UNIVERSITY OF CALIFORNIA
RIVERSIDE

The Effects of a Paraphrasing Intervention on Word Problem Solving Accuracy of
English Learners at Risk for Mathematic Disabilities

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Education

by

Jennifer Eun Re Kong

March 2017

Dissertation committee:
        Dr. H. Lee Swanson, Chair
        Dr. Rollanda O'Connor
        Dr. Austin Johnson

The Dissertation of Jennifer Eun Re Kong is approved:

_____

_____

_____

                                                    Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

DEDICATION

I would like to dedicate this dissertation to my family. First, to my loving husband, Richard – as with anything else in life, we did it together. To my beautiful son, Jacob – I hope to always be the person that you believe I am.

To my biggest cheerleaders, my parents and brother – thank you for your unconditional love and support. You taught me how to work hard and love well. I hope I make you proud.

My deepest gratitude and love to the Kong family, who have taken me in as their own. Thank you for your encouragement and guidance throughout the years.

Above all, everything I have is only because of God's abundant grace. In your presence, there is fullness of joy. Soli Deo gloria.

ABSTRACT OF THE DISSERTATION

The Effects of a Paraphrasing Intervention on Word Problem Solving Accuracy of
English Learners at Risk for Mathematic Disabilities

by

Jennifer Eun Re Kong

Doctor of Philosophy, Graduate Program in Education
University of California, Riverside, March 2017
Dr. H. Lee Swanson, Chairperson

Mathematical word problems require complex processes beyond basic math skills,

such as the use of linguistic information, identifying relevant information, and

constructing the appropriate problem statement (Fuchs et al., 2006; Swanson, 2006).

English learners (ELs) in particular may experience more difficulty with math problem

solving because of the need to preserve information while at the same time processing

information in a second language (Swanson, Kehler, & Jerman, 2010). In light of the

significance of math problem solving to children's mathematical achievement, research

to identify effective instructional strategies for problem solving is critical. Few studies

have investigated the effectiveness of a word problem solving intervention for students

who are both ELs and at risk for mathematic disabilities (MD). Paraphrasing intervention

has been found to be an effective intervention towards improving problem solving

accuracy with monolingual children (e.g., Moran, Swanson, Gerber, & Fung, 2014), but

its application to children learning a second language has not been tested. This study

aims to fill this gap in research by utilizing a multiple baseline design to assess the

effectiveness of a paraphrasing intervention on the problem solving performance for 9

third grade students who are ELs and at risk for MD. This study also investigated the extent to which the paraphrasing intervention facilitated transfer to calculation and reading comprehension measures. Tau-U effect sizes were calculated to determine improvement between baseline and intervention phases, as well as positive trends during the intervention phase. In general, positive gains occurred as a function of the treatment condition, supporting the hypothesis that paraphrasing word problem solving intervention improved students' one- and two-step word-problem solving skills. However, the magnitude of the effect sizes was in the low to moderate range on the targeted measures. The magnitude of the effect sizes were small for the transfer measures, suggesting that the paraphrasing intervention had minimal influence on transfer to calculation and reading comprehension measures. The results suggest that although positive effects in problem solving accuracy can occur with EL children, further research is necessary to identify more robust intervention procedures.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

Mathematic skills are necessary for academic success, everyday problem solving, future career options, and earning potential (McIntosh & Vignoles, 2000; Rivera-Batiz, 1990; Shapka, Domene, & Keating, 2006). Mathematic skills are required in elementary years to demonstrate proficiency on standardized high stakes testing. As many math concepts are cumulative, basic skills in numeracy, calculation, and problem solving are necessary for future academic success. Further, early mathematic proficiency has been found to lead to career aspirations of higher prestige and overall higher likelihood of full time employment (Rivera-Batiz, 1990; Shapka, Domene, & Keating, 2006). Unfortunately, when compared to their monolingual English speaking peers, English Learners (ELs) often perform poorly in mathematics. ELs encounter unique academic challenges, including cultural and linguistic acclimation in addition to the pressures of achieving academically, often resulting in disproportionately low achievement (Garcia & Cuéllar, 2006).

As ELs are the most rapidly growing demographic in U.S. public schools, providing appropriate instruction for these students should be a critical concern. In the 2012–2013 school year, approximately 9.2 percent (4.4 million) ELs were enrolled (National Center for Education Statistics, 2015). Additionally, 22.8 percent of students in California public schools (1.4 million) are ELs. According to the National Center for Education Statistics (NCES, 2013), an astounding 86 percent of EL students do not demonstrate proficiency in mathematics. Approximately 41 percent of EL fourth graders

scored below basic level in mathematics, whereas only 15 percent of non-EL students performed below basic level. By eighth grade, 69 percent of English Learners scored below basic level in mathematics compared to 24 percent non-ELs. This poor math achievement places many ELs at risk for persisting math difficulties.

Mathematical Disabilities

The term mathematical learning disability (MD), also referred to as acalculia, dyscalculia, arithmetic disorder, specific learning disabilities in mathematics, and mathematic difficulty, implies that the disorder stems from a biological base with underlying cognitive deficits (Mazzocco, 2007). Mathematic ability is supported by a wide selection of cognitive abilities such as central executive functions, attentional and inhibition control, manipulation in the language system, information representation, and visuo-spatial skills. Deficits in a combination of these areas may result in MD (Geary & Hoard, 2005). In addition, MD is characterized by weak working memory and language skills, higher incidence of reading disabilities (as a group), and other neuropsychological correlates (e.g., Mazzocco, 2007).

Compared to research on reading disability, research on MD is still emerging, despite the relatively parallel prevalence rates for these disorders. One large-scale study indicated that 6.4 percent of children were identified as having a math disability, compared to 4.9 percent with a reading disability (Badian, 1983). More recent research (e.g. Barbaresi, Katusic, Colligan, Weaver, & Jacobsen, 2005; Geary, Hoard, Byrd-Craven, & DeSoto. 2004; Mazzocco & Myers, 2003) has estimated prevalence rates of

MD were between 5 percent and 8 percent. Some of the challenges that researchers face in the field of MD are identifying a common definition of math learning disability.

Variability exists in how students are identified with MD in previous research and practice. Historically, children have been identified as LD given a significant discrepancy between IQ and achievement (Discrepancy model; Hallahan, Pullen, & Ward, 2013). The growing consensus among researchers has indicated that a cutoff score on achievement is more appropriate to determine MD than a discrepancy between achievement and IQ (e.g., Fletcher et al., 1989; Swanson, Moran, Lussier, & Fung, 2013; Mazzocco, 2007). Researchers have utilized the term "at risk for MD" to identify children who may be at risk for academic failure and benefit from intervention, but have not yet been identified as MD. For example, 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). For this study, students were identified as at risk for MD when scores of mathematic problem solving were below the 25th percentile.

Word problem solving

Mathematical word problems are linguistically presented arithmetic problems that require students to generate a solution (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. As students progress through school, instruction in math programs increasingly emphasizes word problem solving.

Procedural execution and selection for math problem solving are supported by conceptual understanding of the problem. Competency in a given area of math will be determined by the interrelated association between conceptual understanding of the problem and procedural knowledge (Geary, 1994). As children develop arithmetic proficiency, they also begin to discover more efficient procedures, or strategies, to use during problem solving (Geary, 2004; Geary, Hoard, Nugent, & Byrd-Craven, 2007). Additionally, reading comprehension has been found to be highly predictive of problem solving accuracy (Swanson, Cooney, & Brock, 1993).

Students with MD experience significant difficulty with word problems because complex processes beyond basic math skills are involved (Swanson, 2006). Additionally, students with MD perform significantly lower in math than age-equivalent peers, with the gap widening as each academic year passes (Cawley, Parmar, Foley, Salmon, & Roy, 2001). Children with MD tend to utilize more immature strategies (e.g., counting based) when compared to typically achieving peers (Geary, Hoard, Byrd-Craven, & DeSoto, 2004; Geary et al., 2007). Additionally, children with MD commit more errors while counting and solving math word problems. MD children also differ from typically achieving children in the frequency and accuracy of retrieving answers from long-term memory (Barrouillet, Fayol, & Lathuliere, 1997; Geary, 1993).

Math word problems also present unique problems for children who are English learners. Linguistic complexity of word problems is one important factor that presents more difficulty for ELs when compared to their English proficient peers (Martiniello, 2008). Martiniello (2008) identified items on a standardized achievement assessment that

posed more reading comprehension challenges for ELs utilizing differential item functioning (DIF) procedures. Results indicated that syntactic features and vocabulary were examples of items that were more challenging for ELs than for non-ELs with comparable math ability. This implies that word problems' language should be modified for ELs while keeping the content the same.

Identifying the main idea from a text is central to reading comprehension. Students with learning disabilities may experience difficulty with reading comprehension stemming from the inability to locate main ideas. Paraphrasing information has been identified an effective strategy to improve reading comprehension for students with learning disabilities (Bakken, Mastropieri, & Scruggs, 1997; Gajria & Salvia, 1992; Jitendra & Gajria, 2011). Teaching students how to paraphrase or summarize information facilitates reading comprehension because students are taught to identify key information while inhibiting information that is irrelevant. Paraphrase training has also been identified as an effective strategy in group design studies for teaching students with MD to comprehend and solve math word problems (Moran, Swanson, Gerber, & Fung, 2014; Swanson, Moran, Lussier, & Fung, 2013). Although group design studies have identified paraphrase training as effective for children with MD, its effectiveness for children who are English Learners is unknown. Single-case research design may be valuable for close analysis of students who may not respond well to intervention in the development of effective instruction.

Single-Case Design

Single-case research design can be a useful tool for researchers. In single-case designs, researchers can examine the repeated measures on individuals to track growth over training sessions. Additionally, compared to group design studies in which outcomes are averaged (thus obscuring individual differences) researchers can study differences between individuals. Further, research has indicated that single-case research is valuable in identifying evidence-based practices for students with disabilities (Horner et al., 2005). Compared to group design, single-case design does not require a control group (a group that does not receive treatment) as individuals act as their own controls. Individual performance before intervention is compared to performance during and after. Thus, all individuals who participate and are in need of valuable treatment will have access to it.

One characteristic of single-case research is experimental control, which is achieved through systematic manipulation of the independent variable (Kratochwill et al., 2010). This control allows for inference of causal relations between the independent variable (e.g., intervention) and the dependent variable (e.g., word problem solving achievement). In single-case research, individuals serve as the unit of intervention and data analysis, with the outcome measured before, during, and after intervention (Kratochwill & Levin, 2014). A baseline measure is documented to control for confounding variables. When stability is achieved prior to administering intervention, changes in performance after intervention can be attributed to treatment effects.

Internal and external validity. According to the What Works Clearinghouse single-case design criteria (Kratochwill et al., 2010), some possible threats to internal validity in single-case research include: ambiguous temporal precedence, selection, history, maturation effects, statistical regression, attrition, testing, instrumentation, and additive and interactive effects of threats to validity. Threats to internal validity can be addressed through replication and/or randomization (Kratochwill & Levin, 2010; 2014). According to Kratochwill and Levin (2010), randomization can improve internal validity at a higher level than replication. Incorporating randomization into single-case designs strengthens the ability to draw causal inferences. Two single-case randomization schemes are randomized phase-order designs and randomized phase start-point designs. Randomized phase-order designs are within-series sequences of baseline (A) and treatment (B) phases. Randomized phase start-point designs refer to the specific time points at which baseline and treatment phases begin. Additionally, individuals can be randomly assigned to intervention conditions in addition to randomization schemes. In this study, students from each class were randomly assigned to staggered intervention start points. Additionally, specific start time points at which intervention began for each group were randomly determined prior to the intervention.

Threats to external validity include generalizability to the population of interest and the baseline phase possibly having an effect on the treatment condition. Threats to external validity are addressed through careful design of the study, which can be replicated by other researchers. This design would include careful description of the

7

treatment, procedures used to measure behavior, characteristics of the subject, and other possible confounding factors.

If the study design is found to meet evidence standards or meet evidence standards with reservations, a visual analysis is conducted for each outcome variable to determine evidence for causal relations (Kratochwill et al., 2010). This is determined by the following features used to examine within and between phase patterns: (1) demonstration of consistency of level, (2) trend, (3) variability within each phase, (4) immediacy of the treatment effect, (5) proportion of overlapping data, and (6) consistency of the data across phases. A causal relation can be determined if differences in the data in in the intervention phase are more than what would be expected when compared to the data pattern in the baseline phase. In this study, visual analysis was conducted to evaluate the consistency of level, trend, and variability in each phase.

Effect size calculation. Traditionally, the interpretation of outcomes on single-subject design studies has relied on visual analysis. More recently, the reporting of outcomes has relied on the computing of effect sizes (Parker, Hagan-Burke, & Vannest, 2007; Parker & Vannest, 2009; Parker, Vannest, Davis, & Sauber, 2011). There are a number of ways to calculate effect sizes for single-case research: mean difference effect sizes, regression-based effect sizes, the percent of nonoverlapping data, the percent of all nonoverlapping data, and Tau-U. When calculating mean difference or regression-based effect sizes for single-case design, the assumption of observation independence is not tenable as data are collected on the same individual over time (in the same setting, conditions, response definitions, and recording procedures) (Wolery, Busick, Reichow, &

8

Barton, 2010). That is, the data are likely to be serially dependent. Additionally, the

parametric assumptions of normality and homogeneity of variance are not satisfied by

single-case research data (Parker, Hagan-Burke, & Vannest, 2007).

One alternative method of calculating effect is the percent of nonoverlapping data

(PND). PND is the percentage of data that is more extreme in the treatment phase than

the single most extreme data point in the baseline phase. If a treatment is very effective,

PND will near 100%. The percent of all nonoverlapping data (PAND) is similar to PND

but uses all data from both phases. However, even with the use of PAND, several

limitations still exist (Parker, Hagan-Burke, & Vannest, 2007). PAND lacks sensitivity at

the upper end of the scale. Similar to PND, 100% will be awarded to nonoverlapping

data regardless of the distance between two data phases. Also, PND and PAND do not

control for possible trends in the baseline phase. If positive trends in baseline exist, it is

difficult to attribute change to the intervention.

To address these limitations of calculating effect for single-case designs, a

nonparametric index has been recommended (Parker et al., 2011). Tau-U is an analytic

method for calculating effect size which combines nonoverlap between phases (baseline

and intervention) with calculating trend within the intervention phase. Additionally, Tau-

U allows for control of trends within the baseline phase. This index is derived from

Kendall's rank correlation and the Mann-Whitney U test.

The Mann-Whitney U (MW-U) test is an index of nonoverlap between phases

(Bowman-Perrott et al., 2013; Parker et al., 2011). MW-U measures the proportion of

pairwise comparisons that improve from baseline to intervention (percentage of

nonoverlapping data). The algorithm for MW-U combines the scores from two phases

for a cross-group ranking, which are then separated and statistically compared for mean

differences in ranks. MW-U produces two U values (U_L and U_S), which is the number of

times data points in one phase precede data points in the other in ranking (Parker &

Vannest, 2009). The formula for U is as follows:

$$U = R_1 - \frac{n_1(n_1 + 1)}{2}$$

where n_1 and n_2 the sample sizes of the two groups (phases) and R_1 is the sum of ranks

for phase 1 (Cliff, 1993). Kendall's S is the larger U value subtracted by the smaller U

value ($S = U_L - U_S$). So, MW-U = ($U_L - U_S$) / ($U_L + U_S$). The denominator ($U_L + U_S$) is

the total number of comparisons possible (between baseline and intervention phases),

which is calculated by multiplying the two group N number of data points (n_1 x n_2).

Kendall's rank correlation (KRC) is an index of trend within a phase (Bowman-

Perrott et al., 2013; Parker et al., 2011). KRC is an analysis of time and score, which is

calculated by an algorithm that compares scores that are time ordered and all possible

pairs of data. The total number of pairs is $N(N - 1) / 2$, where N is the number of original

scores. Each pairwise contrast is then coded one of the following: (a) positive or

increasing over time; (b) negative or decreasing over time; or (c) tied. Kendall's Tau is S

(positive codes – negative codes) divided by the total number of pairs (Tau = S / #pairs),

or the percentage of data that improves over time. In single-case research, KRC then

calculates the trend of the data in the intervention phase.

By formula, MW-U and Tau are the same (Parker et al., 2011):

$$\text{MW-U} = (U_L - U_S) / (U_L + U_S) \qquad \text{Tau} = (\#pos - \#neg) / \#pairs$$

$$\text{MW-U} = S / \#pairs \qquad \text{Tau} = S / \#pairs$$

Also, both MW-U and KRC utilize the sampling distribution of S as the test statistic for inference testing. For studies with N number of data pairs larger than 10, the S distribution approaches normal. Therefore, $z = S / SE_s$ can be used (Graney, 1979; Parker et al., 2011). This study will have 153 pairs [18(18-1) / 2].

Tau-U produces the following indices: (a) baseline versus intervention phase nonoverlap, (b) nonoverlap and intervention phase trend together, (c) nonoverlap with baseline trend controlled, and (d) nonoverlap and intervention trend with baseline trend controlled (Parker et al., 2011). This study utilized the Tau-U indices to calculate effect size. In short, the formula was:

$$\text{Tau} = N_c - N_d / (n[n-1]/2),$$

where c = concordant pairs, d = discordant pairs, and n = possible pairs (Parker et al., 2011).

Problem Statement and Purpose of the Study

Many EL students who are at risk for MD experience significant and persisting difficulties with mathematical word problem solving. Although some research studies have examined the effectiveness of problem solving interventions for students who are MD, very few studies have addressed interventions for students who are both EL and MD. The development of effective and validated interventions that focus on problem

solving for these students is necessary to provide proper instruction as early as possible to alleviate academic achievement gaps.

The purpose of this study was to investigate the effectiveness of a word problem solving intervention for third grade EL students who are at risk for MD. This study addressed three questions:

1. To what extent does a paraphrasing word problem solving intervention improve students' one- and two-step word-problem solving skills?

