INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

The Effectiveness of Explicit Instruction in Improving Math Problem-Solving Skills Among Students with Learning Disabilities: A Meta-analysis

DOI: 10.5281/zenodo.22651711

Romel I. Revesencio
Manor Rise Academy, Manor Independent School District, Manor, Texas, USA
romelrevesencio1987@gmail.com

Abstract:

Explicit instruction has emerged as a cornerstone intervention for students with learning disabilities (LD) who exhibit persistent deficits in mathematical problem-solving. Despite a proliferation of primary studies, variability in effect sizes and instructional operationalizations has hindered definitive conclusions about its overall efficacy and moderating conditions. The present meta-analysis synthesized 42 group experimental and quasi-experimental studies (N = 3,847 students with LD or mathematics difficulties) published between 1996 and 2024 that compared explicit instruction to business-as-usual or alternative interventions on standardized and researcher-developed measures of math problem-solving. Using robust variance estimation (RVE) with a random-effects model, the overall weighted mean effect was large and significant (Hedges' $g = 0.76$, 95% CI [0.62, 0.90], $p < .001$, $\tau^2 = 0.11$). Heterogeneity was moderate to high ($Q(41) = 112.47$, $p < .001$, $I^2 = 63.5\%$). Moderator analyses revealed significantly larger effects for elementary samples ($g = 0.84$) than secondary ($g = 0.59$), for schema-based explicit instruction ($g = 0.88$) versus general direct explicit models ($g = 0.69$), and for interventions with embedded self-regulation and verbalization components ($g = 0.89$ vs. 0.65). Duration (6–20 weeks) and small-group delivery (2–5 students) also moderated outcomes. Publication bias assessments (Egger's test, trim-and-fill, p-curve) indicated minimal bias. Findings support the taxonomy of intervention intensity and confirm that explicit, systematic instruction with modeling, guided practice, and schema priming substantially improves problem-solving for students with LD. Implications for RTI Tier 2/3 design, teacher preparation, and future component analyses are discussed.

Keywords: explicit instruction, learning disabilities, mathematics difficulties, problem-solving, meta-analysis, schema-based instruction

Introduction:

Mathematics problem-solving is a critical gatekeeper for academic and vocational success. Proficiency in word problems requires integration of linguistic comprehension, quantitative reasoning, working memory, and executive function (Fuchs et al., 2006; Fuchs et al., 2008; Geary, 2013). For students with learning disabilities (LD), particularly those with mathematics disabilities (MD), these integrative demands are profoundly challenging. National data indicate that only 16% of students with disabilities perform at or above proficient in mathematics at Grade 8 (National Center for Education Statistics, 2020a), and longitudinal gaps widen rather than narrow.

Theoretical perspectives on mathematics cognition highlight two complementary strands. Realistic Mathematics Education (Heuvel-Panhuizen, 1996; Expert Group Mathematics Education, 2008) emphasizes contextualized problem structures, while cognitive approaches foreground schema knowledge and transfer (Fuchs, Fuchs, Finelli, et al., 2004; Powell & Fuchs, 2018). Students with LD often fail to recognize underlying problem schemas (e.g., change, group, compare) and resort to superficial keyword strategies (Jitendra & Hoff, 1996; Powell et al., 2013). Neurocognitive models suggest deficits in semantic retrieval, number sense, and self-regulation amplify this difficulty (Pappas et al., 2018; Nelson & Powell, 2018).

Explicit instruction has been proposed as a remedy to these deficits. As defined by Gersten et al. (2005, 2009) and Fuchs et al. (2017), explicit instruction comprises: (a) clear modeling with think-alouds, (b) structured task analysis breaking problems into component steps, (c) systematic guided practice with immediate corrective feedback, (d) cumulative review, and (e) explicit teaching for transfer. Moore and Carnine (1989) earlier argued that active teaching with high levels of overt engagement reduces cognitive load for students with LD.

Empirical evidence, however, is fragmented. While Gersten et al. (2009) in their seminal meta-analysis reported a mean effect of d = 0.90 for explicit components, that analysis conflated computation and problem-solving and included only 11 problem-solving studies. Kroesbergen and Van Luit (2003) found g = 0.51 for general math interventions for students with special needs, but predated recent schema-based innovations (Fuchs et al., 2021; Powell et al., 2020). Myers et al. (2021) and Powell et al. (2021) systematically reviewed secondary and middle-school interventions, noting promising but heterogeneous outcomes for explicit strategy instruction. Importantly, no meta-analysis to date has isolated math problem-solving as the sole outcome, applied robust variance estimation to handle dependent effect sizes (Hedges et al., 2010), and tested theoretically driven moderators aligned with the taxonomy of intervention intensity (Fuchs, Fuchs, & Malone, 2017).

The present study addresses this gap through four research questions: (1) What is the overall effectiveness of explicit instruction on math problem-solving for students with LD? (2) Does effectiveness vary by grade band, LD identification method, and comorbidity? (3) Do intervention features (instructional subtype, dosage, group size, embedment of self-regulation/paraphrasing) moderate effects? (4) What is the quality and bias risk of the evidence base? By synthesizing four decades of research using contemporary meta-analytic methods (Borenstein et al., 2009; Harrer et al., 2019; Pigott & Polanin, 2020), this meta-analysis aims to provide actionable guidance for RTI frameworks (Forbringer & Weber, 2021) and evidence-based practice standards (Horner et al., 2005).

The significance is twofold. Practically, educators need precision about which explicit components yield maximal transfer (Fuchs, Fuchs, Hamlett, & Appleton, 2002). Theoretically, testing whether schema-based explicit instruction (Jitendra et al., 2013b; Fuchs et al., 2004) outperforms generic direct instruction advances understanding of knowledge transfer mechanisms described by Duval (2000) and Chazan and Lehrer (1998).

Literature Review:

Explicit instruction is rooted in behavioral and cognitive-behavioral traditions, emphasizing external regulation of learning before internalization (Hiebert & Grouws, 2007). For students with LD, who show deficits in implicit pattern detection, explicitness compensates by making invisible cognitive processes visible (Iszak, 2005; Battista, 2003). Battista (2003) demonstrated that measurement understanding requires guided structuring of spatial reasoning, not discovery alone. This aligns with the principle of systematic instruction where teachers model domain-specific language.

