3-1-2021

Word-Problem-Solving Interventions for Elementary Students With Learning Disabilities: A Selective Meta-Analysis of the Literature

Jennifer E. Kong
Chapman University, jekong@chapman.edu

Christy Yan
California State University, Fullerton

Allison Serceki
Chapman University

H. Lee Swanson
University of New Mexico

Recommended Citation

Word-Problem-Solving Interventions for Elementary Students With Learning Disabilities: A Selective Meta-Analysis of the Literature

Comments

Copyright

Word Problem Solving Interventions for Elementary Students with Learning Disabilities:

A Selective Meta-analysis of the Literature

Jennifer E. Kong

Christy Yan

Allison Serceki

H. Lee Swanson

Author Note

Jennifer E. Kong, Ph.D., Attallah College of Educational Studies, Chapman University, jekong@chapman.edu; Christy Yan, Ph.D., Department of Child and Adolescent Studies, California State University, Fullerton, cyan@fullerton.edu; Allison Serceki, M.S., Attallah College of Educational Studies, Chapman University, aserceki@chapman.edu; H. Lee Swanson, Ph.D., Educational Psychology, University of New Mexico, hlswanson@unm.edu.

Correspondence concerning this article should be addressed to Jennifer E. Kong, One University Drive, Orange, CA 92866. Email: jekong@chapman.edu

Abstract

This meta-analysis assessed the effect of word problem solving interventions on the word problem solving accuracy of students identified as having a learning disability or at risk for a learning disability in kindergarten to the sixth grade. Eighteen randomized-control group designed studies met the inclusion criteria. Overall, word problem solving interventions yielded a significant positive effect on the word problem solving accuracy of students in elementary grades with LD (*ES*=1.08). Instructional components that underlie effective studies were also identified. Results suggest that peer interaction and transfer instructions yielded large effects on treatment outcomes. Results also suggested that intensive interventions (50-minute sessions, 34 total sessions) in grade 3 regardless of instructional setting yielded the largest effect sizes. These findings support the need to develop and implement quality evidence-based instruction in classroom settings (Tier 1 instruction) prior to utilizing additional resources for more intensive and individualized intervention.

Keywords: word-problem-solving intervention, elementary, at risk, math disabilities

Word Problem Solving Interventions for Elementary Students with Learning Disabilities:

A Selective Meta-analysis of the Literature

Math word problems are linguistically presented arithmetic problems that require students to construct a problem model to solve the problem (Fuchs et al., 2006; Fuchs & Fuchs, 2007). Word problems require students to use linguistic information to identify relevant information for solution accuracy, construct the appropriate number sentence, and calculate the problem accurately. Students with or at risk for learning disabilities (LD) experience considerable difficulty with word problems as they involve processes beyond basic math skills (Swanson, 2006). Additionally, students with LD perform significantly lower in math than age-equivalent peers, with the gap widening as each academic year passes (Cawley, Parmar, Foley, Salmon, & Roy, 2001).

Given the considerable difficulty with word problem solving (WPS) students with LD face, it is important to identify effective instructional practices. One approach to identifying valuable instructional practices is to conduct a synthesis of WPS intervention studies for students with LD. Meta-analysis allows for the comparison of treatment effect sizes across studies to address specific research questions in addition to examining studies by instructional variables (Glass, 1977). Two previous meta-analyses (Gersten et al., 2009; Kroesenberg & Van Luit, 2003) investigated the effect of *general* math intervention (e.g., calculation, mathematics proficiency, basic skills, problem solving strategies) to enhance the math achievement of students with mathematics difficulties. To the authors' knowledge, only four meta-analyses (Lein, Jitendra, & Harwell, 2020; Xin & Jitendra, 1999; Zhang & Xin 2012; Zheng, Flynn, & Swanson, 2013) to date have investigated specifically *word problem solving* interventions for students with learning disabilities in grades K to 12 broadly.

Xin and Jitendra (1999) investigated WPS interventions for students in elementary to postsecondary grades with "learning problems" at risk for math failure. Learning problems (LP) were defined as mild disabilities such as learning disabilities, mild mental retardation, and emotional disabilities and at risk for mathematics failure. A total of 25 intervention studies (14 group-design, 12 single-subject) were included in the study. One study included both group and single-subject design. Moderator variables that may have affected the overall treatment effect were also examined. These variables included student characteristics (IQ, grade groups – elementary, secondary, post-secondary, and classification groups – LD, mixed disabilities, and at-risk), instructional characteristics (intervention approach, length of intervention, deliverer of intervention), and methodological features (published/unpublished, group assignment). Computer-assisted instruction in group-design studies was found to be most effective, yielding a mean effect size of 1.80, followed by representation techniques ($d=1.77$), and strategy training ($d=.74$). An analysis of moderator variables revealed a low mean effect size for studies with elementary students ($d=.78$) and a large effect size for the post-secondary group ($d=1.68$), but no significant difference between the groups. The authors also presented a median PND (percentage of non-overlapping data) score to determine intervention effectiveness. Results from the single-subject design studies indicated a median PND of 89% (range 11% - 100%). That is, a median of 89% of data from the intervention phases were higher than any data point in the baseline phases. Analysis of moderator variables in these single-subject studies revealed a significant advantage for interventions teaching representation techniques. The intervention effect sizes did not differ significantly across grade groups.

As a follow-up to the meta-analysis conducted by Xin and Jitendra, Zhang and Xin (2012) included studies that were published from 1996 to 2009 in their meta-analysis of word

problem solving interventions for students with math difficulties. As in the previous review, studies that included students with "learning problems in mathematics" from kindergarten to twelfth grade were included in the meta-analysis. A distinction in this study from their previous analysis was that students who were identified as those with learning problems were operationally distinguished from students identified with a learning disability via the discrepancy model. Twenty-nine group-design studies and ten single-subject studies (39 total studies) were included in the follow-up meta-analysis. Moderator variables were analyzed in group studies only. Moderator variables included LD definition (at risk, discrepant LD), class setting (inclusive, special education), instructional features (intervention strategy, type of assessment), and types of word problems (algebraic thinking/arithmetic, real-world/simple-structured). The researchers found that word problem solving interventions had a large effect on students' performance in problem solving accuracy, with an overall effect size of 1.85. Single-subject designs yielded a PND of 95%. Moderator analyses on group studies indicated that interventions provided in inclusive settings were more effective than in special education settings. Results also indicated that while all intervention strategies (problem structure representation, cognitive strategy training, strategies involving assistive technology) produced positive effects, problem structure representation techniques yielded the highest effect sizes. Problem structure representation techniques include schema-based explicit instruction. Additionally, no significant differences were found between simple-structured problem solving and real-world problem solving. Finally, there were no significant differences between effect sizes from students diagnosed with discrepant LD and at-risk students.

