Hammill Institute on Disabilities

Remedial and Special Education
2018, Vol. 39(6) 327–340
© Hammill Institute on Disabilities 2017
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/0741932517731887
rase.sagepub.com
SAGE

Mathematics Interventions for Upper Elementary and Secondary Students: A Meta-Analysis of Research

Elizabeth A. Stevens, MAEd[1], Melissa A. Rodgers, MEd[1], and Sarah R. Powell, PhD[1]

Abstract

The purpose of this review was to conduct a meta-analysis of 25 years of mathematics interventions for students with mathematics difficulty or disability in Grades 4 through 12. A search of the literature yielded 25 studies that met the inclusion criteria. Studies were coded to extract important study information (e.g., participant information, research design, description of treatment, and comparison groups) and data needed to calculate Hedge's *g*. We used robust variance estimation (RVE) to address dependence resulting from multiple outcomes per study. The RVE random-effects model estimated a treatment effect of 0.85. After adjusting for small-study effects, the final model estimated an underlying, moderate effect of 0.49 with a large amount of unexplained heterogeneity between studies. Studies with more than 15 hr of treatment and those focused on fraction content significantly moderated mathematics outcomes. Findings are limited by extreme variability across study estimates, the lack of standardized mathematics measures, and a limited number of studies across 25 years of research.

Keywords

mathematics, intervention, learning disability, mathematics difficulty, middle school, high school

Proficiency with mathematics in secondary settings sets the stage for whether students enroll in postsecondary institutions or pursue other options after high school (Lee, 2012), and these decisions are directly related to the income that students earn as adults (Baum, Ma, & Payea, 2010; Deming, Cohodes, Jennings, & Jencks, 2013; Dougherty, 2003). Because of the influence of school mathematics on adulthood outcomes, it is necessary to provide all students with effective mathematics instruction during the elementary and secondary grade levels. High-stakes assessment data from the National Assessment of Educational Progress (NAEP; U.S. Department of Education, 2015) indicate students may not be making adequate progress in current mathematics instructional settings. For example, only 40% of fourth-grade students and 33% of eighth-grade students perform at or above proficient levels in mathematics. For 12th-grade students, this percentage drops to 25%. From this data, it is clear that the majority of students may not be prepared for postsecondary mathematics or be able to pursue postsecondary options because of low mathematics scores. It is also apparent that the mathematics performance of students decreases from elementary to middle to high school.

The 2015 NAEP scores also indicate that students with school-identified disabilities perform significantly lower than students without disabilities, with average scores falling at or below basic levels. This trend is consistent across grade levels (Wei, Lenz, & Blackorby, 2013). For students with disabilities or mathematics difficulty (that is, without a formal disability diagnosis), there is a critical need for effective mathematics interventions to improve mathematics proficiency levels, which, in turn, may promote stronger adulthood outcomes. To understand the impact of such mathematics interventions, we conducted a meta-analysis of mathematics interventions for students in Grades 4 through 12 with or at risk for a mathematics learning disability. Information gained from this meta-analysis may improve mathematics instruction for students with disabilities and frame pathways for future research.

In this introduction, we discuss two categories of students who struggle with mathematics: Students with mathematics learning disability or students with mathematics difficulty. We briefly review the importance of mathematics interventions for promoting stronger mathematics outcomes. Then, we discuss previously published meta-analyses conducted in

[1]The University of Texas at Austin, USA

Corresponding Author:
Elizabeth A. Stevens, College of Education SZB 228, The University of Texas at Austin, 1912 Speedway, D4900, Austin, TX, 78712-1284, USA.
Email: elizabeth.stevens@utexas.edu

mathematics, and the results of these analyses. Finally, we explain the purpose of this meta-analysis and how it fills a void in the research literature. We conclude by proposing the research questions that guided this meta-analysis.

Mathematics Disability or Difficulty

The percentage of school-age students with a diagnosed mathematics learning disability typically ranges from 3% to 8% (Desoete, Roeyers, & De Clercq, 2004). In the research literature, this group of students may be diagnosed with dyscalculia (e.g., Butterworth, 2010; Szucs, Devine, Soltesz, Nobes, & Gabriel, 2013), although this term is more popular outside of the United States. For research related to mathematics difficulty, researchers often focus on an additional sample of students: students at risk for developing a mathematics learning disability (e.g., Fuchs et al., 2010; Jitendra et al., 2013; Swanson, Lussier, & Orosco, 2015). Because specific learning disability is not formally diagnosed until the later elementary grades (O'Connor, Bocian, Beach, Sanchez, & Flynn, 2013), if at all, and because of the high numbers of students who struggle with mathematics without a formal learning disability diagnosis (U.S. Department of Education, 2015), expanding the research base to include students with mathematics disability and difficulty allows for a greater understanding of students who struggle with mathematics. For the remainder of this article, we use the term *mathematics difficulty* (MD) to describe students with a specific learning disability in mathematics or persistent low mathematics performance without a disability diagnosis.

In the elementary grades, students with MD perform lower than peers without MD on tasks related to counting (Stock, Desoete, & Roeyers, 2010), comparison of quantities (De Smedt & Gilmore, 2011), and mathematics fluency (Tolar, Fuchs, Fletcher, Fuchs, & Hamlett, 2016). Students with MD also exhibit low word problem-solving performance compared with students without MD (Fuchs et al., 2008; Kingsdorf & Krawec, 2014). Many of these mathematics skills continue to cause problems for students in secondary settings (Calhoon, Emerson, Flores, & Houchins, 2007), especially as the mathematics expectations increase and involve a greater number of mathematics skills as well as reasoning.

On a positive note, students with MD have demonstrated improved mathematics performance when educators implement interventions targeted at improving mathematics (e.g., Bryant et al., 2008; Mancl, Miller, & Kennedy, 2012; Swanson, Moran, Lussier, & Fung, 2014). Some interventions changed the learning trajectories of students with MD, but without intervention, the outcomes may not be as promising.

Previous Meta-Analyses

Specifically related to students with MD, research teams have conducted several meta-analyses. The majority of meta-analyses focused on one mathematical area, such as mathematics fact fluency (Burns, Codding, Boice, & Lukito, 2010; Codding, Burns, & Lukito, 2011; Joseph et al., 2012) or word problem solving (Xin & Jitendra, 1999; Zhang & Xin, 2012; Zheng, Flynn, & Swanson, 2013) for students across the elementary and secondary levels. With a focus on fractions, Shin and Bryant (2015) conducted a meta-analysis of students with MD across Grades 3 through 12. Results from these content-specific meta-analyses indicated that students benefited from focused and explicit instruction.

Chodura, Kuhn, and Holling (2015), Kroesbergen and Van Luit (2003), and Swanson and Jerman (2006) conducted meta-analyses with more generalized types of intervention (i.e., interventions not focused specifically in one mathematics area). Again, results demonstrated the efficacy of mathematics intervention for improving mathematics outcomes, but the overwhelming majority of studies included in these meta-analyses focused on students in the elementary grades. For example, Chodura et al. (2015) and Kroesbergen and Van Luit (2003) targeted elementary-age students in their reviews. Our search produced one meta-analysis conducted at the secondary level (Hughes, Witzel, Riccomini, Fries, & Kanyongo, 2014), but this analysis was limited to algebra interventions.