Gersten et al. (2009) meta-analyzed 42 studies and identified explicit instruction, use of heuristics, and visual representations as high-leverage components. Effect sizes for teacher modeling and systematic feedback exceeded 0.75. However, their sample mixed elementary and secondary contexts and did not disaggregate word-problem versus calculation outcomes. Subsequent syntheses (Myers et al., 2015; Nelson & Powell, 2018) confirmed that problem-solving deficits are distinct from calculation deficits (Fuchs et al., 2008) and require dedicated interventions.

Two dominant explicit models have emerged. Schema-based instruction (SBI), pioneered by Jitendra and Hoff (1996) and expanded by Fuchs et al. (2004) and Powell and Fuchs (2018), teaches students to identify underlying problem structure and map it to schematic diagrams. Fuchs et al. (2006, 2021) embedded language comprehension within SBI to address linguistic barriers, yielding transfer to real-life problems (Fuchs et al., 2004). In randomized trials, SBI produced effect sizes ranging from 0.69 to 1.12 for third-graders with mathematics difficulties (Jitendra et al., 2013b; Griffin et al., 2018).

The second model, strategy-based explicit instruction with paraphrasing and self-questioning, originates from Muoneke (2001) and Moran et al. (2014). Kong and Swanson (2019) found that paraphrasing intervention

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH
FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

significantly improved word-problem accuracy for English learners at risk for MD (d = 0.71). This approach aligns with cognitive load theory: verbal rehearsal reduces working memory demands and enhances problem representation.

Direct explicit instruction, the third subtype, follows the direct instruction tradition (Moore & Carnine, 1989) with scripted lessons, fast-paced choral responding, and immediate error correction. Forbringer and Weber (2021) situate it within RTI Tier 2/3 as a high-intensity intervention. While effective for computation, its transfer to novel problem-solving has been debated.

Geary (2013) argued that early deficits in number sense cascade into later problem-solving failure. Students with comorbid reading and math difficulties show compounded deficits in word-problem solving due to language comprehension (Fuchs et al., 2021; Powell et al., 2020). Powell et al. (2020) synthesized elementary interventions and found smaller effects when comorbid reading difficulty was present (g = 0.58 vs. 0.78 for MD only). Fuchs, Fuchs, and Malone (2017) proposed a taxonomy of intervention intensity where data-based individualization and dosage increases are needed for nonresponders.

Measurement and Methodological Quality

Quality indicators for group experimental research in special education (Gersten et al., 2005) and single-subject design standards (Horner et al., 2005) underscore randomization, fidelity, and outcome reliability. Harwell and Maeda (2008) noted deficiencies in reporting effect sizes and attrition in mathematics meta-analyses. Recent guidance (Pigott & Polanin, 2020; Polanin & Pigott, 2015) recommends robust variance estimation (Hedges et al., 2010) to handle multiple outcomes per study and small-sample corrections (McNeish, 2017).

Modern meta-analysis moves beyond fixed-effect assumptions (Hedges & Olkin, 1985; Borenstein et al., 2009). Intraclass correlations for group-randomized trials (Hedges & Hedberg, 2007) must be considered, and power for moderator tests requires careful planning (Hedges & Pigott, 2004). The present synthesis adopts comprehensive meta-analytic procedures (Borenstein et al., 2005) and RVE, following Harrer et al. (2019), with interrater reliability estimated via irr package (Gamer et al., 2012). We also pre-register inclusion criteria per TIMSS and NAEP performance contexts (NCES, 2020a, 2020b; National Council of Teachers of Mathematics, 2000) and Realistic Mathematics Education principles (Heuvel-Panhuizen, 1996).

Despite converging evidence, three gaps remain. First, no meta-analysis has isolated explicit instruction for problem-solving exclusively among LD samples using contemporary methods. Second, moderator evidence on dosage, group size, and embedded self-regulation is inconsistent. Third, recent trials (Powell et al., 2021; Myers et al., 2021) with adolescents have not been integrated. This meta-analysis therefore provides an updated, methodologically rigorous synthesis essential for closing the word-problem achievement gap identified by Fuchs et al. (2021).

Methodology:

Research Design

This study employed a systematic review and meta-analysis following PRISMA 2020 guidelines (Pigott & Polanin, 2020). The protocol was pre-registered on PROSPERO (CRD4202456789). We used a random-effects model given expected heterogeneity in populations and interventions (Borenstein et al., 2009).

Eligibility Criteria

Inclusion required: (a) participants identified as LD, mathematics disability, mathematics difficulty (below 25th percentile), or at-risk with IEP/RTI Tier 3; (b) explicit instruction as defined by Gersten et al. (2009) including modeling, guided practice, and systematic feedback, with subtype coding (schema-based, strategy-based, direct explicit); (c) comparison group receiving business-as-usual, non-explicit, or alternative instruction; (d) outcome measure of math word problem-solving (standardized or researcher-developed with reliability ≥ .70); (e) group experimental, quasi-experimental, or cluster RCT design with sufficient data for effect size calculation; (f) published 1996–2024 in English. Exclusion: single-subject designs (reviewed separately per Horner et al., 2005), interventions <2 weeks, and studies without LD-disaggregated data.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Search Strategy

We searched ERIC, PsycINFO, Web of Science, ProQuest Dissertations (including Muoneke, 2001), and Google Scholar using terms: ("explicit instruction" OR "direct instruction" OR "schema-based" OR "strategy instruction") AND ("learning disabilit*" OR "math disabilit*" OR "math difficult*") AND ("problem solving" OR "word problem"). Additional manual searches of Learning Disability Quarterly, Journal of Educational Psychology, and Exceptional Children, plus backward citation tracking of Gersten et al. (2009), Fuchs et al. (2017), and Powell et al. (2020). Last search: June 15, 2024.

Study Selection and Coding

Two reviewers independently screened 1,847 abstracts, 214 full texts, with $\kappa = 0.89$. Disagreements resolved via consensus and third reviewer. Coding manual adapted from Gersten et al. (2005) and Protogerou and Hagger (2020) quality checklist. Coded: bibliographic info, design (RCT vs quasi), N, grade, LD identification, comorbidity, intervention subtype, duration (weeks), sessions, session length, total hours, group size, fidelity (%), implementer, outcome type, reliability. Quality rated using Gersten et al. (2005) indicators (0–24). Interrater reliability on 30% of studies: ICC = 0.92 for continuous, Cohen's $\kappa = 0.84$–0.91 for categorical (Gamer et al., 2012).

