Is Mathematical Representation of Problems an Evidence-Based Strategy for Students With Mathematics Difficulties?

Asha K. Jitendra[1], Gena Nelson[1], Sandra M. Pulles[1], Allyson J. Kiss[1], and James Houseworth[1]

Abstract

The purpose of the present review was to evaluate the quality of the research and evidence base for representation of problems as a strategy to enhance the mathematical performance of students with learning disabilities and those at risk for mathematics difficulties. The authors evaluated 25 experimental and quasiexperimental studies according to the Gersten et al. (2005) guidelines for group research studies. Results suggest that the representation of mathematical problems as a strategy is an evidence-based practice based on the criteria set by Gersten et al. Implications for research are discussed.

Students with learning disabilities (LD) and those at risk for mathematics difficulties (MD) often have difficulty with many aspects of mathematics. Frequently reported manifestations of math-related LD include deficits in early number concepts, number combinations, and place value that lead to errors in procedural computation and higher-level mathematics (e.g., Cirino, Ewing-Cobbs, Barnes, Fuchs, & Fletcher, 2007; Geary, Hoard, Byrd-Craven, Nugent, & Numtee, 2007; Jordan, Hanich, & Kaplan, 2003; Zheng, Swanson, & Marcoulides, 2011). Mathematics is challenging for students with LD who also demonstrate deficiencies in domain-general abilities, such as working memory, language, attentive behavior, and simultaneous storage and processing speed of information (e.g., Fuchs et al., 2010; Geary et al., 2007) that negatively affect their learning. As such, schools need to meet the needs of students both with LD and those at risk for MD (hereafter, *students with MD*) and identify evidence-based teaching practices that improve mathematics achievement. An effective approach for improving the mathematics achievement of students with MD is representation of problems as a strategy (Gersten, Beckmann, et al., 2009; Gersten,

Chard, et al., 2009; Xin & Jitendra, 1999; Zhang & Xin, 2012). Representations have been used for many years with students with MD to "illustrate solution strategies for mathematical problems" (Gersten, Chard, et al., 2009, p. 1229).

The focus of our review was to evaluate the body of research on representation of problems as a strategy for students with MD by applying the quality indicators and standards for evidence-based practices proposed by Gersten et al. (2005).

The Role of Representations in Developing Mathematics Thinking and Students With MD

The use of models (representations) to communicate mathematical thinking is one of

---

[1]University of Minnesota

Corresponding Author:
Asha K. Jitendra, PhD, University of Minnesota,
245 Educational Sciences Building, 56 E River Rd.,
Minneapolis, MN 55455, USA.
E-mail: jiten001@umn.edu

several Standards for Mathematical Practice in the Common Core State Standards (National Governor's Association Center for Best Practices [NGA] & Council of Chief State School Officers [CCSO], 2010). The National Council of Teachers of Mathematics (NCTM, 2000) recommended that students be provided access to a variety of representations to develop mathematical understanding. *Representations* refer to any configuration of characters, images, or concrete objects that symbolizes an abstract idea (Goldin & Kaput, 1996) and may include manipulative materials (physical objects), pictures or diagrams, real-life situations, spoken language, or written symbols (Lesh, Post, & Behr, 1987). Although *internal representations* (i.e., mental imagery) are known to play an important role in mathematical reasoning (see Hegarty & Kozhevnikov, 1999), most research has focused on external representations, given the difficulty in studying mental imagery.

> *When instructional strategies include opportunities for students to . . . make connections between the mathematical concept and the real world, there are positive benefits to student learning.*

A well-developed literature exists that explains the role of external representations in mathematics achievement (see Goldin, 2003). When instructional strategies include opportunities for students to interact with the objects (e.g., algebra tiles) to develop mathematical understanding, use appropriate diagrams (e.g., strip or tape, schema, percentage bar) that depict the mathematical relations among key elements in the problem, or make connections between the mathematical concept and the real world, there are positive benefits to student learning (NCTM, 2000; What Works Clearinghouse [WWC], 2012). Further, many recommend that mathematics instruction incorporate a variety of representations, given that multiple representations allow students to

make connections within and between mathematical ideas to facilitate learning (Cuoco & Curcio, 2001; NCTM, 2000).

Although concrete (or enactive) manipulative materials have been used as a teaching tool in school mathematics, less attention has been given to the use of visual (pictorial) representations in mathematics instruction (Martin & Schwartz, 2005). Research suggests that visual representations support students' visualization of mathematical ideas, especially when instruction is provided that meaningfully connects the abstract mathematical concepts with the visual representations (Ng & Lee, 2009). Further, there is evidence that compared to pictorial (iconic) representations (focus on visual appearance of the objects described in a problem), schematic representations (focus on relevant data and relations in the problem) are more meaningful, and their use is positively correlated with successful problem-solving performance (Hegarty & Kozhevnikov, 1999). The importance of mathematics visualization and schematic representations has been highlighted in two studies that examined students of different achievement levels and their use of visual imagery and spatial visualization (van Garderen, 2006; van Garderen & Montague, 2003). Results of these studies indicated that gifted students demonstrated the highest levels of spatial visualization compared to students with LD and average achievers, and the ability to visualize was strongly associated with the ability to understand mathematics. Further, gifted students tended to use more schematic representations than did average achievers and students with LD; students with LD relied more on pictorial representations than did the other two groups (van Garderen, 2006; van Garderen & Montague, 2003).

Visual representations have been widely used in mathematical problem solving and serve a variety of purposes, such as "summarizing problem information, recording and reasoning about situation/story elements, offloading memory storage, coordinating the

results of intermediate calculations, representing numerical or functional relationships via graphs, and making abstract relationships concrete" (Zahner & Corter, 2010, p. 180). The notion that complete and accurate representations are essential to increasing students' mathematical understanding has its antecedents in the literature on cognitive load theory. Cognitive load theory suggests that the "amount of cognitive processing required for learning at any one time does not exceed the learner's processing capacity" (DeLeeuw & Mayer, 2008, p. 223). In short, the use of appropriate and meaningful representations is an effective approach for reducing working memory demands and supporting the development of abstract reasoning. External representations serve to reduce the attention required to mentally represent the mathematical ideas and allow students to focus more on the abstract mathematical concepts.

Representational thinking is an essential skill for all students. Unfortunately, students with MD often demonstrate difficulties in successfully using representations to express their mathematical thinking for a number of reasons. First, these students not only have difficulty in understanding and using representations on their own, but also their self-constructed representations may not be adequate. Although experimenter-provided representations (diagrams) to illustrate the problem text may not be effective for students with MD (Booth & Koedinger, 2012), such representations can promote learning when instruction supports understanding the representations before using them to elucidate mathematical concepts (Butler, Miller, Crehan, Babbitt, & Pierce, 2003; van Garderen, 2007; Witzel, Mercer, & Miller, 2003). Similarly, encouraging these students to merely visualize the problem or draw a diagram to solve a problem may not be effective (see van Garderen, 2006; van Garderen & Montague, 2003). A representation needs to be understood for it to be useful. For students with MD to effectively generate their own diagrams during problem solving, instruction must be "multifaceted, focus on conceptual understanding of diagrams, diagram generation, and use of diagrams

as tools to reason with" (van Garderen, 2007, p. 541).

