		WRAT-3 Arithmetic, WJ-III Letter-Word Identification
Vukovic and Siegel (2010)	Canada	99	<25th, at least 2 years (MD-persistent; *n* = 26)	<25th, 1 year (MD-transient; *n* = 27)		≥40th, all 4 years (TA; *n* = 46)		WRAT-3 Arithmetic

Note. All cutoffs are reported as percentiles unless otherwise noted. Student risk categories in table are reported as in the studies by authors. MD/MLD/MMD/SMD = math difficulty/disability/serious math problem-solving difficulties; AD = arithmetic difficulty; MDRD/COM = math and reading difficulty/disability/comorbid; RD/MRD = reading difficulty/disability; LD = learning difficulty/disability; LA = low achieving; TA/TD = typically achieving/developing; CTOPP = *Comprehensive Test of Phonological Processing*; ECLS-K = Early Childhood Longitudinal Survey–Kindergarten; TEMA-2 = *Test of Early Mathematics Achievement–Second Edition*; TOWRE = *Test of Word Reading Efficiency*; WIAT-II = *Wechsler Individual Achievement Test–Second Edition*; WISC-III = *Wechsler Intelligence Scale for Children–Third Edition*; WJ-III = *Woodcock-Johnson–Third Edition*; WJ-R = *Woodcock-Johnson–Second Edition*; WRAT-3 = *Wide Range Achievement Test–Third Edition*.

[a]Specific category name provided by authors. [b]Authors presented information for 25th–50th, 51st–75th, and >75th; we collapsed them into one category for analysis. [c]Main analytic sample includes the low-achieving students. [d]Includes the 22 MD-persistent students. [e]Study included 2 other comparison groups with below-average IQs; groups not coded for analyses. [f]Calculated by adding AD, LA, and TA participants due to a discrepancy in the manuscript. [g]Students could not be identified if the two time points were only at kindergarten and Grade 1.

performance was, at a minimum, >25th percentile. Finally, three studies included a second comparison group of students without math difficulty, and the majority of researchers described this group as high achieving.

Summary of Math Performance

We organize the following sections as they relate to the growth of math skills, the longitudinal predictors of math difficulty and achievement, and the longitudinal stability of math difficulty. First, we discuss the growth of math achievement to illustrate differences of specific skills over time between students who experience math difficulty and students who do not. Then, we discuss predictors of math difficulty to later math to evaluate if specific skills at an earlier time point can accurately predict if students will experience consistent and severe deficits in math. Finally, we examine the stability of math difficulty as it relates to variables such as the restrictiveness of the cutoff.

Differences in growth. The majority of studies examined differences in initial achievement of math and rate of growth over time. Most math skills fell into one of the following categories: early numeracy skills, computation, rational numbers, or broad math (i.e., math performance across content areas). We report results according to the skills measured, and we report growth rates for students with math difficulty as compared with typically achieving peers.

Early numeracy skills. Across studies, many early numeracy skills were measured, including magnitude comparison, counting, number knowledge, and number combinations (Geary et al., 2012; Jordan & Hanich, 2003; Navarro et al., 2012; Vukovic, 2012). Of the four studies that measured growth in early numeracy over time, three used screening measures to determine math difficulty that focused on numeracy (e.g., *Early Numeracy Test*) or at least included items related to numeracy (e.g., Numerical Operations subtest). Therefore, there was an alignment between screening measures for determination of math difficulty and the measures used to observe growth.

Three studies reported that after the initial achievement gaps were identified in kindergarten and first grade, growth on skills was essentially parallel for the early numeracy skills of composing and decomposing (Geary et al., 2012); counting, digit placement, and number identification (Jordan & Hanich, 2003); and number sequences and number identification (Vukovic, 2012). Two studies reported faster growth for students with math difficulty for skills such as number line estimation (Geary et al., 2012) and broad numeracy skills measured with the *Early Numeracy Test* (Navarro et al., 2012). Slower growth by students without math difficulty on a broad measure, however, may have been the result of ceiling effects, as accuracy at the start of the study was >70%, compared with 20% for students with math difficulty (i.e., students without math difficulty had less opportunity for growth). Of importance, with both trends, students with math difficulty continued to perform worse than peers.

Studies reported parallel growth for students with and without math difficulty regardless of restrictiveness of the criteria used to determine math difficulty. Math difficulty criteria for this group of studies ranged from performance 1 *SD* below the mean to <25th percentile. One study also required that students have performance <25th percentile and "slow growth" to be identified as having math difficulty (Geary et al., 2012), and another required that students have performance <25th percentile for at least 2 years of the study (meaning that students identified as having math difficulty likely did not have fast growth; Vukovic, 2012). The third study had more restrictive criteria for determining math difficulty (<15th percentile; Jordan & Hanich, 2003). Therefore, the parallel growth observed across these three studies might be a result of the criteria for math difficulty aligning with the students who are least likely to have faster growth. Interestingly, the study that reported faster growth for students with math difficulty in early numeracy was the only study that also used a comprehensive early numeracy screening measure to identify students with math difficulty (Navarro et al., 2012).

Though not specific to growth rates, differences in the early numeracy skill of counting were consistently observed over time (Desoete & Grégoire, 2006; Geary et al., 2000; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Murphy et al., 2007). For example, students who scored ≤15th percentile on number knowledge and mental arithmetic in first grade had significantly lower counting abilities in kindergarten (Desoete & Grégoire, 2006), and students with math difficulty in first grade had more difficulty recognizing correct counting or recognizing counting errors (Geary et al., 2000; Geary et al., 2007; Murphy et al., 2007). Mazzocco, Feigenson, and Halberda (2011) reported number sense differences measured as late as ninth grade, with performance on approximate number sense acuity differentiating students identified with math difficulty.