Although technically not about specific mathematics interventions, Gersten et al. (2009) conducted a meta-analysis about effective instructional practices (e.g., think-alouds), classroom structures (e.g., classwide peer tutoring), and the use of formative data designed to enhance mathematics achievement for students with learning disabilities. Explicit instruction and the use of heuristics demonstrated significant improvement in the mathematics performance of students with learning disabilities. Although not explicitly defined by the authors, our review of the studies included in Gersten et al. (2009) indicated that the majority focused on students in the elementary grades. For teachers interested in instruction, grouping practices, or formative data for elementary students with MD, the findings that emerged from the Gersten et al. (2009) study would be helpful. For researchers and teachers searching for evidence-based interventions for students in Grades 4 through 12, our meta-analysis fills the void.

Across these meta-analyses, students with MD benefited from specially designed instruction. What is absent in the research literature is a comprehensive meta-analysis related to mathematics intervention for students beyond the early elementary grades (i.e., after third grade). For these students, the mathematics demands grow exponentially from the early elementary grades, and the high-stakes assessments administered in late elementary, middle, and high school have implications for college and career readiness. Understanding which interventions demonstrate efficacy and the qualities of such interventions (e.g., treatment duration, dosage, and mathematics content) is necessary to frame mathematics instruction for students with MD.

Purpose and Research Questions

To improve mathematics outcomes for students with MD beyond the early elementary grades, there is a need to understand the effect of mathematics interventions for students with MD. Educators must provide efficacious interventions to alleviate difficulty with mathematics and adequately prepare students for the rigors of mathematics in adulthood. To understand interventions focused on mathematics improvement for students with MD, we asked the following research questions:

Research Question 1: What are the effects of mathematics interventions on mathematics outcomes for students with MD in Grades 4 through 12?
Research Question 2: Do intervention outcomes differ as a function of (a) treatment duration or dosage, (b) mathematics content, (c) study characteristics, and (d) study quality?

To answer these questions, we conducted a meta-analysis of peer-reviewed research conducted over the past 25 years.

Method

Operational Definitions

In this article, a *mathematics disability* refers to a student with a school-identified learning disability. Mathematics difficulty is defined as low mathematics performance determined by an educator or scores on a mathematics measure below the 40th percentile. We selected this cutoff percentile because scores above the 40th percentile are often used to designate typical mathematics performers from students with mathematics difficulty (e.g., Hecht & Vagi, 2010; Jordan, Kaplan, & Hanich, 2002; Martin et al., 2013). From this point forward, we refer students with mathematics disability or difficulty as experiencing MD. We define a *mathematics intervention* as specialized instruction in a particular domain of mathematics that is not provided to a whole classroom of students (i.e., instruction provided in small group or individual settings).

Search Procedures

We conducted a search of four electronic databases: Academic Search Complete, Education Source, Educational Resources Information Clearinghouse (ERIC), and PsycINFO. The search was limited to peer-reviewed journals published between January 1990 and December 2015. We selected 1990 as the start date of the search because it followed the 1989 release of the National Council of Teachers of Mathematics curriculum standards (1989), which substantially changed the direction of mathematics standards in the United States. We used the following search terms: *math*, arithmetic, geometry, algebr*, problem solving, word problem, division*, OR *calculus*, and *disabilit*, at risk, dyscalculia, OR special education*, and *intervention, strateg*, program*, train*, instruction, Tier 2, OR Tier 3.*

Figure 1 displays a PRISMA diagram describing the search process (Moher, Liberati, Tetzlaff, & Altman, 2009). The initial search yielded 7,621 abstracts. We also conducted an ancestral search using the reference lists from relevant mathematics syntheses and meta-analyses conducted since 2000 (Baker, Gersten, & Lee, 2002; Chodura et al., 2015; Hughes et al., 2014; Kroesbergen & Van Luit, 2003; Maccini, Mulcahy, & Wilson, 2007; Myers, Wang, Brownell, & Gagnon, 2015; Shin & Bryant, 2015; Zhang & Xin, 2012; Zheng et al., 2013). This ancestral search resulted in an additional 32 studies for review. Finally, we completed a 2-year hand search of the following journals: *Exceptional Children, Intervention in School and Clinic, Journal of Learning Disabilities, Journal of Special Education, Learning Disabilities Quarterly, Learning Disabilities Research & Practice*, and *Remedial and Special Education*. We selected these journals because they contain the most relevant empirical research in the field of intervention research and special education. We identified no additional articles through the hand search. After removing duplicates, we screened the abstracts of 5,955 records (78%) and identified 375 articles (5%) that met initial inclusion criteria (i.e., mathematics intervention in Grades 4 through 12). We reviewed the full text of these 375 articles and identified 25 experiments (i.e., 24 publications; 0.3%) that met inclusion criteria for this meta-analysis.

Inclusion Criteria

We included studies that met the following inclusion criteria:

1. The study employed an experimental design that assigned participants randomly to groups, a quasi-experimental design in which participants were not randomly assigned to groups, or a multiple treatment design (i.e., studies comparing treatments rather than a treatment with a business-as-usual group). We included multiple treatment designs if one of the treatments served as a contrast to the treatment of interest (i.e., mathematics intervention). We excluded single group, single subject, qualitative, and case study designs.
2. Participants included students identified with MD in Grades 4 through 12. Studies with additional participants (i.e., those in kindergarten through Grade 3) were included if at least 50% of the sample included the targeted grades (i.e., Grades 4 through 12) or results were disaggregated for the targeted grades. Studies with combined samples of participants (i.e., students with and without MD or students with MD and other disabilities) were