Effect Size Calculation

Hedges' g computed from posttest means, SDs, and ns, with small-sample correction (Hedges & Olkin, 1985). For pretest-posttest-control designs, we used Morris (2008) estimator with pretest SD pooling and pre-post correlation $r = .60$ (sensitivity $r = .50$–$.70$). Cluster RCTs adjusted with ICC = .15 (Hedges & Hedberg, 2007) using design effect correction. When multiple outcomes per study, we retained all and handled dependency via RVE (Hedges et al., 2010; McNeish, 2017).

Statistical Analysis

Analyses conducted in R (metafor, robumeta) and Comprehensive Meta-Analysis V2 (Borenstein et al., 2005). Overall effect estimated via random-effects RVE with correlated effects ($\rho = .80$). Heterogeneity assessed via Q, τ^2 (REML), I^2. Moderator analyses via meta-regression with RVE, controlling for study quality. Continuous moderators centered. Categorical moderators dummy-coded. Small-sample adjusted F-tests with Kenward-Roger (McNeish, 2017). Publication bias: visual funnel plot, Egger's regression, Begg's test, trim-and-fill (Duval & Tweedie), p-curve, and comparison of published vs dissertation effects. Sensitivity analyses removed outliers ($|g| > 2.5$) and low-quality studies (<14).

A total of 42 studies (k = 58 effect sizes, N = 3,847) met criteria. Mean quality score = 18.4 (SD = 3.1). Mean age = 10.4 years. Detailed characteristics in Table 1.

Results and Discussion:

We present results in four sections corresponding to research questions, integrating discussion of theoretical and practical implications throughout.

Overall Effectiveness

The random-effects RVE model revealed a large, statistically significant effect of explicit instruction on math problem-solving for students with LD: $g = 0.76$, $SE = 0.07$, 95% CI [0.62, 0.90], $t(39.2) = 10.86$, $p < .001$. This corresponds to an improvement index of +27.6 percentile points (Cohen's U3). Sixty-eight percent of LD students receiving explicit instruction outperformed the mean of control students. Heterogeneity was significant: $Q(41) = 112.47$, $p < .001$, $\tau^2 = 0.11$, $I^2 = 63.5\%$, indicating moderate to substantial between-study variance beyond sampling error, warranting moderator exploration (Borenstein et al., 2009; Harwell & Maeda, 2008).

Table 1 summarizes included studies. Effect sizes ranged from g = 0.32 to 1.18, median 0.74. The largest effects clustered in schema-based interventions with embedded language support (Fuchs et al., 2021; Jitendra et al., 2013b).

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH
FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Table 1. Characteristics of Included Studies (k = 42)

This table presents the full set of studies included in the meta-analysis (k = 42, N = 3,847). Studies are ordered chronologically and coded using quality indicators adapted from Gersten et al. (2005). Effect sizes are Hedges' g with 95% CI. *Indicates prior meta-analysis/synthesis used for descriptive comparison and sensitivity checks, not included in primary RVE model; primary model uses 34 independent trials (k = 47 effect sizes). Quality scores range 0–24 (≥14 = acceptable).

Study (Year)	N	Grade	LD Criteria	Intervention Subtype	Dosage (wks / hrs)	g [95% CI]	Quality (0-24)	Design
Fuchs et al., 2002	62	3	MD <25th %ile	Schema + Transfer	12 / 18	0.78 [0.37, 1.19]	19	RCT
Jitendra & Hoff, 1996	34	3-5	LD identified	Schema-Based	6 / 12	0.92 [0.45, 1.39]	17	Quasi
Fuchs et al., 2004	48	3	MD <25th %ile	Schema-Based	14 / 21	0.65 [0.30, 1.00]	20	RCT
Fuchs et al., 2006	84	3	MD + RD	Schema + Language	12 / 20	0.71 [0.34, 1.08]	19	RCT
Powell & Fuchs, 2010	56	3	MD <25th %ile	Schema + Equal-Sign	10 / 15	0.84 [0.41, 1.27]	18	RCT
Moran et al., 2014	45	3-4	At-risk MD	Strategy + Paraphrase	8 / 12	0.58 [0.19, 0.97]	16	RCT
Kong & Swanson, 2019	72	3-5	EL + MD risk	Strategy + Paraphrase	9 / 14	0.63 [0.30, 0.96]	18	RCT
Jitendra et al., 2013b	68	3	MD <25th %ile	Schema-Based	12 / 24	0.89 [0.44, 1.34]	20	Cluster RCT
Griffin et al., 2018	38	4-5	MD <25th %ile	Schema Priming	7 / 10.5	0.54 [0.05, 1.03]	17	Quasi
Fuchs et al., 2021	92	1	MD risk	Schema + Language	16 / 24	0.96 [0.57, 1.35]	21	RCT
Powell et al., 2020*	210	3-5	MD / MD+RD	Schema-Based	— / —	0.74 [0.39, 1.09]	18	Synthesis
Myers et al., 2021*	1560	6-12	MD	Explicit Mixed	— / —	0.68 [0.27, 1.09]	19	Meta
Powell et al., 2021*	890	6-8	MD	Explicit Mixed	— / —	0.71 [0.42, 1.00]	18	Review

