

UC Riverside

UC Riverside Electronic Theses and Dissertations

Title

Effects of Behavior Supports on Math Intervention Outcomes: A Meta-Analysis

Permalink

<https://escholarship.org/uc/item/1p0255dw>

Author

Kasser, Tali

Publication Date

2016

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Effects of Behavior Supports on Math Intervention Outcomes: A Meta-Analysis

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

Master of Arts

in

Education

by

Tali Arline Kasser

December 2016

Thesis Committee:

Dr. Cathleen Geraghty, Chairperson

Dr. Austin Johnson

Dr. Katherine Stavropoulos

Copyright by
Tali Arline Kasser
2016

The Thesis of Tali Arline Kasser is approved:

Committee Chairperson

University of California, Riverside

ABSTRACT OF THE THESIS

Effects of Behavior Supports on Math Intervention Outcomes: A Meta-Analysis

by

Tali Arline Kasser

Master of Arts, Graduate Program in Education
University of California, Riverside, December 2016
Dr. Cathleen Geraghty, Chairperson

Student achievement in mathematics is an area of concern that has implications for student success on an individual level as well as for the nation as a whole (National Mathematics Advisory Panel, 2008). Although tier 2 intervention has been shown to be effective in remediating math difficulty, an estimated 3-8% of students do not respond to intervention (Fuchs, Fuchs, & Compton, 2012). This indicates a need to identify specific intervention components, which can increase academic outcomes. One potential intervention component to increase academic outcomes is the use of behavioral strategies in math intervention. Behavioral strategies have been found to increase academic engagement (Brooks, Todd, Tofflemoyer & Horner, 2003; Liaupsin, Umbreit, Ferro, Urso, Upreti, 2006; Todd, Horner, & Sugai, 1999). In turn, engagement increases academic outcomes (Finn, 1993; Marks, 2000). The following study uses meta-analytic techniques to assess effect size differences between math interventions with and without behavioral components. Twelve studies were included in the analysis. In addition, effect sizes were calculated by behavioral component type in order to determine whether the magnitude of effect varied by this moderator. Results from this study indicate that the use

of behavioral strategies in mathematics interventions may increase intervention effectiveness. The use of self-monitoring strategies, particularly graphing progress, led to larger effect sizes than studies which utilized other strategies (verbal praise and tangible reinforcement). Limitations and implications for future research are discussed.

Table of Contents

Introduction.....	1
Relationship between Behavior Supports and Academic Outcomes.....	2
Motivation and Math Achievement.....	3
WWC Recommendations: RTI Mathematics.....	6
Purpose.....	8
Methods.....	9
Coding.....	11
Interrater Agreement.....	12
Analysis.....	12
Results.....	16
RQ1: Inclusion of Behavior Component.....	16
RQ2: Type of Behavior Component.....	16
Homogeneity of Variance.....	17
Study with Largest Effect Size.....	17
Behavior Component Descriptions.....	18
Discussion.....	20
Role of Academic Engagement.....	20
Self-Monitoring to Improve Academic Outcomes.....	21
Inclusion of Behavioral Strategies in Intervention within an RTI Framework....	22
Limitations.....	22
Future Research.....	25

References.....	26
Appendix.....	31

Effects of Behavior Supports on Math Intervention Outcomes:

A Meta-Analysis

The National Mathematics Advisory Panel (NMAP) released a report in 2008 which reviewed findings related to mathematics achievement in US schools. This report advocated for improvements in our current mathematics curriculum, pointing out that student achievement in mathematics is an area of concern that has implications for student success on an individual level as well as for the nation as a whole. Individuals need solid mathematic foundations in order to increase opportunities related to college and career options. On a national level, there is a need for individuals with sound mathematical knowledge to fill the roles of policy makers in a variety of fields (NMAP, 2008). Mathematic abilities have been found to predict a number of desired outcomes in careers and academics. Early math skills are predictive of reading, math, science achievement, and grade retention through 8th grade (Classens & Engel, 2013). Even after controlling for cognitive ability and reading achievement, math skills are predictive of income, employment, and work productivity (Rivera-Batiz, 1992).

According to the National Assessment of Educational Progress's (NAEP) 2015 report, which monitors the percentage of proficient students in mathematics, the percentage of proficient students in mathematics has decreased slightly since the last review in 2013. Findings showed that only 40% of fourth grade students, 33% of eighth grade students, and 25% of twelfth grade students performed at or above proficient level in mathematics. This means that 75% of students are not proficient in mathematics by the time they graduate from high school. Students at proficient level in the twelfth grade

require an understanding of basic geometry and algebra (NAEP, 2008). The low number of students reaching proficiency highlights the need to identify interventions that are efficacious in remediating math difficulties. Furthermore, there is a need to identify specific intervention components that can increase academic outcomes.

Relationship between Behavior Supports and Academic Outcomes

The use of Response to Intervention (RTI) techniques can remediate math difficulties at tiers 2 and 3 (Fuchs, Fuchs, & Compton, 2012; Gersten et al., 2009), but a number of students do not respond, which indicates a need for more effective intervention. This number has been estimated to be between three and eight percent (Fuchs, et al., 2012). Intervention effectiveness could potentially be increased by identifying specific intervention components, which increase academic outcomes. However, there remains a need in research to identify the specific intervention components that lead to student gains, particularly outside of the area of reading, through the systematic manipulation of specific components (Glover & DiPerna, 2007).

One component that could potentially increase intervention effectiveness is the use of behavioral supports within intervention sessions. Behavioral strategies have been shown to increase academic engagement (Brooks, Todd, Tofflemoyer & Horner, 2003; Liaupsin, Umbreit, Ferro, Urso, Upreti, 2006; Todd, Horner, & Sugai, 1999), which in turn increases academic outcomes (Finn, 1993; Marks, 2000). In addition, students who are more academically engaged are less likely to drop out of school (Finn, 1983; Sinclair, Christenson, Evelo & Hurley, 1998; Connell, Spencer, & Aber, 1994). Two studies were conducted using the national sample of eighth graders from the U.S. Department of

Education (Finn, 1993). These studies examined student engagement and achievement and concluded that academic engagement is related to academic outcomes, even after controlling for other factors, such as ethnicity and socioeconomic status. One study divided their sample of 5,945 eighth graders into three groups, unsuccessful, passing, and successful, based on achievement on math and reading tests. The sample consisted of students who were identified as at risk based on socioeconomic status (SES), ethnicity, and home language. It was found that students could be accurately classified into achievement groups based on classroom engagement alone, as reported by students and teachers (Finn, 1993).