Second, the presence of multiple representations may be counterproductive for many students with MD, as they often have difficulty choosing among and switching representations as needed to solve a problem. The lack of representational fluency could be one reason that students with MD do not always benefit from multiple representations. Results from a research synthesis related to mathematics interventions for struggling students suggest that for representations to be effective, teachers should provide high levels of instructional guidance (see Gersten, Beckmann, et al., 2009). Instruction should support students by providing explicit guidance to select the appropriate representation and effectively use the representation to visualize data in the problem.

> *Representations can promote learning when instruction supports understanding the representations before using them to elucidate mathematical concepts.*

The purpose of our review was to evaluate the quality of the research and evidence base for representation of problems as a strategy to promote the mathematics performance of students with MD using the evidence standards for group research proposed by Gersten et al. (2005). We focused on representations (concrete and visual) for the following reasons. First, reports from national organizations (NCTM, 2000; NGA & CCSO, 2010) strongly recommend the use of representations to develop students' mathematical thinking. As such, an underlying assumption is that sound scientific evidence exists for the recommendation. Second, researchers have conducted numerous studies that have used representation of problems as a strategy to enhance the mathematics performance of students with MD. Several research syntheses (see Gersten, Beckmann, et al., 2009; Gersten, Chard, et al., 2009; Xin & Jitendra, 1999; Zhang & Xin, 2012) have provided evidence to support the practice. Gersten, Chard, et al. (2009) reported

an effect size of 0.46 for teacher and student use of visual representations. Third, in the field of special education, the use of multiple representations (concrete-representational-abstract [CRA]) as a learning progression is prescribed as an efficacious teaching strategy for increasing students' mathematical competency (Misquitta, 2011). However, the research on representation of problems as a strategy has not been evaluated against the quality standards to determine the practice as evidence based. Therefore, our review evaluated the quality of the research and evidence base for representation of problems as a strategy for students with MD using criteria for group research proposed by Gersten et al. (2005). In addition, we reviewed studies that demonstrated meeting the evidence criteria to examine associated study elements and sample characteristics (i.e., disability or difficulty status of participants, grade level, types of representations, mathematical domain, instructional arrangements, interventionists involved in the delivery of the intervention) related to the use of representations.

Method

We conducted a search of the literature, ending in December 2014, using a three-step process. First, we searched the electronic databases of ERIC and PsycINFO using keywords such as *representations, visuals, concrete, manipulatives, schema*, and *diagrams*. Each keyword was combined with the search terms describing the population (*disabilities, at risk, remedial, learning problems, mathematics difficulties*) and mathematics (*instruction, intervention, training, treatment*). Second, we searched the citation lists of relevant review articles (e.g., Gersten, Beckmann, et al., 2009; Gersten, Chard et al., 2009; Misquitta, 2011; Zhang & Xin, 2012; Zheng, Flynn, & Swanson, 2013) as well as conducted an ancestral search of reference lists of eligible studies. In this process, we identified researchers (e.g., Bottge and colleagues; Fuchs and colleagues) whose work on representations was not identified in the initial phase using keywords. We searched the databases further for studies by these authors. Third, we hand-searched the following journals: *Exceptional Children, Journal of Learning Disabilities, The Journal of Special Education, Learning Disability Quarterly, Learning Disabilities Research & Practice*, and *Remedial and Special Education*.

For the purpose of the current review, we focused on published group-design mathematics intervention studies. We reviewed the abstracts (and narratives of studies by specific authors identified in the second phase of the search) of 394 identified studies to exclude articles that were nonexperimental (i.e., no control group) or which did not include a mathematics outcome measure. We examined a total of 85 studies that met these criteria further to determine whether each study would be included in the present review, using five eligibility criteria:

1. The study focused on school-age children who were identified as having LD and/or were struggling with mathematics. Students struggling with mathematics were considered at risk for mathematics if their scores were below the 25th percentile (standard score of 90) on a standardized test. Typically, the 25th percentile cutoff has been used to identify children at risk (Geary, Hoard, Nugent, & Bailey, 2012; Murphy, Mazzocco, Hanich, & Early, 2007), whereas the 10th percentile cutoff has been used to identify children with mathematics LD (Mazzocco, Myers, Lewis, Hanich, & Murphy, 2013). We excluded studies that used a cutoff higher than the 25th percentile to identify students at risk for MD. When studies included both students with and without LD or at risk for MD, the data for these students had to be disaggregated or more than 50% of the participants in the studies were students with LD or at risk for MD.

2. The study used an experimental or quasiexperimental design, and the study provided sufficient information (e.g., means, standard deviations, t or F statistics) to calculate effect sizes.

3. The study included an intervention that focused on representations. We use the term *representations* to refer to materials, visual sketches, diagrams,

or pictures that symbolized an abstract mathematical idea.

4. The study included at least an experimenter-designed or a norm-referenced mathematics measure.
5. The study was published in English in a peer-reviewed journal. A total of 25 group-design studies met the screening criteria and were included in the present review.

Quality Indicators and Interrater Agreement

We evaluated the methodological quality of the included studies by applying the quality indicators (QIs) outlined by Gersten et al. (2005). The QIs comprise both essential and desirable indicators, which we applied in sequence. The essential QIs address 10 components of four categories of methodological quality (i.e., description of participants, implementation of intervention and description of comparison conditions, outcome measures, data analysis) and were coded on a 1-to-3 scale (3 = *indicator fully met*, 2 = *indicator partially met*, 1 = *indicator not met*) using a rubric, modified from Jitendra et al. (2015) primarily with regard to the description-of-participants category. Specifically, we clarified the quality of student disability diagnoses procedures such that to receive a score of 3, a study had to describe LD or at-risk-for-MD status by linking the definition to that used in state or district criteria or in prior research and document that participants met the requirements of the definition (e.g., IQ and achievement score discrepancy, had an individualized education plan in mathematics, scored below the cutoff on a standardized mathematics test) in addition to providing quality student demographic information.

In accordance with the Gersten et al. (2005) standards, a study was determined to meet the essential QIs if it met at minimum all but one of the 10 components of the four essential QIs. Any component that received a score of 1 was considered unacceptable. We derived scores on the four essential QIs from the totals of the components associated with each QI. We then evaluated the studies that met the criteria for rigorous research to determine whether they were high quality or acceptable by applying the desirable QIs. Studies were coded on a dichotomous scale as present or not for the desirable QIs. We classified each individual study as high quality (met all but one of the essential QIs and at least four of eight desirable QIs) or acceptable quality (met all but one of the essential QIs and at least one desirable quality indicator) in accordance with the Gersten et al. standards.

We calculated interrater agreement following coder training and coding of several articles on representations that did not meet our criteria (e.g., data were not disaggregated for students with LD or MD) and were excluded from this review. Training occurred until at least 90% agreement between raters was reached. Any discrepancies between raters were discussed and resolved before all 25 studies were double-coded by two raters.

Across studies, the mean percentage agreement index (i.e., the total number of agreements divided by the total number of agreements and disagreements multiplied by 100) for the essential QIs was 89%. For each of the four essential QI categories, the mean interrater agreement was 87% for description of participants, 88% for intervention and comparison conditions, 96% for outcome measures, and 92% for data analysis. Most discrepancies (82%) involved one rater scoring a 3 for one component and the second rater awarding a 2 for the same component such that both ratings would indicate the presence of the essential QIs and thus considered acceptable. The mean interrater agreement for desirable QIs was 97%.