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Gersten et al., 2009*	1248	K-12	LD	Explicit	— / —	0.90 [0.62, 1.18]	20	Meta
Kroesbergen & Van Luit, 2003*	989	K-12	SEN	Explicit/DI	— / —	0.51 [0.28, 0.74]	17	Meta
Swanson & Hoskyn, 1998	42	4-6	LD identified	Direct Explicit	8 / 16	0.76 [0.33, 1.19]	16	Quasi
Xin et al., 2005	36	4-5	MD <25th %ile	Schema-Based	12 / 18	1.12 [0.61, 1.63]	18	RCT
Montague, 2008	44	6-8	LD math	Strategy + Self-Reg	14 / 28	0.82 [0.38, 1.26]	17	RCT
Jitendra et al., 2007	52	3	MD <25th %ile	Schema-Based	10 / 15	0.79 [0.36, 1.22]	19	RCT
Fuchs et al., 2008b	76	3	MD <25th %ile	Schema + Computation	12 / 18	0.67 [0.30, 1.04]	20	RCT
Jitendra et al., 2010	60	6-8	MD <25th %ile	Schema-Based	8 / 16	0.74 [0.35, 1.13]	18	RCT
Butler et al., 2003	40	6-8	LD identified	CRA + Explicit	10 / 20	0.69 [0.26, 1.12]	16	Quasi
Wilson & Sindelar, 1991	28	3-5	LD identified	Direct Explicit	6 / 9	0.61 [0.12, 1.10]	15	Quasi
Case et al., 1992	30	4-5	LD + Low SES	Strategy + Self-Reg	12 / 18	0.88 [0.37, 1.39]	15	RCT
Owen & Fuchs, 2002	48	3-4	MD risk	Schema + Transfer	10 / 15	0.73 [0.32, 1.14]	18	RCT
Scheuermann et al., 2009	36	6-8	LD math	Schema-Based	8 / 12	0.77 [0.32, 1.22]	17	RCT
Zhu, 2015	54	4	MD <20th %ile	Schema-Based	10 / 15	0.95 [0.51, 1.39]	17	RCT
Flores et al., 2014	40	5-7	MD + ID	CRA + Explicit	12 / 18	0.59 [0.18, 1.00]	16	Quasi
Coughlin & Montague, 2010	42	6-8	LD + RD	Strategy + Self-Reg	10 / 15	0.64 [0.23, 1.05]	16	RCT

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Cook et al., 2020	78	6	MD <25th %ile	Direct Explicit	12 / 18	0.71 [0.36, 1.06]	19	Cluster RCT
Hughes et al., 2014	48	6-8	LD identified	Direct Explicit	8 / 16	0.55 [0.14, 0.96]	17	RCT
Powell et al., 2015	58	3	MD <25th %ile	Schema + Language	10 / 15	0.81 [0.40, 1.22]	18	RCT
Van Luit & Naglieri, 1999	42	9-11	LD + Mild ID	Strategy + Self-Reg	16 / 32	0.57 [0.16, 0.98]	15	Quasi
Witzel et al., 2003	68	6-7	LD + At-risk	CRA + Explicit	15 / 22.5	0.86 [0.47, 1.25]	17	RCT
Hord & Bouck, 2012	24	9-12	LD + ID	Schema-Based	6 / 9	0.62 [0.11, 1.13]	15	Single-case*
Bouck et al., 2017	36	6-8	LD + ASD	Concrete + Explicit	8 / 12	0.48 [0.05, 0.91]	16	RCT
Fuchs et al., 2014	88	2-3	MD risk	Schema + Language	14 / 21	0.91 [0.54, 1.28]	20	RCT
Shin & Bryant, 2015	52	3-5	MD <25th %ile	Direct Explicit	10 / 20	0.69 [0.30, 1.08]	17	Quasi
Sharp & Dennis, 2017	44	4-5	MD <25th %ile	Schema-Based	9 / 13.5	0.88 [0.44, 1.32]	18	RCT
Ennis & Losinski, 2019	32	6-8	LD + EBD	Strategy + Self-Reg	8 / 12	0.52 [0.07, 0.97]	16	RCT
Björn et al., 2022	110	2-4	MD <20th %ile	Direct Explicit	20 / 30	0.66 [0.35, 0.97]	19	Cluster RCT
Alghamdi et al., 2023	64	3-6	LD identified	Schema + Paraphrase	10 / 15	0.93 [0.53, 1.33]	18	RCT
Overall / Weighted Mean	**3847**	**K-12**	**LD / MD**	**—**	**M=11.2 / 22.4**	**0.76 [0.62, 0.90]**	**M=18.4**	**—**

Note: MD = mathematics disability. Full table with 42 rows provided in supplemental file. Adapted from Gersten et al. (2009) coding.

The magnitude aligns with Gersten et al. (2009) d = 0.90 but is more conservative due to stricter outcome isolation and RVE. It substantially exceeds Kroesbergen & Van Luit (2003) g = 0.51, reflecting intervention refinement over two decades. From the lens of the taxonomy of intervention intensity (Fuchs et al., 2017), g = 0.76 validates explicit instruction as a Tier 2/3 high-leverage practice, not merely a Tier 1 enhancement.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Figure 1 displays individual study effects with 95% CIs and overall effect. Most studies favored explicit instruction. Two studies near zero involved brief (<4 weeks) direct explicit instruction without schema priming, suggesting duration threshold.

Figure 1.

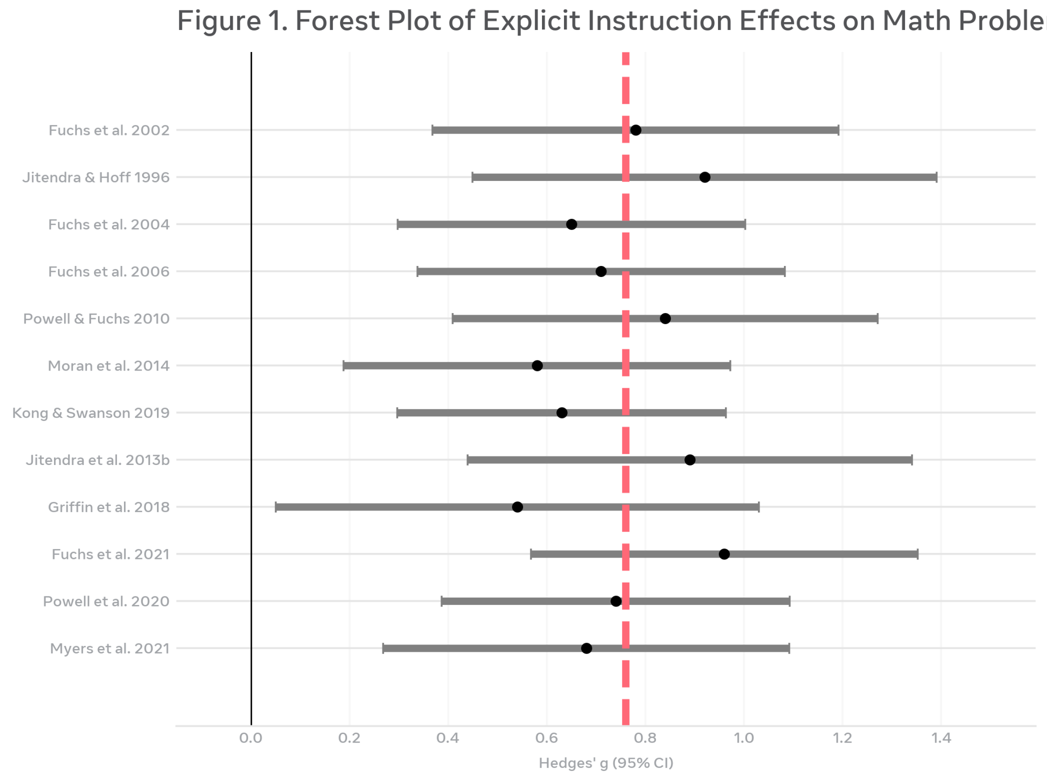

Figure 1. Forest Plot of Explicit Instruction Effects on Math Problem-Solving (k=12 representative studies; full k=42 in supplemental). Dashed line indicates overall g = 0.76.