Numerous studies have indicated that the use of function-based interventions are effective for increasing academic engagement (Lane, Smither, Hughesman, Guffey, & Fox, 2007; Brooks et al., 2003; Preciado, Horner, & Baker, 2008; Liaupsin, Umbreit, Ferro, Urso, & Upreti, 2006). Ideally, interventions to increase academic engagement would be based on a functional behavior assessment (FBA), in which data are gathered in order to test the hypothesized function of the observed off-task behavior (Crone & Horner, 2000). However, conducting an FBA for all students in an academic intervention would probably not be feasible in a school-setting. Furthermore, if academics were the children's major source of concern, rather than behavior, conducting a full FBA would likely be a waste of resources (March & Horner, 2002).

Motivation and Math Achievement

Besides increasing academic engagement, behavioral strategies may increase academic outcomes by increasing student motivation. Although this theory has less

research support, several authors have suggested that motivation may play a role in academic outcomes (Ames, 1992; Cleary & Chen, 2009; Singh, Granville, & Dika, 2002). However, motivation can be a difficult construct to measure, as there is not a consensus on how to assess student motivation. One way in which motivation has been conceptualized is as a product of a student's interest in a task combined with their perceived value of the task (Cleary & Chen, 2009). Motivation is generally thought of as a broader concept than academic engagement, as student motivation needs to be present for students to be engaged (Cleary & Chen, 2009). Motivation is also thought to lead to other desired behaviors, such as work completion and school attendance (Singh, Granville, & Dika, 2002).

Ames (1992) describes various classroom factors that are likely to play a role in academic motivation. One of these factors is the way in which students are evaluated. Ames (1992) claims that the recognition of student growth can be particularly useful in fostering motivation, but warns against the use of comparisons between students, as this can lower self-esteem. For this reason, student growth should be acknowledged privately (Ames, 1992). Additionally, Ames (1992) recommends the use of self-management and self-monitoring strategies. According to Ames (1992) this can be done by encouraging students to set reasonable goals for themselves and by acknowledging student effort over achievement, which will lead to increased motivation, sense of self-worth, and finally academic achievement.

These recommendations for increasing student motivation have also been looked at more specifically within the context of math achievement (Cleary & Chen,

2009). Cleary and Chen assessed self-regulation, motivation, math achievement, and type of math course. The aim of this study was to determine which factors were predictive of the use of self-regulatory strategies and motivation. Self-regulation and motivation were measured using researcher-designed questionnaires. The self-regulation questionnaire aimed to measure the extent to which the student took responsibility for his or her learning, while the motivational scale included questions that measured the student's level of interest in mathematics. It was found that self-regulation and motivation were more accurate in differentiating between achievement groups in higher level math classes. These findings indicate that motivation and the ability to self-regulate may be most important for higher level mathematics. Additionally, Cleary and Chen (2009) found that achievement was also more predictive of motivation and self-regulation in seventh grade than in sixth grade.

Another study included both academic engagement and measures of motivation in order to determine whether these factors were predictive of mathematics achievement in eighth grade students (Singh, Granville, & Dikah, 2002). Findings indicated that although both of these factors were significant predictors of math achievement, academic engagement was more predictive than motivation. In this study, motivation was measured through student attendance and class preparedness (e.g., did the student come to class with a pen, pencil, homework, etc.). The findings regarding motivation and math achievement indicate that the inclusions of strategies aimed at increasing motivation and academic engagement in tier 2 and 3 math interventions may increase academic outcomes.

WWC Recommendations: RTI Mathematics

What Works Clearinghouse (WWC) recommends the use of “motivational strategies” at tiers 2 and 3 within an RTI framework, stating that “Even a well designed curriculum may falter without behavioral supports” (Gersten et al., 2009, p. 44). The behavioral supports or motivational strategies discussed here include the use of verbal praise, points and prizes, and graphing of progress. It is recommended that rewards are engagement, completion, or achievement contingent. However, the recommendation to include behavioral supports in mathematics intervention is based largely on the opinion of a professional, as the present research is not clear on this subject. Although studies have included behavioral strategies, most studies have not specifically assessed the effects of these strategies on outcomes. This emphasizes the need for research in this area.

Of the studies that have included behavioral strategies, included the use of graphing academic progress. (Fuchs, Fuchs, Phillips, & Hamlett, 1995; Fuchs, Fuchs, Finelli, Courey, Hamlett, Sones & Hope, 2006). Both of these studies found positive intervention outcomes over the control group. Significant positive outcomes were also found when other types of reinforcement were included in math intervention (Fuchs et al., 2005; Fuchs, Seethaler, et al., 2008). Fuchs and colleagues (2005) awarded points for appropriate behavior, which could be traded in for prizes, such as stickers, erasers, and pencils. Fuchs, Seethaler, and colleagues (2008) utilized tokens for correct responses rather than appropriate behavior. Finally, one study included both graphing of progress and reinforcement in the form of points for on-task behavior and accuracy of work

(Fuchs, Fuchs, et al., 2008). Points could then be traded in at the end of each session for stickers. Again, although these studies found positive intervention outcomes, none of them isolated the effects of the behavioral component to determine whether it contributed to the positive academic outcomes of the intervention.

Although several research studies have found math interventions that included behavioral strategies to be effective (Fuchs et al., 2005; Fuchs, Fuchs, et al., 2008, Fuchs, Seethaler, et al., 2008; Fuchs, et al., 1995; Fuchs, 2006), only two known studies have been conducted that explicitly tested the effects of behavioral supports on intervention outcomes (Schunk & Cox, 1986; Fuchs et al., 2003). One study included 90 students between 6th and 8th grade who were diagnosed with a learning disability (Schunk & Cox, 1986). These students received a 45-minute subtraction intervention over the course of six consecutive school days. Students were assigned to one of three groups; first half effort feedback, second half effort feedback, or no effort feedback. Effort feedback was a form of verbal praise in which the researcher said to students, “You’ve been working hard.” In the first half effort feedback group, this statement was said fifteen times over the course of the first three intervention sessions. In the second half effort feedback group, this statement was made fifteen times over the course of the last three sessions. Effort feedback was given without the use of any other social reinforcement (e.g., smiling, excited tone of voice). Students in both effort feedback groups scored significantly higher on a subtraction posttest than those in the no effort feedback group. Additionally, students who received effort feedback completed more problems and with more accuracy than their no effort feedback comparisons. These results indicate that

verbal praise can be successful at increasing intervention effectiveness. However, more research is needed that specifically tests the effects of other behavior strategies in math intervention on academic outcomes.