Determining Evidence-Based Practice

Consistent with the recommendations of Gersten et al. (2005), we determined whether representation of problems as a strategy is an evidence-based or a promising practice for students with MD. A practice is deemed *evidence based* if there are at least two high-quality or four acceptable-quality studies supporting the practice and the weighted effect size with a 95% confidence interval is significantly greater than zero; a practice is considered promising when at least two high-quality or four acceptable-quality studies

support the practice, and the weighted effect size with an 80% confidence interval is significantly greater than zero.

When calculating effect sizes, we used posttest measures (experimenter designed and norm referenced) that assessed mathematical outcomes. In the case where a study used multiple outcome measures, we calculated the mean effect size across measures. When studies involved multiple treatments all of which comprised representations of problems as a strategy, we selected the treatment for comparison that would serve to isolate the independent effect of representations. In addition, when studies included multiple treatments that did not all include representations as an instructional component (e.g., Manalo, Bunnell, & Stillman, 2000), we selected the comparison group (i.e., direct instruction) that did not include the target component (representations) but included intervention components that overlapped the most with the treatment (i.e., process mnemonics).

We applied the most commonly used effect size index, Hedges' *g*, with an adjustment for small samples. The effect size was calculated as the difference between the mean mathematics score for the intervention group and the mean mathematics score for the comparison group, divided by the pooled within-group standard deviation of the mathematics measure. We also used the WWC (2014) additional guidelines to calculate effect sizes (e.g., when studies reported unadjusted group means, when studies used different pretest and posttest measures). We then applied Hedges' bias correction for small sample sizes to each of the effect sizes and calculated the weighted mean effect size and 95% confidence interval.

Study Characteristics

We reviewed the high-quality and acceptable-quality studies meeting the evidence standards to determine the conditions under which representation of problems was shown to be effective. We coded the studies for the disability or difficulty status of participants (e.g., LD, MD), grade level (e.g., elementary, middle), types of representations used (e.g., visuals, manipulative materials), mathematical domain (e.g., basic facts and computation, word problems, fractions), instructional arrangements (i.e., one-on-one, small groups, whole class), interventionists involved in the delivery of the intervention (e.g., teachers, researchers, researcher and teachers), and type of outcome measure used (standardized norm referenced versus nonstandardized or experimenter developed). We also calculated effect sizes for each study element category.

Results

Table 1 provides a summary of the essential QIs as applied to the database of representation studies reviewed; Table 2 presents data related to applying the desirable QIs to those studies that demonstrated meeting the minimum criteria for rigorous research.

Participants

For this category, studies were evaluated for their description of (a) participants with regard to procedures for identifying disability or at-risk-for-MD status as well as quality of demographic information, (b) equivalency of groups across conditions, and (c) information about interventionists and comparability of their characteristics or credentials (e.g., teaching experience) across conditions. More than half the studies (*n* = 19; 76%) met or exceeded the minimum criteria (i.e., an average score of 2) across all subcomponents for this category. All six studies that did not meet the minimum criteria did not provide sufficient information about the interventionists implementing the intervention (i.e., Fede, Pierce, & Matthews, 2013; Ives, 2007; Owen & Fuchs, 2002; Peterson, Mercer, & O'Shea, 1988; Van Luit & Schopman, 2000; Woodward & Brown, 2006).

Intervention and comparison conditions. We evaluated studies for the description and implementation of the intervention, fidelity of implementation procedures, and description of the comparison condition. Again, most studies (*n* = 17; 68%) met or exceeded the minimum criteria across all subcomponents for this category. Of the studies

Table 1. Essential Quality Indicators' Mean Ratings and Effect Sizes for Research Studies on Mathematical Representations.

Study	Description of participants	Intervention/ comparison conditions	Outcome measures	Data analysis
Butler, Miller, Crehan, Babbitt, & Pierce (2003)	2.67	3.00	2.50	2.00
Fede, Pierce, & Matthews (2013)	2.00[a]	2.00[a]	3.00	3.00
Fuchs et al. (2005)	2.67	2.33[a]	2.00[a]	3.00
Fuchs et al. (2008)	2.67	2.67	3.00	3.00
Fuchs et al. (2009)	2.67	2.67	3.00	3.00
Fuchs et al. (2013)	2.67	3.00	3.00	3.00
Fuchs et al. (2014)	2.67	3.00	3.00	3.00
Hutchinson (1993)	2.33	1.67[a]	3.00	2.00[a]
Ives (2007)	2.00[a]	2.00[a]	2.50	2.50
Jitendra et al. (1998)	2.67	3.00	3.00	3.00
Jitendra et al. (2013)	3.00	3.00	3.00	3.00
Kelly, Gersten, & Carnine (1990)	2.00	3.00	2.00[a]	2.00[a]
Manalo, Bunnell, & Stillman (2000)	2.00	2.33	2.50	2.50
Owen & Fuchs (2002)	2.00[a]	2.33[a]	2.00[a]	2.50
Peterson, Mercer, & O'Shea (1988)	2.33[a]	2.33	2.00[a]	2.00[a]
Powell & Fuchs (2010)	2.67	2.67	2.50	3.00
Tournaki, Bae, & Kerekes (2008)	2.67	2.00[a]	1.50[a]	2.00[a]
Van Luit & Schopman (2000)	2.33[a]	1.67[a]	3.00	1.50[a]
Walker & Poteet (1989-1990)	2.00	1.67[a]	2.50	1.50[a]
Wilson & Sindelar (1991)	2.33	3.00	2.00[a]	2.00[a]
Witzel, Mercer, & Miller (2003)	2.00	3.00	2.00[a]	2.00[a]
Woodward (2006)	2.67	3.00	3.00	3.00
Woodward & Brown (2006)	2.33[a]	3.00	2.50	2.50
Xin, Jitendra, & Deatline-Buchman (2005)	3.00	3.00	3.00	3.00
Xin et al. (2011)	3.00	3.00	3.00	3.00

[a]Study received a 1-point score for at least one subcomponent in this methodological category.

that did not meet minimum criteria for at least one or more subcomponent, one study (Walker & Poteet, 1989–1990) did not adequately describe the intervention condition, four studies (Fede et al., 2013; Hutchinson, 1993; Tournaki, Bae, & Kerekes, 2008; Van Luit & Schopman, 2000) did not report the process for determining fidelity of implementation, and four studies (Fuchs et al., 2005; Hutchinson, 1993; Ives, 2007; Owen & Fuchs, 2002) did not describe the comparison condition.

Outcome Measures

For this category, studies were evaluated for documenting use of multiple measures and for measuring outcomes at appropriate times. The majority of studies ($n = 19$; 76%) used multiple outcome measures, and more than half the studies ($n = 15$; 60%) included measures of generalized performance or experimenter-developed outcome measures (Butler et al., 2003, Fuchs et al., 2005, 2008, 2009, 2013, 2014; Jitendra et al., 1998, 2013; Van Luit & Schopman, 2000; Woodward, 2006; Woodward & Brown, 2006; Xin et al., , 2011; Xin, Jitendra, & Deatline-Buchman, 2005). With the exception of one study (Fuchs et al., 2005) that did not report when measures were administered at pretest and posttest, the remaining studies met the minimum criteria for documenting collection of data at appropriate times.