2. To what extent does a paraphrasing word problem solving intervention influence transfer to calculation skills?

3. To what extent does a paraphrasing word problem solving intervention influence transfer to reading comprehension?

Chapter 2

Literature Review

The National Mathematics Advisory Panel (2008) has recommended identifying and validating instructional math practices on students with or at risk for MD. Legislation such as No Child Left Behind (NCLB) mandates the use of evidence-based practices to improve educational outcomes for students with disabilities (Simpson, Lacava, & Graner, 2004). In addition, English learners at risk for MD face unique challenges in math, such as acquiring the formal mathematical linguistic register and lack of exposure to discussions which develop higher order thinking skills (Janzen, 2008). Appropriate intervention for EL students at risk for MD may provide valuable support to prevent potential academic failure (Gersten, Jordan, & Flojo, 2005; Griffin, 2007). The current study is a mathematics word problem solving intervention for English learners at risk for MD utilizing a single-case design. Thus, literature on effective problem solving interventions for these students will be reviewed. First, the chapter will provide an overview of validated instructional types and components for students with or at risk for MD. Second, the literature on group design problem solving interventions with at least 40% of the sample including ELs will be reviewed. Third, the literature on single-case design problem solving intervention studies for students with or at risk for MD will be addressed. Finally, single-case design intervention studies for EL students with or at risk for MD will be presented.

Studies that addressed mathematics word problem solving interventions for students identified as EL or at risk for MD were identified for this literature review in

published peer review journals. The reference lists of previous studies which reviewed problem solving interventions (e.g., Gersten et al., 2009; Xin & Jitendra, 1999; Zhang & Xin, 2012; Zheng et al., 2013) were systematically scanned.

Gersten et al. (2009) reviewed general mathematic (not specifically problem solving) interventions for school aged (grades K-12) students with learning disabilities. A total of 42 group-design intervention studies that included students with identified LD according to state regulations were included in the review. Gersten and colleagues identified four major categories of instructional components that contributed to positive effects in math performance for students with learning disabilities: (a) approaches to instruction and/or curriculum design, (b) formative assessment data and feedback to teachers on students' mathematic performance, (c) formative data and feedback to students with LD on their performance, and (d) peer-assisted mathematics instruction. Approaches to instruction and/or curriculum design included explicit instruction, the use of heuristic strategies, student verbalizations or thinking aloud of mathematical reasoning, visual representations to solve problems, a sequence of examples, and other instructional components. Formative assessment data and feedback to teachers included feedback on student progress in addition to recommendations for addressing student instructional needs. Feedback to students included providing students with information regarding their effort or performance from teachers, peers, or software programs. Finally, peer-assisted mathematics instruction included one-on-one peer tutoring, cross grades and within a class. All instructional components yielded significant positive effects with the exception of: (a) student setting a goal and measuring achievement of that goal, and (b)

peer-assisted learning within a class. Gersten and his colleagues recommended the use of the following instructional components to improve math instruction: (a) explicit instruction, (b) visual representations, (c) sequence and/or range of examples, (d) student verbalizations, and (e) providing ongoing feedback to both teachers and students. This meta-analysis included students from a wide range of age (grades K-12) in addition to both problem solving and calculation interventions. This meta-analysis did not include single-case studies as part of the review of the literature.

In a meta-analysis with more stringent criteria, Xin and Jitendra (1999) investigated specifically problem solving interventions for students in grades 1 to 12 with learning problems at risk for math failure. Participants were considered at risk for math failure if they were receiving remedial math intervention or diagnosed with a learning problem. Learning problems were defined as mild disabilities such as learning disabilities, mild mental retardation, and emotional disabilities. 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. The researchers classified students as LD, mixed disabilities, or at risk. Computer-assisted instruction in group-design studies was found to be most effective. In single-case studies, computer aided instruction, representation techniques (e.g., diagramming manipulatives, schema based instruction, linguistic training), and strategy training (e.g., heuristic, cognitive strategies) was found to be effective in facilitating problem solving skills. Additionally, long-term interventions (those lasting for more than 1 month) were found to be more effective than short-term interventions.

As a follow-up to the meta-analysis conducted by Xin and Jitendra (1999), 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. Studies that included students with learning problems in math from kindergarten to twelfth grade were included in the meta-analysis. Twenty-nine group-design studies and ten single-subject studies (39 total studies) were included in the follow-up meta-analysis. An interesting feature of this meta-analysis is that the researchers investigated differences in effect sizes between students who were identified as LD by the discrepancy model and students who were identified as "at risk" (e.g., low achievement scores). Results indicated that interventions provided in inclusive settings were more effective than in special education settings for students with learning problems. Further analysis indicated that typically achieving students and students with learning problems performed similarly in inclusive settings. Results also indicated that while all intervention strategies (problem structure representation, cognitive strategy training, and strategies involving assistive technology) produced positive effects, problem structure representation techniques yielded highest effect sizes for students with math difficulties. Problem structure representation techniques refer to explicit instruction including schema-based diagramming. Cognitive strategies include sequential cognitive or metacognitive strategy instruction. There were no significant differences between effect sizes from students diagnosed with discrepant LD and at-risk students, suggesting that these students responded similarly to interventions.

Finally, Zheng, Flynn, and Swanson (2013) also conducted a selective meta-analysis of word problem solving intervention studies for school aged students with MD. A total of 15 studies (7 group-design, 8 single-subject) were included in the study. 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 problem solving instruction) and 0.90 for single-subject studies. 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. Finally, based on the magnitude of the effect size, students with math difficulties benefitted more from intervention than students with math and reading difficulties.

Additionally, the PsycINFO, Science Direct, and ERIC online databases were systematically scanned for studies from 1990 to 2015. 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: *efficacy, strategy instruction, schema-based instruction, scaffolded instruction,* and *peer interaction.* To be eligible for this review, each study had to meet the following criteria: (a) included students with or at risk for learning disabilities or were identified as English Learners; (b) tested an intervention to improve word-problem-solving; (c) assessed

students' word-problem-solving accuracy (quality measures using standardized assessments); (d) involved a single-case design or group design true-experiment 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 quantitative data on problem solving performance; and (f) was published in English. Studies investigating the effectiveness of instruction improving only math calculation or other related math skills were not included. However, studies that included instruction in math calculation or related skills in addition to word problem solving intervention were considered, provided that results were shown specifically for word problem solving. This initial search generated approximately 233 items. Abstracts for all items were read prior to selecting studies to eliminate articles that clearly did not meet the inclusion criteria. This procedure narrowed the search to 13 documents.

Group-Design Word Problem Solving Interventions for English Learners

The mathematics word problem solving intervention literature on single-case research studies for students who are both English learners and at risk for MD is extremely limited. To broaden the search, intervention studies (including group design intervention studies) for ELs and children at risk for MD were included. Word problem solving intervention studies with exclusively EL participants were not found. However, studies that included at least 40% of the sample ELs were considered.

Jitendra and her colleagues (2013) investigated the efficacy of a problem solving intervention utilizing schema-based instruction (SBI) in small groups for third grade students at risk for math difficulties. Of the 136 participants, 47% (63 students) were

English learners. Children were identified as at risk for MD if their scores on a

standardized district assessment (*Measures of Academic Progress*) were below the 40th

percentile. Of the 71 students receiving SBI, 51% (36 students) were English learners.

Additionally, approximately 14% of the students receiving SBI were identified as special

education students. Students received intensive supplemental intervention for 12 weeks

(5 days a week, 30 minutes). Students in the SBI condition were taught three problem

types (Change, Group, Compare), schematic diagramming, and story checklists to solve

one- and two-step word problems. During the first phase of the SBI intervention,

students were taught to identify problem schema. Problems indicating the Change

schema included three pieces of information: (1) initial quantity, (2) explicit action

indicating change (e.g., gave away), and (3) ending quantity. Group type problems

included two subgroups and a larger group, activating the part-whole schema. Finally,

compare problems involved comparing two separate groups. The second phase of the

intervention (problem solution phase) instructed students to use a metacognitive strategy

utilizing a checklist, schematic diagramming, and select the correct operation for accurate

problem solution. Students were also directed to check and justify their answers. Results

indicated that students in the SBI group who had higher scores at pretest outperformed

students with higher pretest scores in the standards-based curriculum (SBC) group.

Conversely, students with lower scores at pretest in the SBC group performed better than

students with lower scores in the SBI group. This indicates that SBI may be more

effective for students at risk for MD who have already mastered basic computational

skills compared to students who require additional instruction in basic skills.

In another group design intervention study, Moran, Swanson, Gerber, and Fung (2014) examined the effect of a paraphrasing intervention for third grade students at risk for MD. Students were identified as at risk for MD if their scores on measures of fluid intelligence (*Raven Colored Progressive Matrices Test;* Raven, 1976) were above the 25th percentile *and* scores on either the *Test of Mathematical Abilities* (TOMA; Brown, Cronin, & McIntire, 1994) or *Key Math* Revised Problem Solving (Connolly, 1998) was below the 25th percentile. Additionally, of the 72 participating students, 56% (40 students) were English learners. Students were then randomly assigned to one of four conditions: *Restate*, *Relevant*, *Complete*, or control. The Restate proposition condition involved students paraphrasing and rewriting the question in their own words. The Relevant proposition condition taught students to paraphrase all relevant information, including the question and numbers necessary to solve. The Complete proposition condition directed students to paraphrase the question and separate relevant and irrelevant information. Results indicated that students in the Relevant and Complete conditions improved on measures of word problem solving when compared to the students in the Restate and Control conditions. Paraphrasing relevant and irrelevant information in a word problem may be effective for EL students at risk for MD.

Similarly, Swanson, Moran, Lussier, and Fung (2013) investigated the effectiveness of generative strategies for third grade students at risk for MD. Students were identified as at risk for MD if their scores of fluid intelligence (*Raven Colored Progressive Matrices Test*) were above the 16th percentile *and* scores on a standardized math assessment (*Test of Mathematical Abilities;* Brown, Cronin, & McIntire, 1994) were

below the 25th percentile. In this sample of 82 students, 66% (54 students) were English

learners but were considered English proficient based on their scores on the California

English Language Development Test (CELDT). Again, students were assigned to one of

four conditions: *Restate*, *Relevant*, *Complete*, or control. In this study, students' working

memory capacity (WMC) was also measured to investigate if generative strategies were

moderated by WMC. Results indicated that students with high WMC in the Complete

condition performed higher on measures of word problem solving at posttest. This

significant advantage in scores at posttest was not found in the regression modeling when

pretest WMC was set to a low value. This suggests that generative strategies that involve

paraphrasing all propositions are more likely to be effective for students at risk for MD

with larger WMC than those with a smaller WMC.

Single-Case Design Word Problem Solving Intervention Studies

This study proposes the importance of a word problem solving intervention for

ELs who are at risk for MD utilizing a single-case design. Again, the mathematics

problem solving intervention literature on single-case research studies for students who

are English learners and at risk for MD is limited. The following studies are all single-

case design intervention studies that included students who were at risk for or identified

as having MD. However, the students were not English learners.

Word problem solving intervention studies utilizing single-case design were

reviewed and grouped by the central instructional element. The following instructional

components were reviewed: visual diagrams, Self-Regulated Strategy Development

(SRSD), schema-based instruction, cognitive strategy training, and Dynamic Strategic Math (DSM).

Visual diagrams. Visual representations are models that are created to link verbal information with symbolic comprehension (van Garderen & Montague, 2003). Examples of visual representations include generating pictorial representation, manipulatives, and visual diagrams.

The effectiveness of diagram generation to solve one- and two-step word problems was assessed with three eighth grade students with learning disabilities (van Garderen, 2007). All students were English dominant and were identified as LD based on district eligibility criteria of significant differences between IQ and achievement. A multiple baseline across participants was utilized the evaluate effects of the intervention. The intervention was divided into three phases: (1) instruction for generating diagrams, (2) strategy instruction for one-step word problems, and (3) instruction for two-step word problems. Students progressed along the instructional phases when the previous phase was mastered. Students were taught how to draw visual representations to show parts of the word problem via a line and part/whole diagram. Students were also taught the "Visualize" strategy, based on cognitive-metacognitive strategy for solving word problems (Montague, 1997). Results indicated that all students increased in diagram generation and word problem solving performance following intervention. A limitation of this study included the possibility that student' problem solving improved due to another part of the strategy or corrective feedback.

Self-regulated strategy development. In a single-case intervention study, four

students with learning disabilities were taught a word problem solving strategy by way of

the self-regulated strategy development procedures (Case, Harris, & Graham, 1992). The

participants were in the fifth or sixth grade and struggled with executing the proper

operation to solve word problems, performing at least two years below grade level and

IQ. Students were taught a problem solving strategy with the following steps: (a) read

the problem out loud, (b) find and circle important words, (c) draw a picture to show

what is happening, (d) write down math sentence, and (e) write the answer. Instructors

explicitly demonstrated self-regulated strategy use by utilizing "think-alouds" to (a)

define the problem, (b) plan, (c) use the strategy, (d) self-evaluate, and (e) self-reinforce.

Results indicated that problem solving performance in both addition and subtraction

problems improved after strategy instruction. Specifically, the number of errors resulting

from incorrect operation use reduced. Maintenance of these skills varied, with two

students maintaining the improvements eight to 12 weeks after instruction. Limitations

of this study include the simultaneous use of word problem solving strategy intervention

and self-regulated strategy development.

Cassel and Reid (1996) also investigated the effectiveness of self-regulated

strategy instruction on four students. Of the four students, two were identified as having

learning disabilities (based on significant discrepancies between IQ and achievement) and

two students with "mild mental retardation (MMR)." Students were taught a problem

solving strategy with the following steps: (a) read the problem out loud, (b) find and

highlight the question, (c) ask what are the parts of the problem and circle numbers, (d)

set up the problem by writing and labeling numbers, (e) reread problem and decide operation, (f) discover the sign (recheck operation), (g) read number problem, (h) answer problem, and (i) write the answer and check. The mnemonic "FAST DRAW" was utilized to assist students in remembering the strategy steps. Students were also taught to self-regulate by utilizing checklists to (a) define the problem, (b) plan, (c) use the strategy, (d) self-monitor, (e) self-evaluate, and (f) self-reinforce. Results indicated that all students mastered the problem solving strategy, displaying an immediate increase in performance for both addition and subtraction problem. A limitation of this study is that the problem solving strategy contains many (nine) steps, which may be difficult to remember and possibly unmanageable for some students.

Schema-based instruction. Jitendra and Hoff (1996) investigated the effect of schema-based strategy instruction on the word problem solving performance of three students with learning disabilities. Students were identified as LD by significant discrepancies between aptitude and achievement. Third- and fourth-grade students from a private school for students with LD participated in the study. All students experienced difficulty with word-problem solving, but had adequate computational skills. Students were taught how to identify word problems' schema (Change, Group, Compare problem types) using schematic diagramming. Then, students were instructed how to design a solution strategy, or action schema, and choose the appropriate operation. Results indicated that all students improved on word problem solving performance following SBI instruction. Students were also able to generalize the strategy to novel problems. A limitation of this study is that only three students were included.

As a follow-up to the previous study, Jitendra, Hoff, and Beck (1999) examined the effect of schema-based instruction on problem solving skills of four sixth- and seventh-grade students with learning disabilities. Students were identified as LD based on significant discrepancies between ability and achievement. All students were able to compute addition and subtraction problems with at least 90% accuracy, but experienced difficulties with one-step word problems. Students were taught how to identify problem types and design an action schema for one- and two-step word problems. Following instruction, average scores for both one- and two-step addition and subtraction word problems increased from the baseline to treatment condition. In addition, after instruction in one-step word problem solving, students were able to generalize the strategy to two-step problems spontaneously. Improvements over baseline were maintained on follow-up assessments for both one- and two-step word problems two and four weeks after instruction. However, improvements over instruction phases were inconsistent.

Lastly, Jitendra, DiPipi, and Perron-Jones (2002) conducted another single-case study investigating the effects of schema-based instruction on word problem solving skills of four eight-grade students with LD. Participants were identified as LD based on a significant discrepancy between achievement and intellectual ability. All students were able to adequately compute all four operations, but experienced significant difficulty with one-step multiplication and division problems. Students were taught how to identify problem schema and design a plan for problem solution. This study differed from the previous SBI studies in that the researchers introduced two new word problem types to solve multiplication and division problems. The Vary schema included four pieces of

information: (1) a constant per unit, (2) four quantities, two of which were subject units, and two object-units, (3) the association that paired each subject-object unit, and (4) an if-then relationship. The Multiplicative Comparison schema focused on compared and referent sets and their relative quantities, in addition to highlighting part-whole relationships. Following intervention, all students were able to identify and discuss key characteristics of each problem schema which contributed to gains in problem solution. Since SBI was taught in conjunction with visual diagramming, it is difficult to distinguish if gains were the result of schema instruction, visual diagramming, or both.

  Cognitive strategy training. Montague and Bos (1986) investigated the effect of cognitive strategy training on the word problem solving of six high school students with LD. Students were identified as LD based on a significant discrepancy between grade placement and grade equivalent scores on math achievement tests. Students were taught a problem solving strategy with the following steps: (1) read the problem aloud; (2) paraphrase the problem aloud; (3) visualize; (4) state the problem; (5) hypothesize; (6) estimate; (7) calculate; and (8) self-check. Following intervention, five out of six participants made considerable improvements from baseline to the treatment condition in correct two-step word problem solving responses. Four of six participants were able to general this strategy to three-step problems. All students maintained improved math problem solving ability two weeks following discontinuation of the intervention. All four students who were tested again three months following the intervention displayed maintenance of these skills.

Montague assessed the effectiveness of cognitive and metacognitive strategy instruction on the word problem solving of six middle school students with learning disabilities (1992). Students were identified as LD by the school district criteria based on a discrepancy between ability and achievement. All participants were able to solve calculation problems adequately, but experienced difficulty with word problem solving. In the first phase of instruction, three students were taught a cognitive strategy with the following steps: (1) read; (2) paraphrase; (3) visualize; (4) hypothesize; (5) estimate; (6) calculate; and (7) self-check while the remaining three students were taught metacognitive activities associated with each step of the strategy: SAY, ASK, and CHECK. In the second phase of instruction, all students received cognitive strategy instruction (CSI) in conjunction with metacognitive strategy instruction (MSI). Results indicated that three sessions of CSI alone did not significantly improve students' problem solving performance. However, three following sessions of added metacognitive strategy instruction improved the problem solving skills of all students who received this instruction. In contrast, three sessions of MSI alone resulted in slight improvements in students' problem solving performance. Following the second level of instruction, five of six students displayed considerable increases in number of correct responses. Two of five students were able to maintain their skills two weeks to two months after intervention. No students maintained these skills the following school year. Results suggest that cognitive and metacognitive strategy instruction may be more effective when taught together for middle school students with LD than either strategy instruction alone. A limitation of this study is that the instructional package of both cognitive strategy

27

instruction and metacognitive strategy instruction always followed either CSI or MSI alone. Students may have improved based on maturation effects and not as a direct result of the instructional package. Another limitation of both studies (Montague & Bos, 1986; Montague, 1992) is that the cognitive strategy involves employing eight strategy steps to attain a correct answer, which may be unwieldy for many EL students with learning disabilities.