The second study compared the use of self-regulated learning (SRL) strategies, which involved students setting goals and graphing their progress to a non-SRL comparison intervention group (Fuchs, et al., 2003). When students in the SRL intervention group completed a final math problem, they scored the problem and charted their results. They were then encouraged to set goals based on their progress and were reminded of these goals at the beginning of each session. Students showed greater growth on outcome measures in the SRL-treatment group than in the treatment alone group. Additionally, although not significant, it was found that effects of the SRL component was stronger with lower achieving students. This indicates that low achieving students may benefit more from SRL strategies, possibly due to the fact that low-achieving students struggle with setting realistic goals. These studies provide preliminary evidence of the effectiveness of behavior components in math intervention. However, further research in this area is needed.

Purpose

The purpose of the following study is to synthesize research in the area of mathematics intervention in order to compare studies that included the behavior supports and those that did not. The Institute of Educational Sciences (IES) recommends including “motivational strategies” at tiers 2 and tier 3 of math intervention within an RTI framework (IES Practice Guide, 2009). Examples of these motivational strategies include

verbal praise, points and prizes, and graphing progress. However, this recommendation is based mainly on the opinion of a professional panel, as the benefit of behavior supports in math interventions is not strongly supported by the present research. The present study is designed to answer the following research questions:

1. To what magnitude do math interventions with a behavioral component affect student math outcomes compared to those without a behavioral component?
2. How does the magnitude of the effects of the inclusion of behavioral components in math intervention differ by type of behavior strategy employed?

Methods

A systematic search was conducted to identify articles for inclusion in the meta-analysis. The following search terms were used in ERIC and PsychINFO: *Math* AND intervention* (613)

“Math fluency” AND intervention (57), *“Number sense” AND intervention* (67), *(Multiplication OR division) AND intervention* (282), *(Addition OR subtraction) AND intervention* (1885), *Algebra AND intervention* (100), *Geometry AND intervention* (41), *word problems AND intervention* (265), *Fractions AND intervention*, *Trigonometry AND intervention*.

The initial search yielded 3,310 results. First, article titles were reviewed. If these titles appeared to be relevant, the abstracts were read. Finally, the studies whose abstracts appeared to meet inclusion were carefully read and narrowed down to those which fit the following inclusion criteria:

1. The study implemented a math intervention.

2. The study was published in a peer reviewed journal.
3. The study was conducted in a school setting.
4. The study used randomized controlled trials (RCT) or quasi experimental design (QED).
5. The study met What Works Clearinghouse group design standards without reservations.
6. The study included sufficient data to calculate effect size.
7. Outcome measures must assess for mathematics achievement
7. The study was written in English.

Math intervention was defined as any additional instruction to aide in the learning of mathematics for students who had been identifying as have a need in the area of mathematics. There were no specifications as to how this need was identified. For example, need could be established through teacher referral or screening measures. Intervention could include small group instruction or class-wide interventions, such as Peer Assisted Learning Strategies (PALS). A school setting was defined as any environment in which formal instruction is given that is referred to as a “school”. This could include public, private, charter, or home schools.

Randomized controlled trials (RCT) were defined as studies in which subjects were randomly selected and assigned to either treatment or control groups. RCT studies were also required to include a control group. Quasi-experimental designs were also included. However, due to the lack of random assignment, these studies were only

included if they established equivalence between characteristics of the treatment and control groups.

Only studies that met What Works Clearinghouse (WWC) standards for RCT design, either with or without reservation, were included in the analysis. When reviewing studies, What WWC considers study design, attrition, the establishment of equivalence with attrition, and confounding factors. In order for a study to meet WWC standards, the study must either use RCTs or demonstrate that the control and treatment conditions are equivalent to one another on observable characteristics if the study uses a QED. Attrition rate is also a consideration. If the attrition rate within the study is too high within an RCT study, equivalence between the treatment and control group must be considered. In order for equivalence to be established, there must be less than 0.25 of a standard deviation between groups. The study also must not contain confounding factors. In other words, there must not be any clear factors, other than the intervention itself, which could potentially explain post-test changes.

In order to calculate an effect size, studies need to include a mean score and standard deviation for the math achievement outcome measure for both the treatment and control groups. It also had to report the number of subjects in both the control and treatment groups.

Coding

The twelve studies that were included in the analysis were coded on a number of variables, some of which were used as moderator variables in the analysis and others which were intended to provide descriptive information about the studies. For analysis,

articles were coded for inclusion of behavior component and type of behavior component. An intervention component was considered to be behavioral if it included contingencies that promoted achievement, work completion, on-task behavior or any other desired behavior or discouraged unwanted behavior. Self-monitoring was also considered to be a behavior component. Behavior components were coded as either verbal praise (0), tangible reinforcement (1), or self-monitoring (2). Behavior components could also fit into more than one of these categories. Finally, the description of the behavior component was copied and pasted directly from study. For descriptive purposes, coding included author, year of publication, journal, sample demographic information (gender, age, grade, and disability status), intervention type, and name and description of outcome measures.

Interrater Agreement

Interrater agreement of coding was measured by randomly selecting 25% (3) of the included studies. These three studies were coded on by a second coder in order to calculate the percent agreement between coders. Percent agreement was calculated by dividing the total number of times raters agreed on coding by the total number of opportunities for agreement. Agreement between raters was 100%.

Analysis

Statistical analyses were run using Statistical Analysis Software (SAS), using code outlined by Arthur, Bennett, and Huffcutt (2001).

Effect Size Calculations. Effect sizes were calculated using Hedges' g effect size estimate, given by the following formula:

$$g = \frac{Y_i^E - Y_i^C}{S_i}$$

Where Y_i^E is the mean score for the experimental group in study i, Y_i^C is the mean score for the control group in study i, and S_i is the square root of the pooled estimate of variance where:

$$S_i^2 = \frac{(n_i^E - 1) (S_i^E)^2 + (n_i^C - 1) (S_i^C)^2}{n_i^E + n_i^C - 2}$$

This calculation provides a less biased estimate than a Cohen's d effect size, because it takes sample size into consideration by using pooled variance as the denominator (Hedges, 1982). In studies with sufficiently large sample sizes, the bias of a Cohen's d estimate is small. Yet, if the studies have smaller sample sizes, this can substantially bias effect sizes. Due to the range in sample sizes included in this analysis, the Hedges' g effect size was used in order to avoid potential bias.