Table 2. Desirable Quality Indicators' Mean Ratings for Research Studies on Mathematical Representations.

Study	1	2A	2B	3	4	5	6	7	8
Butler, Miller, Crehan, Babbitt, & Pierce (2003)	x				x				x
Fuchs et al. (2008)	x	x			x				x
Fuchs et al. (2009)	x	x			x				x
Fuchs et al. (2013)	x	x	x		x				x
Fuchs et al. (2014)	x	x	x		x				x
Jitendra et al. (1998)	x	x	x	x			x		x
Jitendra et al. (2013)	x	x	x	x		x	x		x
Manalo, Bunnell, & Stillman (2000)	x	x		x	x				x
Powell & Fuchs (2010)	x	x							x
Woodward (2006)	x	x		x			x		x
Woodward & Brown (2006)		x				x	x		x
Xin, Jitendra, & Deatline-Buchman (2005)	x	x	x	x			x		x
Xin et al. (2011)	x	x		x			x		x

Note. 1 = low attrition of participants across conditions documented; 2A = reliability data provided; 2B = data collectors unfamiliar with conditions and participants; 3 = outcome measures beyond immediate posttest; 4 = validity data provided; 5 = procedural fidelity examined quality of implementation; 6 = nature of instruction documented through direct observations in control conditions; 7 = audio or videotape excerpts capture nature of intervention; 8 = results presented in a clear, coherent fashion; x = met the indicator.

Data Analysis

Studies were evaluated with regard to (a) using appropriate data analytic techniques that were aligned with the research questions and using the appropriate unit of statistical analysis when analyzing data and (b) reporting and interpreting effect sizes. Although the majority of studies ($n = 19$; 76%) fully met this criterion (a score of 3), five studies (20%) received a score of 2 (i.e., Butler et al., 2003; Ives, 2007; Owen & Fuchs, 2002; Walker & Poteet, 1989–1990; Woodward & Brown, 2006). One study (Van Luit & Schopman, 2000) did not meet this criterion in that the data analytic techniques were not aligned with the research questions, and the unit of statistical analysis used to analyze data was not appropriate. Most studies ($n = 18$; 72%) met the criterion of reporting effect sizes. Seven studies (28%) failed to report effect sizes (Hutchinson, 1993; Kelly, Gersten, & Carnine, 1990; Peterson et al., 1988; Tournaki et al., 2008; Walker & Poteet, 1989–1990; Wilson & Sindelar, 1991; Witzel et al., 2003). In summary, 13 studies (52%) met the Gersten et al.

(2005) standards for rigorous research with regard to the essential QIs (i.e., a minimum score of 2 on at least nine of the 10 QIs).

Desirable QIs

Based on our analysis, eight studies (61.5%) provided evidence of meeting four (Fuchs et al., 2013, 2014; Manalo et al., 2000; Woodward, 2006; Xin et al., 2011) or more (Jitendra et al., 1998, 2013; Xin et al., 2005) desirable indicators. As such, the research reported in these studies is considered high quality. The remaining studies (Butler et al., 2003; Fuchs et al., 2008, 2009; Powell & Fuchs, 2010; Woodward & Brown, 2006) were deemed acceptable quality; these studies documented meeting three desirable indicators.

The effect sizes for the eight high-quality studies ranged from +0.27 to +1.60. The mean weighted effect size was +0.72, with a 95% confidence interval ranging from a low of +0.55 to a high of +0.90. The effect sizes for the five acceptable-quality studies ranged from +0.19 to +1.35; the average weighted effect size was +0.68 [+0.43, +0.92]. These

confidence intervals do not include a value of 0. As such, the set of studies meets Gersten et al. (2005) criteria for being evidence based for students with LD and at risk for MD.

Study Characteristics

The analysis of each study characteristic across the high- and acceptable-quality studies is shown in Table 3. Most studies included students at risk for MD (*n* = 6; 46.2%), four studies consisted of students with LD (30.8%), and the remaining three studies (23.1%) included a heterogeneous group of students with LD and those at risk for MD. The results demonstrate that outcomes (ESs) are moderate in magnitude for students with

LD (ES = 0.62) and those at risk for MD (ES = 0.70); the effect size of 0.95 for both students with LD and MD combined is substantial according to Cohen's criterion. The majority of the studies involved elementary school students in Grades 2 to 5 (*n* = 9; 69.2%). The average effect sizes of 0.69 and 0.79 for elementary and middle school students are moderate to large.

Regarding the types of representations, about half the studies (*n* = 7; 53.8%) used visual models, such as schema diagrams and number lines, with positive and large effects (ES = 0.87). The remaining studies (*n* = 6; 46.1%) used manipulative materials (e.g., fraction circles, tiles, counters) and visual models, with moderate effects (ES = 0.64).

Table 3. Weighted Effect Sizes, Standard Errors, and Confidence Intervals by Study Characteristics Across High- and Acceptable-Quality Studies.

Study characteristic	Weighted ES	*SE*	CI
LD/MD status			
LD (*n* = 4)	0.62	0.18	[0.27, 0.97]
MD (*n* = 6)	0.70	0.09	[0.53, 0.86]
Mixed (*n* = 3)	0.95	0.23	[0.50, 1.40]
Grade level			
Elementary (*n* = 9)	0.69	0.08	[0.53, 0.84]
Middle (*n* = 4)	0.79	0.18	[0.44, 1.13]
Representation			
Visuals (*n* = 7)	0.87	0.14	[0.59, 1.14]
Manipulative materials and visuals (*n* = 6)	0.64	0.09	[0.48, 0.81]
Mathematical domain			
Basic facts and computation (*n* = 2)	0.32	0.35	[−0.37, 1.01]
Word problems (*n* = 7)	0.73	0.12	[0.50, 0.96]
Fractions (*n* = 3)	0.64	0.10	[0.44, 0.84]
Other (*n* = 1)	1.35	0.31	[0.75, 1.95]
Instructional arrangement			
One-on-one (*n* = 3)	0.65	0.15	[0.35, 0.94]
Small group (*n* = 8)	0.68	0.09	[0.51, 0.85]
Whole class (*n* = 2)	1.09	0.26	[0.58, 1.60]
Interventionist			
Researcher (*n* = 7)	0.68	0.09	[0.51, 0.85]
Teacher (*n* = 3)	0.68	0.19	[0.30, 1.05]
Researcher and teacher (*n* = 2)	1.08	0.31	[0.48, 1.68]
Other (*n* = 1)	0.69	0.27	[0.15, 1.22]
Outcome measure			
Experimenter developed (*n* = 13)	0.75	0.07	[0.61, 0.89]
Standardized norm-referenced (*n* = 7)	0.62	0.08	[0.45, 0.78]

Note. ES = effect size; *SE* = standard error; CI = confidence interval; LD = learning disabilities; MD = mathematics difficulties.