Dynamic Strategic Math. The only studies that could be found that assessed the effectiveness of a word problem solving strategy specifically for ELs using single-case design utilized Dynamic Strategic Math (DSM; Orosco, 2014; Orosco, Swanson, O'Connor, & Lussier, 2013).

Orosco and colleagues (2013) tested the effectiveness of DSM with six second grade students who were Latino English Learners. A changing criterion multiple baseline across subjects design was utilized. DSM intervention involved the researcher providing strategy instruction based in the Dynamic Assessment framework. Vocabulary was systematically modified based on the students' level of understanding utilizing a four-level linguistic modification procedure. The intervention was taught is three steps: (1) preteaching concepts and terminology, (2) comprehension strategies instruction, and (3) dynamic testing. During strategy instruction, the instructor modeled the process and demonstrated the use of the strategy. Students were directed to find the question, important words (pretaught in Step 1), numbers, set up/solve the problem, and check their answers via scaffolding and explicit cueing. Results indicated that all students' ability to solve word problems with increasingly difficult vocabulary increased as a result of the

intervention. A limitation to this study includes the possible confounding factor of using linguistically modified word problems. Also, the intervention could have been described more thoroughly so as to be replicated by other researchers or practitioners.

The effectiveness of DSM was also investigated with six third grade Latino ELs at risk for MD (Orosco, 2014). Risk status was determined by teacher recommendation and if students performed below the 25th percentile on district math tests and on the Applied Problems test from the Woodcock-Johnson NU Tests of Achievement (WJ NU III-ACH; Woodcock, McGrew, & Mather, 2007). Again, a changing criterion multiple baseline across subjects was used to evaluate effects of the DSM intervention. This intervention incorporated three phases: (1) preteaching concepts and vocabulary, (2) teaching five common word problem solving strategies (What Do I Know, What Can I Find, What Is the Set-Up, Solve It, and Check for Understanding), and (3) cooperative learning via student pairing. An appendix was provided detailing intervention procedure. Results indicated that all students' word problem solving ability increased as a result of DSM intervention. One limitation of this study is that although the importance of visual representation techniques were outlined as a proposed practice to support the problem solving of children at risk for MD, the DSM intervention incorporated only verbal techniques.

Research Questions and Hypotheses

To date, few studies have addressed the efficacy of word problem solving interventions for students who are ELs and at risk for MD. It is important to identify appropriate and effective interventions for this population to begin to address existing

achievement differences. The purpose of this study was to investigate the effect of a

paraphrasing strategy intervention for EL students at risk for MD. The following

research questions and hypotheses were considered:

1. To what extent does a paraphrasing word problem solving intervention improve

    students' one- and two-step word-problem solving skills?

It is hypothesized that the paraphrasing intervention will improve the word

problem solving skills of third grade EL students at risk for MD. Paraphrasing activities

are assumed to assist students in cognitive and metacognitive processing, and thus foster

greater comprehension of the information by mentally connecting the new material with

existing schema (Wittrock, 1991). Paraphrasing propositions have been shown to

improve reading comprehension in children with learning disabilities because

summarization of central information is required and focuses the student's attention to

specific information essential for understanding (Bakken, Mastropieri, & Scruggs, 1997).

Therefore, paraphrasing and writing out the components of a word problem is

hypothesized to improve problem comprehension and facilitate accurate problem solving.

2. To what extent does a paraphrasing word problem solving intervention influence

    transfer to calculation skills?

Some researchers have postulated a "bottleneck" hypothesis, in which difficulty

with basic fact fluency represents a bottleneck for attaining other math skills, such as

more complex computation (Fleishner et al., 1982; Geary et al., 1987). The research

supporting calculation transfer from basic fact remediation is inconsistent (Fuchs et al.,

2009; Fuchs, Powell, et al., 2008; Powell et al., 2009). Further, these studies did not find

support for this "bottleneck" hypothesis on word problem solving outcomes.

Additionally, problem solving and computation have been found to represent distinct

areas of mathematical cognition although significantly correlated (Fuchs, Fuchs, et al.,

2008; Swanson & Beebe-Frankenberger, 2004). Problem solving requires a variety of

skills for problem solution, including using linguistic information and computational

accuracy. Since instruction in computation has not supported transfer to problem solving

outcomes, the investigation of problem solving influencing transfer to calculation is

warranted.

Transfer is defined as having generalized effects of performance outcomes across

various stimuli (e.g., tasks, settings) (Kendall, 1981; Kennedy, 2005). In this study,

transfer is operationally defined as the paraphrasing intervention influencing performance

outcomes across tasks (calculation, reading comprehension).

3. To what extent does a paraphrasing word problem solving intervention influence

transfer to reading comprehension?

This proposed problem solving intervention teaches students how to identify key

information and inhibit extraneous information, which is central to reading

comprehension. Reading comprehension has been found to be a strong predictor of

problem solving accuracy (Swanson et al., 1993). Utilizing a strategy targeting reading

comprehension and paraphrasing skills may increase word problem solving as well as

reading comprehension. Thus, students' reading comprehension skills are also

hypothesized to improve as a function of the paraphrasing intervention.

Finally, while students may be able to maintain these skills for a short while, prolonged times without access to continued intervention will result in the decrease of strategy usage and problem solving accuracy. Children who experience severe difficulties may require intense intervention over an extended period of time to remediate such difficulties (Vaughn, Denton, & Fletcher, 2010).

Chapter 3

Methods

Setting and Participants

Nine third-grade EL students at risk for MD participated in this study. The children were selected from four classrooms from an elementary (K-8) school in southern California. The school's population consisted of 700 students (73% Hispanic, 11% Black/African American, 9% White [non-Hispanic], 2% Asian, and 5% Other [two or more races]). Thirty-four percent of these students were English learners. Additionally, 65% of the school's population qualified for free or reduced lunch prices.

For the purposes of this study, children were identified as at risk for MD based on the following considerations: (a) teacher recommendation for intervention based on students receiving general math instruction for at least 2 years; (b) student who continued to experience difficulties solving word problems in the general education classroom; and (c) student performed at or below the 25th percentile on a norm-referenced math assessment, the story problem subtest from the *Test of Math Ability – 2 (*TOMA-2; Brown, Cronin, & McIntire, 1994). Students already receiving special education services were not included in the study.

The story problem subtest from the TOMA-2 (Brown et al., 1994) is a 25-item word problem solving assessment. Students are required to read and solve the word problems individually while recording answers in their test booklets. The items increase in difficulty and involve all four mathematics calculation areas. Testing is discontinued after 10 minutes. Reliability coefficient for this subtest exceeded .80.

EL status was determined by the presence of the California English Language Development Test (CELDT) score. The CELDT is an assessment used to determine and monitor the progress of children who are limited English proficient on listening, speaking, and writing in English. Table 1 provides descriptive and school-related information for the participating students.

General Procedures

A graduate student (licensed special education teacher) administered the intervention utilizing an instructional protocol (See Figure 1). This study was conducted in small groups (3 students) in the general classroom setting for 21 sessions over an 8 week period. Each intervention session averaged 30 minutes and was a supplementary intervention to the general education math curriculum students received (50 minutes/day). Two follow-up sessions were conducted following the conclusion of the study.

Research has indicated that students in small groups of three students made more gains than larger groups of ten students (Vaughn et al., 2010). To control for possible classroom teacher effects, students from each of the four classrooms were randomly assigned to small groups. That is, no small group consisted of students from one classroom teacher.

The paraphrase intervention served as the independent variable. Measures of word problem solving accuracy were the primary dependent variable of interest to answer the research questions. Secondary (transfer) dependent variables were measures of math calculation and reading comprehension accuracy.

In this study, word problems were modified from the classroom text (EngageNY; Expeditionary Learning, 2013). Word problems in the intervention included one- and two-step addition and subtraction word problems with the following elements: (a) a question, (b) relevant information and numbers required to solve the problem, and (c) irrelevant information or numbers. The following example illustrates the components of a one-step word problem.

David has 52 baseball cards. (relevant information)

David gave 19 baseball cards to Nick. (relevant information)

David also collects football cards. (irrelevant information)

How many baseball cards does David have left? (question)

Experimental design

A multiple baseline across subjects design was utilized to evaluate the effects of a paraphrasing intervention on the word problem solving performance of nine EL students at risk for MD (Kennedy, 2005). A functional relationship between the intervention and problem solving performance can be demonstrated by an increase in students' problem solving performance when compared to children left in baseline status. Causality is demonstrated by a systematic change in a dependent variable (i.e., positive changes in children's problem solving behavior) relative to the baseline phase with the introduction of the intervention (Kratochwill et al., 2010).

During the baseline phase, preinstructional performance of word-problem solving that each student could accurately solve without assistance was established. Each baseline measure consisted of 10 one- and two-step addition and subtraction word

problems. The problems were selected and modified from classroom text (EngageNY; Expeditionary Learning, 2013). Seventy problems were selected and then randomly assigned to sessions. Students were placed in three groups of three students. The first group received three baseline measures. Five baseline measures were administered to the second group. Finally, seven baseline measures will be administered to the third group.

After data was collected in the baseline phase, the independent variable was staggered across subjects, clustered by groups. Problem solving instruction directed students to apply a paraphrasing strategy to word problems. Each intervention session presented seven word problems in total - one word problem to deliver explicit instruction, one word problem to solve with teacher assistance, and five word problems to be solved independently. An example of instructional materials (i.e., student notebook) is presented in Figure 2.

Word problem difficulty (i.e., number of sentences, number of steps require to solve the problem) increased across intervention sessions. Lessons 1 – 4 taught one step word problems. Lessons 5 – 15 included multiple step word problems. Additionally, the number of sentences and complexity of sentences increased across lessons. Lessons 1 – 9 included word problem instruction with three to five sentences. Lessons 10 – 15 taught word problems with five to seven sentences.

Each session, measures of problem solving, calculation, and/or reading comprehension were administered to each student to evaluate progress. In addition, curriculum based measures were administered each week (alternating sessions) to evaluate generalization of these skills to novel problems. Two weeks after completion of

the intervention phase, maintenance probes were collected to verify maintenance of treatment skills. Two months after completion of the intervention phase, maintenance tests were administered again to verify continuation of skills after extended time periods without treatment.

Instructional Procedures

A paraphrasing strategy intervention was designed to improve problem solving accuracy. The intervention directed students to paraphrase and *write out* components of a word problem in their own words. The elements of word problem included a question, relevant information and numbers required to solve the problem, and irrelevant information or numbers. The strategy was taught during each intervention session during the course of the study. The instructional stages during each session are described in the subsequent sections.

Phase 1: Warm-up activity. During this phase, students participated in brief warm-up activities alternating between reading comprehension and math calculation exercises. Warm-up activities did not exceed five minutes per session.

For the reading comprehension warm-up activity, students read short paragraphs adapted from student texts (Engage NY; Great Minds, 2015) and answered multiple choice questions regarding the content of the paragraph. These exercises included literal reading comprehension questions, which will require students to recall characters, main events, or main ideas from the paragraphs. Students were also asked to generate a sentence identifying the main idea of the text.

For the calculation warm-up activity, students were provided (one, two, or three digit) addition and subtraction problems. The problems were selected from an experimental calculation measure, developed for the use of this intervention. Students were encouraged to complete as many problems accurately in the given time.

Phase 2: Explicit instruction. During this phase of the intervention, students were taught the paraphrasing strategy through direct instruction. The intervention incorporated the following four steps:

1. Know – "What do I know about the question" occurred after the teacher read the word problem aloud. The teacher identified the question for the group. Then, the teacher modeled how to paraphrase the question by writing a sentence.

2. Find – "Find the relevant information" occurred during reading the word problem aloud for a second time. Again, using think-alouds, the teacher modeled how to find and paraphrase important information to answer the question.

3. Cross-out – "Cross-out irrelevant information" occurred after finding relevant information. Students were guided in eliminating information that was not relevant to solve the problem. This information was not paraphrased.

4. Solve and check – After gathering necessary relevant information by summarizing important propositions, the teacher modeled how to set up and solve the problem. Finally, the teacher checked if the answer stated in a complete sentence addresses the initial question by stating the question again.

During this phase, the instructor demonstrated each step of the strategy through visual, explicit instruction. Visuals to aid instruction included a checklist to serve as reminders for each step as well as the teacher's written sentences.

Phase 3: Guided practice. During this phase of intervention, students answered word problems using the paraphrasing strategy. The teacher prompted students to apply each step of the strategy utilizing an instructional protocol. Instructors checked students' answers for each step of the strategy. If difficulty persisted in a step of the strategy, the instructor provided corrective feedback. If after two attempts of guided instruction the student was not able to apply a step in the strategy, the teacher modeled the answer.

Phase 4: Independent practice. Finally, students solved word problems independently. If a student asked for help during this phase of intervention, the teacher encouraged him/her to solve the problems independently before offering assistance.

Maintenance phase. All instructional phases (Phases 1 – 4) concluded after the predesignated intervention session number for each group (15 sessions, 13 sessions, 11 sessions). Following the completion of the intervention phase (i.e., Session 18), all students were administered three maintenance measures of one- and two-step word problems.

Follow-up phase. Two months after the completion of the maintenance phase, students were administered two follow-up measures of one- and two-step word problems.

Dependent measures

Word problem solving accuracy. The primary or targeted dependent measure included word problems solving accuracy. Each session included the administration of 5

one- and two-step addition and subtraction word problems. Each word problem included: (a) a question, (b) relevant information and numbers required to solve the problem, and (c) irrelevant information or numbers. These word problems were utilized for minority students at risk for MD in previous study (Kong & Orosco, 2015). The coefficient alpha for these problems was acceptable (.77). Accuracy was measured as the percentage correct (number correct divided by number attempted). Solution accuracy was recorded whether the student used the intervention strategy or not.

In addition, to assess generalization of the treatment condition to other problem solving measures, the AIMSweb Math Concepts and Applications (M-CAP) was administered. During the treatment phase, the M-CAP AIMSweb measure was administered every third intervention session. The M-CAP is a general outcome measure of typical math curriculum including problem solving, reasoning, and analytical skills. The M-CAP is group administered and does not exceed 8 minutes to administer. Students read and solve the word problems while recording answers on their test sheet. The alternate-form reliability coefficient for the third grade form is .81 (AIMSweb technical manual, 2012). The reliability of the rate of improvement (ROI) for third grade students whose progress had been monitored at least 10 times over the course of the school year was acceptable (.78; AIMSweb technical manual, 2012). The criterion validity of M-CAP scores with the North Carolina End of Grade Test was consistently above .60 for fall, winter, and spring scores.

Calculation. To assess generalization of skills to calculation, an experimental measure of calculation was administered. Each researcher created test consisted of 20

addition and subtraction problems. Calculation problems were scored by an independent

researcher and accuracy percentages were recorded for each session.

Additionally, to assess generalization of computation skills as a function of

treatment condition, the AIMSweb Mathematics Computation (M-COMP) test was

administered every *third* intervention session. The M-COMP test was administered on

alternate weeks of the M-CAP. The M-COMP is a timed calculation assessment that can

be group administered. Students are allowed 8 minutes to solve as many problems as

possible on their worksheet. The third grade test involves all four calculation procedures.

The alternate-form reliability coefficient for the third grade form is .91 (AIMSweb

technical manual, 2012). The reliability of the rate of improvement (ROI) for third grade

students whose progress had been monitored at least 10 times over the course of the

school year was .75. The criterion validity of M-COMP scores with the *Group*

Mathematics Assessment and Diagnostic Evaluation (G-MADE) was .73.

Reading comprehension. To assess transfer of skills to reading comprehension,

experimental measures of reading comprehension were administered. Each researcher

created test consisted of a short passage with five multiple choice comprehension

questions. Problems were scored by an independent researcher and accuracy percentages

were recorded.

Finally, to assess generalization treatment effects to reading comprehension, the

AIMSweb reading MAZE was administered. The MAZE requires students to read

leveled passages for 3 minutes and select an appropriate word (out of three choices) for

every seventh word. The MAZE was scored with the accompanying key. Number of

words correct and errors were recorded for each student. The reported test- retest

reliability for third grade passages was acceptable (.70; AIMSweb technical manual,

2012). Further, the median correlation between third grade MAZE scores and end of the

year reading tests for 11 different states was .59.

Interscorer reliability and treatment fidelity. At the end of the study, 25

percent of the data was rescored by an independent observer. Interrater reliability was

calculated by dividing the total number of agreements by the total number of agreements

and disagreements.

To ensure consistency of delivery of instruction, all intervention and assessment

sessions were scripted. However, to encourage natural teaching, interaction, and

questions from students, the scripts served as an outline for instruction. A treatment

fidelity checklist (Figure 3) based on the paraphrasing strategy for each phase of

intervention was applied by a classroom observer for 26.67 percent of all intervention

sessions. The observer coded for fidelity via a checklist and score "Yes" or "No" for

each behavior observed. A percentage of presence of intervention behaviors for all

sessions was calculated at the conclusion of the study.

Data analysis

Visual analysis was conducted to evaluate evidence of a causal relationship

between the intervention and each outcome variable (Kratochwill et al., 2010).

According to the What Works Clearinghouse criteria for interpreting single subject

designs (Kratochwill et al., 2010), four steps should be followed in conducting visual

analysis to determine a causal relation between the intervention and each outcome

variable. First, a predictable pattern of data should be documented in the baseline phase. Second, the data within each phase should be assessed for predictable patterns. Third, the data from each phase should be compared to assess whether implementation of the intervention was associated with an effect. In the case of multiple baseline design, changes in word problem solving (WPS) accuracy following the implementation of the intervention in staggered phases were examined. Finally, information from all phases of the study was integrated to determine if there were at least three demonstrations of an effect. Intervention effect is operationalized by consistency of level, trend, and variability in baseline, intervention, and maintenance phases, in addition to the immediacy of the intervention effect.