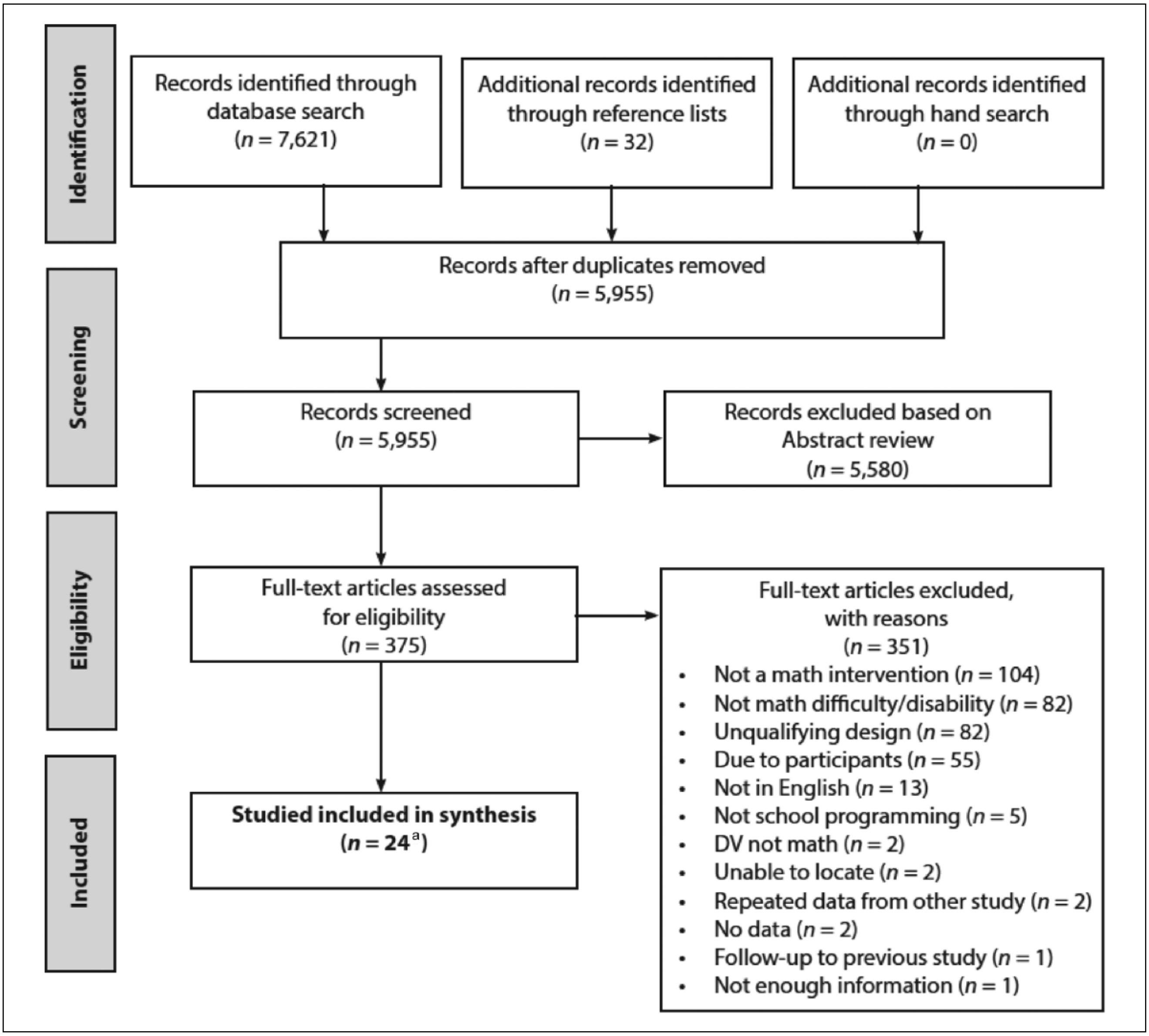

Figure 1. PRISMA statement.
[a]One study included two experiments; manuscript describes 25 studies.

included if data were disaggregated for students identified with MD or at least 50% of the sample included students with MD. We excluded studies targeting students who are deaf or hard of hearing, blind or visually impaired, or identified with autism, Down syndrome, or intellectual disability (i.e., intelligence quotient [IQ] of 70 or below). We excluded interventions conducted with third grade students for two reasons. First and most importantly, the research examining interventions for students in third grade and younger has been synthesized in previous reviews (e.g., Chodura et al., 2015; Gersten et al., 2009; Kroesbergen & Van Luit, 2003; Shin & Bryant, 2015; Zhang & Xin, 2012; Zheng et al., 2013). Baker et al. (2002) conducted the most recent review examining mathematics instruction for older students with MD that was not limited by mathematics content (e.g., problem solving). We expected more research had been generated in the last 15 years, especially with new sets of mathematics standards released in the United States (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). Second, although it is important to synthesize the literature at third grade, we were interested in the effects of

mathematics interventions after the transition to multiplicative knowledge. Mathematics instruction in the early elementary grades focuses on numeracy and additive reasoning; in this review, we aimed to examine mathematics interventions after students' transition to multiplicative thinking, thus targeting multiplicative and higher level mathematics reasoning.
3. The study examined a mathematics intervention provided in English by a teacher or paraprofessional. We excluded peer-mediated interventions because meta-analyses have investigated the effect of such interventions (e.g., Bowman-Perrott et al., 2013; Rohrbeck, Ginsburg-Block, Fantuzzo, & Miller, 2003), and we wanted to understand the impact of educator-led interventions.
4. Intervention was provided as part of the school programming. We excluded home, clinic, and camp settings.
5. At least one dependent variable addressed mathematics outcomes.

Coding Manual

The authors developed a coding protocol and codebook based on recommended guidelines for intervention-related research (Cook et al., 2015; Institute of Education Sciences, 2014); the following sections were developed and collected as part of the coding manual:

- Context and setting (i.e., type of school, type of classroom, curriculum, geographic location, socioeconomic status of school),
- Participant information (age or grade level, gender, ethnicity, socioeconomic, and language status),
- Risk status (i.e., mathematics difficulty or disability of the participants as determined by teacher identification, individualized education program goals in mathematics, a mathematics measure indicating mathematics performance below 40% prior to instruction, an identified mathematics disability or learning disability),
- Research design (i.e., treatment-comparison experiment with random assignment to groups, treatment-comparison quasi-experiment without random assignment to groups, multiple treatment experiment with random assignment to groups, multiple treatment quasi-experiment without random assignment to groups),
- Intervention agent and relevant training (i.e., the role of the individual providing the intervention, relevant background information, and specified training, if any),
- Detailed intervention procedures and materials for each treatment provided,
- Implementation dosage (i.e., minutes per session, total sessions, frequency of sessions, duration of intervention in weeks),
- Implementation fidelity (i.e., did the authors assess and report adherence to the intervention throughout implementation and across interventionists and settings?),
- Description of the comparison group (i.e., What aspects, if any, were reported for the nature of the comparison group's instruction? Did the comparison condition have no or limited access to the treatment?),
- Overall and differential attrition,
- Outcome measures (i.e., description of each measure, administration procedures, and reliability and validity information), and
- Results for each mathematics measure (i.e., means and standard deviations on measures for each group, or other statistical information, such as F values, that allowed for calculation of effect sizes [ESs]).

In addition, we evaluated each study's methodological rigor using the Council of Exceptional Children's quality standards for classifying evidence-based practices in special education (Cook et al., 2015). Studies received a rating (i.e., 0 or 1) according to each of the group comparison indicators within the following areas: context and setting, participants, intervention agent, description of practice, implementation fidelity, internal validity, outcome measures and dependent variables, and data analysis. An overall quality rating (i.e., ranging from 0 to 1) was calculated by averaging a study's scores across all indicators. The quality ratings are reported in Supplemental Table 1S.

Coding Procedure

The code sheet was tested using a sample study prior to conducting reliability. We discussed coding issues, and the protocol and codebook were further revised. The first and third authors independently coded 20% of the included studies and compared codes to achieve reliability. Interrater agreement was calculated as the agreements divided by the number of agreements plus disagreements, multiplied by 100; we achieved 92% interrater agreement across all coding items. The first author coded the remaining studies, which were double-coded by the third author. Any discrepancies in coding were resolved via discussion.

ES Calculation

To quantify intervention effects, we calculated the standardized mean difference and corrected for small sample size bias using Hedge's g (Pigott, 2012). We calculated ESs using R software (version 3.3.0; R Core Team, 2016) for each treatment and comparison contrast on all mathematics-related outcomes.

Meta-Analytic Model

We used a random-effects meta-regression model to account for and explore sources of between-study heterogeneity. As a result of multiple outcome measures or multiple treatment groups per study, 21 of the 25 studies contributed more than one ES. To address this dependency, we used robust variance estimation (RVE; Hedges, Tipton, & Johnson, 2010). RVE uses a mean correlation value (i.e., the correlations between the ESs in each study is usually unreported), ρ, to calculate study weights and between-study variance. This meta-analytic technique provides more precise standard errors and uses all available outcome data (i.e., rather than averaging the ES estimates per study or selecting one ES per study; Hedges et al., 2010). We ran a sensitivity analysis with various ρ values and found the results did not differ based on the selected value; the results reported are based on a ρ value of 0.8. We also used a small sample correction, adjusting each coefficient's degrees of freedom to address inflated Type I error rates for meta-analyses with less than 40 studies (Tipton, 2013).

For each of these analyses, statistical heterogeneity is presented using the I^2 descriptive indices, which is reflective of the percentage of variation, or remaining inconsistency, between studies after accounting for sampling error. Minimal heterogeneity is thought to be present in the population when $I^2 < 25\%$, and high or extreme when $I^2 > 75\%$ (Cooper, Hedges, & Valentine, 2009; Higgins, Thompson, Deeks, & Altman, 2003). To explore the source of detected heterogeneity, additional moderator analyses were run using the following a priori categorical variables: number of total intervention hours (less than 15 hr vs. 15 hr or more), number of total sessions (less than 30 sessions vs. 30 sessions or more), grade level (upper elementary vs. middle and high school), group size (less than eight vs. eight or more), treatment content (problem solving vs. operations vs. fractions), and study quality (0–0.74 vs. 0.75–1.0).