Computation. All studies that measured growth in computation over time used measures that included subtests of computation or measures that focused only on computation (e.g., researcher-developed computation screener) to identify students with math difficulty. Therefore, there was an alignment between the screening measures and the measures used to observe growth across studies. Generally, students with math difficulty were less accurate with computation and used inefficient strategies more often when compared with peers from kindergarten through elementary school (Andersson, 2010; Chong & Siegel, 2008; Jordan & Hanich, 2003; Jordan, Hanich, & Kaplan, 2003a; Jordan, Kaplan, &

Hanich, 2002; Swanson et al., 2008; Vanbinst, Ghesquière, & De Smedt, 2014). The majority of studies reported results for computational fluency, but there were a few studies that also reported growth for computation in word problem contexts. Regarding computational fluency, results were generally consistent across studies in terms of how computation was measured. Authors reported parallel growth (Chong & Siegel, 2008 [timed fact fluency]; Jordan & Hanich, 2003; Swanson et al., 2008) and slightly faster growth (Chong & Siegel, 2008 [untimed written calculation]; Jordan et al., 2002) for students with math difficulty when measuring written computation skills. Authors reported parallel growth between students with math difficulty and those without math difficulty for approximate arithmetic (e.g., selecting the answer closest to the correct answer; Andersson, 2010; Jordan & Hanich, 2003) and exact calculation (e.g., Jordan & Hanich, 2003; Jordan et al., 2003a). Regarding measures of forced retrieval (i.e., examiners read and/or displayed basic facts, and students were required to give an answer within 3 s), two studies observed parallel growth between groups (Andersson, 2010; Jordan & Hanich, 2003), while another study observed slower growth for students with math difficulty (Jordan et al., 2003a). Vanbinst et al. (2014) observed faster growth for students with math difficulty when students were required to answer computation items in a similar format but were not forced to use retrieval (i.e., students were told to answer quickly but were allowed to use strategies other than retrieval).

Studies reported parallel growth for students with and without math difficulty regardless of restrictiveness of the criteria used to determine math difficulty. Math difficulty criteria for this group of studies ranged from 1.5 *SD* below the mean to ≤35th percentile. Interestingly, the two studies that reported faster growth for students with math difficulty in computation had the most and least restrictive criteria for determining math difficulty, ≤10th and ≤35th percentiles, respectively (Chong & Siegel, 2008; Jordan et al., 2002).

Several studies examined strategy use for computation, and results related to strategy use were generally consistent across studies. Students with math difficulty used retrieval strategies less often initially when compared with peers (Geary et al., 2000; Geary et al., 2012; Vanbinst et al., 2014). Two studies reported that students without math difficulty made greater gains in correct computation retrieval from first to second grade (Geary et al., 2000) and from first to fifth grade (Geary et al., 2012). Vanbinst et al. (2014) observed that students with math difficulty more often used procedural skills than retrieval, and Geary et al. (2012) reported that over time, students with math difficulty used more procedural skills to correctly solve simple addition problems but had less accuracy over time with decomposition strategies.

Regarding word problem solving, researchers used either a researcher-developed measure related to computation or subtests from norm-referenced measures of math achievement that included computation and word problem solving to screen students for math difficulty. All studies measured growth with measures that were read orally to students and that required them to solve word problems. Similar to accuracy for computational fluency, the results for word problem solving were inconsistent. Three studies reported substantially similar growth for students with math difficulty (Andersson, 2010; Jordan, Hanich, & Kaplan, 2003a, 2003b), and other studies reported slower growth for students with math difficulty (Cowan & Powell, 2014; Swanson et al., 2008). Jordan and Hanich (2003) reported that the average increase in word problems was faster for students identified as having math difficulty only, while participants identified as having comorbid reading and math difficulty had slower rates of growth. Moreover, we did not observe consistent trends regarding criteria for identifying math difficulty or age of participants regarding growth rates. For example, one study that reported slower growth for students with math difficulty had a restrictive cutoff (<10th percentile), while another study that had a similar restrictive cutoff (<15th percentile) reported faster growth for students with math difficulty.

Rational numbers. Only three studies (Hecht & Vagi, 2010; Mazzocco & Devlin, 2008; Mazzocco et al., 2013) examined more advanced math concepts, such as the understanding and use of fractions and decimals. Although researchers evaluated different components of performance with fractions and included participants from different grade levels, authors consistently reported that students with math difficulty performed worse than peers over time. For example, Hecht and Vagi (2010) examined emerging fraction skills from fourth to fifth grade, and students with math difficulty showed consistently lower performance and smaller gains in performance. Meanwhile, Mazzocco et al. (2013) evaluated the fraction comparison performance of participants in fourth through eighth grade; students with math difficulty did not reach ceiling-level performance by the end of the study as typically achieving students did. Finally, Mazzocco and Devlin (2008) measured student understanding of decimals and representing fractions with pictures and numerals from sixth through eighth grade; students who performed ≤10th percentile had significantly lower scores than students who performed between the 11th and 25th percentiles.

Unlike studies that examined growth for skills related to other areas of mathematics, all studies that examined growth with fractions did not use screening measures aligned to fraction concepts. Studies identified students with math difficulty using norm-referenced measures of mathematics achievement and measured growth with researcher-developed tools. Due to the small number of studies that evaluated fraction growth, we were not able to report on themes related to age or the restrictiveness of criteria for determining math difficulty.

Broad math achievement. Finally, researchers observed group differences over time on performance in broad math. We defined broad math as assessments with a variety of math content, concepts, and procedures. All but one study used measures of broad math achievement to identify students with math difficulty; one study used a "forced retrieval of number facts" measure, as the purpose of the study was to identify students with poor fact mastery (Jordan et al., 2003a). Of the six studies that measured growth in broad math over time, four used the same measure to screen and measure growth over time. Generally, there was alignment between screening measures and measures used to observe growth across studies. Students with math difficulty consistently displayed parallel or slower growth over time and maintained achievement gaps. For example, according to two studies based on the ECLS-K data set, students with math difficulty in the fall of kindergarten yielded lower overall gains through third grade (Bodovski & Farkas, 2007), and students with math difficulty in the fall and spring of kindergarten displayed the lowest achievement and slowest growth rates during the next 5 years when compared with other groups (Morgan et al., 2009). Three other studies observed parallel growth between students with math difficulty and typically achieving students, also maintaining the initial achievement gap (Jordan & Hanich, 2003; Jordan et al., 2003a; Vukovic, 2012).