Finally, Tau-U effect size was calculated to determine overall improvement between baseline and intervention phases and during the intervention. Tau-U also adjusts for positive trends in the baseline phase, which is first determined by comparing the baseline phase with itself (Vannest & Ninci, 2015). Then, the ratio of pairwise comparisons for the baseline phase is calculated for positive trend. Baseline trends above 0.20 were corrected by subtracting the trend in the baseline phase from the Tau score in the baseline to intervention phase comparison (Vannest & Ninci, 2015).

Researchers have created a stand-alone statistical application, which calculates Tau-U and p-values (Vannest, Parker, & Gonen, 2011). As there are more than 10 data pairs, the proportion of improvement was tested for statistical significance using the z distribution. Additionally, the Tau-U calculator allowed for the selection of the adjustment of baseline trends.

The proportion of improvement in the intervention phase was then tested for statistical significance using the *z* distribution. Effect sizes were deemed small (0 to 0.65), medium (.66 to .92), large (above .93) based on recommended ranges for non overlap of all pairs (Parker, Vannest, & Brown, 2009; Soares, Harrison, Vannest, & McClelland, 2016)

Chapter 4

Results

Visual analysis was conducted to determine evidence for causal relations between the paraphrasing intervention and each outcome variable. The recommended steps for conducting visual analysis to document at least three demonstrations of intervention effect were followed (Kratochwill et al., 2010). Tau-U effect sizes were also calculated to determine improvement during the intervention phase. The results are presented into three sections (word problem solving, calculation, and reading comprehension), each addressing one of three research questions that directed this study.

Word Problem Solving

Question 1: To what extent does a paraphrasing word problem solving intervention improve students' one- and two-step word-problem solving skills?

This question was addressed through two separate measures: researcher created word problem solving test (experimental measure) and the M-CAP.

Experimental Word Problem Solving Measures

Figure 4 displays word problem solving accuracy percentage for each student as a function of baseline, intervention, maintenance, and follow-up sessions. Visual analysis indicated that seven out of nine students displayed increases in problem solving accuracy after the staggered implementation of the paraphrasing intervention. There was also a clear consistency of level, trend, and variability in all phases, suggesting a causal relationship (Kratochwill et al., 2010).

The word problem solving accuracy mean scores for all students in the baseline, intervention, maintenance, and follow-up phases, respectively, are reported in Table 2. Mary was administered three baseline session, with a baseline mean score of 16.67 percent accuracy. Mary received 15 intervention sessions with an intervention mean score of 44 percent. Mary's maintenance mean score was 20 percent. Finally, Mary's two follow-up measures' mean score was 15 percent.

James also received three baseline sessions with a mean score of 20 percent accuracy. James was absent for two intervention sessions, resulting in a total of 13 intervention sessions. James's mean intervention score was 40 percent. James's maintenance average score was 36.67 percent. James did not return to this school the following year, and did not participate in follow-up measures.

Edgar received three baseline sessions resulting in a mean score of 20 percent accuracy. Edgar received 15 intervention sessions, with an average intervention score of 42.67 percent. Edgar's maintenance and follow-up mean scores were 66.67 and 25 percent, respectively.

Alex received five baseline measures with mean score of 22 percent accuracy. Alex was absent for one intervention session, resulting in a total of 12 intervention sessions. Alex's mean intervention score was 38.33 percent. Finally, Alex's maintenance and follow-up measures' mean scores were 73.33 and 30 percent, respectively.

Jane also received five baseline sessions resulting in a mean score of 4 percent accuracy. Jane received 13 intervention sessions, increasing her mean score to 26.15

percent accurate. Jane's maintenance mean score was 60 percent. Finally, after two months without access to intervention, Jane's follow-up mean score was 5 percent accuracy.

Brian received five baseline measures with a mean score of 38 percent accuracy. Brian received 13 intervention sessions, increasing his mean score to 67.69 percent. Immediately following intervention, Brian's maintenance mean score across three sessions was 93.33 percent. Finally, Brian's follow-up mean score was 55 percent.

Diana received seven baseline measures with a mean score of 21.43 percent accuracy. Diana was absent for one intervention session, resulting in a total of 10 intervention sessions. Diana's intervention mean score was 32 percent. Diana's maintenance average score was 46.67. At follow-up, Diana's mean score was 5 percent.

Daria also received seven baseline measures resulting in a mean score of 20 percent accuracy. Daria received 11 intervention sessions with an average score of 41.82 percent. Daria's maintenance and follow-up mean scores were 46.67 and 10 percent, respectively.

Lastly, Mateo received seven baseline measures with an average score of 32.86 percent accuracy. Mateo was absent for one intervention session and received a total of 10 intervention sessions. Mateo's average intervention score was 60 percent. Mateo's maintenance and follow-up mean scores were 66.67 and 30, respectively.

Tau-U effect size. Tau-U effect size was calculated to determine overall effect of the paraphrasing intervention for each student, and a combined effect for all students. Tau-U was calculated for baseline versus intervention contrasts (A versus B) for all nine

students, controlling for baseline trend for four students (Mary, Edgar, Daria, and Mateo).

The Tau-U effect size for Mary with a corrected baseline was 0.71, $p = 0.06$ ($SD = 16.88$;

90% CI [0.09, 1]. The Tau-U for James was 0.61, $p = 0.11$ ($SD = 14.87$; 90% CI [-0.01,

1]. The Tau-U effect size for Edgar with a corrected baseline was 0.60, $p = 0.11$ ($SD =$

16.88; 90% CI [-0.02, 1]. The Tau-U for Alex was 0.53, $p = 0.09$ ($SD = 18.97$; 90% CI

[0.01, 1]. The Tau-U for Jane was 0.89, $p = 0.00$ ($SD = 20.29$; 90% CI [0.38, 1]. The

Tau-U for Brian was 0.78, $p = 0.01$ ($SD = 20.28$; 90% CI [0.27, 1]. The Tau-U effect

size for Diana was 0.34, $p = 0.24$ ($SD = 20.49$; 90% CI [-0.14, .82]. The Tau-U effect

size for Daria with a corrected baseline was 0.52, $p = 0.07$ ($SD = 20.49$; 90% CI [0.15, 1].

The Tau-U effect size for Mateo with a corrected baseline was 0.63, $p = 0.03$ ($SD =$

20.49; 90% CI [0.15, 1]. Finally, the weighted average Tau-U of the paraphrasing

intervention on word problem solving accuracy was 0.62, $p = .00$ ($SE = 0.11$; 95% CI

[0.41, 0.84]). This indicated a small, but significant, effect of the paraphrasing

intervention on word problem solving accuracy.

Summary. As predicted, students made gains in problem solving accuracy after

the administration of the intervention and in maintenance sessions immediately following

the conclusion of the intervention. However, all students displayed considerable

decreases in accuracy percentages after prolonged periods of time without access to the

intervention (two months). The Tau-U effect sizes varied from .34 to .89 with an overall

mean of 0.62. According to Parker et al. (2016), the overall mean reflects a small effect

size. However, it is important to note three of the participants yielded effect sizes in the

moderate range (Mary, Jane & Brian).

M-CAP

Figure 5 displays M-CAP accuracy points for each student as a function of baseline, intervention, maintenance, and follow-up sessions. Visual analysis indicated that all students demonstrated a predictable pattern of data in the baseline phase. The variability within all phases was low. All students, with the exception of Alex, displayed increases in problem solving accuracy after the staggered implementation of the paraphrasing intervention.

AIMSweb has presented default cut scores for each of their measures predicting probabilities of success on state tests (Pearson, 2011). These cut scores are associated with 50 and 80 percent probability of passing the state test in math. The first cut score is the lowest scoring 15 percent of the nationally normed sample, indicating severe risk in math (needing intensive intervention). The second cut score is the lowest 45 percent of students, indicating moderate risk (defined as "at-risk" or strategic). The M-CAP cut off scores for third graders in the Spring semester are below 8 for severe risk (below the 15th percentile), and 8-14 for moderate risk (15th – 45th percentile).

The mean MCAP scores for each student in the baseline, intervention, maintenance, and follow-up phases, respectively, are reported in Table 3. Mary was administered three baseline session, resulting in a baseline mean score of 7.00 points. Mary's baseline mean score indicated the need for intensive intervention. Mary received four M-CAP measures during the intervention phase with an intervention mean score of 11.5 points. Mary's intervention mean score fell within the moderate risk range. Mary

received one maintenance measure, with a score of 12. Finally, Mary received two follow-up measures with a mean score of 12.5.

James received three baseline measures with a baseline mean score of 6.33. James's baseline score fell below the cut off score severe risk. James received four intervention measures, resulting in an intervention mean score of 12. James's intervention mean score fell within the moderate risk range. James's maintenance average score was 12.5 points. James did not return to this school the following school year, and did not participate in follow-up measures.

Edgar also received three baseline measures, resulting in a mean baseline score of 7 points. Edgar's mean baseline score indicated severe risk. Edgar's four intervention measures resulted in a mean intervention score of 13.75. Edgar's mean score fell within the moderate risk range. Edgar's maintenance average score was 15. Finally, at follow-up, Edgar's mean score was 16.5 points. Edgar's maintenance and follow-up scores were above the moderate risk cut off score.

Alex received four baseline measures with a mean baseline score of 6.5 points. Alex's mean score fell below the cut off score for severe risk. Alex received three intervention measures, resulting in a mean intervention score of 8. Alex's maintenance mean score was 7 points. At follow-up, Alex's average score was 8.5 points.

Jane received four baseline measures with an average score of 4.5. Jane's baseline mean score indicated severe risk, needing intensive intervention. Jane received three intervention sessions, with a mean score of 8.67. Jane's maintenance average score

was 10. At follow-up, Jane's mean score was 8.5. Jane's intervention, maintenance, and follow-up scores fell within the moderate risk range.

Brian received four baseline measures with an average score of 7.25. Brian's baseline mean score fell below the severe risk cut off score. Brian received three intervention measures, with an average score of 7.33. Brian's maintenance and follow-up mean scores were 11 and 12.5, respectively. Brian's maintenance and follow-up scores fell within the moderate risk range.

Diana was absent for one baseline session, resulting in a total of four baseline measures. Diana's mean baseline score was 5 points, indicating severe risk. Diana was also absent for one intervention session, resulting in one intervention measure with a score of 11. Diana's maintenance mean score was 9. At follow-up, Diana's average score was 9.

Daria received five baseline sessions, with a mean score of 6 points. Daria received two intervention measures with an average score of 8, falling within the lower end of the moderate risk range. Daria's maintenance mean score was 11.5. Daria's follow-up mean score was 8.5.

Finally, Mateo received five baseline measures with a mean score of 11.4, indicating moderate risk. Mateo received two intervention measures, resulting in an average of 16.5. Mateo's maintenance and follow-up scores were also above the cut off score for moderate risk at 18.5 and 21.5, respectively.

Tau-U effect size. Tau-U effect size was calculated to determine overall effect of the paraphrasing intervention on M-CAP accuracy scores for each student, and a

51

combined effect for all students. Tau-U was calculated for baseline versus intervention contrasts (A versus B) for all nine students, controlling for baseline trend for two students (Mary and Daria). The Tau-U effect size for Mary with a corrected baseline was 0.91, p = 0.05 (SD = 5.66; 90% CI [0.14, 1]. The Tau-U for James was 1.00, p = 0.03 (SD = 5.66; 90% CI [0.32, 1]. The Tau-U effect size for Edgar was 1.00, p = 0.03 (SD = 5.66; 90% CI [0.32, 1]. The Tau-U for Alex was 0.41, p = 0.38 (SD = 5.66; 90% CI [-0.36, 1]. The Tau-U for Jane was 1.00, p = 0.03 (SD = 5.66; 90% CI [0.32, 1]. The Tau-U for Brian was 0.17, p = 0.72 (SD = 5.66; 90% CI [-.61, .94]. The Tau-U effect size for Diana was 1.00, p = 0.16 (SD = 2.82; 90% CI [-0.16, 1]. The Tau-U effect size for Daria with a corrected baseline was -0.20, p = 0.70 (SD = 5.16; 90% CI [-1, 0.65]. The Tau-U effect size for Mateo was 0.70, p = 0.18 (SD = 5.16; 90% CI [-0.15, 1]. Finally, the weighted average Tau-U of the paraphrasing intervention on M-CAP accuracy was 0.66, p = .00 (SE = 0.17; 95% CI [0.33, 1]). This indicated a medium and significant effect of the paraphrasing intervention on M-CAP accuracy.

Summary. Overall, with the exception of Mateo, all students' baseline mean scores on the M-CAP fell below the cut off score indicating severe risk status. This indicated that students were in need of intensive intervention. With the exception of Brian and Mateo, all students' intervention mean scores were in the moderate risk range. Mateo's mean intervention score fell above the cut off score, while Brian remained in the severe risk category. At maintenance and follow-up, all students, with the exception of Alex, scored at least within the moderate risk range. Edgar and Mateo's scores were above the cut off score for risk. The Tau-U effect sizes varied from -.20 to 1.0 with an

overall mean of 0.66. According to Parker et al. (2016), the overall mean reflects a moderate effect size.

Word Problem Solving Summary

The general pattern of the results was a positive increase in problem solving performance. However, the magnitude of effect sizes on the experimental measure was in the low range (Tau-U = .62, p = .00). All students displayed a drop in problem solving performance after prolonged periods of time without the intervention. On general outcomes measures of typical grade level problem solving curriculum, the magnitude of effect sizes were in the moderate range (Tau-U = .66, p = .00). In contrast to the experimental problem solving measure, students did not display significant decreases in performance after an extended period of time without the intervention.

<div align="center">

Calculation

</div>

Question 2: To what extent does a paraphrasing word problem solving intervention influence transfer to calculation skills?

This question was addressed through two measures: researcher created calculation tests and the M-COMP.

Experimental Calculation Measures

Figure 6 displays calculation accuracy percentage for each student as a function of baseline, intervention, maintenance and follow-up sessions. Visual analysis indicated that all students displayed a predictable pattern of data in the baseline phase. Most students reached high accuracy percentage in the baseline phase (revealing ceiling effects). Additionally, many students displayed negative trends and low variability in the

baseline phase. Taken together, increases in calculation after the staggered implementation of the paraphrasing intervention were not observed.

The mean calculation accuracy percentage scores for each student in the baseline, intervention, maintenance, and follow-up phases, respectively, are reported in Table 4. Mary was administered three baseline measures with a resulting mean score of 85 percent accuracy. Mary received seven intervention measures with an average score of 65.71 percent. All students were administered three maintenance and two follow-up measures. Mary's maintenance and follow-up mean scores were 79.67 and 60 percent accuracy, respectively.

James was administered three baseline measures with an average score of 56.67. James was absent for one intervention session, resulting in a total of six intervention measures. James's intervention mean score was 45 percent accuracy. Following the intervention, James's maintenance mean score was 35.33. James did not return to this school the following year, and did not participate in follow-up measures

Edgar was also administered three baseline measures with an average score of 58.33. Edgar was administered seven intervention sessions with a mean score of 57.85 percent. Edgars' maintenance and follow-up mean accuracy percentage scores were 61 and 45 percent, respectively.

Alex received four baseline measures with an average score of 87.5 percent. Alex received six intervention measures with a mean score of 74.17 percent. Alex's maintenance and follow-up mean scores were 84.33 and 75 percent, respectively.

Jane also received four baseline measures with a mean accuracy score of 55 percent. Jane received six intervention measures with a resulting average score of 52.50 percent. Jane's maintenance mean score was 88 percent. At follow-up, Jane's mean score dropped to 47.5 percent accuracy.

Brian received four baseline measures with an average score of 82.5. Brian's six intervention measures resulted in an average of 79.17. Brian's maintenance and follow-up mean accuracy scores were 96.67 and 92.50 percent, respectively.

Diana received five baseline measures with a mean score of 52 percent accuracy. Diana was absent for one intervention session, resulting in a total of four intervention measures. Diana's mean intervention score was 35 percent. Diana's maintenance average accuracy score was 56 percent. At follow-up, Diana's mean score was 47.5 percent.

Daria also received five baseline measures with a mean score of 69 percent. Daria received five intervention measures with an average score of 75 percent. Daria's maintenance and follow-up scores were 86.33 and 70 percent accuracy, respectively.

Finally, Mateo received five baseline measures with a mean score of 73 percent. Mateo was absent for one intervention session, resulting in a total of four intervention measures collected. Mateo's intervention session average was 78.75 percent accuracy. Following the intervention, Mateo's maintenance mean score was 83 percent. Mateo's follow-up average score was 75 percent accuracy. The Tau-U effect sizes varied from -0.85 to 0.16 with an overall mean of -0.31.

Tau-U effect size. Tau-U effect size was calculated to determine overall effect of the paraphrasing intervention on calculation measures for each student, and a combined effect for all students. Tau-U was calculated for baseline versus intervention contrasts (A versus B) for all nine students, controlling for baseline trend for two students (Mary and Edgar). The Tau-U effect size for Mary with a corrected baseline was -0.81, $p = 0.05$ ($SD = 8.78$; 90% CI [-1, -.01]. The Tau-U for James was -0.22, $p = 0.61$ ($SD = 7.75$; 90% CI [-0.93, 0.49]. The Tau-U effect size for Edgar with a corrected baseline was -0.14, $p = 0.73$ ($SD = 8.78$; 90% CI [-0.83, 0.57]. The Tau-U for Alex was -0.67, $p = 0.08$ ($SD = 9.38$; 90% CI [-1, -0.02]. The Tau-U for Jane was -0.17, $p = 0.67$ ($SD = 9.38$; 90% CI [-0.81, 0.48]. The Tau-U for Brian was 0.00, $p = 1.00$ ($SD = 9.38$; 90% CI [-0.64, 0.64]. The Tau-U effect size for Diana was -0.85, $p = 0.04$ ($SD = 8.17$; 90% CI [-1, -0.18]. The Tau-U effect size for Daria was 0.16, $p = 0.68$ ($SD = 9.57$; 90% CI [-0.47, 0.79]. The Tau-U effect size for Mateo was -0.10, $p = 0.81$ ($SD = 8.17$; 90% CI [-0.77, 0.57]. Finally, the weighted average Tau-U of the paraphrasing intervention on calculation accuracy was -0.31, $p = .02$ ($SE = 0.13$; 95% CI [-0.57, -0.04]). This indicated a small negative effect of the paraphrasing intervention on calculation accuracy.