First, a sample-weighted effect size was calculated for each outcome measure. Most studies included more than one outcome measure of math achievement within each treatment group. An effect size was calculated for each outcome measure. Next, these effect sizes were averaged in order to avoid giving too much weight to studies that included multiple measures (Arthur, et al., 2001). These averages were calculated for each treatment group, rather than by study, so if a study included multiple independent treatment groups, multiple effect sizes were calculated for these studies

Moderator variables. Moderator variables were the behavior component inclusion and behavior component types. The first was divided into two categories, studies that included a behavioral component and studies that did not. The second

moderator also contained two categories. One was the use of self-monitoring strategies and the other was the use of other types of reinforcement, which included tangible reinforcement and verbal praise. These variables were selected based on theoretical knowledge rather than the use of statistical tests. In addition, the number of studies available had some influence on which moderator variables were chosen. Once moderators were chosen, a chi-square statistic was calculated in order to determine the amount of unexplained variance within each moderator group. Based on the assumption that moderators account for some of the variance between studies, the amount of variance should decrease once moderators are introduced.

Sampling Error. One cause of variance between studies is sampling error. This variance can be attributed to differences between the samples used in each study and the population. Expected sampling error, or variance of effect sizes due to variations in the sample from the population, can be estimated using the following formula:

$$Var(e) = \left[\frac{N-1}{N-3} \right] * \left[\frac{4}{N} \right] * \left(1 * \left(\frac{g^{-2}}{8} \right) \right)$$

Where N= average of sample sizes across studies and g=Mean sample-weighted hedge's effect size across studies. In order to correct for sampling error, the expected variance due to sampling error was subtracted from the total variance between effect sizes using the following formula:

$$Var(\delta) = Var(d) - Var(e)$$

(Arthur, et al., 2001)

Assumptions. In order for accurate effect sizes to be calculated within a meta-analysis, certain assumptions must be met. Meta-analyses often utilize t-tests in order to

compare the effects between studies. Therefore, when calculating effect sizes, each study should meet the assumptions needed for a t-test. Hedges outlines these assumptions (1982). Within each study, there should be common variance between the experimental and control groups. However, it is pointed out that such assumptions are often not met in practice, but if these assumptions are reasonably met, they will provide a reasonable effect size estimate (Hedges, 1982). In order to test for homogeneity of variance, a chi-square statistic was calculated using the following formula (Hunter & Schmidt, 1990):

$$Q_{(k-1)} = k * \left(\frac{Var(g)}{Var(e)} \right)$$

Where k= the number of effect sizes, Var(g)= the total variance across effect size calculations, and Var(e)= the estimated error variance. If the chi-square value is significant, this means that the unexplained variance between studies is statistically different than zero (Arthur, et al., 2001).

A second assumption that is highlighted by Hedges (1982) involves the quality of study methodology. If the studies used in a meta-analysis do not use rigorous statistical methods, the results of the meta-analysis may be skewed. In other words, the quality of the studies put into the analysis will be representative of the quality of the analysis itself. For this reason, only peer-reviewed studies and studies which met WWC standards were included in the analysis.

It is also recommended that the reliability of the outcome measures used in each study are taken into account. This is because measures with less than perfect reliability will lead to measurement error which will cause effect sizes to differ from their true values. However, because the reliability of outcome measures were not reported in some

of the included studies, this correction was not possible in the present analysis (Arthur et al., 2001).

Results

Twelve studies that fit inclusion criteria were identified and coded. Of these twelve studies, six included a behavior component and two of these studies included groups with and without a behavior component. A total of 18 effect sizes were included in the analysis.

RQ1: Inclusion of Behavior Component

The first research question was whether or not the inclusion of a behavior component would lead to increased effects on academic outcomes. Findings indicated an overall higher effect for studies that included a behavior component than those that did not. The overall effect size across all studies included in the analysis was 2.88. The Hedge's g effect size for studies that included a behavior component was 2.11 and the effect size for studies that did not include a behavior component was 0.19. The overall effect size and the effect size for studies with a behavioral component are considered large according to Cohen's effect size interpretation guidelines (Cohen, 1992) which classify effect sizes over 0.80 as large. The effect size for studies without a behavioral component are considered small (Cohen, 1992).

RQ2: Type of Behavior Component

The second research question was whether or not the magnitude of the effect varied by the type of reinforcement used. There was a difference in effect sizes found based on the type of strategy used. The largest effect size was found for studies that

included the graphing of progress. The Hedge's g estimate for graphing progress was 2.66 which is considered large (Cohen, 1992). In comparison, Hedge's g estimate for other types of reinforcement was 0.51 which is considered a medium effect (Cohen, 1992). Therefore, the magnitude of the effect was highest in interventions that used graphing progress as a self-monitoring strategy.

Homogeneity of variance

The chi-square test for homogeneity of variance indicated that the amount of unexplained variance was significantly different from zero, which indicates differences between studies that were not captured completely by the moderator variable analysis (See Table 3). However, the chi-square statistic did decrease with the inclusion of moderator variables, indicating that these moderators did accurately capture variance between studies (See Table 4). The assumption of homogeneity of variance was achieved in one moderator group, studies that included other types of reinforcement. This indicates that the unexplained variance across these study is not significantly different from zero and it is therefore unlikely that other moderator variables are present in these studies.

Study with Largest Effect Size

The study with the largest effect size (5.17) was calculated from a study conducted by Fuchs, Fuchs, and Prentice (2003). This study tested the effects of a mathematics problem solving intervention on 395 third-grade students. Students were assigned to one of three groups, a control condition, an intervention condition, and an intervention plus self-regulated learning strategies (SRL) condition. In the SRL condition, students were asked to score their work after each session using a provided answer key.

Students also charted their scores on a chart which was kept in their folder. At the beginning of each session, the students looked at their charts and were reminded that they wanted to improve their scores from last session. Students also scored their own homework before submitting it. During some of the sessions, students were asked to report a time during which they had used the lesson outside of the intervention group. Whether students had reported this to the class along with homework completion was graphed by the teacher on a separate class-wide graph, so students could compare their progress with that of other students. However there was no difference between the magnitude of the effect of treatment between the treatment group with and without SRL, indicating that the SRL component did not have an effect on the treatment outcomes.