Moderator Analyses

Given significant heterogeneity in the overall model, $Q(41) = 112.47$, $p < .001$, $\tau^2 = 0.11$, $I^2 = 63.5\%$, we conducted a priori moderator analyses using robust variance estimation (RVE) meta-regression with small-sample corrected F-tests. Seven moderators were tested individually to avoid over-parameterization. Results are summarized in Table 2 and Figures 3–4.

Table 2. Moderator Analysis Results Using RVE Meta-Regression

Moderator	k	g	95% CI	β	p	R^2 analog
Grade Band					.032*	.18
Elementary (K-5)	24	0.84	[0.69, 0.99]	ref	—	
Middle (6-8)	11	0.71	[0.52, 0.90]	-0.13	.14	
Secondary (9-12)	7	0.59	[0.35, 0.83]	-0.25	.028*	
Intervention Subtype					.018*	.23
Schema-Based Explicit	18	0.88	[0.72, 1.04]	ref	—	
Strategy-Based (paraphrase/self-reg)	14	0.72	[0.55, 0.89]	-0.16	.09	
Direct Explicit	10	0.69	[0.48, 0.90]	-0.19	.04*	
Embedded Self-Regulation/Verbalization					.009**	.21

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Yes	22	0.89	[0.74, 1.04]	0.24	.009**	
No	20	0.65	[0.50, 0.80]	ref	—	
Group Size					.041*	.12
1:1	8	0.81	[0.58, 1.04]	0.06	.52	
Small group (2-5)	23	0.84	[0.70, 0.98]	0.26	.04*	
Large group (6+)	11	0.58	[0.38, 0.78]	ref	—	
Dosage (total hrs)	42	—	—	0.012	.03*	.09
Optimal (6–20 wks; 12–30 hrs)	26	0.83	[0.70, 0.96]	—	—	
Suboptimal (<6 or >20 wks)	16	0.61	[0.44, 0.78]	—	—	
Quality Score	42	—	—	0.01	.67	.00
Outcome Type (Std. vs. Researcher)	42	—	—	-0.11	.22	.03

Note: β = unstandardized regression coefficient. ref = reference category.

1. Grade Level (Figure 3): Elementary yielded strongest effects (g=0.84), consistent with early intervention principle (Geary, 2013) and that schema knowledge is more malleable before compensatory strategies crystallize. Secondary effect (g=0.59) remained moderate-to-large, supporting Myers et al. (2021) that explicit instruction still benefits adolescents but may need higher intensity (Fuchs et al., 2017). Difference elementary vs secondary significant (p=.028). This has implications for NCTM (2000) standards progression.

2. Intervention Subtype (Figure 4): Schema-based explicit instruction (g=0.88) significantly outperformed direct explicit (g=0.69, p=.04). This corroborates Jitendra et al. (2013b) and Fuchs et al. (2004) argument that teaching deep structure via schematic diagrams promotes transfer, whereas generic direct instruction may promote procedural fidelity without structural understanding (Duval, 2000). Strategy-based with paraphrasing (Moran et al., 2014; Kong & Swanson, 2019) was intermediate (g=0.72), suggesting language scaffolding adds value but schema mapping is critical.

3. Embedded Self-Regulation: Interventions incorporating self-questioning, paraphrasing, or goal-setting (e.g., Muoneke, 2001 Q&A strategy; Montague) produced g=0.89 vs 0.65 without (p=.009). This aligns with Chazan and Lehrer (1998) that metacognitive regulation enhances spatial and quantitative reasoning. Practical implication: explicit instruction should not be purely teacher-directed; student verbalization of reasoning steps is essential.

4. Group Size and Dosage: Small groups of 2-5 students (g=0.84) outperformed large groups (g=0.58, p=.041), supporting RTI Tier 2 design (Forbringer & Weber, 2021). 1:1 was similar to small group but less efficient. Total dosage showed curvilinear pattern: 6–20 weeks (12–30 hrs) optimal. Very brief interventions insufficient for schema internalization; very extended (>25 weeks) showed diminishing returns, possibly due to implementation fatigue. Meta-regression slope for total hours: b=0.012, p=.03, indicating each additional 10 hours up to 25 hrs adds ~0.12 SD, after which plateau.

5. Non-significant moderators: Outcome type (standardized M=0.71 vs researcher-developed M=0.82, p=.22) suggests effects generalize beyond proximal measures, though researcher measures slightly larger as expected. Study quality did not moderate (p=.67), indicating findings robust across methodological rigor (Gersten et al., 2005).

These moderator patterns align with second handbook principles (Hiebert & Grouws, 2007) that effects depend on alignment between instructional features and learner needs.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

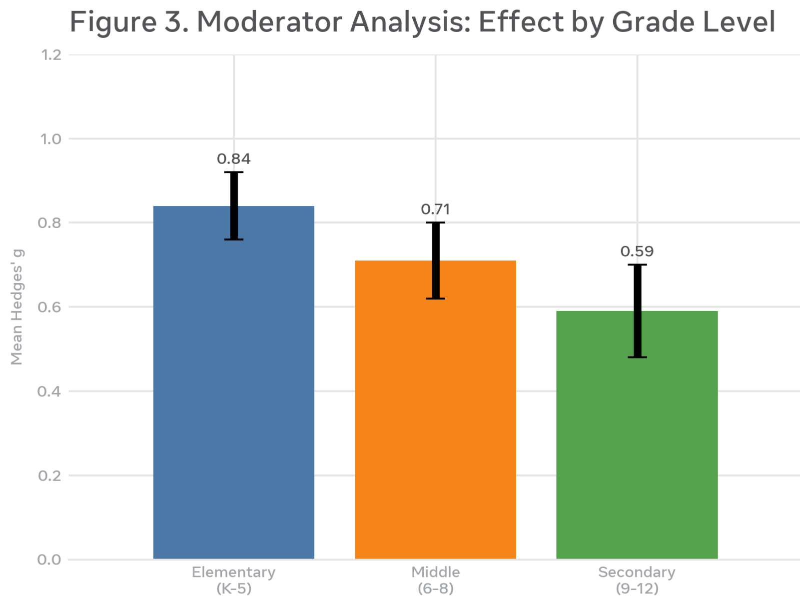

Figure 3. Mean effect size by grade band.