Summary. Students did not make significant gains in calculation after the administration of the intervention. Students maintained similar accuracy scores throughout baseline and intervention phases. However, all but one student (Brian) displayed marked decreases in calculation accuracy at follow-up. The Tau-U effect sizes varied from -.85 to .16 with an overall mean of -0.31, $p = .02$. This indicated that the paraphrasing intervention did not facilitate positive gains in calculation accuracy.

M-COMP

Figure 7 displays M-COMP number correct points for each student as a function of baseline, intervention, maintenance, and follow-up sessions. Visual analysis indicated that Brian and Mary displayed positive trends in the baseline phase. All other students displayed low variability and stable trends in the baseline phase. Generally, increases in M-COMP scores were not noted after the implementation of the intervention.

AIMSweb has presented default cut scores for each of their measures predicting probabilities of success on state tests. The first cut score is the lowest scoring 15 percent of the nationally normed sample, indicating severe risk in math (needing intensive intervention). The second cut score is the lowest 45 percent of students, indicating moderate risk (defined as "at-risk" or strategic). The M-COMP cut off scores for third graders in the Spring semester are below 31 for severe risk status (below the 15th percentile), and 32-53 for moderate risk status (15th – 45th percentile).

The mean M-COMP scores for each student in the baseline, intervention, maintenance, and follow-up phases are reported in Table 5. Mary received three baseline measures with a mean baseline score of 44. Mary's baseline mean score indicated moderate risk. Mary received four M-COMP measures during the intervention phase with an average score of 50.5 points. Mary was absent for one maintenance session, resulting in one maintenance measure of 56 points. Finally, Mary received two follow-up measures with an average score of 38.5 points.

James also received three baseline measures with a mean score of 30. James's baseline mean score indicated severe risk. James was absent for two intervention

measures resulting in a total of two intervention measures. James's intervention mean score was 38.5. James received two maintenance measures with a mean score of 43. James did not return to this school the following school year, and did not participate in follow-up measures.

Edgar received three baseline measures with an average score of 40. Edgar received four intervention measures. Edgar's intervention mean score was 42.5 points. Edgar's maintenance and follow-up mean scores were 44 and 36 points, respectively.

Alex received four baseline measures with a mean score of 43.5, indicating moderate risk. Alex received three intervention measures for an average score of 45 points. Alex's maintenance mean score was 53 points. At follow-up, Alex's mean score was 49 points.

Jane received four baseline measures with an average score of 37.75 points. Jane received three intervention measures. Jane's intervention mean score was 46 points. Jane's maintenance and follow-up mean scores were 50.5 and 41.5 points respectively. Jane's mean scores in all phases fell within the moderate risk range.

Brian also received four baseline measures with a mean score of 53 points. Brian's baseline mean score fell within the upper end of the moderate risk range. Brian received three intervention measures with a mean score of 59 points. Brian's maintenance and follow-up scores were 64.5 and 64 points, respectively. Brian's mean scores in the intervention, maintenance, and follow-up phases were above the cut of score indicating moderate risk.

Diana received five baseline sessions with a mean score of 34.4 points. Diana's baseline mean score fell within the range of moderate risk. Diana was absent for one day of intervention data collection. Diana received one intervention measure with a score of 37. Diana received two maintenance sessions, resulting in a mean score of 40. Finally, Diana's follow-up mean score was 27 points. Diana's scores in the intervention and maintenance phases fell within the range of moderate risk; however, Diana's mean score at follow-up indicated severe risk.

Daria received five baseline sessions with an average score of 44.6 points. Daria was absent on both days data was collected for intervention sessions. Daria received two maintenance measures with a mean score of 58 points. Finally, at follow-up, Daria's mean score was 54.5 points. Daria's maintenance and follow-up mean scores fell above the cut off score for risk.

Mateo received five baseline measures with a mean score of 52.6, indicating moderate risk. Mateo received two intervention measures, resulting in a mean score of 65.5 points. Mateo's maintenance and follow-up mean scores were 64.5 and 57 points, respectively. Mateo's intervention, maintenance, and follow-up mean scores were above the cut off score indicating risk.

Tau-U effect size. Tau-U effect size was calculated to determine overall effect of the paraphrasing intervention on M-COMP accuracy scores for each student, and a combined effect for all students. Tau-U was calculated for baseline versus intervention contrasts (A versus B) for eight students, controlling for baseline trend for five students (Mary, Edgar, Alex, Jane, and Brian). The Tau-U effect size for Mary with a corrected

59

baseline was 0.08, *p* = 0.86 (*SD* = 5.66; 90% CI [-0.69, 0.86]. The Tau-U for James was

1.00, *p* = 0.08 (*SD* = 3.46; 90% CI [0.05, 1]. The Tau-U effect size for Edgar with a

corrected baseline was 0.50, *p* = 0.28 (*SD* = 5.66; 90% CI [-0.28, 1]. The Tau-U for Alex

with a corrected baseline was -0.08, *p* = 0.86 (*SD* = 5.66; 90% CI [-0.28, 1]. The Tau-U

for Jane with a corrected baseline was 0.75, *p* = 0.11 (*SD* = 5.66; 90% CI [-1, 0.44]. The

Tau-U for Brian with a corrected baseline was -0.33, *p* = 0.48 (*SD* = 5.66; 90% CI [-1,

0.44]. The Tau-U effect size for Diana was 0.60, *p* = 0.38 (*SD* = 3.42; 90% CI [-0.52, 1].

The Tau-U effect size for Daria was not calculated because there was no data collected in

the intervention phase. The Tau-U effect size for Mateo was 1.00, *p* = 0.05 (*SD* = 5.16;

90% CI [0.15, 1]. Finally, the weighted average Tau-U of the paraphrasing intervention

on M-COMP accuracy was 0.41, *p* = .03 (*SE* = 0.18; 95% CI [0.05, 0.77]). This

indicated a small and significant effect of the paraphrasing intervention on M-CAP

accuracy.

Summary. With the exception of James, all students' baseline mean scores fell

below the cut off score indicating moderate risk status. James's baseline mean score

indicated performance in the severe risk status range. With the exception of Brian and

Mateo, all students' intervention mean scores were in the moderate risk range. During

the intervention phase, Brian and Mateo's scores were above the cut off score indicating

risk. All other students remained in the same risk category. At maintenance,

approximately half of the students (Mary, Brian, Daria, and Matthew) scored above the

cut off score indicating risk. Decreases in scores were noted for all students at follow-up,

with more severe drops in scores for students who scored lower at maintenance. The

Tau-U effect sizes varied from -0.33 to 1.00 with an overall mean of 0.41, $p = .03$.

According to Parker et al. (2016), the overall mean reflects a small effect size. However, it is important to note three of the participants yielded effect sizes in the moderate to large ranges (James, Jane & Mateo).

Calculation Summary

The results indicated low magnitude of effect size on transfer measures of calculation. On the experimental calculation measure, this study did not support the intervention facilitating significant increases in addition and subtraction problems. Students displayed significant increases in calculation on the computation curriculum based measure, but the magnitude of effect sizes was in the low range. Overall, this study does not support the hypothesis that this paraphrasing intervention influences transfer to calculation skills.

<div align="center">

Reading Comprehension

</div>

Question 3: To what extent does a paraphrasing word problem solving intervention influence transfer to reading comprehension?

This question was addressed through two measures: researcher created reading comprehension tests and the MAZE.

Experimental Reading Comprehension Measures

Figure 8 displays reading comprehension accuracy percentage for each student as a function of baseline, intervention, maintenance, and follow-up sessions. Visual analysis indicated moderate variability in the baseline phase and high variability during the intervention phase. Three of nine students displayed positive trends in the baseline

phase. Generally, students did not display immediate increases in reading comprehension following the implementation of the intervention.

The mean reading comprehension accuracy scores for each student in the baseline, intervention, maintenance, and follow-up phases are reported in Table 6. Mary was administered three baseline measures with a mean score of 20 percent accuracy. Mary received eight intervention measures with a mean score of 50 percent accuracy. All students received three maintenance and two follow-up measures. Mary's maintenance and follow-up mean scores were 26.67 and 40 percent, respectively.

James received three baseline measures with an average score of 40 percent accuracy. James was absent for one intervention measure, for a total of seven intervention sessions. James's intervention mean accuracy score was 42.86 percent. James's maintenance mean score was 20 percent accuracy. James did not return to this school the following year, and did not participate in follow-up measures.

Edgar also received three baseline measures with a mean score of 33.33 percent. Edgar received eight intervention measures with a mean score of 52.5 percent accuracy. At maintenance, Edgar's mean score was 60 percent. Edgar's follow-up average score was 40 percent accuracy.

Alex received three baseline measures resulting in a mean score of 33.33 percent accuracy. Alex was absent for one intervention session and received six intervention measures. Alex's intervention mean score was 43.33 percent. Alex's maintenance and follow-up average scores were 13.33 and 10 percent accuracy, respectively.

Jane also received three baseline measures with a mean score of 20 percent accuracy. Jane received seven intervention sessions with an average score of 45.71 percent. At maintenance and follow-up, Jane's mean scores were 40 and 30 percent accuracy, respectively.

Brian received three baseline sessions with a mean score of 60 percent accuracy. Brian received seven intervention sessions with an average score of 34.29 percent. Brian's maintenance average score was 53.33 percent. Brian's follow-up mean score was 20 percent accuracy.

Diana received four baseline sessions with a mean score of 45 percent. Diana received six intervention sessions, resulting in a mean score of 33.33 percent accuracy. Diana's maintenance and follow-up scores were 26.67 and 30 percent accuracy, respectively.

Daria received four baseline measures with a resulting mean score of 50 percent accuracy. During the intervention phase, Daria received six intervention sessions with a mean score of 66.67 percent. Daria's maintenance average score was 60 percent. Finally, at follow-up Daria's mean score was 50 percent accuracy.

Mateo received four baseline sessions with a mean score of 60 percent. Mateo received six intervention sessions. Mateo's intervention mean score was 63.33 percent. Mateo's maintenance and follow-up scores were 66.67 and 50 percent accuracy, respectively.

Tau-U effect size. Tau-U effect size was calculated to determine overall effect of the paraphrasing intervention on reading comprehension measures for each student, and a

combined effect for all students. Tau-U was calculated for baseline versus intervention contrasts (A versus B) for all nine students, controlling for baseline trend for four students (Mary, Edgar, Daria, and Mateo). The Tau-U effect size for Mary with a corrected baseline was 0.54, $p = 0.18$ ($SD = 9.80$; 90% CI [-.13, 1]. The Tau-U for James was 0.28, $p = 0.49$ ($SD = 8.78$; 90% CI [-0.40, 0.97]. The Tau-U effect size for Edgar with a corrected baseline was 0.58, $p = 0.15$ ($SD = 9.80$; 90% CI [-0.09, 1]. The Tau-U for Alex was 0.07, $p = 0.88$ ($SD = 6.71$; 90% CI [-0.67, 0.80]. The Tau-U for Jane was 0.67, $p = 0.12$ ($SD = 7.75$; 90% CI [-0.04, 1]. The Tau-U for Brian was -0.61, $p = 0.14$ ($SD = 8.78$; 90% CI [-1, 0.07]. The Tau-U effect size for Diana was -0.25, $p = 0.52$ ($SD = 9.38$; 90% CI [-0.89, 0.39]. The Tau-U effect size for Daria with a corrected baseline was 0.38, $p = 0.34$ ($SD = 9.38$; 90% CI [-0.27, 1]. The Tau-U effect size for Mateo with a corrected baseline was 0.00, $p = 1.00$ ($SD = 9.38$; 90% CI [-0.64, 0.64]. Finally, the weighted average Tau-U of the paraphrasing intervention on reading comprehension accuracy was 0.18, $p = .19$ ($SE = 0.14$; 95% CI [-0.09, 0.45]). This indicated no significant effect of the paraphrasing intervention on reading comprehension accuracy.

Summary. Compared to the baseline phase, students made minor gains in reading comprehension after the administration of the paraphrasing intervention and in maintenance sessions. The Tau-U effect sizes varied from -0.61 to 0.67 with an overall weighted average Tau-U of 0.18. According to Parker et al. (2016), the overall mean reflects a small effect size.

MAZE

Figure 9 displays MAZE accuracy points for each student as a function of baseline, intervention, maintenance, and follow-up sessions. Visual analysis indicated that all students demonstrated a predictable pattern of data in the baseline phase. There was a clear consistency of level, trend, and variability in all phases. Six out of nine students demonstrated increases in MAZE accuracy scores after the implementation of the intervention.

AIMSweb has presented default cut scores for each of their measures predicting probabilities of success on state tests (Pearson, 2011). The first cut score is the lowest scoring 15 percent of the nationally normed sample, indicating severe risk in math (needing intensive intervention). The second cut score is the lowest 45 percent of students, indicating moderate risk status (defined as "at-risk" or strategic). The MAZE cut off scores for third graders in the Spring semester are below 9 for severe risk (below the 15th percentile), and 10-15 for moderate risk status (15th – 45th percentile).

The mean MAZE scores for each student in the baseline, intervention, maintenance, and follow-up phases are reported in Table 7. Mary received three baseline sessions with a mean score of 9 points. This indicated severe risk. Mary received four intervention measures with an average score of 9.75. Mary was absent for one maintenance measure, resulting in one maintenance session with a score of 7. At follow-up, Mary received two measures with a mean score of 9.50 points.

James received three baseline sessions with a mean score of 10 points. James received four intervention measures with a mean score of 9.75. James received three

maintenance sessions resulting in an average score of 13. James did not return to this school the following year, and did not participate in follow-up measures.

Edgar also received three baseline measures with an average score of 10 points. Edgar received four intervention measures with a mean score of 16 points. This was above the cut off score for risk. Edgar's mean score during the maintenance phase was 22.50, well above the cut off for risk. Finally, at follow-up, Edgar's mean score was 11.5, which fell within the range of moderate risk.

Alex was absent for one baseline session and received a total of three baseline measures. Alex's baseline mean score was 4.33, well below the cut off score for severe risk. Alex received three intervention measures, for a mean score of 11.33 points. This placed him within the moderate risk category. Alex's maintenance and follow-up scores were 11 and 10.50 points, respectively.

Jane received four baseline measures with a mean score of 9.5 points. This fell at the lower end of the moderate risk range. Jane received three intervention measures with an average score of 16.33. Jane's intervention mean score was above the cut off score indication moderate risk. Jane's maintenance and follow-up scores were 23.50 and 20 points, respectively.

Brian also received four baseline measures with an average score of 7. Brian's baseline mean score fell below the cut off score for severe risk. Brian received three intervention measures with a mean score of 14.67 points, falling within the upper end of the moderate risk range. Brian's maintenance mean score was 26.50, well above the cut off score for risk. Finally, Brian's follow-up mean score was 12.50.

Diana received five baseline measures with an average score of 9.80 points. Diana received two intervention measures for a mean score of 12 points. Following intervention, Diana's maintenance mean score was 18 points. At follow-up, Diana's mean score dropped significantly to 9 points.

Daria also received five baseline measures with a mean score of 8.4 points. This is below the cut off score indicating severe risk. Daria received two intervention measures for a mean score of 5.5. Following intervention, Daria's maintenance and follow-up scores were 26 and 13.5 points, respectively.

Finally, Mateo was absent for one baseline measure for a total of four baseline sessions. Mateo's baseline phase mean score was 12.25 points. Mateo received two intervention measures with a mean score of 15 points. Mateo's maintenance mean score was 19 points, above the cut off score for moderate risk. At follow-up, Mateo's average score was 11 points.

Tau-U effect size. Tau-U effect size was calculated to determine overall effect of the paraphrasing intervention on MAZE accuracy for each student, and a combined effect for all students. Tau-U was calculated for baseline versus intervention contrasts (A versus B) for all nine students, controlling for baseline trend for seven students (Mary, Edgar, Alex, Jane, Brian, Daria, and Mateo). The Tau-U effect size for Mary with a corrected baseline was 0.08, $p = 0.86$ ($SD = 5.66$; 90% CI [-0.69, 0.86]. The Tau-U for James was 0.08, $p = 0.86$ ($SD = 5.66$; 90% CI [-0.69, 0.86]. The Tau-U effect size for Edgar with a corrected baseline was 0.92, $p = 0.05$ ($SD = 5.66$; 90% CI [0.14, 1]. The Tau-U for Alex with a corrected baseline was 0.67, $p = 0.19$ ($SD = 4.58$; 90% CI [-0.17,

1]. The Tau-U for Jane with a corrected baseline was 0.50, $p = 0.29$ ($SD = 5.66$; 90% CI [-0.28, 1]. The Tau-U for Brian with a corrected baseline was 0.42, $p = 0.38$ ($SD = 5.66$; 90% CI [-0.36, 1]. The Tau-U effect size for Diana was 0.20, $p = 0.70$ ($SD = 5.16$; 90% CI [-0.65, 1]. The Tau-U effect size for Daria with a corrected baseline was -0.90, $p = 0.08$ ($SD = 5.16$; 90% CI [-1, -0.05]. The Tau-U effect size for Mateo with a corrected baseline was 0.13, $p = 0.82$ ($SD = 4.32$; 90% CI [-0.76, 1]. Finally, the weighted average Tau-U of the paraphrasing intervention on MAZE accuracy was 0.24, $p = .14$ ($SE = 0.16$; 95% CI [-0.08, 0.56]). This indicated minimal effect of the paraphrasing intervention on MAZE accuracy. Table 8 displays weighted average Tau-U effect sizes for each measure.

  Summary. All students were at risk for at least moderate risk status at baseline. Four of nine students fell below the cut off score indicating severe risk during the baseline phase. During the intervention phase, two students scored above the cut off score. Daria's score remained at severe risk, and five of nine students were at moderate risk. Following intervention, six of nine students scored above the cut off score for risk. With the exception of Mary, decreases in scores were noted for all students at follow-up. The Tau-U effect sizes varied from -0.90 to 0.92 with an overall mean of 0.24. According to Parker et al. (2016), the overall mean reflects a small effect size.

Reading Comprehension Summary

  Finally, the paraphrasing intervention phase had minimal influence on reading comprehension measures. The magnitude of effect sizes for both the experimental and curriculum based measures were in the low range. It is important to note, however, that

six out of nine students no longer fell within any risk category on the curriculum based reading measure following the intervention.