Interestingly, trends emerged regarding the rate of growth on broad math measures and the restrictiveness of the criteria used to determine math difficulty. Half of the studies reported parallel growth; however, the only study to report that students with math difficulty had faster rates of growth utilized the least restrictive cutoff (<35th percentile), while two of the three studies reporting that students with math difficulty had significantly slower growth than peers used the most restrictive criteria (≤10th percentile). The other study that reported slower growth used performance <25th percentile as a cutoff. The three studies that reported slower growth used the same measure at screening as was used to measure growth. These studies followed students from kindergarten through third and fifth grades, while most studies that reported parallel or faster rates of growth followed students from only second to third grade.

To understand the complexity of math difficulty comorbid with reading difficulty, several researchers included students who exhibited math and reading difficulty, in addition to students who exhibited math difficulty only, as a way to compare growth in broad math achievement. For example, Vukovic (2012) observed student performance beginning in kindergarten. Results indicated significant group differences in initial math achievement. That is, typically achieving outperformed math difficulty, which in turn outperformed math and reading difficulty. Interestingly, the growth trajectories of each group were similar, and the achievement gap remained through third grade. Jordan and Hanich (2003) examined the performance of students with math difficulty only and comorbid math and reading difficulty from second through third grade. In the fall of second grade, the two math difficulty groups did not significantly differ, but at the second, third, and fourth data collection points (through spring of third grade), the group with math difficulty only outperformed the group with comorbid math and reading difficulty. So, although both math difficulty groups performed lower than the students with only reading difficulty and typically achieving students, the students with only math difficulty who were good readers grew faster in math than did the students with math difficulty who were poor readers.

Predictors of math difficulty. Several studies (Desoete et al., 2012; Geary et al., 2009; Locuniak & Jordan, 2008; Mazzocco & Thompson, 2005; Stock et al., 2010; Vukovic & Siegel, 2010) evaluated specific variables that predicted group membership (math difficulty vs. no math difficulty). Two studies used discriminant function analyses (Stock et al., 2010; Vukovic & Siegel, 2010). Vukovic and Siegel (2010) reported that significant predictors of persistent math difficulty (at least 2 years) compared with math difficulty displayed during only 1 year and typically achieving students were math concepts, number series, and number naming. Stock et al. (2010) observed that two of the strongest predictors of group membership in kindergarten through second grade were early numeracy skills in magnitude comparison and seriation. Stock et al. also reported that conceptual counting measured in kindergarten correctly identified 31% of students with math difficulty in first and second grades and that 44% of students identified with math difficulty in first and second grades had severe deficits in magnitude comparison.

In contrast, Mazzocco and Thompson (2005) evaluated whether scores on broad math measures accurately predicted math difficulty status at a later time. Kindergarten scores on the *KeyMath* (Connolly, 2007) and *Test of Early Mathematics Ability* (Ginsburg & Baroody, 2003) accurately predicted math difficulty (≤10th percentile) in second and third grades with 83.5% correct classification. Researchers examined the relationship between specific items and math difficulty, and interestingly, they reported results similar to other studies. Mazzocco and Thompson (2005) reported that early numeracy items that addressed counting, reading one-digit numerals, addition of one-digit numerals, magnitude judgments, and numeral comparison were significantly related to math difficulty. Two studies reported that specific number sense skills, such counting, number knowledge, and composing and decomposing, as measured in kindergarten and first grade predicted difficulty with arithmetic in second grade with 52% accuracy (Locuniak & Jordan, 2008) and predicted broad math difficulty in third grade with 66% accuracy (Geary et al., 2009). Desoete et al. (2012) reported that proficiency with symbolic (numerals) and nonsymbolic (dots) comparison as measured in kindergarten was significantly related to group membership (math difficulty vs. no math difficulty) in first and second grades, as determined by achievement in computation, with students with math difficulty underperforming as comparing with peers. In summary, studies that evaluated predictors of math difficulty regularly reported that early numeracy skills, such as counting and magnitude comparison, predicted math difficulty; however, with only three studies reporting such results and with each study focusing on participants in fourth grade and younger, it is difficulty to generalize and compare the results of these studies with other skills and other grade levels.

Stability of math difficulty. A subset of studies ($n = 11$) reported the longitudinal stability of math difficulty for participants. Generally, more restrictive cutoffs for math difficulty resulted in higher proportions of students retaining the same math difficulty status. Morgan et al. (2009) examined the stability of math difficulty as a function of performing ≤10th percentile at specific time points. Stability of math difficulty from kindergarten to fifth grade ranged from 28% to 65%, with the higher proportion of students retaining math difficulty status being those who performed ≤10th percentile during both the fall and spring of kindergarten, as compared with math difficulty during either the fall or spring of kindergarten. In the same study, the highest proportion of students (70%) who retained the same classification were those who experienced math difficulty in both the fall and spring of kindergarten and math difficulty in the spring of third grade.

In two other studies, Geary and colleagues reported that of the students identified with math difficulty by performance <15th percentile in first grade, approximately 68% of students had performed similarly the year before in kindergarten (Geary et al., 2007), and 75% retained math difficulty status in third grade (Geary, Bailey, & Hoard, 2009). Mazzocco and Myers (2003) reported similar retention rates (63%) for math difficulty for students who performed <10th percentile from kindergarten to third grade. Although, Jordan and Hanich (2003) reported that only 18% of students identified as experiencing math difficulty by performing ≤15th percentile in second grade retained the same status in third grade; however, they also reported that 88% of students identified as experiencing comorbid math and reading difficulty retained the same classification. Schwenck et al. (2015) reported a similar pattern between second and fourth grade: 18% of students with math difficulty retained the same status, while 43% of students with comorbid math and reading difficulty maintained status. In contrast, studies with less restrictive cutoffs to determine math difficulty reported smaller proportions of students who retained math difficulty status. For example, Martin et al. (2013) employed a cutoff of performance <32nd percentile to identify third- and fourth-grade students with math difficulty, but 2 years later only 40% of students retained the same classification; this decrease was significant ($p < .001$).