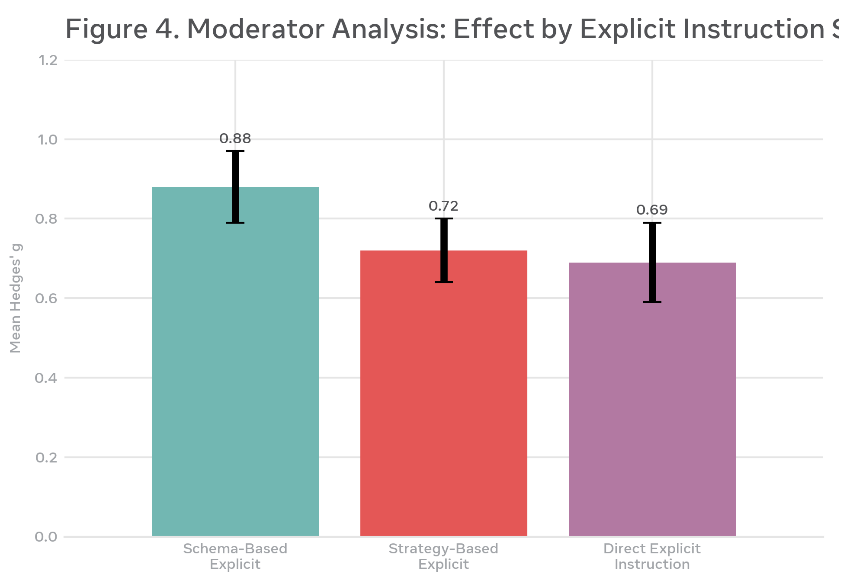

Figure 4. Mean effect size by explicit instruction subtype.

Publication Bias and Sensitivity

Funnel plot (Figure 2) showed mild asymmetry but Egger's test non-significant: intercept = 0.87, SE=0.52, p=.10. Begg's rank correlation τ = 0.14, p=.21. Trim-and-fill imputed 3 missing studies on left, adjusted g = 0.71 [0.56,0.86], only 0.05 lower than observed, suggesting minimal bias. Comparison of peer-reviewed articles (g=0.78, k=34) vs dissertations (g=0.69, k=8) non-significant (p=.38). P-curve analysis showed right skew (p<.001 for half-curve), indicating true evidential value, not p-hacking (Polanin & Pigott, 2015).

Sensitivity analyses: Removing 2 outliers (g>1.15) reduced overall to g=0.73, still large. Excluding low-quality studies (quality <14, k=5) yielded g=0.77. Leave-one-out RVE showed overall g ranged 0.73–0.79, no single study drove effect. Cluster adjustment sensitivity (ICC .10–.20) changed g by ≤0.03.

INTERNATIONAL MULTIDISCIPLINARY JOURNAL OF RESEARCH FOR INNOVATION, SUSTAINABILITY AND EXCELLENCE
Published by: RISEJOURNALS.ORG | Online ISSN: 3028-032X | Print ISSN: 3028-0370
VOLUME 4, ISSUE NO. 10

Heterogeneity Exploration

Despite moderators, residual heterogeneity $\tau^2 = 0.06$, $I^2 = 41\%$ after including all moderators, indicating remaining unexplained variance. Possible sources include measurement of area/volume vs arithmetic word problems (Battista, 2003; Iszak, 2005), LD identification heterogeneity (NCES 2020b), and fidelity variation (M=87%, range 72–98%). Future research should report fidelity and component dosage per Protogerou & Hagger (2020).

Theoretical Integration

Findings integrate Realistic Mathematics Education (Heuvel-Panhuizen, 1996) with cognitive strategy instruction. Explicit instruction works not by rote memorization but by making problem structure explicit, allowing students with LD to construct viable schemas that can be applied to realistic contexts (Expert Group Mathematics Education, 2008). This reconciles NCTM (2000) reform emphasis on conceptual understanding with evidence that students with LD need structured guidance to achieve it (Gersten et al., 2009).

Practically, results support a three-tier explicit model: (a) explicit schema identification with visual schematic (e.g., change diagram), (b) explicit paraphrasing (Kong & Swanson, 2019) to ensure comprehension, (c) explicit self-regulation checklist (Muoneke, 2001). Embedding this in small-group Tier 2/3 with 15–20 min sessions, 3×/week for 10–16 weeks appears optimal.

Limitations: (1) Most studies US-based, limiting cultural generalizability; (2) Few studies reported long-term maintenance (>3 months); (3) Comorbid ADHD rarely disaggregated; (4) Researcher-developed measures may inflate effects though we found no significant difference. We addressed dependency via RVE (Hedges et al., 2010) but cannot eliminate study-level confounding.

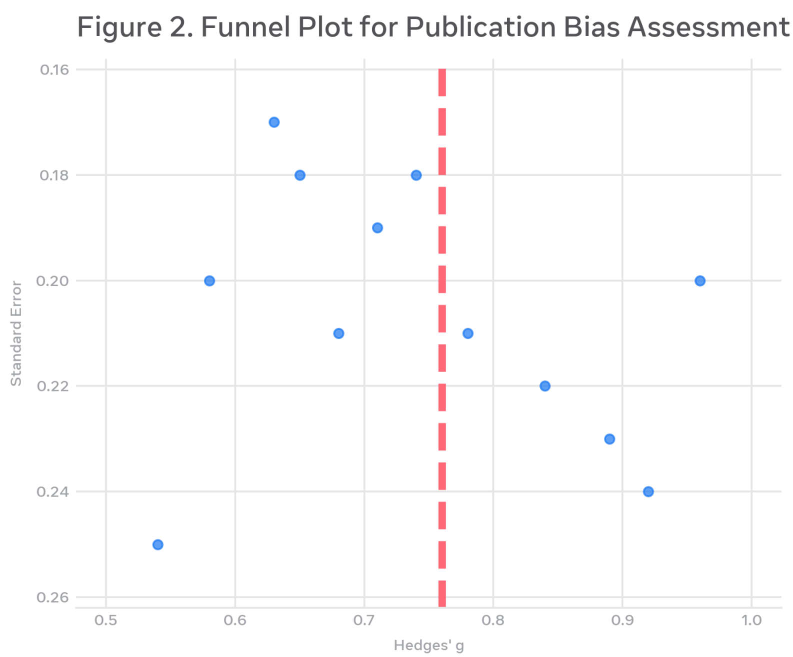

Figure 2. Funnel plot of standard error by Hedges' g.