Zheng, Flynn, and Swanson (2013) also conducted a selective meta-analysis of intervention studies on WPS for students with math disabilities (MD). Students were identified

as MD if participants' scores fell below the 25th percentile on a standardized math test (e.g.,

Wechsler Individual Achievement Test [WIAT], *Wide Range Achievement Test-Third Edition*

[WRAT-3]). A total of 15 studies (7 group-design, 8 single-subject) were included in the meta-

analysis. Word problem solving interventions were determined to be effective for students with

MD, yielding an effect size of 0.78 for group-design studies (compared to students with MD who

did not receive instruction). The average *ES* for single-subject studies across participants after

removing outliers Rosenthal's (1994) formula was 0.90. All studies were also coded for the

occurrence of various instructional components. Studies that significantly improved students'

WPS skills included instructional components that incorporated advanced organizers, skill

modeling, explicit practice, task difficulty control, elaboration, task reduction, questioning, and

providing strategy cues. Also, small-group instruction was found to be an effective approach for

students with MD.

In a more recent study, Lein, Jitendra, and Harwell (2020), reviewed current studies on

WPS interventions for K-12 students with learning disabilities and math difficulties. LD was

defined as students who were identified based on a discrepancy model or through the school

district evaluation (non-responsiveness to intervention). Math difficulties were defined as

students who scored at or below the 35th percentile on a standardized mathematics test. A total of

31 group-design studies were included in this meta-analysis, which found that word problem

solving interventions yielded a moderate mean effect size (g=0.56). This study found that there

was no significant difference in the magnitude of effect sizes by LD and at risk status.

Additionally, interventions for students in elementary grades (g=0.63) were found to yield higher

effect sizes than those in secondary grades (g=0.33). Finally, this study investigated intervention

models as a moderating variable of overall effects. The results indicated that interventions that

included a schema broadening and transfer instruction model yielded highest effect sizes ($g=1.06$).

The present meta-analysis extends upon the previous reviews in several ways. All prior meta-analyses examined a broad range of ages, from kindergarten to postsecondary. This may be problematic for making instructional recommendations for a specific age or grade group. Xin and Jitendra's meta-analysis included 11 studies with participants in kindergarten to grade 6 (5 group-design, 6 single-subject) out of 25 studies. Xin and Jitendra reported large effect sizes for post-secondary students ($d = 1.68$), moderate effects for secondary students ($d = .78$), and low effect size for the elementary students ($d = .47$). An analysis of specific grades as a significant moderator of treatment outcomes was not conducted. Zhang and Xin's study included 18 elementary experimental studies (15 group-design, 3 single-subject) out of a total of 39 studies. Zheng and colleagues' meta-analysis included 10 elementary studies (5 group-design, 5 single-subject) out of 15 total studies. Lein et al.'s study included 5 studies conducted on elementary students and 6 with secondary studies. Further, the meta-analyses conducted by Zhang and Xin (2012) and Zheng et al. (2013) did not report information on how intervention effects may have diverged for elementary and secondary students. Lein et al. (2020) also did not investigate specific instructional components within interventions, but rather utilized general schema and models of intervention. Thus, specific conclusions about treatment effects of word problem solving interventions and instructional recommendations drawn from these meta-analyses may not be appropriate for elementary students specifically.

Additionally, as evidenced by the studies above, variability exists in how students are identified with LD in research and practice. Earlier studies have included students with "learning problems" more broadly, including students with learning disabilities (Xin & Jitendra, 1999;

Zhang & Xin, 2012) or have relied on a model that includes a discrepancy between IQ and

achievement (Discrepancy model; e.g., Hallahan, Pullen, & Ward, 2013). Researchers have also

utilized the term "at risk for LD" to identify children who may be at risk for academic failure and

benefit from intervention, but have not yet been identified as LD. For example, performance

below the 25th percentile cut-off score on standardized measures has been commonly used to

identify children at risk (e.g., Fletcher et al., 1989; Siegel & Ryan, 1989; Swanson, Lussier, &

Orosco, 2013). With the growing need to deliver interventions to students who are most at risk as

early as possible, it will be important to clarify the role definitions play in instructional

outcomes. Specifically, the present meta-analysis will examine interventional components such

as the intensity of intervention, setting, and specific instructional features in relation to student

characteristics (e.g., grade, LD identification).

Finally, earlier studies have not investigated the possible moderating effect of

interventions for students who are English learners (ELs). ELs, in particular, may experience

more difficulty in comparison to monolingual children with math problem solving because of the

need to preserve information while at the same time processing information in a second language

(e.g., Swanson, Kong, & Petcu, 2019). The National Center for Education Statistics (NCES,

2019) reports that 41% of English learners score below basic in mathematics, compared to 16%

of their non-EL peers scoring below basic. Given that ELs are a rapidly increasing demographic

in U.S. public schools, research to identify effective instructional strategies for problem solving

is critical.

The present meta-analysis will focus on group-designed intervention studies conducted

with elementary-aged (K-6) students in an attempt to make more detailed recommendations for

effective interventions for this age group. The current study will add to the current research by

including samples of students who have been identified as LD via the discrepancy model and "at

risk" for math disabilities and investigating the effects of specific instructional components

within interventions rather than global procedures on experimental studies conducted with

elementary participants.

This study will address the following three research questions:

1. Are word problem solving interventions effective for kindergarten to sixth-grade students

    with LD? Effective outcomes will be based on the magnitude of the ESs. The average

    ESs among the group designed studies in the previous syntheses was 1.18 for students in

    grades 1 to 12.

2. Do specific effect sizes in word problem solving interventions vary as a function of

    moderator variables such as participant characteristics (EL status, LD definition, grade

    level)? Some of the previous syntheses have not reported the impact of sample

    characteristics on treatment outcomes, and therefore generalizations to children with

    specific learning difficulties is unclear. This meta-analysis attempts to characterize the

    sample found to benefit from problem solving interventions.

3. Do effect sizes in word problem solving interventions vary as a function of specific

    instructional components? Previous synthesis have found that general instructional

    approaches, such as computer assisted instruction, problem structure representation (i.e.,

    schema-based explicit instruction), and instructional scaffolding (organizers, modeling,

    task reduction) contributed to significant improvements in students' performance. These

    studies have focused on an array of children with learning problems (LD, mild mental

    retardation, emotional disabilities) in grades K – 12. This study extends the literature by

    identifying instructional components of word problem solving interventions that are

directed to elementary aged (K – 6) students identified as having a learning disability or at risk for a specific learning disability in math.

Method

Data Collection

The PsycINFO, Science Direct, and ERIC online databases were systematically scanned for studies from 1990 to 2019 that met the inclusion criteria. Search terms describing word problem solving (*word problem solving instruction* or *word problem solving intervention* or *problem solving instruction* or *story problem* or *math intervention*), the population (*special education* or *learning disabled* or *learning disabilit** or *at risk for math difficulty*), word-problem-solving outcomes were combined with these keywords: *elementary school, efficacy, strategy instruction, schema-based instruction, scaffolded instruction,* and *peer interaction.* This initial search generated approximately 1,592. Of these, 239 studies were selected for further review based on title and abstract review. The reference lists of prior meta-analyses (e.g., Gersten et al., 2009; Kroesenbergen & Van Luit, 2003; Xin & Jitendra, 1999; Zhang & Xin, 2012; Zheng et al., 2013) were also systematically scanned.