Discussion

We conducted this systematic review of longitudinal research related to math difficulty to understand the growth of students with and without math difficulty on math measures, if math difficulty was related to earlier or later math performance, and if math difficulty varied according to factors such as participant age and method of classification. The goal of our first research question was to describe the characteristics of longitudinal studies related to math difficulty. Although we conducted an electronic search of the data from the last 30 years, the majority of studies (77%) were published in the last 10 years, and most were conducted in the United States. The recency of the literature is likely due to improved educational funding for research projects and the increase in focus on math education, especially at earlier grade levels. With the push for high-quality math instruction at earlier grade levels (National Council of Teachers of Mathematics, 2006), it is not surprising that researchers shifted their attention to evaluating longitudinal differences between students with and without math difficulty at earlier ages. The majority of studies focused data collection at the early elementary grade levels, with only a few studies collecting data beyond sixth grade.

We also examined how math difficulty was determined, and we observed a lack of consistency across studies and within research teams regarding the cutoff for classification of math difficulty. Interestingly, only one study (Andersson, 2010) used special education services in math to identify those students with math difficulty. Of the remaining studies, the majority identified students with math difficulty by employing a less restrictive cutoff (performance <25th vs. <10th percentile); furthermore, only a few studies (e.g., Mazzocco & Myers, 2003; Vanbinst et al., 2014) required that students meet criteria for math difficulty more than one time during the study. A deviation from more restrictive cutoff values and the lack of requiring for more than one measurement of low performance to identify students with math difficulty may have critical implications for determining how students are selected to receive additional supports and interventions and for the effectiveness of those interventions. For example, previous intervention research indicates that even when the average effectiveness of an intervention is reported as moderate, the subgroup performances of students with more severe math difficulty may in fact be much lower than both the average effect and the effect with less severe math difficulty (Toll & Van Luit, 2012).

In summary, in the 35 studies included, more than 15 methods of identification were used to identify students in just the first category of math difficulty (i.e., the most restrictive category). Research teams with works included in this synthesis, as well as researchers conducting interventions with students with math difficulty, have noted the absence of a "gold standard" for determining math difficulty (Fuchs et al., 2005; Mazzocco et al., 2013; Vanbinst et al., 2014). When researchers use different methods to identify students as having math difficulty, the irregularity may make it difficult for practitioners to confidently identify students who may need more intensive academic supports. Moreover, inconsistent criteria and methods of identification make it quite difficult to compare results across studies. For these reasons, practitioners and researchers would benefit from a consensus on what defines *math difficulty.*

With our second research question, we investigated the longitudinal growth of students with and without math difficulty on math measures, and although results were somewhat inconsistent, students with math difficulty performed lower than typically achieving peers and generally had parallel or slower growth on measures of early numeracy, computation, rational numbers, and broad math. Due to the nature of identifying students with math difficulty, it was expected that students with math difficulty would have lower initial math achievement; however, it was discouraging that students with math difficulty failed to catch up to peers without math difficulty. The results of this synthesis were consistent with what is known about students with math difficulty regarding difficulty with skills such as counting and computation (Geary, 2004).

Many studies followed students longitudinally through early elementary school, and a few studies continued tracking students through fifth grade and into middle school. The results of a review of longitudinal predictors of math achievement determined that early numeracy skills—such as quantity comparison, counting, and understanding of the number line (as measured in kindergarten and first grade) and computation (as measured in early elementary school)—were instrumental to student performance in broad math in fifth and sixth grades (Powell, Nelson, & Peng, 2017). If broad math achievement in later grades can be predicted as early as kindergarten with early numeracy measures, then the results of this review suggest that students with math difficulty have lower achievement in skills such as early numeracy. Therefore, students with math difficulty may lag behind peers throughout school unless targeted interventions are implemented to address deficits in math early. Moreover, our results shed light on the range of math skills in which students persistently have difficulty, which may inform necessary next steps for instruction and intervention in math across grade levels. Students with math difficulty consistently displayed deficits in counting, computation, use of retrieval strategies, fraction comparison and estimation, and applied problem solving. The results of this review also highlight the lack of research regarding growth for students with math difficulty in specific areas of math. Of the studies included, very few examined growth rates related to word problem solving and fractions. As all students are expected to set up and solve word problems on high-stakes standardized assessments and as understanding fractions is instrumental to success in algebra, more research related to problem solving and rational numbers is necessary. Furthermore, because no studies in this systematic review measured achievement beyond ninth grade, we could not determine growth patterns for students with math difficulty related to other skills, such as algebra or geometry. Finally, although the primary purpose of this synthesis did not focus on math language, only one study measured growth in understanding of specific math language (Toll & Van Luit, 2014). Because comorbidity rates of math and reading difficulty are high, researchers may consider achievement and growth in understanding specific math language a primary focus in future research related to longitudinal achievement of students with math difficulty.

Our third research question aimed to identify whether math difficulty was related to earlier or later math performance. Generally, studies reported that early numeracy skills, such as counting, number naming, seriation, and magnitude comparison, were predictive of math difficulty later on, with other skills, such as computation (Geary et al., 2009; Locuniak & Jordan, 2008; Stock et al., 2010; Vukovic & Siegel, 2010). What is interesting is that these early numeracy skills are typically learned informally prior to beginning school or early in formal schooling upon kindergarten entry. Furthermore, these skills are significantly related to broad math performance through fifth grade (Powell, Nelson, & Peng, 2017). Thus, the results of this review suggest that it is critical to address deficits for students who exhibit math difficulty early, such as use of procedural skills over retrieval skills for simple arithmetic, inability to identify correct or incorrect counting, and deficits in magnitude comparison and number naming. Practitioners must be aware of the skills that predict later difficulty, determine which students may be at risk for math difficulty in later grades, and intervene with targeted instruction in those skills in hopes of curtailing future math deficits.

Our fourth research question addressed the stability of math difficulty according to different factors. Unfortunately, the number of studies that reported stability of math difficulty did not allow for analyses to examine the potential relationships between stability and type of measure and initial age of identification for participants. The results of this review, however, report themes regarding the restrictiveness of cutoffs and the proportion of students who retained the math difficulty classification. These findings have critical implications for how researchers and practitioners continue to identify students with math difficulty. As discussed, there is a lack of consistency regarding the cutoff used to determine math difficulty status; the results of this synthesis suggest that more restrictive cutoffs (e.g., <10th percentile) result in higher proportions of the same students exhibiting math difficulty in later grades. Therefore, practitioners may consider employing stricter cutoffs to determine which students are at the greatest risk for later math failure. As resources in schools are limited, it is critical to provide services to the students who have the highest probability of math failure without extra support. Furthermore, the limited research on the stability of math difficulty sheds light on a gap in the research base in early intervention and special education. Of the 11 studies that reported on the stability of math difficulty over time, only one study (Martin et al., 2013) reported on stability for participants beyond fourth grade. Little is known regarding the stability of math difficulty over time beyond elementary school. More research related to the stability of math difficulty that spans other grade levels (e.g., middle school and beyond) and considers the retention of math difficulty status for more than 2 or 3 years will move the field toward a consensus on the criteria for identifying students with math difficulty.