Treatment Fidelity

A treatment fidelity checklist (Figure 3) based on the paraphrasing strategy for each phase of intervention was applied by a classroom observer for 26.67% of all intervention sessions. The observer coded for fidelity via a checklist and score "Yes" or "No" for each behavior observed. The percentage of presence of intervention behaviors for all sessions was 94.92%. The total agreement calculation method for each session indicated stable presence of intervention behaviors across intervention sessions. When intervention behaviors were not presented in a particular observation, results were shared immediately following the intervention session. Additionally, the instructor was retrained on the intervention procedures and protocol.

Interscorer reliability

At the end of the study, 25% of the data was rescored for interscorer reliability. Six word problem solving measures, five calculation measures, five reading comprehension measures, three M-COMP, three M-CAP, and three MAZE measures were rescored. Reliability was calculated by dividing the total number of agreements by the total number of agreements and disagreements. Interscorer agreement was 100% across all measures.

Chapter 5

Discussion

The purpose of this study is to investigate the effectiveness of a word problem solving intervention on word problem solving accuracy for third grade EL students who are at risk for mathematic disabilities. Additionally, the study explored whether a paraphrasing intervention influenced transfer to calculation and reading comprehension measures.

Overall, the current study provides positive support for the effectiveness of the paraphrasing intervention on the word problem solving accuracy of these English learners at risk for mathematic disabilities. However, the magnitude of the effect sizes for the majority of participants was small. In addition, the results related to transfer were mixed. The magnitude of the effect sizes suggested that students did not make substantial gains on the experimental calculation measure. However, some students did yield effect sizes in the moderate range in accuracy scores on curriculum based measures of computation. Additionally, the results showed minimal effects related to the paraphrasing intervention on measures of reading comprehension. The results of this study are summarized by addressing the three major research questions.

Question 1: To what extent does a paraphrasing word problem solving intervention improve students' one- and two-step word-problem solving skills?

Students displayed gains in one- and two-step word problem solving following the administration of the paraphrasing intervention and in maintenance sessions after the conclusion of the intervention. Tau-U effect sizes ranged from the small to moderate

ranges (Parker, Vannest, & Brown, 2009; Soares, Harrison, Vannest, & McClelland, 2016). The weighted average Tau-U effect size of the paraphrasing intervention on problem solving accuracy was 0.62. That is, 62% of the intervention phase data showed improvement when compared to the baseline phase. This improvement trend was a positive increase compared to baseline performance ($p = .00$). However, all students displayed decreases in problem solving accuracy percentages after an extended period of time without access to the intensive intervention.

In addition, students displayed positive increases on a general outcome measure used to test performance on typical grade level curriculum (M-CAP). Tau-U effect sizes ranged from small to large effects. The weighted Tau-U effect size of the paraphrasing intervention on M-CAP accuracy was moderate (Tau-U = 0.66, $p = .00$). In contrast to the experimental measure, students did not display decreases in performance in the follow-up phase from the treatment phase, two months following the conclusion of the intervention.

It is important to note that this intervention only included students who were English learners at risk for MD. English learners face the significant and unique challenge of concurrently developing proficiency in a new language and acquiring academic vocabulary/content (Lesaux & Harris, 2013). Additionally, the level of parental involvement and assistance in tasks that require English academic vocabulary is unknown. This may indicate that ELs at risk for MD are a unique group with several risk factors, such as lack of exposure to academic language, math difficulties, and cultural and

71

linguistic acclimation. For these reasons, it may be more difficult to attain major changes for this specific group.

Even with a sample of only ELs, the results of this study were generally consistent with the literature regarding the positive effects of paraphrasing interventions on the word problem solving skills of students with MD (Moran, Swanson, Gerber, & Fung, 2014; Swanson, Moran, Lussier, & Fung, 2013). Swanson and colleagues (2013) found that students at risk for MD with relatively higher working memory capacity benefitted from interventions that emphasized paraphrasing all propositions in a word problem. However, paraphrasing only parts of the problem was not found to significantly improve problem solving performance. This previous study included both English proficient students and English learners whose CELDT scores indicated English proficiency. ELs in particular may experience more difficulty with math problem solving because of the need to preserve information while at the same time processing information in a second language.

Additionally, although it is difficult to compare results directly due to different samples and methodology, Moran and colleagues (2014) found that paraphrasing relevant propositions produced an effect size of 0.93. However, this study included both English learners and English proficient students. Additionally, students received intervention two times a week for 25-30 minutes over the course of 10 weeks (20 intervention sessions). The current study delivered a maximum of 15 intervention sessions over the course of five weeks (approximately three times a week, 30 minutes each session).

Finally, the literature in terms of single subject designs for word problem solving interventions for ELs at risk for MD is extremely limited. The only single subject design studies for problem solving interventions exclusively for ELs at risk for MD that could be identified were conducted by Orosco and his colleagues (2013; 2014). Again, it is difficult to compare the current study directly since the previous studies included the use of word problem solving level, as opposed to accuracy percentage per session, to measure growth. However, both of these studies have emphasized the importance of focusing on academic language and comprehension strategies to address the word problem solving skills of ELs at risk for MD. This current study extends the literature base by focusing on the use of a paraphrasing comprehension strategy intervention to improve the math problem solving skills of ELs at risk for MD.

Question 2: To what extent does a paraphrasing word problem solving intervention influence transfer to calculation skills?

This study provided mixed support for the positive influence of paraphrasing intervention on calculation measures. The magnitude of the Tau-U effect sizes were within the low range. The weighted Tau-U effect size of the paraphrasing intervention on the experimental calculation measure was not positive (Tau-U = -0.31, *p* = .02). Visual inspection indicated that many students had high scores in the baseline phase, suggesting that a ceiling effect may have been an artifact in the treatment outcomes.

In contrast to the experimental measures, students displayed some gains in calculation as measured by the M-COMP, which involved all four calculation procedures (addition, subtraction, multiplication, and division). The magnitude of the Tau-U effect

73

sizes ranged from the small to moderate. The weighted average Tau-U effect size of the

paraphrasing intervention on M-COMP accuracy scores was small, but significant (Tau-

U = 0.41, p = 0.03). Finally, students displayed decreases in M-COMP scores after two

months without intervention. Students with lower scores demonstrated most severe

declines, suggesting that a certain level of proficiency in calculation may help in loss of

skill without intervention over time.

Swanson (2006) found that computation and problem solving required different

cognitive resources. The best predictor of problem solving was identified as working

memory's executive system, whereas the best predictors of computation were reading,

inhibition, vocabulary, and visual-spatial working memory. Additionally, Fuchs and

colleagues (2008) identified computation and math problem solving as distinct cognitive

dimensions. While this study did not explicitly instruct students in the area of

computation, the effectiveness of an intervention that emphasized a reading

comprehension strategy while explicitly teaching students to inhibit irrelevant

information was investigated. The extent to which an intervention designed to target

problem solving skills in conjunction with calculation practice would facilitate transfer to

calculation abilities warrants further study.

Question 3: To what extent does a paraphrasing word problem solving intervention

influence transfer to reading comprehension?

Overall, this study did not provide support for the assumption that paraphrasing

intervention influences reading comprehension performance. Students' scores were

highly variable on the experimental measure of reading comprehension. The majority of

Tau-U effect sizes were in the low range. The weighted average Tau-U of the

paraphrasing intervention on the experimental reading comprehension measure was low

(Tau-U = 0.18, p = .19), suggesting no significant effect.

Additionally, in general, students did not display significant increases in MAZE

accuracy scores from baseline to intervention. However, Tau-U effect sizes ranged from

the low to high range. The weighted average Tau-U effect size of the paraphrasing

intervention on MAZE accuracy scores was low (Tau-U = 0.24, p = .14). It is important

to note, however, that following the intervention, six of nine students no longer fell

within the risk category.

The results of this study qualify the notion that generative problem solving

interventions facilitating transfer to novel or real-world tasks is challenging (Fuchs et al.,

2002). However, since research has indicated that one of the strongest predictors of

problem solving in young children is reading comprehension (Swanson et al., 1993),

further research is needed to investigate whether these are concurrent skills or if one skill

can be targeted by teaching the other.

Limitations

Despite some encouraging outcomes of this study, results should be interpreted

with caution. There were several limitations of the current study. First, this was a small

scale study (N = 9) in which data for individuals were collected for a duration of 23

sessions. Thus, the extent to which this paraphrasing intervention could have mediated

word problem solving skills in other EL students at risk for MD for this duration of time

is unknown. Thus, generalization of intervention effectiveness to other populations of students who are ELs at risk for MD is limited.

In addition, this paraphrasing intervention was delivered in a highly intensive manner. Students were taught the intervention in small groups with an average of three times a week over the course of six weeks. The extent to which frequency and duration of the intervention would impact the efficacy of the intervention should be investigated.

Finally, when students were absent for any session, data were not able to be collected for that particular session. Since single subject design studies examine repeated measures on an individual over sessions, the extent to which missing data affected the results is unclear.

Implications for practice and research

The results of this study offer implications for practice and future research. Intervention that focused on a reading comprehension strategy involving paraphrasing helped English learners who were at risk for math disabilities improve problem solving performance. Teachers may consider incorporating strategies that focus on comprehension and language to address the math problem solving difficulties of students who are English learners. It is imperative to provide targeted interventions for ELs at risk for MD to prevent academic achievement gaps.

However, the results also suggested that students did not maintain skills after an extended time without access to the intervention. Students may need continual support until a certain level of mastery or proficiency is reached. Additionally, this finding

emphasizes the need to include techniques into the instructional protocol that may promote and explicitly teach generalization.

Finally, students' scores were variable depending on the skill and throughout the intervention sessions. Few studies have investigated the role of student motivation on daily outcomes. A study examining student motivation across tasks and throughout the duration of the intervention could provide more information about the learning experiences and outcomes of EL students at risk for MD.

Conclusion

In summary, this study found that ELs at risk for MD may have difficulty with word problems because word problems require the use of linguistic information in addition to basic math skills. Paraphrasing interventions that focus on language and vocabulary development may be a critical component in addressing the problem solving difficulties of some students. Clearly, given the small to moderate outcomes related to the intervention, further research should be conducted to identify and develop effective word problem solving interventions for ELs at risk for MD.

References

Abedi, J, Courtney, M, & Leon, S. (2003). *Research-supported accommodation for English language learners in NAEP.* Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Badian, N. A. (1983). Dyscalculia and nonverbal disorders of learning. In H.R. Myklebust (Ed.). *Progress in Learning Disabilities. 5*(235-264). New York: Stratton.

Bakken, J. P., Mastropieri, M. A., & Scruggs, T. E. (1997). Reading comprehension of expository science material and students with learning disabilities: A comparison of strategies. *The Journal of Special Education*, *31*(3), 300-324.

Barbaresi, W. J., Katusic, S. K., Colligan, R. C., Weaver, A. L., & Jacobsen, S. J. (2005). Math learning disorder: Incidence in a population-based birth cohort, 1976-1982, Rochester, Minn. *Ambulatory Pediatrics. 5*(281-289).

Barrouillet, P., Fayol, M., & Lathulière, E. (1997). Selecting between competitors in multiplication tasks: An explanation of the errors produced by adolescents with learning difficulties. *International Journal of Behavioral Development*, *21*(2), 253-275.

Bowman-Perrott, L., Davis, H., Vannest, K., Williams, L., Greenwood, C., & Parker, R. (2013). Academic benefits of peer tutoring: A meta-analytic review of single-case research. *School Psychology Review*, *42*(1), 39-55.

Brown, V. L., Cronin, M. E., & McEntire, E. (1994). *Test of mathematical abilities*. Austin, TX: PRO-ED.

Case, L. P., Harris, K. R., & Graham, S. (1992). Improving the mathematical problem-solving skills of students with learning disabilities: Self-regulated strategy development. *Journal of Special Education, 26,* 1-19.

Cassel, J., & Reid, R. (1996). Use of a self-regulated strategy intervention to improve word problem-solving skills of students with mild disabilities. *Journal of Behavioral Education, 6,* 153-172.

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.

Connolly, A. J. (1998). *KeyMath (revised/normative update)*.Circle Pines, MN: American Guidance

Expeditionary Learning. (2013). *EngageNY curriculum grade 3 English language arts.* Retrieved from https://www.engageny.org/resource/grade-3-english-language-arts

Fleishner, J. D., Garnett, K., & Shepherd, M. J. (1982). Proficiency in calculations BF computation of learning disabled and nondisabled children. *Focus on Learning Problems in Mathematics, 4,* 47-56.

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. doi: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). Baltimore, MD: 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*, 29-43.

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 mathematic disabilities. *Learning Disabilities Research & Practice, 17*(2), 90-106.

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

Fuchs, L. S., Powell, S. R., Hamlett, C. L., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2008). Remediating computational deficits at third grade: A randomized field trial. *Journal of Research on Educational Effectiveness*, *1*(1), 2-32.

Garcia, E., & Cuéllar, D. (2006). Who are these linguistically and culturally diverse students?. *The Teachers College Record*, *108*(11), 2220-2246.

Geary, D. C. (1993). Mathematical disabilities: cognitive, neuropsychological, and genetic components. *Psychological bulletin*, *114*(2), 345-362.

Geary, D. C. (1994). *Children's mathematical development: Research and practical applications.* Washington, DC: American Psychological Association.

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, *37*(1), 4-15.

Geary, D. C., & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics. In J.I.D. Campbell (Ed.), *Handbook of mathematical cognition* (253-267). New York: Psychology Press.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., & DeSoto, M. C. (2004). Strategy choices in simple and complex addition: Contributions of working memory and counting knowledge for children with mathematical disability. *Journal of Experimental Child Psychology*, *88*(2), 121-151.

Geary, D. C., Hoard, M. K., Byrd-Craven, J., Nugent, L., & Numtee, C. (2007). Cognitive mechanisms underlying achievement deficits in children with mathematical learning disability. *Child Development*, *78*(4), 1343-1359.

Geary, D. C., Widaman, K. F., Little, T. D., & Cormier, P. (1987). Cognitive addition: Comparison of learning disabled and academically normal elementary school children. *Cognitive Development*, *2*(3), 249-269.

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., Jordan, N. C., & Flojo, J. (2005). Early identification and interventions for students with mathematical difficulties. *Journal of Learning Disabilities, 38*(4), 293-304.

Graney, M. J. (1979). Nomogram for Obtaining Z-Test Statistic from Kendall's S and Sample Size 10 To 50. *Educational and Psychological Measurement*, *39*(4), 761-765.

Great Minds. (2015). *EngageNY curriculum grade 3 mathematics.* Retrieved from https://www.engageny.org/resource/grade-3-mathematics

Griffin, C. (2007) Early intervention for children at risk for developing mathematical learning difficulties. 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* (373-395). Baltimore: Paul H. Brookes Publishing Co.

Hallahan, D., Pullen, P., & Ward, D. (2013). A brief history of the field of learning disabilities. In H.L. Swanson, K.R. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities*, 2nd ed. (pps. 15-32). New York: Guilford Press.

Hollander, M., & Wolfe, D. A. (1999). Nonparametric Statistical Methods, 2nd ed. John Wiley & Sons. *New York.*

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

Janzen, J. (2008). Teaching English language learners in the content areas. *Review of Educational Research*, *78*(4), 1010-1038.

Jitendra, A. K., DiPipi, C. M., & Perron-Jones, N. (2002). An exploratory study of schema-based word-problem solving instruction for middle school students with learning disabilities: An emphasis on conceptual and procedural understanding. *The Journal of Special Education*, *36*, 22–38. doi:10.1177/00224669020360010301

Jitendra, A., & Gajria, M. (2011). Main idea and summarization instruction to improve reading comprehension. In R. E. O'Connor and P. Vadasy (Eds.), *Handbook of Reading Interventions* (pps. 198-219). New York: Guilford.

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*(4), 422-431.

Jitendra, A. K., Hoff, K., & Beck, M. M. (1999). Teaching middle school students with learning disabilities to solve word problems using a schema-based approach. *Remedial and Special Education*, *20*(1), 50-64.

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

Kendall, P. C. (1981) Assessing generalization and the single-subject strategies. *Behavior Modification 5*(3), 307-319.

Kennedy, C. (2005). *Single-case designs for education research.* Boston, MA. Allyn & Bacon.

Kong, J. E., & Orosco, M. J. (2015). Word-Problem-Solving Strategy for Minority Students at Risk for Math Difficulties. *Learning Disability Quarterly, 39*(3), 171-181. doi: 10.1177/0731948715607347.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M & Shadish, W. R. (2010). Single-case designs technical documentation. Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf.

Kratochwill, T. R. & Levin J. R. (2010). Enhancing the scientific credibility of single-case intervention research: Randomization to the rescue. *Psychological Methods, 15*(2), 124-144.

Kratochwill, T. R. & Levin J. R. (2014). Meta- and statistical analysis of single-case intervention research data: Quantitative gifts and a wish list. *Journal of School Psychology, 52*(2014) 231-235.

Lesaux, N.K., & Harris, J.R. (2013). Linguistically diverse students' reading difficulties: Implications for models of learning disabilities identification and effective instruction. In H.L. Swanson, K.R. Harris, & S. Graham (Eds.), *Handbook of Learning Disabilities*, 2nd ed. (pps. 69-84). New York: Guilford Press.

Marr, D., Rodden, L., & Woods, A. (2009). *Technical report for the California English Language Development Test (CELDT)* (California Department of Education). Monterey, CA: CTB/McGraw-Hill

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review*, *78*(2), 333-368.

Mazzocco, M. M. (2007). Defining and differentiating MLD and MD. In Berch, D. & Mazzocco, M.M (Eds) *Why is math so hard for some children?: The nature and origins of mathematical learning difficulties and disabilities* (29-47). Baltimore: Paul H. Brookes Co.

Mazzocco, M. M., & Myers, G. F. (2003). Complexities in identifying and defining mathematics learning disability in the primary school-age years. *Annals of Dyslexia, 53*(67-85).

McIntosh, S., & Vignoles, A. (2001). Measuring and assessing the impact of basic skills on labour market outcomes. *Oxford Economic Papers*, 453-481.

Montague, M. (1992). The effects of cognitive and metacognitive strategy instruction on the mathematical problem solving of middle school students with learning disabilities. *Journal of Learning Disabilities*, *25*(4), 230-248.

Montague, M. (1997). Cognitive strategy instruction in mathematics for students with learning disabilities. *Journal of Learning Disabilities*, *30*, 164-177.

Montague, M., & Bos, C. S. (1986). The effect of cognitive strategy training on verbal math problem solving performance of learning disabled adolescents. *Journal of Learning Disabilities*, *19*(1), 26-33.

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*(3), 97-105.