For the mathematical domain, half the studies provided instruction in word problem solving (*n* = 7; 53.8%), three studies (23.0%) addressed fractions, and two studies taught basic facts and computation (5.5%). The remaining study in the Other category provides instruction that addresses the entire grade-level content. The effect sizes are moderate for word problems (0.73) and fractions (0.64), whereas the effect size of 1.35 is large for the Other category (number, measurement, geometry, data analysis). The effect size of 0.32 for basic facts and computation is small and nonsignificant (includes 0 in the CI).

With regard to instructional arrangement, more than half the studies provided instruction in small groups (*n* = 8; 62%), and the remaining studies provided one-on-one (*n* = 3; 23%) or whole-class instruction (*n* = 2; 15%). The effect sizes of 0.68 and 0.65 for small-group instruction and one-on-one instruction are moderate. In contrast, the effect size of 1.09 for whole-class instruction is substantial. In terms of interventionists, researchers delivered instruction in the majority of the studies (*n* = 7; 54%), followed by teachers (*n* = 3; 23%), researchers and teachers combined (*n* = 2; 15%), and other (*n* = 1; 8%). A moderate effect size of 0.68 each was found for researchers and teachers, whereas a large effect size of 1.08 was found for researchers and teachers combined. The effect size for the Other category (tutors from the community) is moderate (0.69).

For the type of outcome measure, all studies used experimenter-developed measures, and about half the studies (*n* = 7; 53.8%) also included standardized, norm-referenced mathematics measures. The effect sizes are positive and moderate for both experimenter-developed (ES = 0.75) and standardized, norm-referenced mathematics measures (ES = 0.62).

Discussion

In the current study, we evaluated the quality of published research on representation of problems as a strategy to determine whether the research met Gersten and colleagues' (2005) standards of high quality and rigor.

Thirteen of the 25 studies provided strong evidence of rigorous research across the four categories of the essential QIs; 12 studies met all the essential QIs. One study (Woodward & Brown, 2006) met all but one of the essential QIs; this study failed to provide sufficient information on teachers delivering the intervention. Only eight of the 13 studies that met the essential QIs provided evidence of meeting at least four desirable indicators to be classified as high quality (Fuchs et al., 2013, 2014; Jitendra et al., 1998, 2013; Manalo et al., 2000; Woodward, 2006; Xin et al., 2005, 2011).

The eight high-quality studies documented low or no attrition of participants across conditions and coherently reported the results; however, they differed in their documentation of meeting the other desirable indicators. Although all eight studies provided evidence of reliability (e.g., internal consistency, test-retest reliability, or interrater agreement) for the outcome measures, only five studies (Fuchs et al., 2013, 2014; Jitendra et al., 1998, 2013; Xin et al., 2005) also documented that data collectors were blind to conditions to meet this indicator. Six studies reported results for outcome measures beyond an immediate posttest (Jitendra et al., 1998, 2013; Manalo et al., 2000; Woodward, 2006; Xin et al., 2005, 2011). Three high-quality studies (Fuchs et al., 2013, 2014; Manalo et al., 2000) provided evidence of concurrent validity for at least one of the outcome measures, and only one study (Jitendra et al., 2013) provided information on the quality of implementation. Five high-quality studies (Jitendra et al., 1998, 2013; Woodward, 2006; Xin et al., 2005, 2011) documented the nature of instruction in control classrooms via direct observations. None of the studies used excerpts from audio or videotape recordings to summarize the nature of the intervention.

In sum, 13 studies met the standards for methodological rigor to be considered exemplars of high-quality or acceptable-quality research. The weighted mean effect sizes for high-quality (+0.72) and acceptable-quality (+0.68) studies are considered substantively important based on the WWC standards of interpretation that effect sizes of 0.25 standard

deviations or larger are qualified positive (or negative) effects. Consequently, the set of studies on representation of problems as a strategy met Gersten et al. (2005) standards for an evidence-based practice for students with MD. Further, the findings that the weighted mean effect sizes of +0.62 for students with LD, +0.70 for students at risk for MD, and +0.95 for both students with LD and those at risk for MD support the results of prior research about the benefits of representations for students struggling in mathematics (Gersten, Beckmann et al., 2009; Gersten, Chard et al., 2009).

Implications for Research and Practice

Since the publication of the Gersten et al. (2005) QIs, several reviews have been conducted that applied the QIs for evaluating the methodological rigor of research studies and documented evidence-based practices, such as self-regulated strategy development for writing (Baker, Chard, Ketterlin-Geller, Apichatabutra, & Doabler, 2009), cognitive strategy instruction for text comprehension (Jitendra, Burgess, & Gajria, 2011), and strategy instruction priming the mathematical problem structure (Jitendra et al., 2015). The results from our review add to the research base on the application of QIs to evaluate the quality of research and have implications for future research.

We identified 25 experimental and quasiexperimental studies, and found evidence that 52% of the studies met the standards for methodological rigor. One possible explanation for the large database of 25 identified studies is that the majority of studies (92%) were published in the past 15 years (2000–2014) following the publication of the NCTM (2000) standards, which touted the importance of mathematical representations to develop students' conceptual understanding. Our finding of 13 studies that met the criteria for quality research is similar to the number of quality studies (14) identified by Jitendra et al. (2015) and more than in previous research that

applied the QIs (e.g., Baker et al., 2005; Jitendra et al., 2011). Similar to Jitendra et al. (2015), the majority of the quality studies in this review were published in the past 10 years (2005–2014), which highlights a period of increasing expectations of rigor in educational research (e.g., the use of randomized controlled trials).

Across the 13 studies that met the criteria for methodological rigor, researchers provided ample evidence (a score of 3) such that at least 75% of the studies met eight of the 10 components of the QIs, which is more than the number of components (i.e., five) met in previous research (see Jitendra et al., 2015). Taken together, the studies (a) described the participants, including documenting that participants met the requirements of the disability or difficulty status along with providing student demographic information (85%); (b) established equivalence of groups across conditions (85%); (c) described the intervention (100%); (d) documented the process for measuring fidelity (92%); (e) used multiple measures (77%); (f) measured at appropriate times (92%); (g) linked data analysis techniques to research questions and used appropriate unit of statistical analysis (85%); and (h) reported and interpreted effect sizes (85%). Although these findings are encouraging, it is important for future researchers to address the components that did not met the minimum requirements or were frequently unmet in this database.

First, it is important that researchers adequately describe interventionists (e.g., teachers, research assistants) with regard to salient demographic information and credentials (e.g., teaching experience) as well as establish their equivalence across conditions; this feature was inadequately addressed (a score <3) in 83% of the studies. Second, the QI component of describing the comparison condition is crucial to determining the strength of the treatment in relation to a comparison condition that has few or several overlapping components as the treatment condition. Given that this component was inadequately addressed (a score of <3) in 50% of the studies (including several high-quality and acceptable-quality studies)

suggests the need for researchers to attend to the counterfactual in designing future studies.

The frequencies of the 13 studies meeting the desirable QIs ranged from a low of three desirable indicators to a high of six indicators (i.e., Jitendra et al., 2013). It is worth noting that one of the indicators required not only providing evidence of adequate reliability of outcome measures but also ensuring that data collectors are unfamiliar with the study conditions and participants. As a result, two studies (Fuchs et al., 2008, 2009) that did not meet this indicator did not qualify for high-quality research even though they met three criteria and provided reliability estimates for the outcome measures. Perhaps future research should separate these two components. At the same time, the indicators not adequately met in this review were description of the nature of the intervention using excerpts from audio or videotape recordings and description of the quality of implementation. These indicators were also ones that were not met or infrequently met in previous research applying the QIs (see Baker et al., 2009; Jitendra et al., 2011, 2015), suggesting that researchers are likely to attend to these indicators if journal editors emphasize the importance of these indicators when reviewing studies.