Limitations and Directions for Future Research

As with any study, there are limitations. First, participants from approximately 75% of the studies were likely representative of the same geographic areas because seven research teams conducted the 75% of the studies. This may limit the generalizability of the results to students in other areas of the United States and even in other countries, as well as students from different cultural and ethnic backgrounds. Furthermore, the results of this synthesis may be limited because the research teams that represented the majority of the work may have employed similar methodologies across their studies. Future research should include more longitudinal efforts that include participants from other cultural and regional populations as a way to replicate and extend current research findings.

Second, much of the data collected in these studies was collected prior to the adoption of the Common Core State Standards (CCSS; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Many American states have since adopted the CCSS, which means that standards for math learning and instruction are in flux. Thus, the differences in growth of math skills across time between students with and without math difficulty may change as a result of alternation to math standards. Researchers may want to replicate their work within the era of the CCSS. Moreover, this synthesis did not include studies that examined the effectiveness of interventions, but future research may explore the differences between growth for students with and without math difficulty when students with math difficulty receive extra support that corresponds to national and state standards for math learning.

Third, many of the studies included in this synthesis focused on students in kindergarten through third grade, with a handful of studies extending to fifth grade. We did not review any studies that focused on preschool or high school populations. This limits the generalizability of our findings and the knowledge base regarding what we know about students with math difficulty across school-age years regarding differences in growth, predictability of difficulty, and stability of math difficulty over time. In addition, because more studies focused on early elementary grades, we reported many findings on early numeracy skills, computation, and some findings on fractions. Unfortunately, due to the restricted age of participants in this synthesis, we cannot draw conclusions about other skills, such as word problem solving, algebra, or geometry. Future longitudinal research should include preschool participants and extend into the middle and high school years to gain knowledge about math difficulty in students of all ages.

Implications for Practitioners

The results of this systematic review provide the foundation for implications for practitioners. First, students with more severe and persistent math difficulty may have fundamental math differences from peers with math difficulty identified with less restrictive cutoffs or difficulty that is not persistent. Students identified with math difficulty with a restrictive cutoff have lower initial performance and slower growth than students who are identified with math difficulty with less restrictive criteria, and students with persistent math deficits typically perform lower than peers with inconsistent deficits. These results have implications for practitioners, as there may be distinct differences within a subgroup of students with math difficulty and the effectiveness of an intervention may be moderated by the degree to which students exhibit math deficits. Practitioners may consider results with caution, as students identified as experiencing math difficulty with varying degrees of restriction may actually represent fundamentally different groups of students; thus, interventions that are effective for some students with math difficulty may not be as effective for other students with math difficulty (Powell, Cirino, & Malone, in press). Practitioners can be proactive by frequently monitoring the progress of students with math difficulty, as a method for determining at what rate they are growing in relation to the peers, such that they can make immediate changes to instruction.

Second, practitioners may consider evaluating the instructional methods that they use with students with math difficulty. The results of this review consistently show that students with math difficulty do not catch up to peers across math content areas. This raises questions regarding the effectiveness of the instruction and intervention methods that students with math difficulty are exposed to in the classroom. Previous research has identified high-quality and evidence-based practices in mathematics that practitioners may use with students with learning difficulties, such as explicit and systematic instruction (Gersten et al., 2009) and use of representations (Jitendra, Nelson, Pulles, Kiss, & Houseworth, 2016). Districts and schools may consider the importance of providing their practitioners with more opportunities for professional development and on-site coaching related to evidence-based practices in math for students who struggle. Another consideration for schools and practitioners is to examine the implementation fidelity of instructional programs and the alignment of interventions to specific student deficits. Perhaps students with math difficulty who show slower rates of growth over time do not receive interventions at the recommended dosage or do not receive intensive intervention that is aligned to the specific skill deficit in math. Furthermore, researchers who conduct longitudinal research may consider collecting data regarding the types of instruction that students in their studies are receiving, to provide more insight related to growth rates and stability of math difficulty.

Conclusion

We reviewed longitudinal studies to compare the growth of students with and without math difficulty, examined specific predictors of math difficulty, and evaluated the stability of math difficulty regarding classification criteria. At all grade levels and across skills, students with math difficulty consistently performed lower than students without math difficulty and had similar or slower growth. Moreover, the likelihood of retaining math difficulty across years was high. Thus, students with math difficulty did not catch up to their peers. Without targeted interventions and early determination of difficulty with math, students as early as kindergarten who display math difficulty may be at risk for poor secondary and adulthood outcomes.

Authors' Note

The content is solely the responsibility of the authors and does not necessarily represent the official views of the Institute of Education Sciences or the U.S. Department of Education.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported in part by Grant R324A150078 from the Institute of Education Sciences in the U.S. Department of Education to the University of Texas at Austin. This research was also supported by the Greater Texas Foundation.

References

*Asterisk denotes a study included in this synthesis.

Andersson, U. (2008). Mathematical competencies in children with different types of learning difficulties. *Journal of Educational Psychology*, *100*, 48–66. doi:10.1037/0022-0663.100.1.48

*Andersson, U. (2010). Skill development in different components of arithmetic and basic cognitive functions: Findings from a 3-year longitudinal study of children with different types of learning difficulties. *Journal of Educational Psychology*, *102*, 115–134. doi:10.1037/a0016838

Ashkenazi, S., & Henik, A. (2010). Attentional networks in developmental dyscalculia. *Behavioral and Brain Functions*, *6*(2), 1–12. doi:10.1186/1744-9081-6-2

Baum, S., Ma, J., & Payea, K. (2010). *Education pays 2010: The benefits of higher education for individuals and society*. New York, NY: The College Board.

*Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *The Elementary School Journal*, *108*, 115–130. doi:10.1086/525550

Boets, B., De Smedt, B., & Ghesquiere, P. (2011). Coherent motion sensitivity predicts individual differences in subtraction. *Research in Developmental Disabilities*, *32*, 1075–1080. doi:10.1016/j.ridd.2011.01.024

Branum-Martin, L., Fletcher, J. M., & Stuebing, K. K. (2012). Classification and identification of reading and math disabilities: The special case of comorbidity. *Journal of Learning Disabilities*, *46*, 490–499. doi:10.1177/0022219412468767

Bryant, D. P., Bryant, B. R., Gersten, R. M., Scammacca, N., Funk, C., Winter, A., . . . Pool, C. (2008). The effects of Tier 2 intervention on the mathematics performance of first-grade students who are at risk for mathematics difficulties. *Learning Disability Quarterly*, *31*, 47–63. doi:10.2307/20528817

Bryant, B. R., Bryant, D. P., Porterfield, J., Dennis, M. S., Falcomata, T., Valentine, C., . . . Bell, K. (2016). The effects of a Tier 3 intervention on the mathematics performance of second grade students with severe mathematics difficulties. *Journal of Learning Disabilities*, *49*, 176–188. doi:10.1177/0022219414538516

Calhoon, M. B., Emerson, R. W., Flores, M., & Houchins, D. E. (2007). Computational fluency performance profile of high school students with mathematics disabilities. *Remedial and Special Education*, *28*, 292–303. doi:10.1177/074193250702 80050401

*Chong, S. L., & Siegel, L. S. (2008). Stability of computational deficits in math learning disability from second through fifth grades. *Developmental Neuropsychology*, *33*, 300–317. doi:10.1080/87565640801982387

Cirino, P. T., Fuchs, L. S., Elias, J. T, Powell, S. R., & Schumacher, R. F. (2015). Cognitive and mathematical profiles for different forms of learning difficulties. *Journal of Learning Disabilities*, *48*, 156–175. doi:10.1177/0022219413494239

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*, 415–427. doi:10.1016/j.econedurev.2008.09.003

*Claessens, A., & Engel, M. (2013). How important is where you start? Early mathematics knowledge and later school success. *Teachers College Review*, *115*, 1–29.

Clarke, B., Smolkowski, K., Baker, S. K., Fien, H., Doabler, C. T., & Chard, D. J. (2011). The impact of a comprehensive Tier

1 core kindergarten program on the achievement of students at risk in mathematics. *The Elementary School Journal*, *111*, 561–584. doi:10.1086/659033

Connolly, A. J. (2007). *KeyMath3*. Bloomington, MN: Pearson.

*Cowan, R., & Powell, D. (2014). The contributions of domain-general and numerical factors to third-grade arithmetic skills and mathematical learning disability. *Journal of Educational Psychology*, *106*, 214–229. doi:10.1037/a0034097

De Smedt, B., & Gilmore, C. K. (2011). Defective number module or impaired access? Numerical magnitude processing in first grades with mathematical difficulties. *Journal of Experimental Child Psychology*, *108*, 278–292. doi:10.1016/j.jecp.2010.09.003

*Desoete, A., Ceulemans, A., De Weerdt, F., & Pieters, S. (2012). Can we predict mathematical learning disabilities from symbolic and non-symbolic comparison tasks in kindergarten? Findings from a longitudinal study. *British Journal of Educational Psychology*, *82*, 64–81. doi:10.1348/2044-8279.002002

*Desoete, A., & Grégoire, J. (2006). Numerical competence in young children and in children with mathematics learning disabilities. *Learning and Individual Differences*, *16*, 351–367. doi:10.1016/j.lindif.2006.12.006

Desoete, A., Roeyers, H., & De Clercq, A. (2004). Children with mathematics learning disabilities in Belgium. *Journal of Learning Disabilities*, *37*, 50–61. doi:10.1177/00222194040370010601

Driver, M. K., & Powell, S. R. (2015). Symbolic and nonsymbolic equivalence tasks: The influence of symbols on students with mathematics difficulty. *Learning Disabilities Research and Practice*, *30*, 127–134. doi:10.1111/ldrp.12059

Dyson, N., Jordan, N. C., Beliakoff, A., & Hassinger-Das, B. (2015). A kindergarten number-sense intervention with contrasting practice conditions for low-achieving children. *Journal for Research in Mathematics Education*, *46*, 331–370. doi:10.5951/jresematheduc.46.3.0331

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*, 493–513. doi:10.1037/0022-0663.97.3.493

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., . . . Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*, 29–43. doi:10.1037/0022-0663.98.1.29

Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2010). The effects of strategic counting instruction, with and without deliberate practice, on number combination skill among students with mathematics difficulties. *Learning and Individual Differences*, *20*, 89–100. doi:10.1016/j.lindif.2009.09.003

Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., . . . Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology*, *106*, 499–514. doi:10.1037/a0034341

Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, *74*, 155–173. doi:10.1177/001440290807400202

Garrett, A. J., Mazzocco, M. M. M., & Baker, L. (2006). Development of the metacognitive skills of prediction and evaluation in children with or without math disability. *Learning Disabilities Research and Practice*, *21*, 77–88. doi:10.1111/j.1540-5826.2006.00208.x

Geary, D. C. (2000). From infancy to adulthood: The development of numerical abilities. *European Child and Adolescent Psychiatry*, *9*(Suppl. 2), S11. doi:10.1007/s007870070004

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, *37*, 4–15. doi:10.1177/00222194040370010201

*Geary, D. C., Bailey, D. H., & Hoard, M. K. (2009). Predicting mathematical achievement and mathematical learning disability with a simple screening tool. *Journal of Psychoeducational Assessment*, *27*, 265–279. doi:10.1177/0734282908330592

Geary, D. C., Bailey, D. H., Littlefield, A., Wood, P., Hoard, M. K., & Nugent, L. (2009). First-grade predictors of mathematical learning disability: A latent class trajectory analysis. *Cognitive Development*, *24*, 411–429. doi:10.1016/j.cogdev.2009.10.001

*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. doi:10.1006/jecp.2000.2561

*Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*, 1343–1359.

*Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2012). Mathematical cognition deficits in children with learning disabilities and persistent low achievement: A five-year prospective study. *Journal of Educational Psychology*, *104*, 206–223. doi:10.1037/a0025398

*Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology*, *33*, 277–299. doi:10.1080/87565640801982361

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*, 1202–1242. doi:10.3102/0034654309334431

Ginsburg, H. P., & Baroody, A. J. (2003). *Test of Early Mathematics Ability* (3rd ed.). Austin, TX: Pro-Ed.

*Hecht, S. A., & Vagi, K. J. (2010). Sources of group and individual differences in emerging fraction skills. *Journal of Educational Psychology*, *102*, 843–859. doi:10.1037/a0019824

Jitendra, A. K., Nelson, G., Pulles, S. M., Kiss, A. J., & Houseworth, J. (2016). Is mathematical representation of problems an evidence-based strategy for students with mathematics difficulties? *Exceptional Children*, *83*, 8–25. doi:10.1177/0014402915625062

Jitendra, A. K., Rodriguez, M., Kanive, R., Huang, J.-P., Church, C., Corroy, K. A., & Zaslofsky, A. (2012). Impact of small-group tutoring interventions on the mathematical problem solving and achievement of third-grade students with mathematics difficulties. *Learning Disability Quarterly*, *36*, 21–35. doi:10.1177/0731948712457561

Jordan, N. C., Glutting, J., & Ramineni, C. (2010). The importance of number sense to mathematics achievement in first and third grades. *Learning and Individual Differences*, *20*, 82–88. doi:10.1016/j.lindif.2009.07.004

*Jordan, N. C., & Hanich, L. B. (2003). Characteristics of children with moderate mathematics deficiencies: A longitudinal perspective. *Learning Disabilities Research and Practice*, *18*, 213–221. doi:10.1111/1540-5826.00076

*Jordan, N. C., Hanich, L. C., & Kaplan, D. (2003a). Arithmetic fact master in young children: A longitudinal investigation. *Journal of Experimental Child Psychology*, *85*, 103–119. doi:10.1016/S0022-0965(03)00032-8

*Jordan, N. C., Hanich, L. B., & Kaplan, D. (2003b). A longitudinal study of mathematical competencies in children with specific mathematics difficulties versus children with comorbid mathematics and reading difficulties. *Child Development*, *74*, 834–850. doi:10.1111/1467-8624.00571

*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. doi:10.1037//0022-0663.94.3.586

Kingsdorf, S., & Krawec, J. (2014). Error analysis of mathematics word problem solving across students with and without learning disabilities. *Learning Disabilities Research and Practice*, *29*, 66–74. doi:10.1111/ldrp.12029

*Locuniak, M. N., & Jordan, N. C. (2008). Using kindergarten number sense to predict calculation fluency in second grade. *Journal of Learning Disabilities*, *41*, 451–459. doi:10.1177/0022219408321126

Magnuson, K. A., Ruhm, C., & Waldfogel, J. (2007). Does prekindergarten improve school preparation and performance? *Economics of Education Review*, *26*, 33–51. doi:10.1016/j.econedurev.2005.09.008

Mancl, D. B., Miller, S. P., & Kennedy, M. (2012). Using the concrete-representational-abstract sequence with integrated strategy instruction to teach subtraction with regrouping to students with learning disabilities. *Learning Disabilities Research and Practice*, *27*, 152–166. doi:10.1111/j.1540-5826.2012.00363.x

*Martin, R. B., Cirino, P. T., Barnes, M. A., Ewing-Cobbs, L., Fuchs, L. S., Stuebing, K. K., & Fletcher, J. M. (2013). Prediction and stability of mathematics skill and difficulty. *Journal of Learning Disabilities*, *46*, 428–443. doi:10.1177/0022219411436214

*Mazzocco, M. M. M., & Devlin, K. T. (2008). Parts and "holes": Gaps in rational number sense among children with vs. without mathematical learning disabilities. *Developmental Science*, *11*, 681–691. doi:10.1111/j.1467-7687.2008.00717.x

*Mazzocco, M. M., Feigenson, L., & Halberda, J. (2011). Impaired acuity of the approximate number system underlies mathematical learning disability (dyscalculia). *Child Development*, *82*, 1224–1237. doi:10.1111/j.1467-8624.2011.01608.x

*Mazzocco, M. M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia*, *53*, 218–253. doi:10.1007/s11881-003-0011-7

*Mazzocco, M. M. M., Myers, G. F., Lewis, K. E., Hanich, L. B., & Murphy, M. M. (2013). Limited knowledge of fraction representations differentiates middle school students with mathematics learning disability (dyscalculia) versus low mathematics achievement. *Journal of Experimental Child Psychology*, *115*, 371–387. doi:10.1016/j.jecp.2013.01.005

*Mazzocco, M. M., & Thompson, R. E. (2005). Kindergarten predictors of math learning disability. *Learning Disabilities Research & Practice*, *20*, 142–155. doi:10.1111/j.1540-5826.2005.00129.x

Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Körne, G. (2014). Specific learning disorder: Prevalence and gender differences. *PLoS ONE*, *9*(7), e103537. doi:10.1371/journal.pone.0103537

*Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-year growth trajectories of kindergarten children with learning difficulties in mathematics. *Journal of Learning Disabilities*, *42*, 306–321. doi:10.1177/0022219408331037

Murphy, M. M., & Mazzocco, M. M. M. (2009). The trajectory of mathematics skills and working memory thresholds in girls with fragile X syndrome. *Cognitive Development*, *24*, 430–449. doi:10.1016/j.cogdev.2009.09.004

*Murphy, M. M., Mazzocco, M. M., Hanich, L. B., & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, *40*, 458–478.

Murray, E., & Harrison, L. J. (2011). The influence of being ready to learn on children's early school literacy and numeracy achievement. *Educational Psychology*, *31*, 529–545. doi:10.1080/01443410.2011.573771

National Council of Teachers of Mathematics. (2006). *Curriculum focal points for prekindergarten through Grade 8 mathematics*. Reston, VA: Author.

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common Core State Standards mathematics*. Washington, DC: Authors.