Study Eligibility Criteria

To be eligible for this analysis, each study had to meet the following criteria: (a) included students with or at risk for learning disabilities in grades K to 6; (b) tested an intervention to improve word-problem-solving; (c) assessed students' word-problem-solving accuracy (measure included normed or experimental/researcher developed measures); (d) involved an experimental design with randomization, quasi-experiment with pre- and post-test data, or a within-subjects design (i.e., all students participated in both the treatment and comparison conditions); (e) provided data to permit the calculation of effect sizes and average weighted *ES*s; and (f) was

published in English. Studies investigating the effectiveness of instruction or improving only

math calculation were not included. This procedure narrowed the search to 33 documents, 18 of

which met inclusion criteria. Some studies had more than one WPS intervention, so 113 different

*ES*s were calculated.

Interrater agreement. Two graduate students independently coded 22% of the articles for

inclusion criteria and coding accuracy. Inter-rater agreement was calculated as a number of

agreements divided by the number of agreements plus disagreements multiplied by 100. The

mean interrater agreement for article inclusion was above 95%. The mean interrater agreement

for coding of the twelve instructional components outlined below was also above 95%.

Coding of Study Features

The general categories of coding for each study included: (a) year of publication, (b)

sample characteristics (gender, grade, disability or risk, EL status), (c) intervention

characteristics (number of sessions, number of minutes, group size, who delivered the

instruction), and (d) components of instruction.

Categorization of Treatment Variables. Each study was coded on the occurrence or

nonoccurrence of the following instructional components. These instructional components have

been linked to academic outcomes in earlier meta-analyses that have included students with

learning disabilities (Dennis et al., 2016; Swanson & Hoskyn, 1998; Zheng et al., 2013). The

instructional components coded are as follows:

1. Explicit instruction – statements in the treatment description included characteristics of

    explicit direct instruction (e.g., teacher/researcher directed instruction, administering

    probes).

2. Technology – statements in the treatment description about the use of technology tools such as computers, tablets, or other media to supplement or provide instruction.

3. Strategy cues – statements in the treatment description about using strategies, multi-step procedures, verbalizations of procedures, metacognitive strategies, questioning, and think-alouds by teacher/researcher.

4. Peer interaction – statements in the treatment description about using peer interaction to complete activities to present, model, practice, or review instruction.

5. Instructional feedback – statements in the treatment description about providing participants with frequent instructional feedback and correction.

6. Visual aids – statements in the treatment description about the use of graphics, charts, diagrams, illustrations, visual aids, semantic mapping, or pictorial representations to supplement instruction.

7. Foundational skills – statements in the treatment description about providing participants with instruction and practice in foundational skills such as computation and fact fluency.

8. Schema instruction – statements in the treatment description about providing participants with explicit instruction of underlying structures of the word problem type, basic schema for problem type, and solving specific problem types.

9. Instruction to transfer – statements in the treatment descriptions about explicit instruction to transfer or generalize skills on novel problems.

10. Manipulatives – statements in the treatment descriptions about providing students with concrete materials, manipulatives, or other hands-on materials.

11. Behavioral reinforcement – statements in the treatment description about providing participants with praise, token economy, or reinforcement schedules.

12. Self-regulated learning – statements in the treatment description about students setting

goals for their performance, self-monitoring, or self-evaluation.

Data Analysis

Effect Size Calculation. Effect Sizes (*ES*s) were calculated utilizing pretest and posttest

means and standard deviations. Hedges' *g* was the measure of *ES* for this study, calculated as the

difference between pretest-posttest means for the treatment group and the pretest-posttest means

for the comparison group. This difference score was then divided by the pooled within-group

standard deviation of posttest scores. Hedges' *g* was calculated as

$$\frac{\left(X_{post_1} - X_{pre_1}\right) - \left(X_{post_2} - X_{pre_2}\right)}{\sqrt{\left([n_1-1]s_1^2 + [n_2-1]s_2^2 / [n_1+n_2-2]\right)}},$$

where X_{pre_1} and X_{pre_2} were unadjusted pretest means, X_{post_1} and X_{post_2} were unadjusted posttest

means, n_1 and n_2 were sample sizes, and s_1 and s_2 were unadjusted standard deviations for the

treatment and comparison groups, respectively.

Several planned tests that compared effect sizes as a function of intervention

characteristics and grade levels utilizing a general linear model procedure were computed

(Borenstein et al., 2009). Because of the variance between and within studies, the PROC Mixed

(SAS, 2014) procedure was used to determine effect sizes as a function of instructional

components. For this mixed analysis, the grand mean centered variable of " grade" was used as a

covariate in the analysis. Due to the small sample in these comparison, we employed a restricted

maximum likelihood estimation (REML) with a Bonferroni correction (McNeish, 2017).

<div align="center">

Results

</div>

Question 1: Are word problem solving interventions effective for kindergarten to sixth-

grade students with LD?

To answer Research Question 1, a single weighted *ES* for all 18 studies was calculated. To determine if specific sample characteristics related to excess variability in *ES*s, general linear models categorizing between-class effects were analyzed. Grade, LD definition, and intervention characteristics, such as who delivered the instruction, group size, number of sessions, type of measure, and number of minutes (intensity of intervention), were examined for contributing excess variability in *ES*.

Table 1 provides a descriptive summary of the studies included in this synthesis. The total *n* refers to the total number of students who were included in the studies. The LD (Learning Disability) *n* is the number of LD students who received treatment. The EL (English Learner) *n* is the number of LD students who were identified as English learners. Table 1 also displays whether LD students were identified by the discrepancy model or considered at risk for LD (below specified cutoff score), grade level, and type of research design.

All studies included in this synthesis were published in peer-reviewed journals, with publication dates ranging from 1998 to 2014. Fourteen of the 18 studies focused on only third grade students. Participants' grade levels ranged from 2 to 5. Eight studies included students designated as LD by discrepancy (e.g., Fuchs et al., 2003b), while the other ten studies included at-risk students (e.g., Moran, Swanson, Gerber & Fung, 2014).

Intervention

The number of intervention sessions ranged from 4 (Owen & Fuchs, 2002) to 60 (Jitendra et al., 2013a). The length of each session varied from 20 to approximately 140 minutes (Fuchs et al. 2003a). One study did not report the length of each intervention session (Owen & Fuchs, 2002). Eight studies reported administering the intervention in a whole group, general education setting. An intervention was conducted in a small group setting in 8 studies, and 1 study

individually. Two studies reported multiple interventions with both whole group and small group instruction. General education teachers delivered the intervention in 5 studies, and research assistants/graduate students were responsible for delivering instruction in 8 studies. Two studies reported administration of the intervention by researchers and a community member delivered one. Two studies reported multiple deliverers of intervention: 1 study included an instructional assistant and parent, while another utilized a teacher and graduate student. Finally, 12 studies used researcher-developed measures to assess word problem solving accuracy, and 2 of the 18 studies used norm-referenced tests on pretests, posttests, and transfer tests. Four studies utilized both researcher-developed and norm-referenced measures.