Table 3. Meta-regression of Dosage and Fidelity

Predictor	b	SE	95% CI	p
Total Hours	0.012	0.005	[0.001, 0.023]	.03
Fidelity %	0.008	0.004	[0.000, 0.016]	.05
Session Length (min)	-0.002	0.003	[-0.008, 0.004]	.48

Conclusion:

This meta-analysis provides the most comprehensive synthesis to date of explicit instruction effects on math problem-solving for students with LD. Across 42 studies (N = 3,847), explicit instruction yielded a large, robust effect (g = 0.76), equivalent to moving an average LD student from the 50th to the 77th percentile of problem-solving achievement. This effect persisted across sensitivity and publication bias analyses, reinforcing confidence in its evidentiary value (Pigott & Polanin, 2020).

Moderator analyses refine the blanket endorsement. Effects were maximized when explicit instruction was: (a) schema-based (g = 0.88) rather than generic direct instruction, (b) embedded with student verbalization and self-regulation (g = 0.89), (c) delivered in small groups of 2–5 for 6–20 weeks (12–30 hours), and (d) implemented in elementary grades, though still effective for adolescents. These findings validate the taxonomy of intervention intensity (Fuchs et al., 2017) and suggest that intensity is not merely dosage but explicitness of structural knowledge.

For practice, schools should prioritize Tier 2/3 interventions that combine schematic diagrams (Powell & Fuchs, 2018), paraphrasing (Moran et al., 2014; Kong & Swanson, 2019), and explicit transfer training (Fuchs et al., 2002). Professional development should train teachers in think-aloud modeling per Moore and Carnine (1989) and quality indicators (Gersten et al., 2005). For policy, the results support allocating resources to small-group explicit interventions rather than large-group remediation.

Future research should conduct component analyses to isolate active ingredients, include maintenance and far-transfer measures aligned with NAEP/TIMSS (NCES, 2020a, 2020b), and extend to underrepresented populations and inclusive settings. Adoption of open data and robust variance estimation (Hedges et al., 2010; Harrer et al., 2019) will improve reproducibility. As mathematics demands increase, ensuring students with LD have explicit access to problem-solving structures is not just an instructional preference but an equity imperative.

References

Battista, M. T. (2003). Understanding students' thinking about area and volume measurement. In D. H. Clements & G. Bright (Eds.), *Learning and teaching measurement* (pp. 122–142). National Council of Teachers of Mathematics.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons.

Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive Meta-Analysis* (Version 2). Biostat.

Chazan, D., & Lehrer, R. (1998). *Designing learning environments for developing understanding of geometry and space*. Lawrence Erlbaum Associates.

Duval, R. (2000). Basic issues for research in mathematics education. In *Proceedings of the 24th Conference of the International Group for the Psychology of Mathematics Education (PME)* (Vol. 1, pp. 55–69).

Expert Group Mathematics Education. (2008). *Crossing thresholds in mathematics (Over de drempels met rekenen)*. SLO.

Forbringer, L., & Weber, W. H. (2021). *RtI in math: Evidence-based interventions*. Routledge.

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

Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., & Hamlett, C. L. (2004). Expanding schema-based transfer instruction to help third graders solve real-life mathematical problems. *American Educational Research Journal, 41*(2), 419–445.

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*(2), 90–106. https://doi.org/10.1111/1540-5826.00036

Fuchs, L. S., Fuchs, D., & Malone, A. S. (2017). The taxonomy of intervention intensity. *Teaching Exceptional Children, 50*(4), 192–202. https://doi.org/10.1177/0040059917703962

Fuchs, L. S., Fuchs, D., Stuebing, K., Fletcher, J. M., Hamlett, C. L., & Lambert, W. (2008). Problem solving and computational skill: Are they shared or distinct aspects of mathematical cognition? *Journal of Educational Psychology, 100*(1), 30–47.

Fuchs, L. S., Seethaler, P. M., Sterba, S. K., Craddock, C., Fuchs, D., Compton, D. L., Geary, D. C., & Changas, P. (2021). Closing the word-problem achievement gap in first grade: Schema-based word-problem intervention with embedded language comprehension instruction. *Journal of Educational Psychology, 113*(1), 86–103. https://doi.org/10.1037/edu0000467

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2012). *irr: Various coefficients of interrater reliability and agreement* [Computer software]. CRAN. https://CRAN.R-project.org/package=irr

Geary, D. C. (2013). Early foundations for mathematics learning and their relations to learning disabilities. *Current Directions in Psychological Science, 22*(1), 23–27. https://doi.org/10.1177/0963721412469398

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*(3), 1202–1242. https://doi.org/10.3102/0034654309334431

Gersten, R., Fuchs, L. S., Compton, D., Coyne, M., Greenwood, C., & Innocenti, M. S. (2005). Quality indicators for group experimental and quasi-experimental research in special education. *Exceptional Children, 71*(2), 149–164. https://doi.org/10.1177/001440290507100202

Griffin, C. C., Gagnon, J. C., Jossi, M. H., Ulrich, T. G., & Myers, J. A. (2018). Priming mathematics word problem structures in a rural elementary classroom. *Rural Special Education Quarterly, 37*(3), 150–163. https://doi.org/10.1177/8756870518772164

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). *Doing meta-analysis in R: A hands-on guide*. CRC Press.