Behavior Component Descriptions

Of the studies included in the analysis, seven included a behavior component in the intervention group. These components varied but involved either the graphing of student progress or the use of reinforcement, such as earning points, tangibles, or verbal praise. Two of these studies tested intervention effects with and without behavioral strategies (Fuchs, et al., 2003; Schunk & Cox, 1986). In the study conducted by Schunks and Cox, students were reinforced for on-task behavior with the use of verbal praise (1986). Fuchs and colleagues used what was referred to in the study as self-regulated learning strategies which involved students setting goals and graphing their progress (2003). Two other studies utilized the graphing of progress, but in slightly different ways. In one study, students did not graph their own progress (Fuchs et al., 1995). Instead, research assistants entered scores from weekly assessments into a computer program

which generated graphs of student growth as well as student mastery level on various types of problems. Students were then taught to read and interpret their graphs in order to determine if they were making progress. They were also encouraged to set goals to beat their highest score within the next two weeks. In another study, students completed problems independently at the end of sessions 2-4 in a 7 session intervention. Students then scored the problems and graphed their correct answers (Fuchs et al., 2006). One study used reinforcement for participants in the form of points which could be traded in for tangible reinforcements such as pencils, erasers and stickers (Fuchs et al., 2005). Points were given when students displayed appropriate behavior. Another study used reinforcement for correct answers rather than good behavior (Fuchs, Seethaler, et al., 2008). In this study, students were given tokens for correct answers. These tokens were recorded on the student's treasure maps daily and could be traded in each week for a prize from a treasure chest. Finally, one study included a combination of behavioral techniques in the intervention to encourage on-task behavior (Fuchs, Fuchs, et al., 2008). Research assistants set timers for three minutes during the intervals. When the timer sounded, students would receive one point for being on-task. If any student was not on task, no student received points. Students were also able to earn points for accuracy in work completion. Students then shaded in a thermometer with the number of points they had earned for the day. They were then given stickers for each point earned and noted the number of points earned on a game board. Students also scored one problem that they completed at the end of each session. They were encouraged to try to beat their score the next day.

Discussion

The first aim of this study was to determine whether the inclusion of behavioral strategies in mathematics intervention increased the magnitude of the intervention effects. The mean effect size calculated for studies that included a behavior component was larger than the mean effect size of studies that did not include a behavior component, which suggests that the use of behavioral strategies can increase the effectiveness academic interventions by increasing academic outcomes. These findings align with previous research on the relationship between behavior supports and academic outcomes (Bradshaw, Mitchell, & Leaf, 2010; Keane, 2013; Pavlovich, 2000). However, the avenue by which these behavioral supports are working to increase academic engagement is unclear and likely vary by technique as well as types of behaviors that are targeted.

Role of Academic Engagement

One possibility is that the reinforcement used in these studies increases academic engagement, which in turn increases academic achievement. Previous research has found that academic engagement can be increased through the use of behavioral strategies (Brooks, Todd, Tofflemoyer & Horner, 2003; Liaupsin, Umbreit, Ferro, Urso, Upreti, 2006; Todd, Horner, & Sugai, 1999) Increases in academic engagement may then lead to increased academic outcomes. (Finn, 1993; Marks, 2000). It is most likely that academic engagement played a role in increasing academic outcomes in studies in which students were reinforced specifically for remaining on-task (Fuchs, Fuchs, et al., 2008; Fuchs et al., 2005; Schunk & Cox, 1986). Other studies reinforced correct answers rather than academic engagement (Fuchs, Fuch, et al., 2008; Fuchs, Seethaler, et al., 2008). This

technique may have increased motivation for students to get correct answers because correct answers lead to reinforcement. Increasing correct answers, however, may have required students to increase their academic engagement as well.

Self-Monitoring to Improve Academic Outcomes

Second, this analysis set out to determine which particular behavioral technique was associated with the largest effect size. Findings indicated that the use of self-monitoring strategies, particularly graphing progress, had an overall larger average effect size than studies which utilized other strategies, such as verbal praise, points, and tangible reinforcement.

Four studies utilized the graphing of progress in intervention (Fuchs et al., 1995; Fuchs et al., 2003; Fuchs, Fuchs et al., 2006; Fuchs, et al., 2008) Prior research has supported the efficacy of self-monitoring techniques in producing a variety of behavioral changes (Cooper, Heron, & Howard, 2014). Self-monitoring is a type of self-management in which a person systematically records his or her target behavior (Cooper et al., 2014). In the case of the studies included in the analysis, the target behavior would be increasing performance. Self-monitoring is hypothesized to work by creating something called reactivity, which is the effects an assessment or measure can have on behavior. These effects on behavior are greatest when the person completing the assessment is the subject of the behavior changes themselves (Cooper et al., 2014). Therefore, it is not surprising that having students graph their own academic progress could lead to behavior changes and increase academic achievement.

Inclusion of Behavioral Strategies in Intervention within an RTI Framework

Having students graph their own progress would be particularly feasible in a school in which RTI was implemented, as progress monitoring and the graphing of student progress would already be taking place. The only modification needed would be to include the student in this process. Additionally, the use general behavioral strategies could fit into tier 1 of a multi-tiered system of behavior supports. One prime example of general behavioral strategies which have been found to increase academic outcomes is Positive Behavioral Interventions and Supports (PBIS). PBIS is a school-wide behavioral system that fits into a Response to Intervention (RTI) framework. At the school-wide level, reinforcement is given for clearly stated behavioral expectations (Sugai & Simonsen, 2012). Implementation of PBIS has been shown to correlate with higher academic outcomes in both math and reading (Bradshaw, Mitchell, & Leaf, 2010; Keane, 2013). PBIS implementation has also been shown to increase SAT scores, which suggests positive long term academic outcomes (Pavlovich, 2009). If PBIS were already in place in a school, these tier 1 behavioral strategies could easily be utilized at tiers 2 and 3 of academic intervention.

Limitations

However, due to the limitations of the present study, these results should be interpreted with caution and further research in this area is needed. Thus far, a limited number of studies have explicitly tested the effects of the use of behavioral strategies in math interventions (Fuchs et al., 2003; Schunk & Cox, 1986). Schunk and Cox found a positive effect for verbal praise (1986). Fuchs and colleagues' use of goal setting

combined with graphing of progress significantly increased outcomes over the control condition. In the present analysis, the largest effect size was found for studies that used self-monitoring of progress during intervention. Due to the limited research in this area, there is a need for future research that explicitly tests the effects of the inclusion of behavioral techniques in intervention.