To inspect for the presence of publication bias, we used weighted meta-regression correction method to estimate an empirical pooled ES adjusted for small-study effects (Stanley & Doucouliagos, 2014). Specifically, a hypotheses-based, conditional approach to estimate the most appropriate adjusted pooled ES estimate is a combination of the Funnel Plot Asymmetry Test (FAT), Precision Effect Test (PET), and Precision Effect Estimation With Standard Error (PEESE). We chose to use the FAT–PET–PEESE method for its simplicity in design and interpretation—a linear regression model that controls for the precision of the estimates using the standard error or the variance as a covariate.

The FAT–PET is estimated using a weighted least squares version of Egger's Regression (Egger, Smith, Schnieder, & Minder, 1997), regressing ESs on the standard errors of the estimates. A FAT slope coefficient (b_1) significantly different than zero indicates asymmetry and the presence of small-study effects. The PET intercept (b_0) is the underlying ES estimate after controlling for these small-study effects. The PEESE estimation simply replaces the standard error predictor with the variance of ESs, and again, the intercept coefficient (b_0) is interpreted as an empirical ES adjusted by holding the variance for all studies constant.

The FAT-PET-PEESE method is a conditional procedure that examines the significance of the FAT, or the slope coefficient (b_1). This test is known to be underpowered; if there is evidence of selection bias, it is recommended to proceed to the PET-PEESE interpretations of the underlying effect after adjusting for small-study effects (Egger et al., 1997; Stanley, 2008). The PET intercept (b_0) provides the most reliable estimate if the coefficient is not significantly different than zero, implying that there is an underlying zero effect. If the PET estimate is significant, indicating a genuine nonzero effect, then the PEESE intercept (b_0) is considered the most valid estimate of the pooled ES (Stanley, 2008; Stanley & Doucouliagos, 2014).

The FAT-PET-PEESE is a relatively new approach, and there is debate among recent methodological studies as to whether the variance or standard error results in better estimates after accounting for small-study effects. Consequently, we explored variation in the adjusted ES estimates of both the PET and the PEESE results. We also applied the common method of Trim and Fill that estimates the number of studies missing from the sample as a function of ES and sample size, and then imputes missing effects to estimate a new, pooled ES (Duval & Tweedie, 2000). However, this method has been critiqued for potentially resulting in a biased, adjusted pooled estimate, particularly in the presence of heterogeneity (Moreno et al., 2009; Peters, Sutton, Jones, Abrams, & Rushton, 2007; Terrin, Schmid, Lau, & Olkin, 2003). In addition, it ignores dependency and does not allow for the inclusion of moderators to explain heterogeneity. Recent research indicates that incorporating the simultaneous assessment and adjustment for publication bias into general weighted least squares models (e.g., FAT-PET-PEESE) outperformed other methods when estimating the adjusted ES in the presence of moderate to extreme heterogeneity (Rücker, Schwarzer, Carpenter, Binder & Schumacher, 2011; Stanley & Doucouliagos, 2014).

Results

The 25 studies identified for this meta-analysis were published from 1990 to 2015, were implemented for 2 to 96 hr

Table 1. Summary of Characteristics for Studies (N = 25).

Characteristic	n	%
Publication year		
1990s	8	32
2000s	12	48
2010–2016	5	20
Study location		
The United States	17	68
International	8	32
Grade level[a]		
Elementary (k–5)	15	60
Middle (6–8)	3	12
High (9–12)	8	32
Math content		
Operations	10	40
Fractions	6	24
Problem solving	8	32
General skills	1	4
Instructional group size		
<8	17	68
8 or more	8	32
Sample size		
<25	6	24
25–50	11	44
51–100	3	12
>100	5	20
Total sessions		
<10	2	8
10–20	9	36
21–30	5	20
31–50	5	20
>50	4	16
Total hours		
<10	8	32
10–20	10	40
21–30	3	12
31–50	2	8
>50	2	8

[a]Several studies included more than one category.

and 5 to 128 sessions, and included sample sizes of students with MD ranging from 6 to 259. The interventions provided instruction in one of the following areas: operations, fractions, problem solving, or general mathematics skills. Table 1 presents the descriptive characteristics of the 25 studies included in this meta-analysis, Supplemental Table 1S reports study characteristics, including all moderators used in the exploratory analyses, and Supplemental Table 2S provides Hedge's g ESs by study and treatment comparisons for the 102 ESs included in the analysis. Finally, Tables 2 and 3 present the parameter estimates from the RVE random-effects model, the meta-regression correction method model (FAT-PET-PEESE), and the moderator analyses. An average of four ESs was contributed per study, ranging from 1 to 12. There was great variability in the magnitude and direction of the effects, ranging from −0.66 to 4.65, with a median of 0.71.

The RVE random-effects model estimated a statistically significant treatment effect of 0.85 ($p < .001$); 86.5% of the variance in ESs was between studies ($\tau^2 = 0.48$, $I^2 = 86\%$). The large standard error (0.14) resulted in a wide confidence interval (CI = [0.56, 1.14]).

Meta-Regression Correction Method: FAT-PET-PEESE

Results from the full sample. The FAT slope coefficient was positively significant ($p = .01$), confirming the presence of small-study effects. The PET and PEESE estimates of the underlying, overall ESs were −0.04, CI = [−0.70, 0.62], and 0.51, CI = [0.16, 0.87], respectively. The PET is known to underestimate the underlying effect whereas the PEESE overestimates the underlying effect (Stanley & Doucouliagos, 2014). Figure 2 layers the PET and PEESE regression lines on a funnel plot using the RVE random-effects estimate, showing the underlying ESs after controlling for either the standard error (PET) or variance (PEESE) of the ESs. The PEESE estimated a significant effect whereas the PET did not. The wide and overlapping CIs for the PET and PEESE estimates, however, suggest the need to further explore extreme, potentially influential studies and examine sources of heterogeneity.

Results excluding studies with small samples. The underestimated, pooled PET estimate may have been the result of extreme studies representative of the small-study effects, which explained the variation in estimates between the two models. We conducted a sensitivity analysis by excluding studies with small sample sizes (i.e., treatment or comparison group size was less than 10 for a treatment–comparison contrast), thereby assessing publication bias via multiple methods (i.e., FAT, PET, PEESE) to provide a more informative basis for the true underlying effect. After removing studies that potentially skewed the adjustments for small-study effects (i.e., 16 ESs from four studies where treatment or comparison group size was less than 10 for each treatment–comparison contrast; Bottge, 1999; Manalo, Bunnell, & Stillman, 2000, Experiments 1 and 2; Scarlato & Burr, 2002), the PET estimated an effect of 0.49, CI = [−0.38, 1.36], and the PEESE estimated an effect of 0.54, CI = [0.02, 1.05]. Although both estimates were smaller than the random-effects estimate (ES = 0.85), they were closer in magnitude than the estimates resulting from the full sample of studies.

Comparative Trim and Fill method. Trim and Fill results also confirmed the presence of publication bias, with 15 ESs estimated missing to the left of the pooled estimate

Table 2. Parameter Estimates From RVE Random-Effects Model and Meta-Regression Correction Methods.