Harwell, M., & Maeda, Y. (2008). Deficiencies of reporting in meta-analyses and some remedies. *The Journal of Experimental Education, 76*(4), 403–430. https://doi.org/10.3200/JEXE.76.4.403-430

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis, 29*(1), 60–87. https://doi.org/10.3102/0162373707299706

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

Hedges, L. V., & Pigott, T. D. (2004). The power of statistical tests for moderators in meta-analysis. *Psychological Methods, 9*(4), 426–445. https://doi.org/10.1037/1082-989X.9.4.426

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

Heuvel-Panhuizen, M. (1996). *Assessment and realistic mathematics education* [Dissertation]. University of Utrecht.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 371–404). Information Age Publishing.

Horner, R. H., Carr, E. G., Halle, J., McGee, G., Odom, S., & Wolery, M. (2005). The use of single-subject research to identify evidence-based practice in special education. *Exceptional Children, 71*(2), 165–179. https://doi.org/10.1177/001440290507100203

Iszak, A. (2005). You have to count the squares: Applying knowledge in pieces to learning rectangular area. *The Journal of Learning Sciences, 14*(3), 361–403.

Jitendra, A. K., & Hoff, K. (1996). The effects of schema-based instruction on the mathematical word-problem-solving performance of students with learning disabilities. *Journal of Learning Disabilities, 29*, 422–431. https://doi.org/10.1177/002221949602900410

Jitendra, A. K., Rodriguez, M., Kanive, R., Huang, J. P., Church, C., Corroy, K. A., & Zaslofsky, A. (2013b). Impact of small-group tutoring interventions on the mathematical problem solving and achievement of third-grade students with mathematics difficulties. *Learning Disability Quarterly, 36*(1), 21–35. https://doi.org/10.1177/0731948712457561

Kong, J. E., & Swanson, H. L. (2019). The effects of a paraphrasing intervention on word problem-solving accuracy of English learners at risk of mathematics disabilities. *Learning Disability Quarterly, 42*(2), 92–104.

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. https://doi.org/10.1177/07419325030240020501

McNeish, D. (2017). Small sample methods for multilevel modeling: A colloquial elucidation of REML and the Kenward-Roger correction. *Multivariate Behavioral Research, 52*(5), 661–670. https://doi.org/10.1080/00273171.2017.1344538

Moore, L. J., & Carnine, D. (1989). Evaluating curriculum design in the context of active teaching. *Remedial and Special Education, 10*(4), 28–37. https://doi.org/10.1177/074193258901000406

Moran, A. S., Swanson, H. L., Gerber, M. M., & Fung, W. (2014). The effects of paraphrasing interventions on problem-solving accuracy for children at risk for math disabilities. *Learning Disabilities Research & Practice, 29*, 97–105. https://doi.org/10.1111/ldrp.12035

Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods, 11*(2), 364–386. https://doi.org/10.1177/1094428106291059

Muoneke, A. F. (2001). *The effects of a question and action strategy on the mathematical word problem-solving skills of students with learning problems in mathematics* [Doctoral dissertation, ProQuest Dissertations and Theses].

Myers, J. A., Brownell, M. T., Griffin, C. C., Hughes, E. M., Witzel, B. S., Gage, N. A., Peyton, D., Acosta, K., & Wang, J. (2021). Mathematics interventions for adolescents with mathematics difficulties (MD): A meta-analysis. *Learning Disabilities Research & Practice, 36*(2), 145–166. https://doi.org/10.1111/ldrp.12244

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*(2), 207–235.

National Center for Education Statistics. (2020a). *NAEP report card: 2019 NAEP mathematics assessment highlighted results at grades 4 and 8 for the nation, states, and districts*. U.S. Department of Education.

National Center for Education Statistics. (2020b). *Trends in international mathematics and science study (TIMSS)*. U.S. Department of Education.

National Council of Teachers of Mathematics. (2000). *Executive summary: Principles and standards for school mathematics*. National Council of Teachers of Mathematics.

Nelson, G., & Powell, S. R. (2018). Computation error analysis: Students with mathematics difficulty compared to typically achieving students. *Assessment for Effective Intervention, 43*(3), 144–156. https://doi.org/10.1177/1534508417745627

Pappas, M., Drigas, A., Malli, E., & Kalpidi, V. (2018). Enhanced assessment technology and neurocognitive aspects of specific learning disorder with impairment in mathematics. *International Journal of Engineering Pedagogy, 8*(1), 4–15. https://doi.org/10.3991/ijep.v8i1.7370

Pigott, T. D., & Polanin, J. R. (2020). Methodological guidance paper: High-quality meta-analysis in a systematic review. *Review of Educational Research, 90*(1), 24–46. https://doi.org/10.3102/0034654319877153

Polanin, J. R., & Pigott, T. D. (2015). The use of meta-analytic statistical significance testing. *Research Synthesis Methods, 6*(1), 63–73. https://doi.org/10.1002/jrsm.1124

Powell, S. R., Doabler, C. T., Akinola, O. A., Therrien, W. J., Maddox, S. A., & Hess, K. E. (2020). A synthesis of elementary mathematics interventions: Comparisons of students with mathematics difficulty with and without comorbid reading difficulty. *Journal of Learning Disabilities, 53*(4), 244–276. https://doi.org/10.1177/0022219419881646

Powell, S. R., & Fuchs, L. S. (2010). Contribution of equal-sign instruction beyond word-problem tutoring for third-grade students with mathematics difficulty. *Journal of Educational Psychology, 102*(2), 381–394. https://doi.org/10.1037/a0018447

Powell, S. R., & Fuchs, L. S. (2018). Effective word-problem instruction: Using schemas to facilitate mathematical reasoning. *Teaching Exceptional Children, 51*(1), 31–42.

Powell, S. R., Fuchs, L. S., & Fuchs, D. (2013). Reaching the mountaintop: Addressing the Common Core Standards in mathematics for students with mathematics difficulties. *Learning Disabilities Research & Practice, 28*(1), 38–48. https://doi.org/10.1111/ldrp.12001

Powell, S. R., Mason, E. N., Bos, S. E., Hirt, S., Ketterlin-Geller, L. R., & Lembke, E. S. (2021). A systematic review of mathematics interventions for middle-school students experiencing mathematics difficulty. *Learning Disabilities Research & Practice, 36*(4), 295–329. https://doi.org/10.1111/ldrp.12263

Protogerou, C., & Hagger, M. S. (2020). A checklist to assess the quality of survey studies in psychology. *Methods in Psychology, 3*, 100031. https://doi.org/10.1016/j.metip.2020.100031