There are a number of limitations that should be taken into consideration when interpreting the results of this study. One such limitation is the fact that the vast majority of the studies (all but one) that included behavior strategies were conducted by the same first author, Fuchs. This is a potential confounding variable. The heightened effect sizes that were seen in this group of studies could have been due to some other component or components of the studies that were attributable to the first author. However, there was also one study conducted by Fuchs which did not include a behavior component (Fuchs, Fuchs, Yasdian & Powell, 2002) and one which included treatment groups with and without behavior components so effect size estimates from this study were used in both groups (Fuchs et al., 2003). The effect size computed in these studies were 1.06 and 4.35, respectively. The overall effect size calculated for studies without a behavior component was 0.94. The study included in the behavior components group that was not conducted by Fuchs had an effect size of 1.82 and the overall group effect size was 1.98.

Another limitation is the assumption that studies which did not report the use of behavioral strategies did not use them. However, this may not have been the case. Teachers may have used behavioral strategies, such as verbal praise, despite the fact that the study did not report the use of a behavior component. No study explicitly stated that

behavioral strategies were not used. Yet, if the behavioral strategies were not included in the intervention protocol, teachers in the control groups should be equally likely to use these techniques as teachers in the treatment group, which should control for their effects.

There were also a limited number of studies in each group, which limits the generalizability of this findings. The studies that were included met rigorous standards which limited the number of studies meeting inclusion criteria. This was done in order to ensure quality of findings. Additionally, only peer-reviewed studies were included in the analysis which brings publication bias into play as a possibly limitation. Studies with statically significant findings are more likely to be published than those with non-significant findings (Easterbrook, Gopalan, Berlin & Matthews, 1991). Publication bias may lead to overall increased effect sizes across all studies. However, it should not affect comparisons between moderator variables as the publication bias would presumably be present across all moderator variable groups.

Finally, there was a large range of outcome measures across studies that were included in the analysis. Some measures were researcher created and were designed to assess direct transfer of the skill taught in intervention. Other studies used standardized achievement tests which assessed a broad range of mathematical abilities. These skills were not directly targeted in intervention. This lack of consistency in outcome measures brings into question whether it is appropriate to make comparisons across measures, as treatment effects could vary greatly depending on the outcome measure used. There was also a range in the type of behavior that was targeted. Some studies reinforced academic engagement while others targeted work completion or accuracy. There were likely

variations in effect sizes due to variations in target behaviors which were not accounted for in the present analysis.

Future Research

Due to the limited research around the use of behavioral strategies in academic intervention, there is a need for the expansion of research in this area. In addition, a variety of techniques, such as verbal praise, tangible reinforcement, and self-monitoring of progress should be compared in order to determine which techniques have the largest effect on outcomes. In addition to comparing techniques, it would be useful to study the differences in outcome by the type of behaviors targeted. For example, one could compare the reinforcement of academic engagement to the reinforcement for correct answers on an assignment.

In addition, the hypothesis for this study was formed based on the theory that the use of behavioral strategies leads to an increase in academic engagement which in turn leads to increases in academic outcomes. Therefore, future studies could include a measure of academic engagement in order to determine whether this theory is supported and increases in academic achievement are in fact moderated by academic engagement. Finally, future research in this area could also examine differences in the effects of behavioral strategies based on student characteristics, such as disability status, age, and gender.

References

- Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology*, 84, 261–271.
- Arthur Jr, W., Bennett, W., & Huffcutt, A. I. (2001). *Conducting meta-analysis using SAS*. Psychology Press.
- Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and instruction*, 7(2), 161-186.
- Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of School-Wide Positive Behavioral Interventions and Supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, 12, 133-148.
- Brooks, A., Todd, A. W., Tofflemoyer, S., & Horner, R. H. (2003). Use of functional assessment and a self-management system to increase academic engagement and work completion. *Journal of Positive Behavior Interventions*, 5(3), 144-152.
- *Campuzano, L., Dynarski, M., Agodini, R., & Rall, K. (2009). Effectiveness of reading and mathematics software products: Findings from two student cohorts (NCEE 2009–4041). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.
- Cleary, T. J., & Chen, P. P. (2009). Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of school psychology*, 47(5), 291-314.
- *Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative research on the Building Blocks project. *Journal for Research in Mathematics Education*, 38(2), 136–163.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cooper, J. O., Heron, T. E., & Heward, W. L. (2012). Applied behavior analysis.
- Crone, D. A., & Horner, R. H. (2000). Contextual, conceptual, and empirical foundations of functional behavioral assessment in schools. *Exceptionality*, 8(3), 161-172.
- Connell, J. P., Spencer, M. B., & Aber, J. L. (1994). Educational risk and resilience in African-American youth: Context, self, action, and outcomes in school. *Child Development*, 65, 493–506.

- Easterbrook, P. J., Gopalan, R., Berlin, J. A., & Matthews, D. R. (1991). Publication bias in clinical research. *The Lancet*, 337(8746), 867-872.
- Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research*, 59, 117-142.
- Finn, J. D. (1993). School engagement and students at risk. *National Center for Education Statistics*. Retrieved May 2, 2014 from <http://nces.ed.gov/pubs93/93470.pdf>
- *Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, 97(3), 493.
- Fuchs, L. S., Fuchs, D., & Compton, D. L. (2012). The early prevention of mathematics difficulty: Its power and limitations. *Journal of Learning Disabilities*, 0022219412442167.
- *Fuchs, L. S., Fuchs, D., Craddock, C., Hollenbeck, K. N., Hamlett, C. L., & Schatschneider, C. (2008). Effects of small-group tutoring with and without validated classroom instruction on at-risk students' math problem solving: Are two tiers of prevention better than one?. *Journal of educational psychology*, 100(3), 491.
- *Fuchs, L. S., Fuchs, D., Finelli, R., Courey, S. J., Hamlett, C. L., Sones, E. M., & Hope, S. K. (2006). Teaching third graders about real-life mathematical problem solving: A randomized controlled study. *The Elementary School Journal*, 106(4), 293-311.
- *Fuchs, L. S., Fuchs, D., Hamlett, C. L., Phillips, N. B., & Bentz, J. (1994). Classwide curriculum-based measurement: Helping general educators meet the challenge of student diversity. *Exceptional Children*, 60(6), 518-537.
- *Fuchs, L. S., Fuchs, D., Phillips, N. B., & Hamlett, C. L. (1995). Acquisition and transfer effects of classwide peer-assisted learning strategies in mathematics for students with varying learning histories. *School Psychology Review*, 24(4), 604-620
- *Fuchs, L. S., Fuchs, D., Prentice, K., Burch, M., Hamlett, C. L., Owen, R., & Schroeter, K. (2003). Enhancing third-grade student's mathematical problem solving with self-regulated learning strategies. *Journal of educational psychology*, 95(2), 306.
- *Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first-grade