	Results from full sample (*N* = 25)					Results excluding small sample studies (*n* = 21)				
	B	*SE*	95% CI	*df*	*p*	*B*	*SE*	95% CI	*df*	*p*
RVE random-effects model	0.85	0.14	[0.56, 1.14]	23	<.001	0.73	0.13	[0.46, 0.99]	20	< .001
FAT	2.60	0.72	[0.76, 4.45]	5	.015	0.79	1.20	[−1.84, 3.40]	12	.525
PET	−0.04	0.29	[−0.70, 0.62]	8	.899	0.49	0.38	[−0.38, 1.36]	9	.236
PEESE	0.51	0.17	[0.16, 0.87]	15	.008	0.54	0.23	[0.02, 1.05]	9	.042

Note. RVE random-effects model calculated prior to any adjustments for small-study effects. Meta-Regression Correction Methods included FAT, PET, and PEESE. RVE = robust variance estimation; CI = confidence interval; FAT = Funnel Plot Asymmetry; PET = Precision Effect Test; PEESE = Precision Effect Estimation With Standard Error.

Table 3. Moderator Analyses by Grade, Sessions, Hours, Group Size, and Treatment Content Controlling for Small-Study Effects (PET Model).

	Results using full sample of studies (*N* = 25)					Results excluding small sample studies (*n* = 21)				
	k	*ESw*	*SE*	95% CI	*p*	*k*	*ESw*	*SE*	95% CI	*p*
Grade										
Elementary (3–5)	15	−0.05	0.29	[−0.72, 0.63]	.641	15	0.47	0.39	[−0.42, 1.37]	.660
Middle and high school (6–12)	10	0.08	0.33	[−0.68, 0.83]		6	0.59	0.44	[−0.39, 1.58]	
Treatment sessions[a]										
≤30	18	−0.34	0.24	[−0.91, 0.24]	.046	14	0.06	0.40	[−0.83, 0.94]	.070
>30	9	0.30	0.32	[−0.44, 1.04]		8	0.66	0.32	[−0.10, 1.42]	
Treatment hours provided										
≤15	15	−0.22	0.21	[−0.72, 0.27]	.100	12	0.25	0.30	[−0.44, 0.94]	.013
>15	10	0.26	0.35	[−0.55, 1.06]		9	0.79	0.33	[0.01, 1.56]	
Group size										
<8	17	0.06	0.20	[−0.62, 0.74]	.207	14	0.50	0.40	[−0.40, 1.40]	.644
≥8	8	−0.27	0.29	[−0.93, 0.38]		7	0.38	0.42	[−0.54, 1.30]	
Mathematics content										
Fractions	6	0.41	0.34	[−0.44, 1.27]		5	0.74	0.36	[−0.11, 1.58]	
Problem solving	10	−0.36	0.33	[−1.12, 0.41]	.033	8	0.18	0.45	[−0.89, 1.24]	.127
Operations	8	−0.68	0.24	[−0.85, 0.32]	.065	7	−0.76	0.24	[−0.59, 0.56]	.038
Quality indicator										
<0.75	13	−0.01	0.33	[−0.75, 0.74]	.761	9	0.61	0.41	[−0.33, 1.53]	.340
≥0.75	12	−0.09	0.27	[−0.72, 0.54]		12	0.34	0.29	[−0.35, 1.02]	

Note. PET = Precision Effect Test; ESw = weighted effect size; CI = confidence interval.
[a]Treatment sessions varied within one study; counts for treatment sessions sums to 26.

($p < .001$). The Trim and Fill random-effects model, which included the missing ESs, estimated a new significant overall effect of 0.61 with a significantly large amount of heterogeneity remaining between studies ($I^2 = 93\%$, $p < .001$). As previously stated, we consider the Trim and Fill results a diagnostic tool as they do not perform well given the large amount of between-study heterogeneity, and the naïve random-effects analysis ignores dependancy in multiple ESs per study (Peters et al., 2007).

Moderator analyses. We assessed the potential role of moderators in explaining the large amount of heterogeneity remaining between studies ($I^2 > 80\%$ in all models). Within the FAT-PET model, we ran separate moderator models using the full sample and the subsample (see Table 3). The subsample FAT-PET results showed that group size, number of sessions, grade level, and study quality were not statistically significant predictors of ES, whereas fractions interventions ($p = .04$) and treatment provided for more than 15 hr ($p = .013$) were statistically significant predictors of ES.

Discussion

The purpose of this meta-analysis was to determine the effects of mathematics interventions on mathematics outcomes for students with MD in Grades 4 through 12.

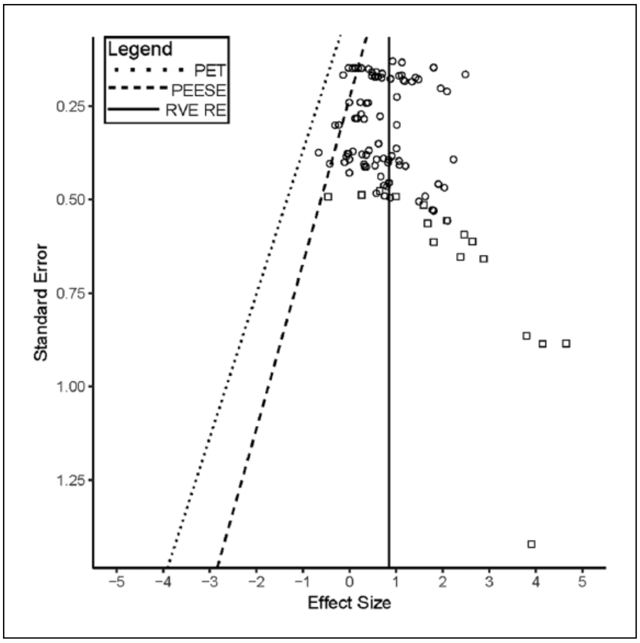

Figure 2. Funnel plot of effect sizes with RVE random-effects, PET, and PEESE model pooled estimates.
Note. The squares denote effect sizes that were removed during the sensitivity analysis to adjust for small-study effects that potentially skewed the pooled estimate. RVE = robust variance estimation; PET = Precision Effect Test; PEESE = Precision Effect Estimation With Standard Error.

Effects of Mathematics Intervention

We learned of extreme variability across study effect estimates, ranging from −0.66 to 4.65. The small sample sizes of the included studies contributed to the variability in ES estimates. Controlling for precision (i.e., the variance of the estimates), the PET estimated a pooled effect of 0.49 after excluding studies with small sample size. We interpreted this estimate as the underlying empirical estimate after accounting for missing studies due to small-study effects. Even though the PET estimate was not significant, the ES suggests practical significance of half a standard deviation. In addition to using this meta-regression method that accounts for small-study effects, we examined study quality using the Cook et al. (2015) quality indicators. Although study quality did not significantly predict ES, the PET model resulted in a lower effect (ES = 0.29) for more rigorous studies (i.e., study quality ≥.75) than for less rigorous studies (i.e., ES = 0.41; study quality <.75). Study quality ranged from 0.48 to 0.92 with almost half of the included studies (k = 12) receiving a rating of less than 0.75. As with any review, the findings should be considered in light of the quality of the studies included.

This PET estimate (0.49) is lower than the mean effects found in prior reviews for interventions targeting primarily elementary students. For example, Chodura et al. (2015) and Gersten et al. (2009) reported large mean effects for mathematics instruction targeting elementary students (M = 0.83 and M range = 1.04–1.56, respectively). This suggests that remediating MD for older students may be more challenging than remediation for younger students. One reason for this might be that older students' mathematics difficulties are more complex, further entrenched, and thus not as easily remediated. These students may have extensive knowledge gaps across multiple mathematics domains; given that mathematics knowledge builds throughout the grades, students may require interventions of greater intensity and duration to make substantial improvements in mathematics performance. Our moderator analysis supports this conclusion, as student performance increased significantly (p = .013) with interventions providing more than a total of 15 hr of intervention. Future studies might further investigate older students' response to varying amounts of time in intervention.