Gersten et al. (2005) provided different sets of indicators as essential or desirable to evaluate the quality of research studies. Some researchers who have applied these QIs to databases of studies have suggested including all or selected desirable indicators (e.g., evidence of reliability and validity of outcome measures, quality of implementation) as essential QIs. Recently, the Council for Exceptional Children (CEC, 2014) released the *CEC Standards for Evidence-Based Practices in Special Education*. The document integrated the standards for both group-design and single-case-design studies to include eight essential indicators. For group designs, a set of eight essential indicators comprises 24 QI components (including Gersten et al. desirable QIs of evidence for reliability and validity of outcome measures). Unlike the Gersten et al. standard (meet all but one of the essential QIs) for studies meeting the methodological rigor, the CEC criterion requires that studies meet all

of the QIs prior to being classified as "evidence-based practices, potentially evidence-based practices, mixed effects, insufficient evidence, or negative effects" (CEC, 2014, p. 6).

The fact that the results are mixed based on applying the different design standards . . . demonstrates the need for further dialogue among research experts around determining criteria for evaluating the methodological quality of individual studies.

Of the 13 studies in our review that met the essential QIs, only six studies (Butler et al., 2003; Fuchs et al., 2008, 2009, 2013, 2014; Manalo et al., 2000) provided some evidence for validity of the outcome measures. For the current review, a study needed to provide validity estimates for any one of several relevant outcome measures (e.g., KeyMath–Revised Problem Solving, as in Fuchs et al., 2008, 2009). As such, the number of individual studies that met this indicator is greater than that in previous reviews, which required providing validity data for the primary outcome measure (Jitendra et al., 2015) or all outcome measures (Jitendra et al., 2011). Despite more studies meeting this indicator, fewer studies in this review would be considered methodologically sound on the basis of CEC's criteria for classifying the evidence base of practices. It is interesting to note that only one study (Jitendra et al., 2013) met the CEC indicator of assessing and reporting implementation fidelity related to dosage. However, Jitendra et al. (2013) did not provide evidence of validity of the outcome measures. Consequently, none of the studies meet CEC's design criteria needed to move forward with the determination of evidence-based practice using CEC criteria. The fact that the results are mixed based on applying the different design standards (CEC, 2014; Gersten et al., 2005) demonstrates the need for further dialogue among research experts around determining criteria for evaluating the methodological quality of individual studies. This dialogue may need to consider including evidence of reliability and validity as essential QIs, because these criteria are

critical components of internal validity and directly relate to the methodological quality of a study.

Notwithstanding the aforementioned limitation of using different standards to determine evidence-based practices, the results of this review demonstrate that the use of mathematical representations is an evidence-based practice in accordance with Gersten et al. (2005) standards. A caveat to the findings is that the majority of studies included representations as part of multicomponent interventions, making it difficult to determine the influence of representations per se. Moving forward, it would be important to design studies that examine the individual effects of representations for students with LD and those at risk for MD.

The specific elements in the interventions that included representations varied across studies, and it is not clear which element or combination of elements was related to the outcome. About half the studies in this database included visual models, such as schema diagrams (e.g., Jitendra et al., 1998, 2013; Xin et al., 2005, 2011) and number lines (e.g., Woodward, 2006, Woodward & Brown, 2006), with positive effects. However, one study on process mnemonics, which presented numbers as characters and operations as situational stories (Manalo et al., 2000) yielded a statistically nonsignificant and small effect (ES = 0.27) compared to instruction that did not use the visuals. Also, this was the only study using visual representations that did not combine representations with other instructional components and resulted in the smallest effect for studies included in this category. Consequently, further research is needed before this representational technique is considered for use with students with LD. Studies in which both teachers and students used visuals resulted in relatively greater effects than studies in which only teachers used visual models (i.e., Manalo et al., 2000; Woodward, 2006). The effects for studies in which both teachers and students used visuals were robust even when the comparison group was exposed to visuals. These positive outcomes may be a function of the specificity of the visuals (illustrated the relations in the problem to make abstract relations concrete) in the treatment

condition and support the findings of Gersten, Chard, et al. (2009).

In the other half of studies (i.e., Butler et al., 2003; Fuchs et al., 2008, 2009, 2013, 2014; Powell & Fuchs, 2010), teachers and students used both manipulative materials and visual models, with equally strong effects as for visuals only. Although all studies included manipulative materials and visual models as part of multicomponent interventions, the comparison group in three studies was also exposed to both manipulative materials and visuals (Fuchs et al., 2013, 2014) or visuals only (Butler al., 2003). One study that resulted in a large effect (Fuchs et al., 2013) was related to the difference in focus on fraction content in the treatment and control conditions. The treatment group focused on the measurement interpretation of fractions, which is based on the recommendations espoused in the NCTM (2000) standards and Common Core State Standards (NGA & CCSO, 2010), whereas the control condition focused on the part–whole interpretation of fractions and on procedures. Further studies could examine which elements of representations are critical (e.g., specificity) and whether these vary based on the particular content emphasis.

The conditions under which the practice of representing problems is effective was also not clear. Which representation or combination of representations is crucial for a particular task? Word problem solving was the most common mathematical content addressed in the majority of studies, and representations included manipulative materials or visuals (i.e., schema diagrams). Researchers (e.g., Fuchs et al., 2008; Jitendra et al., 2013; Xin et al., 2011) provided instruction in problem representation using diagrams that emphasized mathematical modeling to assist students in developing mathematical problem-solving and reasoning skills. The few studies that addressed fractions included the use of both manipulative materials and visual models (e.g., number lines). Visuals only were used in two studies to teach basic facts and computational skills as well in one study that included multiple domains (i.e., number, measurement, geometry, data analysis). Further research that

attends to the effectiveness of different types of representations, visuals or manipulative materials, or a combination thereof is important in the selection of representations for different tasks. Also important is the need for researchers to focus on advanced mathematics, such as algebra, geometry, measurement, and data analysis, needed for school success.

A majority of studies reviewed occurred in elementary grades, specifically Grades 3 and 4. Most studies (*n* = 6) focused on word problem solving (Fuchs et al., 2008, 2009; Jitendra et al., 1998, 2013; Powell & Fuchs, 2010; Xin et al., 2011), two on fractions (Fuchs et al., 2013, 2014), and one on multiplication facts (Woodward, 2006). These studies used visual models, manipulative materials, or a combination of manipulative materials and visuals. Three studies conducted with middle school students used manipulative materials or visuals. In one of these studies (Manalo et al., 2000), visuals were used to teach computation skills involving all four operations to eighth-grade students with LD. Teaching below grade-level content (i.e., computational skills) may be appropriate for some struggling middle school learners who are performing well below their peers. The questions of concern are whether the mathematics taught is developmentally appropriate and whether instruction connects mathematical practices (e.g., making sense of problems, reasoning abstractly and quantitatively) to mathematical content.