Overall, word problem solving interventions had a positive effect on word problem solving accuracy across all studies, Hedge's g = 1.08 (*K*=113, 95% CI of .79 to 1.37). According to Cohen's (1988) criterion, this is a large effect size. A homogeneity statistic Q was computed to determine whether studies shared a common ES. The statistic Q has a distribution similar to the distribution of Chi-square with *k*-1 degrees of freedom, where *k* is the number of ESs. As expected, there was significant heterogeneity in the findings, Q (*df*=112)= 2036.78, p < .001. Because homogeneity was not achieved (which is usually the case), the variability of the ES as a function of moderator variables were analyzed. The results are shown in Table 2. Because the commonly reported Q statistic has been criticized, the I^2 statistic (Higgins & Thompson, 2002) was computed, using the following formula:

$$I^2 = \frac{Q - (k\text{-}1)}{Q}$$

I^2 indices of 25%, 50%, and 75% are classified as low, medium, and high heterogeneity, respectively (e.g., Higgins. & Thomspon, 2002), . The I^2 statistic was .95, suggesting an extremely high percentage of variability across the majority of measures.

Moderator variables. Table 2 shows the Hedge's g mean effect sizes and 95%

confidence intervals for the moderator variables. There were several significant differences when

comparisons were made within the various moderator variables. For example, there were

significant differences in weighted effect sizes (Hedges g weighted by the reciprocal of the

sampling variance) by the number of reported minutes per session, Q_B(df=9)=863.10, $p < .001$.

The Q_B statistic is the weighted between-categories sum of squares of an ANOVA. Fifty-minute

sessions produced the largest effect size relative to the other conditions whereas 25-minute

sessions produced the smallest effect size. There were also significant differences in weighted

ESs as a function of the number of sessions, Q_B(df 9)=79.28, $p < .05$. Interventions with 34

sessions yielded the largest effect size.

There were significant differences in effect sizes by type of measure used

Q_B(df=1)=244.54, $p< .05$. Effect sizes of researcher-developed measures were significantly

larger than that of norm-referenced measures. Further, there were significant differences in the

mean effects by grouping of students in the intervention, Q_B(df=2)=111.18, $p < .05$ and the

deliverer of intervention, Q_B(df=5)=229.57, $p < .05$. Interventions that were delivered in whole

groups (*M*=2.85) and small groups (*M*=2.39) yielded higher effect sizes than interventions

delivered in individual settings (*M*=0.92). Finally, interventions delivered by the classroom

teacher (*M*=2.11) and university/graduate students (*M*=2.80) produced higher effect sizes when

compared to researchers, instructional assistants, parent, or community higher (*M*=0.50, 0.02,

0.07, 036, respectively).

In summary, word problem solving interventions had a positive and large effect for

students who are in grades 2-5. Effect sizes for interventions that were delivered across 34

sessions yielded the highest effect size. Additionally, interventions delivered in small or whole

class instruction by classroom teachers or graduate students yielded the highest effect sizes.

Question 2: Do specific effect sizes in word problem solving interventions vary as a function of participant characteristics (EL status, LD definition, grade level)

To answer Research Question 2, a meta-regression analysis (Borenstein et al., 2009) was

conducted on the moderator variables related to the sample description (LD definition, EL status,

grade level) to determine if three moderators accounted for excess variability in ESs.

There were significant differences in weighted ESs as a function of LD definition, Q_B(df

1)=183.97, p <.001. Mean effect sizes for students at risk for LD (M=1.35) were higher than for

students who were identified as LD through the school district (M=0.74). There were also

significant differences in effect sizes by ratio of students who were English learners

Q_B(df=1)=373.03, p< .001. Effect sizes for interventions that included a higher ratio of students

who were English learners reported higher mean effects (M=1.40) than studies that did not

include students who were English learners (M=0.77). Finally, there were differences in the

weighted ESs as a function of grade, Q_B(df=3)=70.38, p<.001. The majority of the effect sizes

that were computed in this review were for students in third grade (82 effect sizes). Effect sizes

for interventions taught to third grade students reported highest mean effects (M=2.71). The

mean effect sizes, Hedge's g ES, Q and I^2 statistics, and the 95% confidence intervals as a

function of the moderator variables are shown in Table 2.

In summary, effect sizes of word problem solving interventions were highest for students

who are defined as "at risk" and in grade 3. Also, the mean ES of interventions that included

students who are English learners was higher than interventions that did not include or did not

report inclusion of English learners. It is important to note that only 11 out of 18 (61%) studies

included this demographic information, so more research in this area may need to be conducted.

Question 3: Which specific word problem solving interventions/components of WPS

intervention are effective with kindergarten to sixth-grade students with LD?

To answer Research Question 3, a multi-level random effect analysis of covariance was

conducted to determine if significant effects in weighted *ES*s existed between studies that

included instructional components and those that did not (McNeish, 2017). Mean centered grade

was utilized as a covariate in the analysis. A multi-level ANCOVA model included a random

effects variance within and between studies.

Table 3 displays a summary of the occurrence of instructional components in each study

and mean effect sizes for each study. All studies included explicit instruction and strategy cues

as instructional component. Fourteen of the 18 (78%) of the interventions included visual aids,

while seventy-two percent (13 out of 18) of the studies included schema instruction. Twelve out

of 18 studies (67%) included self-regulated learning in descriptions of interventions. Sixty-one

percent of the studies included descriptions of instructional feedback and instruction to transfer.

Peer interaction was reported in 50 percent of the studies. Instruction and practice in

foundational skills such as computation and fact fluency was reported in 39% of the studies.

Twenty-eight percent of the studies included concrete math materials and manipulatives.

Behavior reinforcements were reported in 17 percent of the studies. Finally, only one of the

studies included technology tools in the study.

Table 4 shows the fixed effects of studies that included and did not include each

instructional component. A Bonferroni correction for multiple comparisons was utilized to

determine significance (*p*=.004). A multi-level ANCOVA revealed that studies that included

instructional component 4 - peer interaction ($F_{1,64}$ = 13.50, p=0.0005) and instructional

component 9 - instruction to transfer($F_{1,64}$ = 10.11, p=0.002) yielded significant contrasts when

compared to studies that did not included these components. Studies that included descriptions of

peer interaction in the intervention (M=1.70) yielded significantly higher effect sizes than studies

that did not include peer interaction (M=0.24). Finally, studies that included descriptions of

instruction to transfer (M=1.61) yielded higher effect sizes when compared to studies that did not

include transfer instruction (M=0.42).

Discussion

The purpose of this meta-analysis was to determine if word problem solving interventions

are effective for improving word problem solving accuracy in students with LD in elementary

grades and if so, determine if effect sizes vary as a function of participant and/or instructional

components. Three important finding emerged. First, problem solving interventions had a

positive effect on word problem solving accuracy overall. These results were qualified in that the

largest effect sizes occurred in intensive interventions (50 minute sessions and 34 total sessions).

Second, effect sizes for students at risk for LD were higher than for students who were identified

as LD through the school district. Effect sizes for interventions that included a higher ratio of

students who were English learners yielded higher mean effects than studies that did not include

students who were English learners. Finally, peer interaction and transfer instructions yielded

large effects on treatment outcomes relative to the other conditions.