*Navarro, J. I., Aguilar, M., Marchena, E., Ruiz, G., Menacho, I., & Van Luit, J. E. (2012). Longitudinal study of low and high achievers in early mathematics. *British Journal of Educational Psychology*, *82*, 28–41. doi:10.1111/j.2044-8279.2011.02043.x

Nunes, T., Bryant, P., Burman, D., Bell, D., Evans, D., & Hallett, D. (2009). Deaf children's informal knowledge of multiplicative reasoning. *Journal of Deaf Studies and Deaf Education*, *14*, 260–277. doi:10.1093/deafed/enn040

O'Connor, R. E., Bocian, K. M., Beach, K. D., Sanchez, V., & Flynn, L. J. (2013). Special education in a 4-year response to intervention (RtI) environment: Characteristics of students with learning disability and grade of identification. *Learning Disability Research and Practice*, 28, 98–112. doi:10.1111/ldrp.12013

*Peng, P., Namkung, J. M., Fuchs, D., Fuchs, L. S., Patton, S., Yen, L., . . . Hamlett, C. (2016). A longitudinal study on predictors of early calculation development among young children at risk for learning difficulties. *Journal of Experimental Child Psychology*, *152*, 221–241. doi:10.1016/j.jecp.2016.07.017

Powell, S. R., Cirino, P. T., & Malone, A. S. (in press). Child-level predictors of responsiveness to evidence-based mathematics intervention. *Exceptional Children*.

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2009). Do word-problem features differentially affect problem difficulty as a function of students' mathematics difficulty with and without reading difficulty? *Journal of Learning Disabilities*, *42*, 99–110. doi:10.1177/0022219408326211

Powell, S. R., Nelson, G., & Peng, P. (2017). *Mathematics trajectories from preschool to postsecondary: The predictive nature of mathematics performance*. Manuscript submitted for publication.

Rubinsten, O., & Tannock, R. (2010). Mathematics anxiety in children with developmental dyscalculia. *Behavioral and Brain Functions*, *6*(46), 1–13. doi:10.1186/1744-9081-6-46

Scanlon, D. (2013). Specific learning disability and its newest definition: Which is comprehensive? And which is insufficient? *Journal of Learning Disabilities*, *46*, 26–33. doi:10.1177/0022219412464342

*Schwenck, C., Dummert, F., Endlich, D., & Schneider, W. (2015). Cognitive functioning in children with learning problems. *European Journal of Psychology of Education*, *30*, 349–367. doi:10.1007/s10212-014-0242-5

Shalev, R. S., Auerbach, J., Manor, O., & Gross-Tsur, V. (2000). Developmental dyscalculia: Prevalence and prognosis. *European Child and Adolescent Psychiatry*, *9*, II/58–II/64. doi:10.1007/s007870070009

Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, *110*, 18116–18120. doi:10.1073/pnas.1302751110

*Stock, P., Desoete, A., & Roeyers, H. (2010). Detecting children with arithmetic disabilities from kindergarten: evidence from a 3-year longitudinal study on the role of preparatory arithmetic abilities. *Journal of Learning Disabilities*, *43*, 250–268. doi:10.1177/0022219409345011

Swanson, H. L. (2012). Cognitive profile of adolescents with math disabilities: Are the profiles different from those with reading disabilities? *Child Neuropsychology*, *18*, 125–143. doi:10.1080/09297049.2011.589377

Swanson, H. L., & Beebe-Frankenberger, M. (2004). The relationship between working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *96*, 471–491. doi:10.1037/0022-0663.96.3.471

*Swanson, H. L., Jerman, O., & Zheng, X. (2008). Growth in working memory and mathematical problem solving in children at risk and not at risk for serious math difficulties. *Journal of Educational Psychology*, *100*, 343–379. doi:10.1037/0022-0663.100.2.343

Swanson, H. L., Moran, A., Lussier, C., & Fung, W. (2014). The effect of explicit and direct generative strategy training and working memory on word problem-solving accuracy in children at risk for math difficulties. *Learning Disability Quarterly*, *37*, 111–123. doi:10.1177/0731948713507264

Tolar, T. D., Fuchs, L., Fletcher, J. M., Fuchs, D., & Hamlett, C. L. (2016). Cognitive profiles of mathematical problem solving learning disability for different definitions of disability. *Journal of Learning Disabilities*, *49*, 240–256. doi:10.1177/0022219414538520

Toll, S. W. M., & Van Luit, J. E. H. (2012). Early numeracy intervention for low-performing kindergarteners. *Journal of Early Intervention*, *34*, 243–264. doi:10.1177/1053815113477205

*Toll, S. W. M., & Van Luit, J. E. H. (2014). The developmental relationship between language and low early numeracy skills throughout kindergarten. *Exceptional Children*, *81*, 64–78. doi:10.1177/0014402914532233

*Vanbinst, K., Ghesquière, P., & De Smedt, B. (2014). Arithmetic strategy development and its domain-specific and domain-general cognitive correlates: A longitudinal study in children with persistent mathematical learning difficulties. *Research in Developmental Disabilities*, *35*, 3001–3013. doi:10.1016/j.ridd.2014.06.023

*Vukovic, R. K. (2012). Mathematics difficulty with and without reading difficulty: Findings and implications from a four-year longitudinal study. *Exceptional Children*, *78*, 280–300. doi:10.1177/001440291207800302

*Vukovic, R. K., & Siegel, L. S. (2010). Academic and cognitive characteristics of persistent mathematics difficulty from first through fourth grade. *Learning Disabilities Research and Practice*, *25*, 25–38. doi:10.1111/j.1540-5826.2009.00298.x

Watson, S. M. R., & Gable, R. A. (2012). Unraveling the complex nature of mathematics learning disability: Implications for research and practice. *Learning Disability Quarterly*, *36*, 178–187. doi:10.1177/0731948712461489

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*, 352–360. doi:10.3102/0013189X14553660

Wei, X., Lenz, K. B., & Blackorby, J. (2013). Math growth trajectories of students with disabilities: Disability category, gender, racial, and socioeconomic status differences from ages 7 to 17. *Remedial and Special Education*, *34*, 154–165. doi:10.1177/0741932512448253