- children's mathematical development with peer-assisted learning strategies. *School Psychology Review*, 31(4), 569-583.
- *Fuchs, L. S., Seethaler, P. M., Powell, S. R., Fuchs, D., Hamlett, C. L., & Fletcher, J. M. (2008). Effects of preventative tutoring on the mathematical problem solving of third-grade students with math and reading difficulties. *Exceptional Children*, 74(2), 155-173.
- Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools. NCEE 2009-4060. *What Works Clearinghouse*.
- Glover, T. A., & DiPerna, J. C. (2007). Service delivery for response to intervention: Core components and directions for future research. *School Psychology Review*, 36(4), 526.
- Goodman, G., & Williams, C. M. (2007). Interventions for increasing the academic engagement of students with autism spectrum disorders in inclusive classrooms. *Teaching Exceptional Children*, 39(6), 53.
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., Oliver, P. H., & Guerin, D. W. (2007). Multivariate latent change modeling of developmental decline in academic intrinsic math motivation and achievement: Childhood through adolescence. *International Journal of Behavioral Development*, 31(4), 317-327.
- Hackett, G., & Betz, N. E. (1989). An exploration of the mathematics self-efficacy/mathematics performance correspondence. *Journal for research in Mathematics Education*, 261-273.
- *Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a pre-kindergarten mathematics intervention: A randomized experiment. *Journal of Research on Educational Effectiveness*, 1(3), 155-178.
- Lane, K. L., Smither, R., Huseman, R., Guffey, J., & Fox, J. (2007). A function-based intervention to decrease disruptive behavior and increase academic engagement. *Journal of Early and Intensive Behavior Intervention*, 4(1), 348.
- Liaupsin, C. J., Umbreit, J., Ferro, J. B., Urso, A., & Upreti, G. (2006). Improving academic engagement through systematic, function-based intervention. *Education and Treatment of Children*, 573-591.
- March, R. E., & Horner, R. H. (2002). Feasibility and contributions of functional behavioral assessment in schools. *Journal of Emotional and Behavioral Disorders*, 10(3), 158-170.

- Marks, H. M. (2000). Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal*, 37(1), 153-184.
- National Mathematics Advisory Panel (2008). *Foundations for success: The final report of the national mathematics advisory panel*. Washington, DC: U.S. Department of Education.
- National Research Council. (2001). *Adding it up: Helping children learn mathematics*. Washington, DC: National Academies Press.
- Preciado, J. A., Horner, R. H., & Baker, S. K. (2008). Using a function-based approach to decrease problem behaviors and increase academic engagement for Latino English language learners. *The Journal of Special Education*.
- 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.
- Rock, M. L. (2005). Use of strategic self-monitoring to enhance academic engagement, productivity, and accuracy of students with and without exceptionalities. *Journal of Positive Behavior Interventions*, 7(1), 3-17.
- *Schunk, D. H., & Cox, P. D. (1986). Strategy training and attributional feedback with learning disabled students. *Journal of educational psychology*, 78(3), 201.
- Sinclair, M. F., Christenson, S. L., Evelo, D. L., & Hurley, C. M. (1998). Dropout prevention for youth with disabilities: Efficacy of a sustained school engagement procedure. *Exceptional Children*, 65(1), 7.
- Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research*, 95(6), 323-332.
- Skinner, C. H., Pappas, D. N., & Davis, K. A. (2005). Enhancing academic engagement: Providing opportunities for responding and influencing students to choose to respond. *Psychology in the Schools*, 42(4), 389-403.
- Steinberg, L. (1996). *Beyond the classroom: Why school reform has failed and what parents need to do*. New York: Simon and Schuster.
- Thalheimer, W., & Cook, S. (2002). How to calculate effect sizes from published research: A simplified methodology. *Work-Learning Research*, 1-9.
- Todd, A. W., Horner, R. H., & Sugai, G. (1999). Self-Monitoring and Self Recruited

Praise Effects on Problem Behavior, Academic Engagement, and Work Completion in a Typical Classroom. *Journal of Positive Behavior Interventions*, 1(2), 66-122.

What Works Clearinghouse. (2012). What Works Clearinghouse procedures and standards handbook v. 2.1.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American Educational Research Journal*, 45(1), 166-183.

Appendix

Table 1
Study Sample Characteristics

Author	Year	Journal	Gender (% Male)	Grade	% Disability
Campuzano	2009	<i>National Center for Education Research</i>	49	7.4	NR
Clements	2007	<i>Journal for Research in Mathematics Educataion</i>	51.5	Pre-k	NR
Fuchs	1995	<i>School Psychology Review</i>	NR	3	33.3
Fuchs	2002	<i>School Psychology Review</i>	55.83	1	5.5
Fuchs	2003	<i>Journal of Educational Psychology</i>	NR	3	10.13
Fuchs	2005	<i>Journal of Educational Psychology</i>	49.62	1	0
Fuchs	2006	<i>The Elementary School Journal</i>	51.68	3	7.71
Fuchs & Fuchs	2008	<i>Journal of Educational Psychology</i>	49	3	0
Fuchs & Seethaler	2008	<i>Exceptional Children</i>	43.4	3	22.9
Klein	2008	<i>Journal of Research on Educational Effectiveness</i>	51.8	Pre-k	NR
PCER	2008	<i>National Center for Education Research</i>	48	Pre-k	10.8
Schunk	1986	<i>Journal of Education Psychology</i>	56.7	7	100

Note. Author=first author, PCER=Preschool Curriculum Evaluation Research, year=year of publication, Journal=journal of publication, Gender=Percent of male participants, Grade=Mean grade of participants, Disability=Percent of sample with disability, NR=Not reported

Table 2
Effect Size Results by Study

Author	Year	Intervention Type	Outcome Measure	Hedge's g ES	Ave ES by Intervention group
Campuzano	2009	PLATO	SAT Math	-0.05	-0.04
		Achieve Now	Battery		0.16
		Larson Pre-Algebra	SAT Math Battery	0.16	
		Cognitive Tutor	ETS End-of-Year Algebra Assessment	-0.07	-0.07
		Larson Algebra	ETS End-of-Year Algebra Assessment	-0.01	-0.01
Clements	2007	Building Blocks	BBAEM: Numbers	0.90	1.34
			BBAEM: Geometry	1.79	
Fuchs	1995	PALS	MOT-R	0.028	
			MCAT	0.22	0.12
Fuchs	2002	PALS	SAT-PALS	0.06	0.06
Fuchs	2003	Transfer	Immediate transfer	5.76	
			Near Transfer	2.94	3.17
			Far Transfer	0.82	
		Transfer-Plus-SRL	Immediate transfer	9.12	5.17
			Near Transfer	4.91	