We will now address the three questions that directed this study.

Question 1: Are word problem solving interventions effective for kindergarten to sixth-grade students with LD?

Generally, word problem solving interventions were effective for students with LD in elementary grades, resulting in a weighted *ES* of 1.08 across 18 studies. Two previous studies (Lein et al., 2020; Xin & Jitendra, 1999) that included students in grades K-12 research has reported divergent effect sizes for elementary grades (*g*=0.63 and *d* = .47, respectively). These studies included 11 and 12 studies for elementary aged students in their respective meta-analyses. The current study suggests that recent research in elementary grades have shown that word problem solving interventions are highly effective for students with LD. Additionally, the results indicated that 50-minute sessions and 34 total sessions yielded the highest effect sizes when compared to other reported time durations. Intervention effects were highest in small and whole group instruction (compared to individual instruction). Additionally, interventions delivered by the classroom teacher and university students yielded highest effect sizes. These results should be interpreted with caution however, as the majority of participants were in third grade and a large number of studies utilized researcher-developed measures.

This finding is consistent with previous research (e.g., Gersten et al., 2009; Zheng, Flynn, & Swanson, 2012) that has found that intensive interventions are effective for students with learning disabilities. While the results indicated that interventions that included 34 total sessions and 50-minute sessions yielded the highest effects, these figures are not prescriptive, per se. What this seems to reflect is the sentiment that intensive interventions are effective for students with LD. Gersten and colleagues (2009) found a negative correlation between the number of treatment sessions in general math instruction and effect size, but did not specify the number of sessions. However, these studies may not be directly comparable as the effects of general math instruction and problem solving intervention may differ.

This current study did not find that there was a significant difference in effect sizes for interventions administered in smaller groups or whole class inclusive settings, though either of these settings yielded higher effects that individual instruction. Additionally, the results of this study revealed that effects of interventions delivered by classroom teachers and university students yielded similarly high effects. In previous research (Carlberg & Kavale, 1980; Zhang & Xin, 2012), the issue of administering interventions in special education settings (small group) or inclusive classroom settings (whole class) has been debated. The results of the meta-analysis reveal that for word problem solving interventions specifically, either of these particular settings did not appear superior in terms of yielding higher effect sizes. We speculate that it is possible that the severity of students' disabilities may differ in various instructional settings in schools, with students with more severe needs requiring more intensive interventions (to be discussed below under Question 2). However, these findings support the importance of providing quality evidence-based instruction in Tier 1 general class instruction before the need for intensive interventions in smaller groups is needed. Word problem solving interventions delivered in general class instruction may have great potential for students with learning disabilities and students at risk alike, bolstering the need for quality Tier 1 instruction.

Of the 18 studies included in this study, 12 studies utilized researcher-developed tests, 2 used standardized assessments, and 4 used both. Results indicated that effect sizes on researcher-developed measures were significantly higher than standardized measures, which seems consistent with previous research that have indicated the possibility of alignment of the intervention materials and researcher-developed probes, which mirrors curriculum-based measures that are more sensitive to changes (Zhang & Xin, 2012). This finding, however, is particularly of interest for teachers of students with LD who may be receiving special education

services in schools. This finding affirms the importance of utilizing curriculum-based measures

to monitor progress and evaluate intervention effectiveness for specific skills that are taught in

the classroom.

Question 2: Do specific effect sizes in word problem solving interventions vary as a function

of moderator variables such as participant characteristics (EL status, LD definition, grade

level)?

Eight studies included descriptions of students who were identified as LD via the

discrepancy model and/or through the school district. These studies ranged from 1991 to 2009.

With more recent efforts to address limitations to the discrepancy model of identifying children

with LD, the Response to Intervention (RtI) model has been recommended (IDEA, 2004).

Students who are "at risk" for LD, or achieving below a designated cutoff point (e.g., 25th

percentile) would be eligible to receive intervention to begin to remediate any existing

achievement gaps. Studies that included students "at risk" ranged from 2003 to 2014. The results

of this study indicated word problem solving interventions were more effective for students at

risk for LD than for students identified as LD through a discrepancy model. As mentioned

earlier, it possible that students who are diagnosed as LD via the discrepancy model may have

more extensive needs. However, these findings seem to support the notion that the RtI model

might be a start to differentiate between students who respond to intervention and were merely at

risk for a math disability, and those who do not and may require more intensive support, all the

while providing much-needed instruction to low achieving students (Fuchs et al., 2003).

One of the areas that is particularly difficult for EL students is solving math word

problems (Bumgarner et al., 2013; Powell et al., 2020). The results of this study indicated that

studies that included students who were English learners yielded higher effects than ones that did

not. This supports the emerging research that demonstrates that problem solving interventions

are highly effective for elementary students who are ELs (Gersten & Baker, 2000; Kong and

Swanson, 2019; Orosco, Swanson, & O'Connor, 2011; Swanson, Kong, Moran, & Orosco;

2019). However, these findings should be interpreted with caution, as some studies may have

included participants who were ELs, but were not reported as such in the studies we reviewed. It

is also worth noting the small percentage of students that were reported as ELs in the studies

included in this meta-analysis (5.06%) compared to national averages (9.6% nationally in 2016;

U.S. Department of Education, 2019). Previous meta-analyses on the effects of word problem

solving for K-12 students with LD (Xin & Jitendra, 1999; Zhang & Xin, 2012; Zheng, Flynn, &

Swanson; 2013) have not included EL status as a moderating variable.

Finally, a majority of the studies included participants in the third grade. Effect sizes of

interventions provided for third grade students were significantly higher than the effect sizes for

second, fourth, and fifth grade students. No studies of word problem solving interventions for

young children (K-1) were included in this study. Future studies could investigate story problem

interventions for young children, as well as the continued effectiveness of problem solving

interventions for older elementary students.

Question 3: Which specific word problem solving interventions/components of WPS

intervention are effective with kindergarten to sixth-grade students with LD?

The results indicated that descriptions of peer interaction was reported in 9 out of the 18

studies included in this meta-analysis. Of those 9 studies, 5 included descriptions of students

identified as LD via the discrepancy model, and 4 studies included students at risk. These studies

that included descriptions of peer interaction (to present, model, practice, or review instruction)

in the intervention yielded higher effect sizes than studies that did not include peer interaction.

This indicated that word problem solving interventions for students with LD should include

opportunities for students with LD to collaborate and interact with more skilled peers. This does

not seem to support the existing literature on peer-assisted learning in math for students with LD

(Gersten et al., 2009). Gersten and colleagues' meta-analysis on math instruction for students

with LD found that while studies that included cross-age tutoring yielded high effect sizes,

studies that included peer-assisted learning or peer interaction within the class did not yield high

effect sizes ($g=0.14$). One point to consider, however, is that this previous analysis included

studies in all math interventions broadly and across all grade levels (K-12) and not word problem

solving in elementary-aged students specifically. Further analysis on the moderating effects of

grade and word problem solving interventions specifically were not considered. It may be

possible that interventions of word problem solving that include peer interaction and

mathematical discourse may be better suited for elementary grades, or for word problem solving

specifically. Learning via peer interaction is consistent with the social development theory

(Vygotsky, 1978), in which children acquire knowledge through social and verbal experiences

from a more knowledgeable individual. As suggested by Gersten and colleagues (2009), when

provided explicit and structured guidelines and moderated by teachers, elementary-aged students

with LD may perhaps be able to learn new word problem-solving skills from interaction with

their peers.