National Center for Education Statistics (2013). The nation's report card: A first look: 2013 mathematics and reading (NCES 2014-451). Institute of Education Sciences, U.S. Department of Education, Washington, D. C.

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Retrieved from http://www.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

NCS Pearson, Inc. (2012). *AIMSweb technical manual.* Retrieved from www.aimsweb.com/wp-content/uploads/aimsweb-Technical-Manual.pdf

Orosco, M. J. (2014) Word problem strategy for Latino English language learners at risk for math disabilities. *Learning Disability Quarterly, 37*(1), 45-53.

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*, 96–107.

Parker, R. I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-overlapping data (PAND): An alternative to PND. *The Journal of Special Education*, *40*(4), 194-204.

Parker, R. I., & Vannest, K. (2009). An improved effect size for single-case research: Nonoverlap of all pairs. *Behavior Therapy*, *40*(4), 357-367.

Parker, R. I., Vannest, K. J., & Brown, L. (2009). The "improvement rate difference" for single-case research. *Exceptional Children, 75*, 135–150.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy*, *42*(2), 284-299.

Pearson Education, Inc. (2012) *Aimsweb Technical Manual*. Retrieved from http://www.aimsweb.com/resources/research-articles-and-information

Powell, S. R., Fuchs, L. S., Fuchs, D., Cirino, P. T., & Fletcher, J. M. (2009). Effects of Fact Retrieval Tutoring on Third-Grade Students with Math Difficulties with and without Reading Difficulties. *Learning Disabilities Research & Practice*, *24*(1), 1-11.

Raven, J. C. (1976). *Colored Progressive Matrices Test*. London,England: H. K. Lewis.

Rivera-Batiz, F. L. (1992). Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources*, 313-328.

Shapka, J. D., Domene, J. F., & Keating, D. P. (2006). Trajectories of career aspirations through adolescence and young adulthood: Early math achievement as a critical filter. *Educational Research and Evaluation*, *12*(4), 347-358.

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.

Simpson, R. L., Lacava, P. G., & Graner, P. S. (2004). The No Child Left Behind Act challenges and implications for educators. *Intervention in School and Clinic*, *40*(2), 67-75.

Sisco-Taylor, D., Fung, W., & Swanson, H. L. (2014). Do curriculum-based measures predict performance on word-problem-solving measures? *Assessment for Effective Intervention, 40*(3), 131-142. doi: 10.1177/1534508414556504

Soares, D. A., Harrison, J. R., Vannest, K. J., & McClelland, S. S. (2016). Effect size for token economy use in contemporary classroom settings: A meta-analysis of single-case research. *School Psychology Review*, *45*(4), 379-399.

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, *98*, 265-281.

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

Swanson, H. L., Cooney, J. B., & Brock, S. (1993). The influence of working memory and classification ability on children′s word problem solution. *Journal of Experimental Child Psychology*, *55*(3), 374-395.

Swanson, H. L., Kehler, P., & Jerman, O. (2010). Working memory, strategy knowledge, and strategy instruction in children with reading disabilities. *Journal of Learning Disabilities*, *43*(1), 24-47.

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*(4), 170-183.

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.

van Garderen, D. (2007). Teaching students with LD to use diagrams to solve mathematical word problems. *Journal of Learning Disabilities*, *40*(6), 540-553.

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

Vannest, K. J., & Ninci, J. (2015). Evaluating intervention effects in single-case research designs. *Journal of Counseling & Development*, *93*(4), 403-411.

Vannest, K.J., Parker, R.I., & Gonen, O. (2011). Single Case Research: web based calculators for SCR analysis. (Version 1.0) [Web-based application]. College Station, TX: Texas A&M University. Retrieved Tuesday 1st December 2015. Available from singlecaseresearch.org

Vaughn, S., Denton, C. A., & Fletcher, J. M. (2010). Why intensive interventions are necessary for students with severe reading difficulties. *Psychology in the Schools*, *47*(5), 432-444.

Wolery, M., Busick, M., Reichow, B., & Barton, E. E. (2010) Comparison of overlap methods for quantitatively synthesizing single-subject data. *Journal of Special Education, 44*(1), 18-28.

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

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*(5), 303-318.

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*(2), 97-111.

Table 1

Demographic and School-Related Data

Student	Gender	Ethnicity	Age	CELDT level	CELDT description	TOMA Standard Score	TOMA percentile	DRA Spring
Mary	F	Hispanic	8	1	Beg	6	9	34
James	M	Hispanic	9	2	EI	8	25	38
Edgar	M	Hispanic	9	2	EI	7	16	38
Alex	M	Hispanic	9	1	Beg	7	16	20
Jane	F	Hispanic	9	2	EI	6	9	38
Brian	M	Indonesian	9	3	Int	8	25	50
Diana	F	Hispanic	9	3	Int	7	16	38
Daria	F	Hispanic	9	2	EI	7	16	38
Mateo	M	Hispanic	9	2	EI	7	16	38

Note. CELDT = California English Language Development; Beg = Beginning; EI = Early Intermediate; Int = Intermediate; TOMA = Test of Math Ability (Problem Solving Subtest); DRA = Developmental Reading Assessment

Table 2

Word Problem Solving Mean Percent Accuracy Scores Across Phases

Student	Baseline	Intervention	Maintenance	Follow-up
Mary	16.67	44.00	20.00	15.00
James	20.00	40.00	36.67	
Edgar	20.00	42.67	66.67	25.00
Alex	22.00	38.33	73.33	30.00
Jane	4.00	26.15	60.00	5.00
Brian	38.00	67.69	93.33	55.00
Diana	21.43	32.00	46.67	5.00
Daria	20.00	41.82	46.67	10.00
Mateo	32.86	60.00	66.67	30.00

Table 3

M-CAP Mean Scores Across Phases

Student	Baseline	Intervention	Maintenance	Follow-up
Mary	7.00	11.50	12.00	12.50
James	6.33	12.00	12.50	
Edgar	7.00	13.75	15.00	16.50
Alex	6.50	8.00	7.00	8.50
Jane	4.50	8.67	10.00	8.50
Brian	7.25	7.33	11.00	12.50
Diana	5.00	11.00	9.00	9.00
Daria	6.00	8.00	11.50	8.50
Mateo	11.40	16.50	18.50	21.50

Table 4

Calculation Mean Percent Accuracy Scores Across Phases

Student	Baseline	Intervention	Maintenance	Follow-up
Mary	85.00	65.71	79.67	60.00
James	56.67	45.00	35.33	
Edgar	58.33	57.85	61.00	45.00
Alex	87.50	74.17	84.33	75.00
Jane	55.00	52.50	88.00	47.50
Brian	82.50	79.17	96.67	92.50
Diana	52.00	35.00	56.00	47.50
Daria	69.00	75.00	86.33	70.00
Mateo	73.00	78.75	83.00	75.00

Table 5

M-COMP Mean Scores Across Phases

Student	Baseline	Intervention	Maintenance	Follow-up
Mary	44.00	50.50	56.00	38.50
James	30.00	38.50	43.00	
Edgar	40.00	42.50	44.00	36.00
Alex	43.50	45.00	53.00	49.00
Jane	37.75	46.00	50.50	41.50
Brian	53.00	59.00	64.50	64.00
Diana	34.40	37.00	40.00	27.00
Daria	44.60		58.00	54.50
Mateo	51.60	65.50	64.50	57.00

Table 6

Reading Comprehension Mean Percent Accuracy Scores Across Phases

Student	Baseline	Intervention	Maintenance	Follow-up
Mary	20.00	50.00	26.67	40.00
James	40.00	42.86	20.00	
Edgar	33.33	52.50	60.00	40.00
Alex	33.33	43.33	13.33	10.00
Jane	20.00	45.71	40.00	30.00
Brian	60.00	34.29	53.33	20.00
Diana	45.00	33.33	26.67	30.00
Daria	50.00	66.67	60.00	50.00
Mateo	60.00	63.33	66.67	50.00

Table 7

MAZE Mean Accuracy Scores Across Phases

Student	Baseline	Intervention	Maintenance	Follow-up
Mary	9.00	9.75	7.00	9.50
James	10.00	9.75	13.00	
Edgar	10.00	16.00	22.50	11.50
Alex	4.33	11.33	11.00	10.50
Jane	9.50	16.33	23.50	20.00
Brian	7.00	14.67	26.50	12.50
Diana	9.80	12.00	18.00	9.00
Daria	8.40	5.50	26.00	13.50
Mateo	12.25	15.00	19.00	11.00

Table 8

Weighted Average Tau-U Effect Sizes for Each Measure

Measure	Tau-U	*p* value	*SE*	95% CI
WPS	0.62	0.00	0.11	0.41, 0.84
M-CAP	0.66	0.00	0.17	0.33, 0.99
Calculation	-0.31	0.02	0.13	-0.57, -0.04
M-COMP	0.41	0.03	0.18	0.05, 0.77
Reading Comprehension	0.18	0.19	0.14	-0.09, 0.45
MAZE	0.24	0.14	0.16	-0.08, 0.56

<u>Instructional Protocol</u>
Warm-up activity (5 minutes)

Warm up activities will alternate between brief writing, calculation, and reading comprehension exercises.

Explicit Instruction (5-7 minutes)
Let's look at the word problem together. Please put your finger on the word problem. Follow along as I read the problem aloud.

Question
A word problem asks a question. We know it is a question because it ends with a question mark. [Point to the question mark] *Here is the question mark.*

What is the question? In this word problem, the question is:_____.

Underline the question in your notebooks. Now we know the question in this word problem is: _____. This is your question.

We can say this question or problem in our own words. For example, I might say "The question is _____."

Please write the question in your own words as I did in the space below where it says "The question is:".

Relevant information
Before we can solve this word problem, we need to think about what we already know. In each word problem we need to find important information. This information may include numbers or math words. This is called "relevant" information. The word "relevant" means important information to solve the problem.

What is the information I need? In this word problem, the relevant information is: _____.
Please draw a box around the relevant information in the word problem at the top of the page. Remember, relevant means important information to answer the question. We can say this information in our own words. For example, I might say "The important information is _____."
Please write the relevant in your own words as I did in the space below where it says "Information I need:".
Irrelevant information
Sometimes word problems have information that we do NOT need. This is called "irrelevant information."

In this word problem, the irrelevant information is _____. We can cross out the irrelevant sentence.
This word problem's question is not asking about _____. We cross out the irrelevant sentence.

Solve and check
Now we are ready to solve our problem.
The question asks us _____.
The information I need to solve the problem is _____.
What do we need to do to solve this problem?[add or subtract]

We set up our equation to solve for the answer.
[Teacher models equation]

Our answer is _____. Remember my solution should answer the question. Now I will say my answer in a full sentence to see if it answers my question.

Does it answer my question? [yes or no]
If yes, we are done.
If no, let's go back to the question.

Guided Practice (8-12 minutes)
Question
A word problem asks a question. Please find the question. Write the question in your own words in the space below where it says "The question is:".

Instructor checks student answers. If necessary, provide corrective feedback. After two attempts of guided instruction, if student does not paraphrase question in own words, instructor should model the answer.

Relevant information
In each word problem we need to find important information. This information may include numbers or math words. Please find the relevant information. Write the important information in your own words in the space below where it says "Information I need:".

Instructor checks student answers. If necessary, provide corrective feedback. After two attempts of guided instruction, if student does not paraphrase relevant information in own words, instructor should model the answer.

Irrelevant information
Sometimes word problems have information that we do NOT need. Please cross out the irrelevant information.
Instructor checks student answers. If necessary, provide corrective feedback. After two

attempts of guided instruction, if student does not cross out irrelevant information, instructor should model the answer.

Solve and check

Please set up the equation to solve for the answer.
Instructor checks student answers. If necessary, provide corrective feedback. After two attempts of guided instruction, if student does not set up correct equation, instructor should model the answer.

Please write the answer in a full sentence. Does it answer the question?
Instructor checks student answers. If necessary, provide corrective feedback. After two attempts of guided instruction, if student does not write full sentence, instructor should model the sentence.

Independent Practice (10 minutes)
Students solve problems independently.
If students ask for help:
1st prompt: *Try your best to solve the problems on your own first, and then I will help you.*
2nd prompt: Draw a line up to where students were able to solve independently. Provide guidance below the line.

Figure 1. Instructional protocol used every intervention session.

Robert likes to collect marbles. He has 24 red and 15 purple marbles. Scott likes to collect stamps. How many marbles does Robert have in all?

The question is: _____ _____
Information I need: _____ _____
Cross out information you do not need. Equation:
Answer: _____ _____ Does it answer your question?   Yes   No

Figure 2. Instructional materials used every intervention session.

| Examiner Behavior	Yes	No
Explicit Instruction		
1. Teacher or student reads problem aloud		
2. Teacher asks "What is the question?"		
3. Teacher states "The question is_____."		
4. Teacher asks "What is the information I need?"		
5. Teacher states "The important information is _____."		
6. Teacher states "Cross out the irrelevant information."		
7. Teacher asks "What do we need to do to solve this problem?"		
8. Teacher states the answer in a full sentence.		
Guided practice		
1. Teacher or student reads problem aloud twice		
2. Teacher states "Write the question in your own words."		
3. Teacher states "Write the important information in your own words."		
4. Teacher states "Cross out the irrelevant information."		
5. Teacher instructs students to solve on their own.		
6. Teacher instructs students to write the answer in a full sentence.		
Independent Practice		
1. Teacher directs students to solve the problems independently.		
	_____% Yes	