Intervention content. The results also revealed that fractions interventions significantly improved students' mathematics outcomes more than interventions in operations (p = .04). Only one study investigated the effects of generalized mathematics skills (Ketterlin-Geller, Chard, & Fien, 2008); no studies singularly targeted geometry, measurement and data, statistics and probability, or algebraic expressions and equations, all of which are necessary skills to succeed in middle and high school courses. Students may not respond favorably to interventions targeting a specific content domain (e.g., algebraic expressions and equations) with existing knowledge gaps in previous skills (e.g., number and operations in base 10). Therefore, there may be value in investigating interventions that build in complexity or target multiple content areas. It remains unknown if students respond more favorably to a multicomponent mathematics intervention, targeting multiple, yet related domains, or an intervention that builds in complexity to address gaps in students' prior understanding.

Given more rigorous national and state mathematics standards (e.g., Common Core State Standards; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010), students are required to analyze and to solve a variety of problems, exhibiting higher level mathematics reasoning. Prior reviews found support for explicit instruction (Chodura et al., 2015; Gersten et al., 2009; Kroesbergen & Van Luit, 2003), which provides step-by-step modeling and opportunities for practice with targeted feedback. The results of this review support the use of explicit instruction in problem-solving, fractions, and general mathematics skills (e.g., Fuchs et al., 2016; Ketterlin-Geller et al., 2008; Xin, Jitendra, Deatline-Buchman, 2005). In contrast to explicit instruction, the Gersten et al. (2009) review also found statistically significant results in favor of heuristics, a more generalized approach that involves multiple ways to solve a problem and often includes discussion to evaluate proposed solutions. Several studies in the current corpus compared the effects of explicit instruction and heuristics with a business-as-usual group, resulting in mixed findings (e.g., Kroesbergen & van Luit, 2002; Kroesbergen, Van Luit, & Maas, 2004). Fuchs et al. (2016), however, combined explicit instruction with high-quality explanations, resulting in improved content understanding, particularly for students with limited working memory. Explicit instruction by itself may not allow students the opportunity to benefit from problem-solving discussions with teachers and peers. Researchers might consider ways to embed discussion within step-by-step instruction to provide opportunities for students to explain their mathematical reasoning. High-quality explanations may support content understanding and expose students to various solutions, thus engendering flexibility in mathematical thinking.

Characteristics of Effective Interventions

In addition to determining the effects of mathematics interventions on mathematics outcomes for students with MD in Grades 4 through 12, we also aimed to identify intervention or design features that moderated intervention effectiveness.

Interventionist. Seven studies employed teachers as the interventionist (Bottge, 1999; Bottge, Rueda, LaRoque, Serlin, & Kwon, 2007; Butler, Miller, Crehan, Babbitt, & Pierce, 2003; Ketterlin-Geller et al., 2008; Scarlato & Burr, 2002; Van Luit, 1994; Walker & Poteet, 1990), three used computers (Burns, Kanive, & DeGrande, 2012; Christensen & Gerber, 1990; Fuchs, Fuchs, Hamlett, & Appleton, 2002), and the remaining 15 employed researchers. Chodura and colleagues (2015) found computer-delivered interventions (ES = 0.77) to be just as effective as interventions provided by humans (ES = 0.87) for elementary students with MD. Due to the limited number of studies investigating computer-provided interventions in this corpus, it is unclear whether computers are an effective way to deliver mathematics intervention for older students with MD. Computer-delivered intervention may provide additional practice opportunities and allow teacher-time to be dedicated to explicit instruction. Furthermore, computer-based practice might allow for smaller instructional groupings, thus intensifying teacher-provided intervention time. Achieving smaller groups is particularly difficult in the middle and high school grades as scheduling challenges may preclude the provision of increasingly intensive instruction via small group or one-on-one format. Finally, the feasibility of these interventions is unknown given most were researcher-provided. Future research might explore interventions provided by school-based personnel—to examine feasibility of implementation—and computers—to allow for additional practice opportunities and smaller instructional groupings.

Grade, sessions, and group size. Grade (elementary vs. middle and high school), number of treatment sessions (30 or less vs. more than 30), and group size (less than eight vs. eight or more) did not significantly predict intervention effectiveness. Across 25 years of research, we found three studies that provided 50 or more intervention sessions. One of those studies provided more than 100 sessions, but was excluded from the final PET model due to sample size (n = 6; Scarlato & Burr, 2002). Three studies employed a one-on-one format via computerized training; none of the studies employed educator—or paraprofessional—provided interventions in a one-on-one setting. In the literature, increasing the number of sessions and reducing group size are common methods for intensifying intervention provided across multitiered systems of support. In this corpus, we did not detect research investigating students' response to increasingly intensive levels of interventions within a tiered system of support. This finding highlights our lack of understanding related to increasingly intensive interventions, students' response at varying levels, and the most promising ways to intensify interventions for limited responders.

Limitations and Future Research

The limited number of studies available for this meta-analysis and the variability across ES estimates leave many questions unanswered about the effects of mathematics interventions for students with MD. We did not include unpublished studies in our review; however, we attempted to account for this by estimating a pooled effect robust to small-study effects (i.e., FAT-PET-PEESE regression method). Given the change in special education legislation at the turn of the 21st century (i.e., No Child Left Behind Act, 2002), it is understandable that 48% of selected studies were published between 2000 and 2009. Since 2010, the number of published studies declined, with only five studies investigating mathematics outcomes for students with MD in the last 5 years. The majority of the studies included in this meta-analysis targeted fourth- and fifth-grade students, demonstrating a lack of mathematics research conducted with middle and high school students. As Mann Koepke and Miller (2013) suggested, we simply need more research in this area to definitively summarize the effects of mathematics interventions, and the characteristics that moderate those effects, on mathematics outcomes for students with MD. Furthermore, only six studies included standardized mathematics measures; as such, the results reflect largely the effects of mathematics interventions as measured by researcher-developed measures closely aligned to the intervention. The implication is that researcher-developed measures may inflate the intervention effects (Cheung & Slavin, 2016). Overall, there is a need for high-quality mathematics intervention research that employs rigorous designs, larger sample sizes, and the use of rigorous, standardized mathematics assessments with better reliability and validity.

In an educational environment in which students are targeted for intervention via their response to a multitiered system of support, our knowledge base on intensive mathematics interventions is lacking. In particular, this line of research lags behind reading intervention research, specifically as it relates to intensive interventions. As a point of comparison, consider Wanzek and colleagues' (2013) review of extensive reading interventions for older students with reading difficulties. Across 17 years of research (i.e., 1995–2011), 19 studies investigated extensive reading interventions provided in 75 sessions or more. The duration of these studies lasted between 5 and 9 months, with many provided in a one-on-one or small group setting. Similarly, Solis and colleagues (2014) reported the findings of several longitudinal studies examining adolescent readers' response to increasingly intensified levels of intervention; students in Tier 2 and Tier 3 interventions received 50 min of daily, small group instruction across one school year (i.e., approximately 160 sessions). There is a need to investigate limited responders' response to increasingly intensive levels of mathematics intervention (i.e., reducing group size, increasing the duration, frequency, or length of each session, and providing more explicit, systematic instruction). Research funding should target investigating the implementation of longitudinal mathematics interventions for students with MD within tiered systems of support.