Additionally, students with LD in elementary grades benefitted from explicit instruction

to transfer learned skills to novel problems. This finding supports the existing literature on

instructional components that improve students' WPS skills (Griffin, Case, & Siegler, 1994;

National Research Council, 2001; Zheng, Flynn, & Swanson, 2013). Similar to other academic

skills, it is important for young students to transfer knowledge of skills to novel situations. Word

problem solving may be a crucial medium to select and apply strategies to solve everyday problems.

Limitations

Although this synthesis provided information about students with LD in the elementary grades, the findings should be interpreted with caution. First, the criteria for determining at risk students varied across studies. Though we did attempt to categorize studies based on how students were identified, criteria differed even within those categories. Second, only group studies published in peer-reviewed articles were included, excluding unpublished work, dissertations, and single-subject designs. These selection processes reduce generalization of our findings. Finally, a majority of the studies included in this meta-analysis included participants in third grade, which may limit the generalization of these findings.

Implications for Practice

The present meta-analysis found that word problem solving interventions, specifically those that include peer interaction and explicit instruction to transfer learned skills to novel problems are effective for elementary students with LD. Elementary students with LD or at risk for LD may benefit from word problem solving interventions with opportunities to use language and interact with peers and instructors to transfer skills or schema to new problems. The results of this review suggest that these instructional components are more effective for students who are at risk for LD. Additionally, students may also benefit from intensive intervention regardless of the instructional setting. This supports the significance of delivering evidence-based instruction in the general classroom (Tier 1 instruction) before resources for small group instruction are utilized.

More research is needed to identify effective components of instruction for students in elementary school who are at risk for or identified as LD. Particularly, research should be conducted with students in primary grades (K-2), to identifying possible precursors for word problem solving difficulty and early interventions. Additionally, future studies should consider learner characteristics, particularly for those who are most at risk (ELs, low SES, LD).

References

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. doi:10.1002/9780470743386

Bumgarner, E., Martin, A., & Brooks-Gunn, J. (2013). Approaches to learning and Hispanic children's math scores: The moderating role of English proficiency. *Hispanic Journal of Behavioral Sciences, 35*(2), 241-259. doi:/10.1177/0739986312473580

Cawley, J., Parmar, R., Foley, T. E., Salmon, S., & Roy, S. (2001). Arithmetic performance of students: Implications for standards and programming. *Exceptional Children*, *67*, 311-328. https://doi.org/10.1177/001440290106700302

Cohen, J. (1988). Statistical power analysis for the behavioral sciences, (2nd ed.). Erlbaum.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155-159.

Fletcher, J. M., Epsy, K. A., Francis, P. J., Davidson, K. C., Rourke, B. P., & Shaywitz, S. E. (1989). Comparison of cutoff and regression-based definitions of reading disabilities. *Journal of Learning Disabilities, 22,* 334-338. httsp://doi.org/10.1177/002221948902200603

Fuchs, L. S. & Fuchs, D. (2007) Mathematical problem solving. In D. B. Berch & M.M. Mazzocco (Eds.) *Why is math so hard for some children? The nature and origins of mathematical learning difficulties and disabilities* (397-414). Paul H. Brookes Publishing Co.

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

*Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C.

    (2008a). Effects of small-group tutoring with and without validated classroom instruction

    on at-risk students' math problem solving: Are two tiers of prevention better than one?

    Journal of Educational Psychology, *100*(3), 491. https://doi.org/10.1037/0022-

    0663.100.3.491

*Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., & Hamlett, C. L. (2004a). Expanding schema-

    based transfer instruction to help third graders solve real-life mathematical

    problems. *American Educational Research Journal*, *41*(2), 419-445.

    https://doi.org/10.3102/00028312041002419

*Fuchs, L. S., Fuchs, D., Hamlett, C. L., & Appleton, A. C. (2002). Explicitly teaching for

    transfer: Effects on the mathematical problem-solving performance of students with

    mathematics disabilities. *Learning Disabilities Research & Practice*, *17*, 90-106.

    https://doi.org/10.1111/1540-5826.00036

*Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to mathematical problem-

    solving instruction comparing students at risk of mathematics disability with and without

    risk of reading disability. *Journal of Learning Disabilities*, *37*, 293-306.

    https://doi.org/10.1177/00222194040370040201

*Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., Hosp, M., & Jancek,

    D. (2003a). Explicitly teaching for transfer: Effects on third-grade students' mathematical

    problem solving. *Journal of Educational Psychology*, *95*(2), 293.

    https://doi.org/10.1037/0022-0663.95.2.293

*Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K.

    (2003b). Enhancing third-grade students' mathematical problem solving with self-

regulated learning strategies. *Journal of Educational Psychology*, *95*, 306.

https://doi.org/10.1037/0022-0663.95.2.306

*Fuchs, L. S., Fuchs, D., Prentice, K., Hamlett, C. L., Finelli, R., & Courey, S. J. (2004b).

Enhancing mathematical problem solving among third-grade students with schema-based

instruction. *Journal of Educational Psychology*, *96*(4), 635. https://doi.org/10.1037/0022-

0663.96.4.635

Fuchs, D., Mock, D., Morgan, P. L., & Young, C. L. (2003) Responsiveness-to-intervention:

Definitions, evidence, and implications for the learning disabilities construct. *Learning

Disabilities Research and Practice, 18,*157–171.

*Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M.

(2008b). Effects of preventative tutoring on the mathematical problem solving of third-

grade students with math and reading difficulties. *Exceptional Children*, *74*, 155-173.

https://doi.org/10.1177/001440290807400202

Gersten, R., & Baker, S. (2000). What we know about effective instructional practices for

English-language learners. *Exceptional Children*, *66*(4), 454-470.

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

Glass, G. V. (1977). Integrating findings: The meta-analysis of research. *Review of Research in

Education*, 351-379.

Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual

prerequisites for first formal learning of arithmetic to students at risk for school failure. In

K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 24–49). MIT Press.

*Griffin, C. C., & Jitendra, A. K. (2009). Word problem-solving instruction in inclusive third-grade mathematics classrooms. *The Journal of Educational Research*, *102*(3), 187-202. https://doi.org/10.3200/JOER.102.3.187-202

Higgins, J. P. & Thomspon, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539-1558. https://doi.org/10.1002/sim.1186

Individuals with Disabilities Education Act of 2004, 20 U. S. C. § 1400 et seq. (2004).