For students who struggle to develop mathematics competence, schools can expect moderate to strong improvements when multiple representations . . . are introduced over time and students are provided with high levels of instructional guidance.

Most studies in this review involved researchers (i.e., professors, graduate students) delivering the intervention. Though it is acknowledged that the effectiveness of an intervention is a function of researchers

implementing interventions with a high level of fidelity, it is important that future studies evaluate the efficacy of the intervention when delivered by a teacher. Related to this issue is that instruction in the majority of studies involved individual or small-group tutoring implemented by researchers. Further research that attends to the delivery of representation of problems as a strategy by varied implementers and under varying instructional arrangements is needed for the practice to be widely adopted and used in schools.

About half the studies reviewed used standardized, norm-referenced measures (Butler et al., 2003; Fuchs et al., 2008, 2009, 2013, 2014; Woodward & Brown, 2006; Xin et al., 2011). For word problem solving, the most commonly used standardized measure (*n* = 3) was the KeyMath–Revised Problem Solving (Connolly, 1998), followed by the Problem Solving and Data Interpretation subtest of the Iowa Test of Basic Skills (Hoover, Hieronymous, Dunbar, & Frisbie, 1993). Fraction competence was assessed in two studies using a widely accepted measure (released items between 1990 and 2009 from the National Assessment of Education Progress), whereas another study used the Brigance Comprehensive Inventory of Basic Skills–Revised (Brigance, 1999). One study assessed multiple domains (i.e., number, measurement, geometry, data analysis) using the CTB Terra Nova (McGraw-Hill, 2002). Although research indicates that experimenter-developed measures deliver substantially larger effect sizes than standardized, norm-referenced measures, the positive, moderate effects found for both types of outcome measures is encouraging. However, the need for future studies to include norm-referenced measures that are not tightly aligned with the instructional content of the experimental group is necessary to ensure that apparent differential effects in some cases are not simply a function of which level of variable contained more nonstandardized or experimenter-developed measures. For example, the effect size of 0.30 for word problem-solving instruction on norm-referenced measures was small, whereas experimenter-developed measures yielded a large effect (ES = 0.83).

The research base for representation of problems as a strategy also has implications for practice. It may be useful for schools to consider the reasonable uses of representation of problems as a strategy based on the research studies reviewed. Interventions that include representation of problems as a strategy in general education or special education settings could be a valuable investment, especially in late elementary and middle school grades, in which most of the studies were conducted. The evidence suggests that embedding visual models such as schema diagrams in instructional approaches significantly improves students' mathematical word problem–solving performance. An example of a diagram that models the mathematical relation between quantities in equal-groups or multiplicative-compare problems is the factor-factor-product model (see Xin et al., 2011). Also, visual representations such as number lines can be used to help students build an understanding of mathematical ideas. For example, a derived fact strategy for 6 × 7 can be taught using a number line (see Woodward, 2006). In the studies we reviewed, half of the high-quality research studies included multiple representations. For students who struggle to develop mathematics competence, schools can expect moderate to strong improvements when multiple representations (e.g., number lines, fraction tiles, and fraction circles) are introduced over time and students are provided with high levels of instructional guidance (e.g., Fuchs et al., 2014).

Conclusion

Identifying evidence-based practices in special education is critical when supporting students with disabilities, who typically require highly effective instruction to reduce the achievement gap in mathematics (Dammann & Vaughn, 2001). Per the Gersten et al. (2005) standards, representation of problems as a strategy is an evidence-based practice. In studies that demonstrated evidence of being high quality or acceptable quality, representations included visual models only or both manipulative materials and visuals. Studies also demonstrated

that representations could be effectively used to teach word problem solving and fractions. The positive effect of representations on both norm-referenced and researcher-developed measures is promising. Most implementers were researchers, and small-group instruction was common among these studies. Studies also represented both students with LD and those at risk for MD and most commonly included elementary school–age students. Even as research continues to emerge in this area, representation of problems is an evidence-based strategy with a moderate literature base to support its use with students with MD.

References

References marked with an asterisk indicate studies included in the review.

Baker, S., Chard, D. J., Ketterlin-Geller, L. R., Apichatabutra, C., & Doabler, C. (2009). Teaching writing to at-risk students: The quality of evidence for self-regulated strategy development. *Exceptional Children*, *75*, 303–318.

Booth, J. L., & Koedinger, K. R. (2012). Are diagrams always helpful tools? Developmental and individual differences in the effect of presentation format on student problem solving. *British Journal of Educational Psychology*, *82*, 492–511. doi:10.1111/j.2044-8279.2011.02041.x

Brigance, A. H. (1999). *Brigance Comprehensive Inventory of Basic Skills–Revised*. North Billerica, MA: Curriculum Associates.

*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

Cirino, P. T, Ewing-Cobbs, L., Barnes, M., Fuchs, L. S., & Fletcher, J. M. (2007). Cognitive arithmetic differences in learning disability groups and the role of behavioral inattention. *Learning Disabilities Research & Practice*, *22*, 25–35. doi:10.1111/j.1540-5826.2007.00228.x

Connolly, A. J. (1998). *KeyMath–Revised, normative update: A diagnostic inventory of essential mathematics*. Circle Pines, MN: American Guidance Service.

Council for Exceptional Children. (2014). *Council for Exceptional Children standards for*

evidence-based practices in special education. Retrieved from https://www.cec.sped.org/Standards/Evidence-Based-Practice-Resources-Original

Cuoco, A. A., & Curcio, F. R. (2001). *2001 Yearbook: The roles of representation in school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

Dammann, J. E., & Vaughn, S. (2001). Science and sanity in special education. *Behavioral Disorders*, *27*, 21–29.

DeLeeuw, K. E., & Mayer, R. E. (2008). A comparison of three measures of cognitive load: Evidence for separable measures of intrinsic, extraneous, and germane load. *Journal of Educational Psychology*, *100*, 223–234. doi:10.1037/0022-0663.100.1.223

*Fede, J. L., Pierce, M. E., & Matthews, W. J. (2013). The effects of computer-assisted, schema-based instruction intervention on word problem-solving skills of low-performing fifth grade students. *Journal of Special Education Technology*, *28*, 9–21.

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

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., Seethaler, P. M., . . . Schatschneider, C. (2010). Do different types of school mathematics development depend on different constellations of numerical versus general cognitive abilities? *Developmental Psychology*, *46*, 1731–1746. doi:10.1037/a0020662

*Fuchs, L. S., Powell, S. R., Seethaler, P. M., Cirino, P. T., Fletcher, J. M., Fuchs, D., . . . Zumeta, R. O. (2009). Remediating number combination and word problem deficits among students with mathematics difficulties: A randomized control trial. *Journal of Educational Psychology*, *101*, 561–576.

*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*, 155–174. doi:10.1177/001440290807400202

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematics learning disability. *Child Development*, *78*, 1343–1359. doi:10.1111/j.1467-8624.2007.01069.x

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

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). *Assisting students struggling with mathematics: Response to intervention (RtI) for elementary and middle schools* (NCEE 2009-4060). Washington, DC: National Center for Education Evaluation and Regional Services, Institute of Education Sciences, U.S. Department of Education.

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

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*, 149–164.