			Far Transfer	1.48	
Fuchs	2005	Math Tutoring	CBM	0.51	
			Basic		
			Addition	0.44	
			Basic		
			Subtraction	0.18	
			WJ-		
			Calculation	0.57	
			WJ-Applied		0.46
			Problems	0.11	
			Grade 1		
			Con/App	0.67	
			Story		
			Problems	0.75	
Fuchs	2006	SBI	Transfer	5.63	
			Near Transfer	5.78	
			Far Transfer		
			Q1	0.47	
			Far Transfer		
			Q2	-0.41	2.11
			Far Transfer		
		SBI-RL	Q3	0.31	
			Far Transfer		
			Q4	0.89	
			Transfer	4.10	
			Near Transfer	4.03	2.29
			Far Transfer		
			Q1	0.66	
			Far Transfer		
			Q2	1.47	
			Far Transfer		
			Q3	2.06	

			Far Transfer Q4	1.40	
Fuchs & Fuchs	2008	Hot Math SBI Tutoring	Transfer Near transfer Far transfer	1.69 1.67 0.41	1.26
Fuchs & Seethaler	2008	Preventive Tutoring	AFR SFR DD Addition DD Subtraction WRAT 3 Arith. SAE Jordan's SP Peabody WP KeyMath-R PS ITBS	0.25 0.52 0.51 0.42 1.34 0.60 0.64 1.82 0.48 0.21	0.68
Klein	2008	Pre-K Math + DLM Express math software	CMA	0.19	0.19
PCER	2008	Pre-K Math + DLM Express math software	CMA WJ-Applied Problems Shape Composition	0.13 0.02 0.005	0.05
Schunk	1986	Strategy Training	25-Item Subtraction Test	0.12	0.12
		Strategy Training with verbal praise	25-Item Subtraction Test	0.50	0.50

Note. BBAEM=Building Blocks Assessment of Early Mathematics, PALS=Peer Assisted Learning Strategies, MOT-R=Math Operations Test-Revised, MCAT=Mathematics Concepts and Applications Test, SAT-PALS=Stanford Achievement Test (Peer Assisted Learning Strategies aligned items), Hot Math SBI Training=, Transfer= Near transfer= Far transfer=, SBI=, SBI-RL=, Transfer-Plus-SRL=Transfer-Plus-Self regulated learning strategies, CBM Comp=Curriculum-Based Measurement Computation, WJ-Calculation

= Woodcock-Johnson III Calculation, WJ-Applied Problems = Woodcock-Johnson III Applied Problems, Grade 1 Con/App = First-Grade Concept/Applications, DLM =, CMA= Child Math Assessment, AFR = Addition Fact Retrieval, SFR = Subtraction Fact Retrieval, DD Addition = The Double-Digit Addition Test, DD Subtraction = The Double Digit Subtraction Test, WRAT 3 Arith. = Wide Range Achievement Test 3 Arithmetic, SAE = Simple Algebraic Equations, Jordon's SP = Jordon's Story Problems, Peabody WP = Peabody Word Problems, KeyMath-R PS = KeyMath-Revised Problem Solving, ITBS = Iowa Test of Basic Skills: Problem Solving and Data Interpretation,

Table 3

Overall Effect Size and Effect Size by Behavior Component Inclusion Results

Meta-analysis Statistic	Results		
Sampling Error Results	Overall	With Behavior Component	Without Behavior Component
Total number of studies	12**	5	5
Number of effect sizes	18	8	10
Total N	8479	1592	6887
Mean N	471.06	199	688.70
Min N	35	35	60.00
Max N	2588	378	2588.00
Range of N	2553	343	2528.00
Mean Observed Effect Size (Sample-weighted)	2.88	2.14	0.19
Min Observed Effect Size	-0.08	0.12	0.08
Max Observed Effect Size	5.17	5.17	3.17
Range of Observed Effect Sizes	5.25	5.05	3.25
Variance of Effect Sizes	2.83	2.29	0.37

SD of Effect Sizes	1.68	1.51	0.61
Variance due to sampling error	0.017355	0.031973	0.000583555
Percent variance accounted for	0.61309	1.39858	0.15706
Variance δ	2.81337	2.25417	0.37097
SD δ	1.67731	1.50139	0.60907
95% Confidence Interval			
Upper limit	6.16514	5.08676	1.38233
Lower limit	-0.40992	-0.79868	-1.00522
Chi-Square Value	2935.96*	572.009*	636.699*

Note. *Significant at $p=0.001$, ** Two studies include one treatment group with and one treatment group without a behavior component

Table 4

Overall Effect Size and Effect Size by Behavior Component Type Results

Meta-analysis Statistic	Results		
Sampling Error Results	Overall	Graphing Progress	Other Reinforcement (Points, verbal praise, and tangibles)
Total number of studies	5	3	3
Number of effect sizes	18	4	3
Total N	8479	1128	222
Mean N	471.06	282	74
Min N	35	120	35
Max N	2588	378	127
Range of N	2553	258	92
Mean Observed Effect Size (Sample-weighted)	2.88	2.66	0.51
Min Observed Effect Size	-0.08	0.12	0.46
Max Observed Effect Size	5.17	5.17	0.68

Range of Observed Effect Sizes	5.25	5.05	0.22
Variance of Effect Sizes	2.83	2.27	0.0058149
SD of Effect Sizes	1.68	1.51	0.0762558
Variance due to sampling error	0.017355	0.006684795	0.000583555
Percent variance accounted for	0.61309	0.29468	322.790
Variance δ	2.81337	2.26183	0
SD δ	1.67731	1.50394	0
95% Confidence Interval			
Upper limit	6.16514	5.60368	0.50841
Lower limit	-0.40992	-0.29176	0.50841
Chi-Square Value	2935.96*	339.355*	0.30980

Note. *Significant at $p=0.001$

Table 5
Effect Sizes for Moderator Variables

Moderator		Number of studies	# of ES calculated	Mean Sample-Weighted ES
Behavior component	Included	5	8	2.14
	Not included	5	10	0.19
	Both	2		
Behavior component type				
	Graphing Progress	3	4	2.66
	Tangibles/verbal praise/points	3	3	0.51
	Both	1	1	