*Jitendra, A. K., Dupuis, D. N., Rodriguez, M. C., Zaslofsky, A. F., Slater, S., Cozine-Corroy, K., & Church, C. (2013a). A randomized controlled trial of the impact of schema-based instruction on mathematical outcomes for third-grade students with mathematics difficulties. *The Elementary School Journal*, *114*(2), 252-276. https://doi.org/10.1086/673199

*Jitendra, A. K., Griffin, C. C., Haria, P., Leh, J., Adams, A., & Kaduvettoor, A. (2007). A comparison of single and multiple strategy instruction on third-grade students' mathematical problem solving. *Journal of Educational Psychology, 99*(1), 115–127. https://doi.org/10.1037/0022-0663.99.1.115

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

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 mathematic disabilities. *Learning*

    Disability Quarterly, *42*(2), 92-104.

Kroesenberg, 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. doi:10.1080/00273171.2017.1344538

*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

National Research Council (2001). Adding it up: Helping children learn mathematics. J.

    Kilpatrick, J. Swafford and B. Findell (Eds.). Mathematics Learning Study Committee,

Center for Education, Division of Behavioral and Social Sciences and Education. National Academy Press.

Orosco, M. J., Swanson, H. L., O'Connor, R., & Lussier, C. (2013). The effects of dynamic strategic math on English language learners' word problem solving. *The Journal of Special Education*, *47*(2), 96-107.

*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. https://doi.org/10.1177/07419325020230050201

Powell, S. R., Berry, K. A., & Tran, L. M. (2020). Performance differences on a measure of mathematics vocabulary for English learners and non-English learners with and without mathematics difficulty. *Reading & Writing Quarterly*, *36*(2), 124-141. https://doi.org/10.1080/10573569.2019.1677538

Rosenthal, R. (1994). Parametric measures of effect size. In H. Cooper & L. V. Hedges (Eds.), *The Handbook of Research Synthesis* (pp. 231-244). Sage.

Siegel, L. S., & Ryan, E. B. (1989). The development of working memory in normally achieving and subtypes of learning disabled children. *Child Development, 60*, 973-980. htts://doi/org/10.2307/1131037

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, *98*, 265. https://doi.org/10.1037/0022-0663.98.2.265

Swanson, H. L., & Hoskyn, M. (1998). Experimental intervention research on students with learning disabilities: A meta-analysis of treatment outcomes. *Review of Educational Research*, *68*(3), 277-321.

Swanson, H. L., Kong, J. E., Moran, A. S., & Orosco, M. J. (2019). Paraphrasing Interventions

    and Problem-Solving Accuracy: Do Generative Procedures Help English Language

    Learners with Math Difficulties? *Learning Disabilities Research & Practice*, *34*(2), 68-

    84.

Swanson, H. L., Kong, J., & Petcu, S. D. (2019). Growth in Math Computation among

    Monolingual and English Language Learners: Does the Executive System have a Role?

    Developmental neuropsychology, *44*(8), 566-593.

    https://doi.org/10.1080/87565641.2019.1688328

Swanson, H. L., Lussier, C., & Orosco, M. (2013). Effects of cognitive strategy interventions and

    cognitive moderators on word problem solving in children at risk for problem solving

    difficulties. *Learning Disabilities Research & Practice*, *28*, 170-183.

    https://doi.org/10.1111/ldrp.12019

*Swanson, H. L., Moran, A., Lussier, C., & Fung, W. (2014). The effect of explicit and direct

    generative strategy training and working memory on word problem-solving accuracy in

    children at risk for math difficulties. *Learning Disability Quarterly, 37*(2), 111-123.

    https://doi.org/10.1177/0731948713507264

Swanson, H. L., Orosco, M. J., & Lussier, C. M. (2014). The effects of mathematics strategy

    instruction for children with serious problem-solving difficulties. *Exceptional Children*,

    80, 149-168. https://doi.org/10.1177/001440291408000202

Swanson, H. L., & Sachse-Lee, C. (2000). A meta-analysis of single-subject-design intervention

    research for students with LD. *Journal of Learning Disabilities*, *33*(2), 114-136.

    https://doi.org/10.1177/002221940003300201

U.S. Department of Education (2019). Institute of Education Sciences, National Center for

     Education Statistics, National Assessment of Educational Progress (NAEP) 2019.

van Garderen, D. (2007). Teaching students with LD to use diagrams to solve mathematical

     word problems. *Journal of Learning Disabilities*, *40*, 540-553.

     https://doi.org/10.1177/00222194070400060501

Vygotsky, L.S. (1978). *Mind in society*. Harvard University Press.

*Wilson, C. L., & Sindelar, P. T. (1991). Direct instruction in math word problems: Students

     with learning disabilities. *Exceptional Children*, *57*(6), 512-519.

     https://doi.org/10.1177/001440299105700605

Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word

     problems for students with learning problems: A meta-analysis. *The Journal of Special

     Education*, *32*, 207-225. https://doi.org/10.1177/002246699903200402

*Xin, Y. P., Zhang, D., Park, J. Y., Tom, K., Whipple, A., & Si, L. (2011). A Comparison of

     Two Mathematics Problem-Solving Strategies: Facilitate Algebra-Readiness. The Journal

     of Educational Research, 104, 381-395.

Zhang, D., & Xin, Y. P. (2012) A follow-up meta-analysis for word-problem-solving

     interventions for students with mathematics difficulties. *The Journal of Educational

     Research, 105,* 303-318. https://doi.org/10.1080/00220671.2011.627397

Zheng, X., Flynn, L. J., & Swanson, H. L. (2013). Experimental intervention studies on word

     problem solving and math disabilities: A selective analysis of the literature. *Learning

     Disability Quarterly, 36*, 97-111. https://doi.org/10.1177/0731948712444277

Table 1

Summary of Study Characteristics

	Study	Total n	LD n	EL n	LD Definition	Grade	Design
1	Fuchs, Fuchs, Hamlett, & Appleton (2002)	40	30	0	D	4	RCT
2	Fuchs et al. (2008a)	243	243	4	AR	3	RCT
3	Fuchs et al. (2004a)	351	33	4	AR	3	RCT
4	Fuchs et al. (2004b)	366	57	2	AR	3	RCT
5	Fuchs, Fuchs, & Prentice (2004)	201	35*	4	AR	3	RCT
6	Fuchs et al. (2003a)	375	52	8	AR	3	RCT
7	Fuchs et al. (2003b)	40	23	5	D	3	RCT
8	Fuchs et al. (2008b)	35	16	3	AR	3	RCT
9	Griffin & Jitendra (2009)	30	5	0	D	3	RCT
10	Jitendra et al. (2013a)	109	53	17	AR	3	RCT
11	Jitendra et al. (2007)	45	2	3	D	3	RCT
12	Jitendra et al. (1998)	34	17	0	D	2, 3, 4, 5	RCT
13	Jitendra et al. (2013b)	135	71	63	AR	3	RCT
14	Moran, Swanson, Gerber, & Fung (2014)	72	49	2	AR	3	RCT
15	Owen & Fuchs (2002)	24	16	0	D	3	RCT
16	Swanson, Moran, Lussier, &	82	62	0	AR	3	RCT

	Fung (2014)						
17	Wilson & Sindelar (1991)	62	21	0	D	2, 3, 4, 5	RCT
18	Xin et al. (2011)	29	16	0	D	3, 4, 5	RCT