Goldin, G. A. (2003). Representation in school mathematics: A unifying research perspective. In J. Kilpatrick, W. Gary Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 275–285). Reston, VA: National Council of Teachers of Mathematics.

Goldin, G. A., & Kaput, J. J. (1996). A joint perspective on the idea of representation in learning and doing mathematics. In L. Steffe, P. Nesher, P. Cobb, G. Goldin, & B. Greer (Eds.), *Theories of mathematical learning* (pp. 397–430). Hillsdale, NJ: Lawrence Erlbaum.

Hegarty, M., & Kozhevnikov, M. (1999). Types of visual-spatial representations and mathematical problem solving. *Journal of Educational Psychology*, *91*, 684–689.

Hoover, H. D., Hieronymous, A. N., Dunbar, S. B., & Frisbie, D. A. (1993). *Iowa Test of Basic Skills, Form K*. Itasca, IL: Riverside.

*Hutchinson, N. L. (1993). Effects of cognitive strategy instruction on algebra problem solving of adolescents with learning disabilities. *Learning Disability Quarterly*, *16*, 34–63.

*Ives, B. (2007). Graphic organizers applied to secondary algebra instruction for students with learning disorders. *Learning Disabilities Research & Practice*, *22*, 110–118. doi:10.1111/j.1540-5826.2007.00235.x

Jitendra, A. K., Burgess, C., & Gajria, M. (2011). Cognitive strategy instruction for improving expository text comprehension of students with learning disabilities: The quality of evidence. *Exceptional Children*, *77*, 135–159. doi:10.1177/001440291107700201

*Jitendra, A. K., Dupuis, D. N., Rodriguez, M. C., Zaslofsky, A. F., Slater, S., Cozine-Corroy, K., & Church, C. (2013). A randomized controlled trial of the impact of schema-based instruction on mathematical outcomes for third-grade students with mathematics difficulties. *Elementary School Journal*, *114*, 252–276. doi:10.1086/673199

*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. *Journal of Educational Research*, *91*, 345–355. doi:10.1080/00220679809597564

Jitendra, A. K., Peterson-Brown, S., Lein, A., Zaslofsky, A., Kunkel, A., Jung, P-G., & Egan, A. (2015). Teaching mathematical word problem solving: The quality of evidence for strategy instruction priming the problem structure. *Journal of Learning Disabilities*, *48*, 51–72. doi:10.1177/0022219413487408

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

*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

Lesh, R., Post, T., & Behr, M. (1987). Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (Ed.), *Problems of representation in the teaching and learning of mathematics* (pp. 33–40). Hillsdale, NJ: Lawrence Erlbaum.

*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:10.2307/15111142

Martin, T., & Schwartz, D. L. (2005). Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts. *Cognitive Science*, *29*, 587–625.

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

McGraw-Hill. (2002). *CTB Terra Nova*. Monterrey, CA: CTB McGraw-Hill.

Misquitta, R. (2011). A review of the literature: Fraction instruction for struggling learners in mathematics. *Learning Disabilities Research & Practice*, *26*, 109–119. doi.org/10.1111/j.1540-5826.2011.00330.x

Murphy, M. M., Mazzocco, M. M. M., Hanich, L. B., & Early, M. C. (2007). Cognitive characteristics of children with mathematics learning disability (MLD) vary as a function of the cutoff criterion used to define MLD. *Journal of Learning Disabilities*, *40*, 458–478. doi:10.1177/00222194070400050901

National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.

National Governor's Association Center for Best Practices & Council of Chief State School Officers (2010). *Common Core State Standards*. Washington, DC: Author.

Ng, S. F., & Lee, K. (2009). The model method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education*, *40*, 282–313.

*Owen, R. L., & Fuchs, L. S. (2002). Mathematical problem-solving strategy instruction for third-grade students with learning disabilities. *Remedial and Special Education*, *23*, 268–278. doi:10.1177/07419325020230050201

*Peterson, S. K., Mercer, C. D., & O'Shea, L. (1988). Teaching learning disabled students place value using the concrete to abstract sequence. *Learning Disabilities Research*, *4*, 52–56.

*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*, 381–394. doi:10.1037/a0018447

*Tournaki, N., Bae, Y. S., & Kerekes, J. (2008). Rekenrek: A manipulative used to teach addition and subtraction to students with learning disabilities. *Learning Disabilities: A Contemporary Journal*, *6*, 41–59.

van Garderen, D. (2006). Spatial visualization, visual imagery, and mathematical problem solving of students with varying abilities. *Journal of Learning Disabilities*, *39*, 496–506.

van Garderen, D. (2007). Teaching students with learning disabilities to use diagrams to solve mathematical word problems. *Journal of Learning Disabilities*, *40*, 540–553.

van Garderen, D., & Montague, M. (2003). Visual-spatial representation, mathematical problem solving, and students of varying abilities. *Learning Disabilities Research & Practice*, *18*, 246–254. doi:10.1111/1540-5826.00079

*Van Luit, J. H., & Schopman, E. A. (2000). Improving early numeracy of young children with special education needs. *Remedial and Special Education*, *21*, 27–40.

*Walker, D. W., & Poteet, J. A. (1989–1990). A comparison of two methods of teaching mathematics story problem-solving with learning disabled students. *National Forum of Special Education Journal*, *1*, 44–51.

What Works Clearinghouse. (2012). *Improving mathematical problem solving in Grades 4 through 8. Educator's practice guide (NCEE 2012-4055)*. Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/pdf/practice_guides/mps_pg_052212.pdf

What Works Clearinghouse. (2014). *Procedures and standards handbook (Version 3.0)*. Washington, DC: Author.

*Wilson, C. L., & Sindelar, P. T. (1991). Direct instruction in math word problems: Students with learning disabilities. *Exceptional Children*, *57*, 512–519.

*Witzel, B. S., Mercer, C. D., & Miller, M. D. (2003). Teaching algebra to students with learning difficulties: An investigation of an explicit instruction model. *Learning Disabilities Research & Practice*, *18*, 121–131. doi:10.1111/1540-5826.00068

*Woodward, J. (2006). Developing automaticity in multiplication facts: Integrating strategy instruction with timed practice drills. *Learning Disability Quarterly*, *29*, 269–289. doi:10.2307/30035554

*Woodward, J., & Brown, C. (2006). Meeting the curricular needs of academically low-achieving students in middle grade mathematics. *The Journal of Special Education*, *40*, 151–159.

Xin, Y. P., & Jitendra, A. (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.

*Xin, Y. P., Jitendra, A., & Deatline-Buchman, A. (2005). The 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. *Journal of Educational Research*, *104*, 381–395. doi:10.1080/00220671.2010.487080

Zahner, D., & Corter, J. E. (2010). The process of probability problem solving: Use of external visual representations. *Mathematical Thinking And Learning: An International Journal*, *12*, 177–204.

Zhang, D., & Xin, Y.P. (2012). A follow-up meta-analysis for word-problem-solving interventions for students with mathematics difficulties. *Journal of Educational Research*, *105*, 303–318. doi.org/10.1080/00220671.2011.6273

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.

Zheng, X., Swanson, H. L., & Marcoulides, G. A. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, *110*, 481–498. doi:10.1016/j.jecp.2011.06.001