Figure 3. Fidelity of implementation checklist.
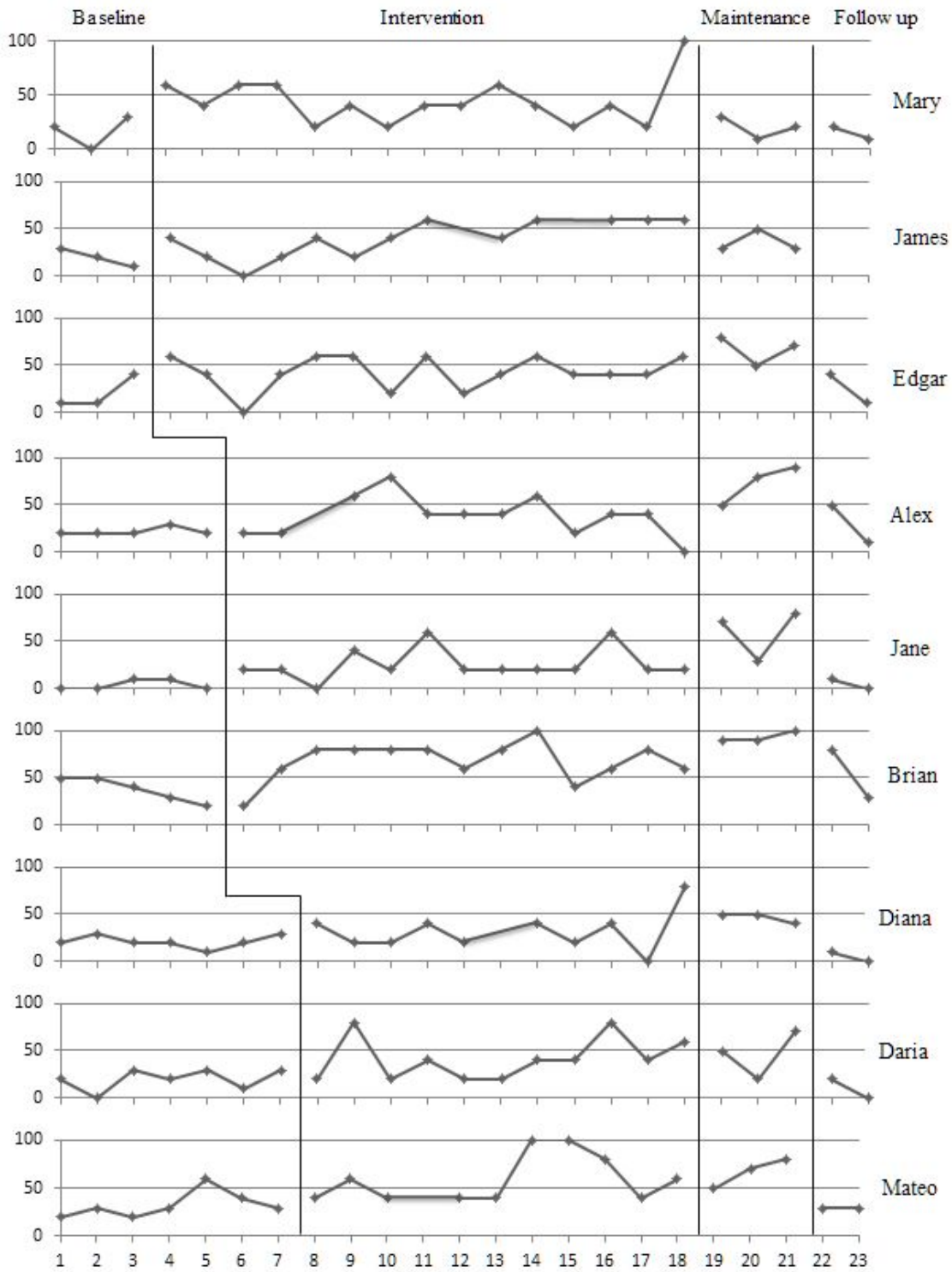
Figure 4. Word problem solving percentage accuracy as a function of baseline, treatment, maintenance, and follow-up.
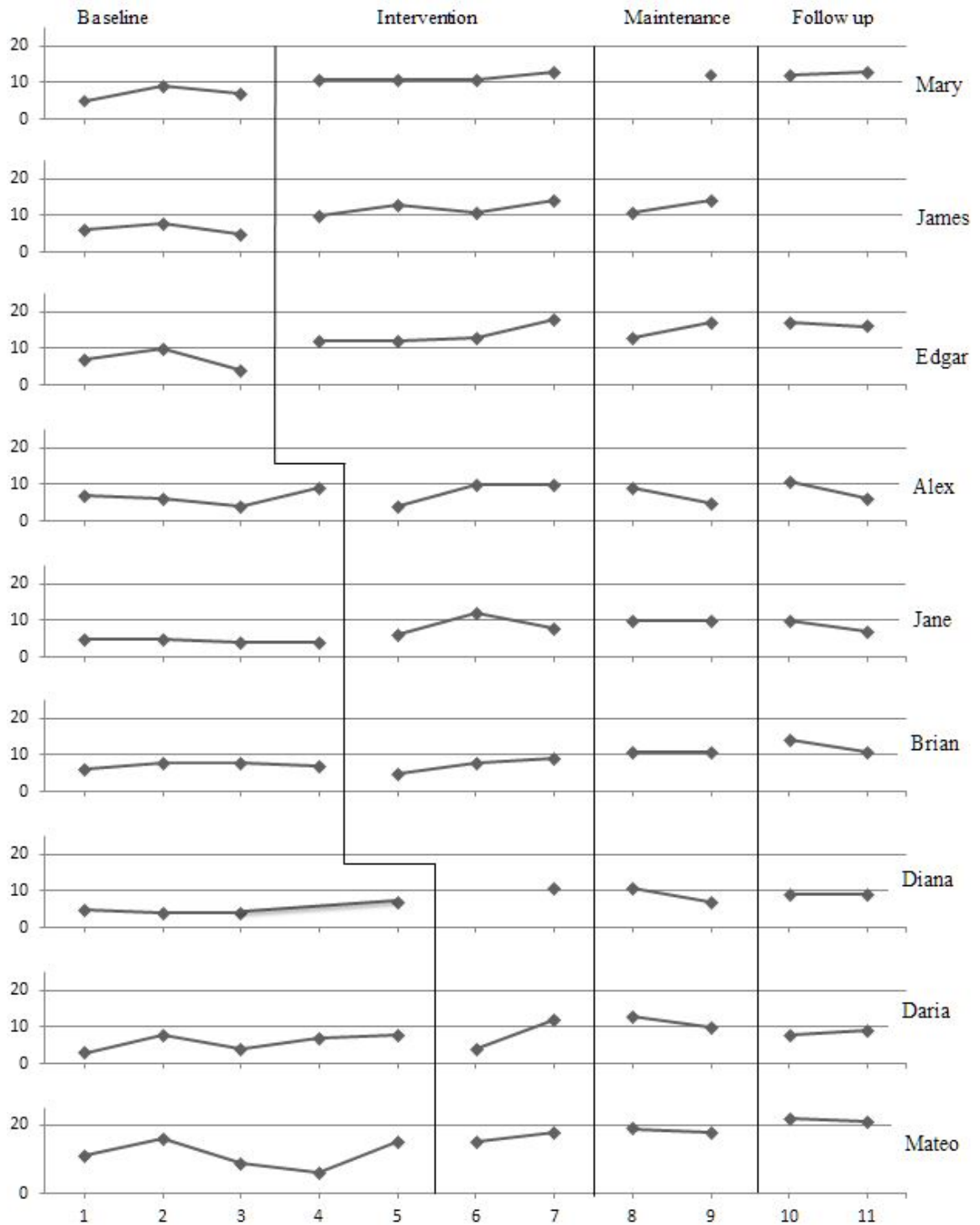
Figure 5. AIMSweb Math Concepts and Applications points accuracy as a function of baseline, treatment, maintenance, and follow-up.
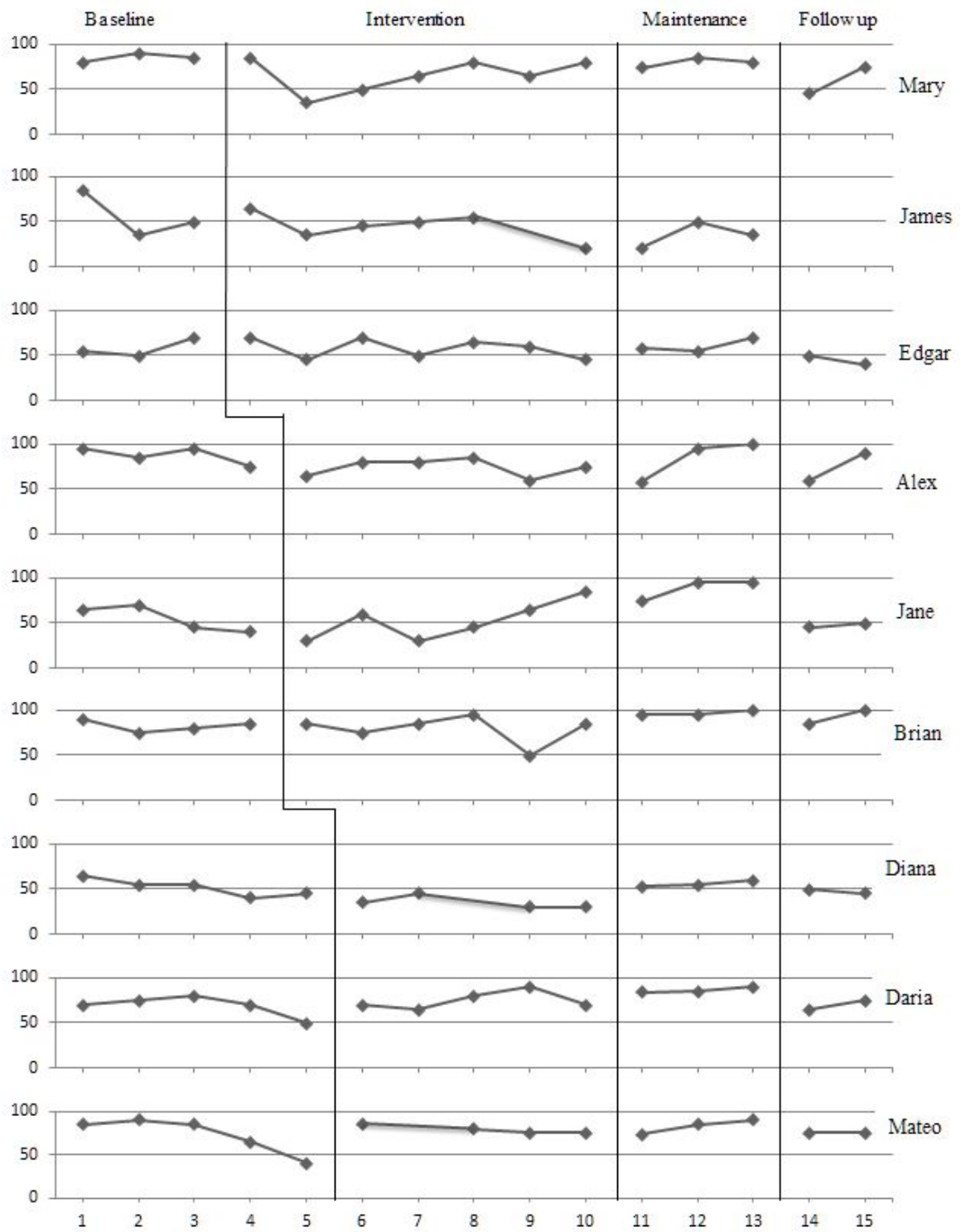
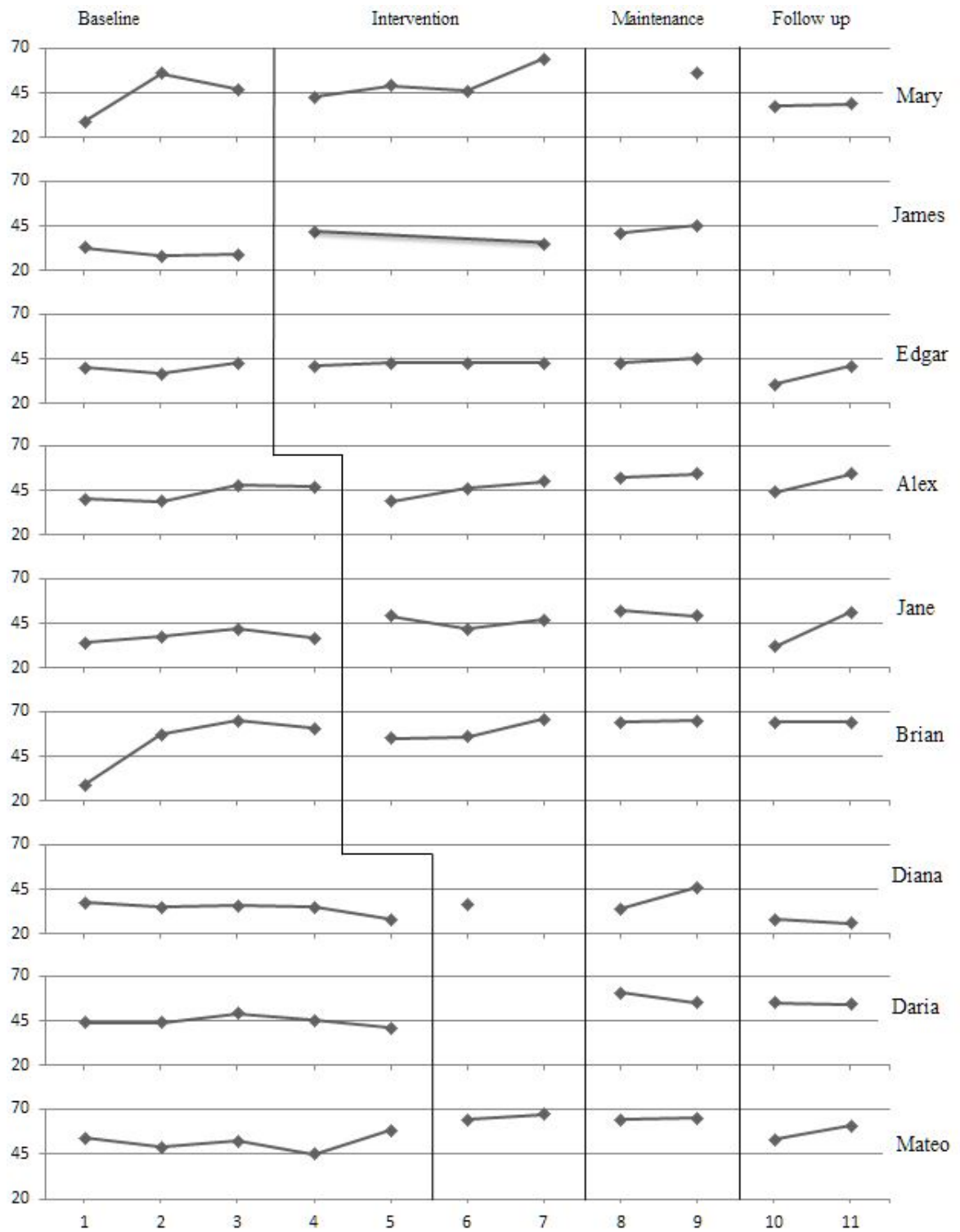*Figure 6.* Calculation percentage accuracy as a function of baseline, treatment, maintenance, and follow-up.

Figure 7. AIMSweb Math Computation points accuracy as a function of baseline, treatment, maintenance, and follow-up.
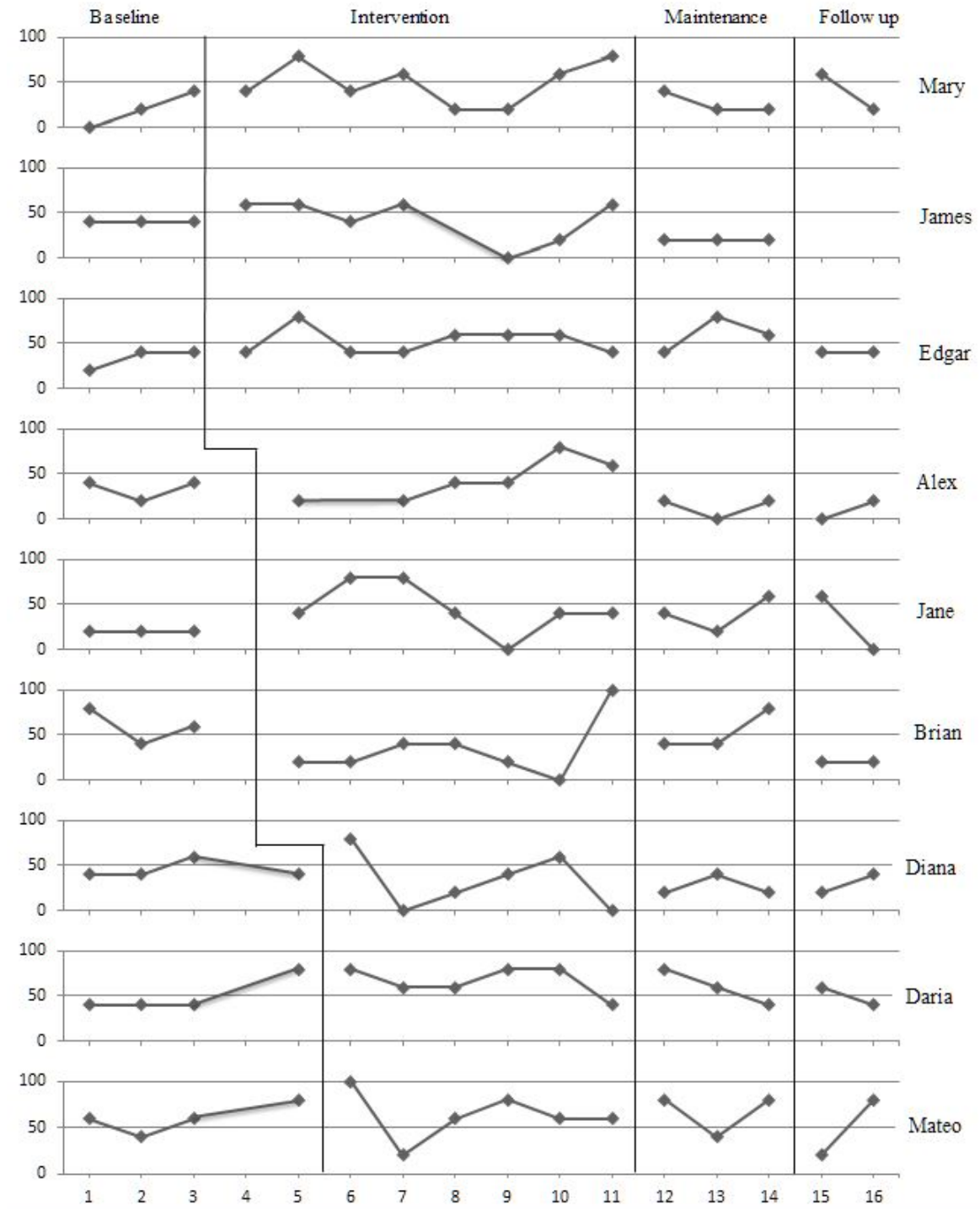
103

Figure 8. Reading comprehension percentage accuracy as a function of baseline, treatment, maintenance, and follow-up.
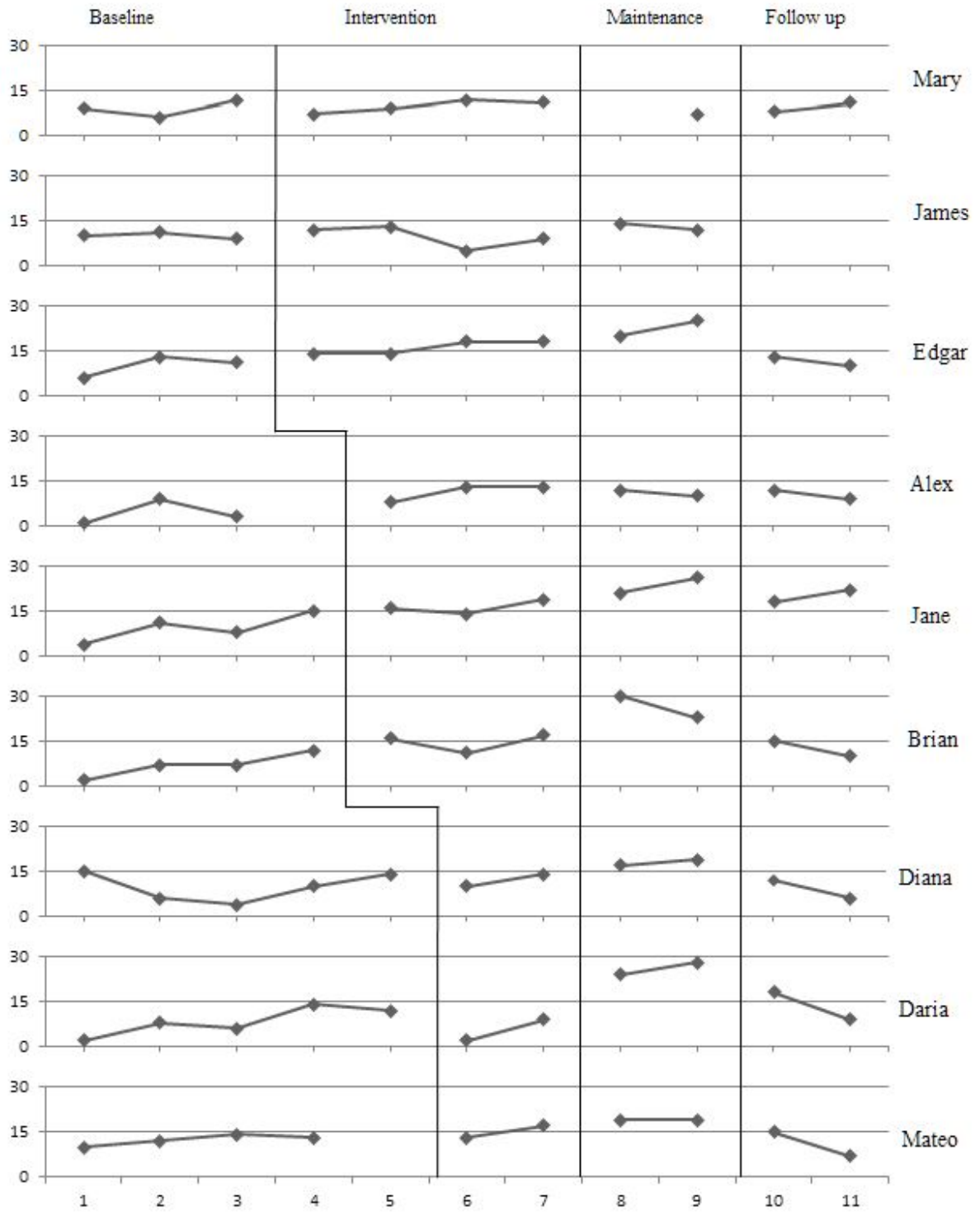
Figure 9. AIMSweb MAZE accuracy points as a function of baseline, treatment, maintenance, and follow-up.

105