Although the final model estimated a moderate, mean effect (ES = 0.49) that was not statistically significant, these results suggest mathematics interventions show promise for improving students' mathematics outcomes. Further high-quality research is needed to fully understand the effects of mathematics interventions on the mathematics outcomes of fourth- through 12th-grade students with MD within tiered systems of support.

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. 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.

Supplemental Material

Supplemental Tables 1S and 2S are available at http://journals.sagepub.com/doi/suppl/10.1177/0741932517731887.

References

*Studies included in the meta-analysis.

Baker, S., Gersten, R., & Lee, D. S. (2002). A synthesis of empirical research on teaching mathematics to low-achieving students. *The Elementary School Journal*, *103*, 51–73. doi:10.1086/499715

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.

*Bottge, B. A. (1999). Effects of contextualized math instruction on problem solving of average and below-average achieving students. *The Journal of Special Education*, *33*, 81–92. doi:10.1177/002246699903300202

*Bottge, B. A., & Hasselbring, T. S. (1993). A comparison of two approaches for teaching complex, authentic mathematics problems to adolescents in remedial math classes. *Exceptional Children*, *59*, 556–566. doi:10.1177/001440299305900608

*Bottge, B. A., Rueda, E., LaRoque, P. T., Serlin, R. C., & Kwon, J. (2007). Integrating reform-oriented math instruction in special education settings. *Learning Disabilities Research & Practice*, *22*, 96–109. doi:10.1111/j.1540-5826.2007.00234.x

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, *42*, 39–55.

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

Burns, M. K., Codding, R. S., Boice, C. H., & Lukito, G. (2010). Meta-analysis of acquisition and fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. *School Psychology Review*, *39*, 69–83. doi:10.1111/j.1540-5826.2010.00323.x

*Burns, M. K., Kanive, R., & DeGrande, M. (2012). Effect of a computer-delivered math fact intervention as a supplemental intervention for math in third and fourth grades. *Remedial and Special Education*, *33*, 184–191. doi:10.1177/0741932510381652

*Butler, F. M., Miller, S. P., Crehan, K., Babbitt, B., & Pierce, T. (2003). Fraction instruction for students with mathematics disabilities: Comparing two teaching sequences. *Learning Disabilities Research & Practice*, *18*, 99–111. doi:10.1111/1540-5826.00066

Butterworth, B. (2010). Foundational numerical capacities and the origin of dyscalculia. *Trends in Cognitive Sciences*, *14*, 534–541. doi:10.1016/j.tics.2010.09.007

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:10.1177/07419325070280050401

Cheung, A., & Slavin, R. (2016). How methodological features affect effect sizes in education. *Educational Researcher*, *45*, 283–292. doi:10.3102/0013189x16656615

Chodura, S., Kuhn, J. T., & Holling, H. (2015). Interventions for children with mathematical difficulties: A meta-analysis.

Zeitschirft Für Psychologie, *223*, 129–144. doi:10.1027/2151-2604/a000211

*Christensen, C. A., & Gerber, M. M. (1990). Effectiveness of computerized drill and practice games in teaching basic math facts. *Exceptionality: A Special Education Journal*, *1*, 149–165. doi:10.1080/09362839009524751

Codding, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice*, *26*, 36–47. doi:10.1111/j.1540-5826.2010.00323.x

Cook, B. G., Buysse, V., Klingner, J., Landrum, T. J., McWilliam, R. A., Tankersley, M., & Test, D. W. (2015). CEC's standards for classifying the evidence base of practices in special education. *Remedial and Special Education*, *36*, 220–234. doi:10.1177/0741932514557271

Cooper, H., Hedges, L. V., & Valentine, J. C. (Eds.). (2009). *The handbook of research synthesis and meta-analysis*. New York, NY: Russell Sage Foundation.

Deming, D. J., Cohodes, S., Jennings, J., & Jencks, C. (2013, September). *School accountability, postsecondary attainment and earnings* (Working Paper No. 19444). Cambridge, MA: National Bureau of Economic Research.

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

Dougherty, C. (2003). Numeracy, literacy and earnings: Evidence from the National Longitudinal Survey of Youth. *Economics of Education Review*, *22*, 511–521. doi:10.1016/S0272-7757(03)00040-2

Duval, S., & Tweedie, R. (2000). A nonparametric "Trim and Fill" method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*, 89–98. doi:10.1080/01621459.2000.10473905

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629–634. doi:10.1136/bmj.315.7109.629

*Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Appleton, A. C. (2002). Explicitly teaching for transfer: Effects on the mathematical problem-solving performance of students with mathematics disabilities. *Learning Disabilities Research & Practice*, *17*, 90–106. doi:10.1111/1540-5826.00036

*Fuchs, L. S., Malone, A. S., Schumacher, R. F., Namkung, J., Hamlett, C. L., Jordan, N. C., . . . Changas, P. (2016). Supported self-explaining during fraction intervention. *Journal of Educational Psychology*, *108*, 493–508. doi:10.1037/edu0000073

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., Long, J., Namkung, J., Hamlett, C. L., Cirino, P. T., . . . Changas, P. (2013). Improving at-risk learners' understanding of fractions. *Journal of Educational Psychology*, *105*, 683–700. doi:10.1037/a0032446

*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

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

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

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65. doi:10.1002/jrsm.5

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analysis. *British Medical Journal*, *327*, 557–560. doi:10.1136/bmj.327.7414.557

Hughes, E. M., Witzel, B. S., Riccomini, P. J., Fries, K. M., & Kanyongo, G. Y. (2014). A meta-analysis of algebra interventions for learners with disabilities and struggling learners. *The Journal of the International Association of Special Education*, *15*, 36–47.

Institute of Education Sciences. (2014). *What works clearinghouse procedures and standards handbook (Version 3.0)*. Retrieved from https://ies.ed.gov/ncee/wwc/Docs/referenceresources/wwc_procedures_v3_0_standards_handbook.pdf

*Jitendra, A. K., Griffin, C. C., McGoey, K., Gardill, M. C., Bhat, P., & Riley, T. (1998). Effects of mathematical word problem solving by students at risk or with mild disabilities. *The Journal of Educational Research*, *91*, 345–355. doi:10.1080/00220679809597564

Jitendra, A. K., Rodriguez, M., Kanive, R., Huang, J.-P., Church, C., Corroy, K. A., & Zaslofsky, A. (2013). 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., 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

Joseph, L. M., Konrad, M., Cates, G., Vajcner, T., Eveleigh, E., & Fishley, K. M. (2012). A meta-analysis review of the cover-copy-compare and variations of this self-management procedure. *Psychology in the Schools*, *49*, 122–136. doi:10.1002/pits.20622

*Kelly, B., Gersten, R., & Carnine, D. (1990). Student error patterns as a function of curriculum design teaching fractions to remedial high school students and high school students with learning disabilities. *Journal of Learning Disabilities*, *23*, 23–29. doi:10.1177/002221949002300108

*Ketterlin-Geller, L. R., Chard, D. J., & Fien, H. (2008). Making connections in mathematics conceptual mathematics intervention for low-performing students. *Remedial and Special Education*, *29*, 33–45. doi:10.1177/0741932507309711

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

*Kroesbergen, E. H., & van Luit, J. E. (2002). Teaching multiplication to low math performers: Guided versus structured instruction. *Instructional Science*, *30*, 361–378. doi:10.1023/a:1019880913714

Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and Special Education*, *24*, 97–114. doi:10.1177/07419325030240020501