Note. Total *N* = total number of students who were included in the study; LD *n* = LD students who received treatment; * = MD students only; EL = reported number of students who were English learners receiving intervention; D = students identified by discrepancy model; AR = students at risk for LD or below percentile cutoff score; RCT = randomized control trials

Table 2

Mean Effect Sizes and Confidence Intervals as a Function of Moderator Variables

Moderator Variable	K	ES	SE	95% CI	Q	I^2
LD definition						
Discrepancy	49	0.74	0.2	0.34 – 1.14	377.56	0.87
At risk	63	1.35	0.21	0.94 – 1.77	1557.13	0.96
EL						
Studies with EL	56	1.4	0.23	0.94 – 1.86	1492.83	0.96
Studies without EL	57	0.77	0.18	0.41 – 1.12	424.60	0.87
Grade						
2	4	0.12	1.09	-1.61 – 1.85	38.48	0.92
3	82	1.31	0.18	0.95 – 1.68	1705.92	0.95
4	15	0.77	0.25	0.24 – 1.31	82.45	0.83
5	6	0.08	0.46	-1.11 – 1.27	56.99	0.91
Duration of study						
12 sessions	18	1.15	0.48	0.15 – 2.16	281.24	0.94
18 sessions	8	0.01	0.34	- 0.80 – 0.81	41.21	0.83
20 sessions	14	0.21	0.17	0.00 – 0.42	4.82	0.00
24 sessions	2	1.38	0.95	-10.63 – 13.39	6.61	0.85
26 sessions	8	1.75	0.55	0.45 – 3.05	211.51	0.97
32 sessions	4	0.65	0.38	-0.57 – 1.87	5.06	0.41
34 sessions	8	3.24	0.73	1-51 – 4.96	374.57	0.98
36 sessions	7	1.45	0.38	0.53 – 2.38	93.70	0.94
60 sessions	6	0.12	0.08	-0.08 – 0.32	0.00	0.00
Deliverer of instruction						
Researcher	6	0.35	0.11	0.05 – 0.64	0.75	0.00
Teacher	37	1.23	0.25	0.73 – 1.74	357.22	0.90
Instructional assistant	2	0	0	-0.31 – 0.31	0.04	0.00
University student	64	1.15	0.21	0.73 – 1.58	1541.54	0.96
Parent	2	0	0.07	-0.86 – 0.86	0.31	0.00
Community hire	2	0.36	0.01	0.28 – 0.44	0.00	0.00
Grouping of students						
Large group	40	1.64	0.27	1.09 – 2.19	703.63	0.94
Small group	69	0.78	0.17	0.44 – 1.13	1216.84	0.94
Individual	4	0.67	0.25	-0.15 – 1.49	5.97	0.50
Type of measure						
Norm referenced	24	0.37	0.16	0.03 – 0.71	69.98	0.67
Researcher developed	89	1.27	0.18	0.92 – 1.63	1858.26	0.95

K=number of effect sizes

Table 3

Summary of Reported Use of Instructional Components

Study	Instructional Components											
	IC 1	IC 2	IC 3	IC 4	IC 5	IC 6	IC 7	IC 8	IC 9	IC 10	IC 11	IC 12
1 **Fuchs, Fuchs, Hamlett, & Appleton (2002)** Mean ES = 1.24	X	X	X	X		X		X	X		X	
2 **Fuchs et al. (2008a)** Mean ES = 0.67	X		X	X	X	X	X	X	X	X	X	X
3 **Fuchs et al. (2004a)** Mean ES = 3.24	X		X	X	X	X	X	X	X			
4 **Fuchs et al. (2004b)** Mean ES = 3.31	X		X	X	X	X		X	X			
5 **Fuchs, Fuchs, & Prentice (2004)** Mean ES = 2.09	X		X			X	X	X	X			X
6 **Fuchs et al. (2003a)** Mean ES = 0.66	X		X	X	X	X		X	X			
7 **Fuchs et al. (2003b)** Mean ES = 1.28	X		X				X	X	X			X
8 **Fuchs et al. (2008b)** Mean ES = 1.05	X		X		X	X	X	X	X	X	X	X
9 **Griffin & Jitendra (2009)** Mean ES = -0.06	X		X	X	X	X	X			X		X
10 **Jitendra et al. (2013a)** Mean ES = 0.00	X		X			X				X		X
11 **Jitendra et al. (2007)** Mean ES = -0.002	X		X	X	X	X		X		X		X
12 **Jitendra et al. (1998)** Mean ES = 0.52	X		X		X	X		X	X			X
13 **Jitendra et al. (2013b)** Mean ES = 0.36	X		X			X		X				X
14 **Moran, Swanson, Gerber, & Fung (2014)** Mean ES = 0.53	X		X		X			X				
15 **Owen & Fuchs (2002)** Mean ES = 3.48	X		X	X		X				X		X
16 **Swanson, Moran, Lussier, & Fung (2014)** Mean ES = 0.16	X		X		X		X					

		IC 1	IC 2	IC 3	IC 4	IC 5	IC 6	IC 7	IC 8	IC 9	IC 10	IC 11	IC 12
17	**Wilson & Sindelar (1991)** Mean ES = -0.01	X		X					X	X			X
18	**Xin et al. (2011)** Mean ES = 0.01	X	X	X	X	X	X						X
		18	2	18	9	11	14	7	13	11	5	3	12

Note: ES = effect size; IC 1 = explicit instruction; IC 2 = Technology; IC 3 = strategy cues; IC 4 = peer interaction; IC 5= instructional feedback; IC 6 = visual aids; IC 7 = foundational skills; IC 8 = schema instruction; IC 9 = instruction to transfer; IC 10 = manipulatives; IC 11 = behavior reinforcement; IC 12 = self-regulated learning

Table 4

Fixed Effects of Instructional Components

	Included		Did not include		Contrast	
	Estimate	*SE*	Estimate	*SE*	*F* Ratio	*p* value
IC1	1.09	0.21	0.68	1.28	0.10	0.75
IC 2	0.54	0.66	1.15	0.22	0.76	0.39
IC 3	1.17	0.21	0.28	0.66	1.66	0.20
IC 4	1.70	0.24	0.39	0.26	13.50	0.0005*
IC 5	0.90	0.29	1.27	0.29	0.84	0.36
IC 6	1.40	0.24	0.53	0.32	4.73	0.03
IC 7	1.10	0.35	1.08	0.26	0.00	0.96
IC 8	1.14	0.24	0.94	0.39	0.19	0.66
IC 9	1.61	0.25	0.42	0.28	10.11	0.002*
IC 10	0.58	0.52	1.18	0.22	1.14	0.29
IC 11	1.02	0.57	1.10	0.22	0.02	0.89
IC 12	0.68	0.42	1.21	0.23	1.22	0.27

Note. *Bonferroni correction; *p*=.004; IC 1 = explicit instruction; IC 2 = Technology; IC 3 = strategy cues; IC 4 = peer interaction; IC 5= instructional feedback; IC 6 = visual aids; IC 7 = foundational skills; IC 8 = schema instruction; IC 9 = instruction to transfer; IC 10 = manipulatives; IC 11 = behavior reinforcement; IC 12 = self-regulated learning