*Kroesbergen, E. H., & Van Luit, J. E. (2005). Constructivist mathematics education for students with mild mental retardation. *European Journal of Special Needs Education*, *20*, 107–116. doi:10.1080/0885625042000319115

*Kroesbergen, E. H., Van Luit, J. E., & Maas, C. J. (2004). Effectiveness of explicit and constructivist mathematics instruction for low-achieving students in the Netherlands. *The Elementary School Journal*, *104*, 233–251. doi:10.1086/499751

Lee, J. (2012). College for all gaps between desirable and actual P–12 math achievement trajectories for college readiness. *Educational Researcher*, *41*, 43–55. doi:10.3102/0013189X11432746

Maccini, P., Mulcahy, C. A., & Wilson, M. G. (2007). A follow-up of mathematics interventions for secondary students with learning disabilities. *Learning Disabilities Research & Practice*, *22*, 58–74. doi:10.1111/j.1540-5826.2007.00231.x

*Manalo, E., Bunnell, J. K., & Stillman, J. A. (2000). The use of process mnemonics in teaching students with mathematics learning disabilities. *Learning Disability Quarterly*, *23*, 137–156. doi:10.2307/1511142

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 & Practice*, *27*, 152–166. doi:10.1111/j.1540-5826.2012.00363.x

Mann Koepke, K., & Miller, B. (2013). At the intersection of math and reading disabilities: Introduction to the special issue. *Journal of Learning Disabilities*, *46*, 483-489. doi:10.1177/0022219413498200

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

*Milo, B. (2005). Math instruction for students with special educational needs: Effects of guiding versus directing instruction. *Educational & Child Psychology*, *22*, 68–80.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(6), e1000097. doi:10.1371/journal.pmed1000097

Moreno, S. G., Sutton, A. J., Ades, A., Stanley, T. D., Abrams, K. R., Peters, J. L., & Cooper, N. J. (2009). Assessment of regression-based methods to adjust for publication bias through a comprehensive simulation study. *BMC Medical Research Methodology*, *9*(2), 1–17. doi:10.1186/1471-2288-9-2

Myers, J. A., Wang, J., Brownell, M. T., & Gagnon, J. C. (2015). Mathematics interventions for students with learning disabilities (LD) in secondary school: A review of the literature. *Learning Disabilities: A Contemporary Journal*, *13*, 207–235.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school 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.

No Child Left Behind Act of 2001, Pub. L. No. 107-110, § 115, Stat. 1425. (2002).

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 Disabilities Research & Practice*, *28*, 98–112. doi:10.1111/ldrp.12013

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, *26*, 4544–4562. doi:10.1002/sim.2889

Pigott, T. D. (2012). *Advances in meta-analysis*. New York, NY: Springer.

R Core Team. (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Available from http://www.R-project.org/

Rohrbeck, C. A., Ginsburg-Block, M. D., Fantuzzo, J. W., & Miller, T. R. (2003). Peer-assisted learning interventions with elementary school students: A meta-analytic review. *Journal of Educational Psychology*, *95*, 240–257. doi:10.1037/0022-0663.95.2.240

Rücker, G., Schwarzer, G., Carpenter, J. R., Binder, H., & Schumacher, M. (2010). Treatment-effect estimates adjusted for small-study effects via a limit meta-analysis. *Biostatistics*, *112*, 122–142. doi:10.1093/biostatistics/kxq046

*Scarlato, M. C., & Burr, W. A. (2002). Teaching fractions to middle school students. *Journal of Direct Instruction*, *2*(1), 23–38.

Shin, M., & Bryant, D. P. (2015). Fraction interventions for students struggling to learn mathematics: A research synthesis. *Remedial and Special Education*, *36*, 374–387. doi:10.1177/0741932515572910

Solis, M., Miciak, J., Vaughn, S., & Fletcher, J. M. (2014). Why intensive interventions matter: Longitudinal studies of adolescents with reading disabilities and poor reading comprehension. *Learning Disability Quarterly*, *37*, 218–229. doi:10.1177/0731948714528806

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effect in the presence of publication selection. *Oxford Bulletin of Economics and Statistics*, *70*, 103–127. doi:10.1111/j.1468-0084.2007.00487.x

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*, 60–78. doi:10.1002/jrsm.1095

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., & Jerman, O. (2006). Math disabilities: A selective meta-analysis of the literature. *Review of Educational Research*, *76*, 249–274. doi:10.3102/00346543076002249

Swanson, H. L., Lussier, C. M., & Orosco, M. J. (2015). Cognitive strategies, working memory, and growth in word problem solving in children with math difficulties. *Journal of Learning Disabilities*, *48*, 339–358. doi:10.1177/0022219413498771

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

Szucs, D., Devine, A., Soltesz, F., Nobes, A., & Gabriel, F. (2013). Developmental dyscalculia is related to visuo-spatial memory and inhibition impairment. *Cortex*, *49*, 2674–2688. doi:10.1016/j.cortex.2013.06.007

Terrin, N., Schmid, C. H., Lau, J., & Olkin, I. (2003). Adjusting for publication bias in the presence of heterogeneity. *Statistics in Medicine*, *22*, 2113–2126. doi:10.1002/sim.1461

Tipton, E. (2013). Robust variance estimation in meta-regression with binary dependent effects. *Research Synthesis Methods*, *4*, 169–187. doi:10.1002/jrsm.1070

Tolar, T. D., Fuchs, L. S., 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 Disability*, *49*, 240–256. doi:10.1177/0022219414538520

U.S. Department of Education. (2015). *The Nation's Report Card*. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics.

*Van Luit, J. E. (1994). The effectiveness of structural and realistic arithmetic curricula in children with special needs. *European Journal of Special Needs Education*, *9*(1), 16–26. doi:10.1080/0885625940090102

*Van Luit, J. E., & Naglieri, J. A. (1999). Effectiveness of the MASTER program for teaching special children multiplication and division. *Journal of Learning Disabilities*, *32*, 98–107. doi:10.1177/002221949903200201

*Walker, D. W., & Poteet, J. A. (1990). A comparison of two methods of teaching mathematics story problem-solving with learning disabled students. *National Forum of Special Education Journal*, *1*, 44–51.

Wanzek, J., Vaughn, S., Scammacca, N. K., Metz, K., Murray, C. S., Roberts, G., & Danielson, L. (2013). Extensive reading interventions for students with reading difficulties after grade 3. *Review of Educational Research*, *83*, 163–195. doi:10.3102/0034654313477212

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

Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. *The Journal of Special Education*, *32*, 207–225. doi:10.1177/002246699903200402

*Xin, Y. P., Jitendra, A. K., & Deatline-Buchman, A. (2005). Effects of mathematical word problem-solving instruction on middle school students with learning problems. *The Journal of Special Education*, *39*, 181–192. doi:10.1177/00224669050390030501

*Xin, Y. P., Zhang, D., Park, J. Y., Tom, K., Whipple, A., & Si, L. (2011). A comparison of two mathematics problem-solving strategies: Facilitate algebra-readiness. *The Journal of Educational Research*, *104*, 381–395. doi:10.1080/00220671.2010.487080

Zhang, D., & Xin, Y. P. (2012). A follow-up meta-analysis for word-problem-solving interventions for students with mathematics difficulties. *The Journal of Educational Research*, *105*, 303–318. doi:10.1080/00220671.2011.627397

Zheng, X., Flynn, L. J., & Swanson, H. L. (2013). Experimental intervention studies on word problem solving and math disabilities: A selective analysis of the literature. *Learning Disability Quarterly*, *36*, 97–111. doi:10.1177/0731948712444277