THE ISOLATED AND COMBINED EFFECTS OF INTERVENTION COMPONENTS ON MATHEMATIC FLUENCY

By

KIMBERLY A. HASTINGS

Bachelor of Science in Psychology
Abilene Christian University
Abilene, Texas
2005

Master of Science in Educational Psychology
Oklahoma State University
Stillwater, Oklahoma
2006

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
July, 2010

THE ISOLATED AND COMBINED EFFECTS OF

INTERVENTION COMPONENTS ON MATHEMATIC

FLUENCY

Dissertation Approved:

Dr. Gary J. Duhon

Dissertation Adviser

Dr. Terry A. Stinnett

Dr. Brian C. Poncy

Dr. C. Robert Davis

Outside Committee Member

Dr. Mark E. Payton

Dean of the Graduate College

ACKNOWLEDGMENTS

I would like to first like to thank all of the people in my life that have stood by me throughout this graduate school process. To my wonderful husband, thank you for the countless hours of formatting excel spread sheets, cheering me on through each hurdle, and most importantly loving me throughout it all. I am so grateful to have you standing by me each step. You should receive an honorary degree for all of your help. To my parents, all four of you, each one of you inspired me to pursue this degree and taught me where hard work could take me. Thank you for your ambition, values, and support. To my loving friends, you got me through the day to day difficulties and kept me motivated to keep going. I am grateful that you rescued me from the stress and rallied around me when it got hard. To my advisor, thank you for your enthusiasm about research and the field as a whole. Your dedication is definitely something to be admired. And finally, to God, through which all things are possible and who listened and answered numerous prayers regarding this document and the entire graduate school journey.

TABLE OF CONTENTS

Chapter Page

LIST OF TABLES

Table Page

CHAPTER I

INTRODUCTION

Research on interventions can be categorized as having one of two primary purposes. The first and most obvious is to identify treatments that are empirically supported. Empirically supported treatments (EST) are treatments that have been determined as effective within the literature by two independent research teams (Hughes, 2000). In determining ESTs it must be demonstrated that a selected treatment be more effective when evaluated against no treatment, a placebo, and/or currently established treatments (Hughes, 2000). Within discussions of EST, there is a need for "sophisticated ways of identifying treatments that are empirical [methodologies], conceptually elegant, and clinically meaningful" (Doll, 2000, p. 332). Hughes (2000) argues that for the treatment to be identified as effective it requires controlled research (i.e., randomization, clinical trials) that demonstrates change that is not due simply to chance but a result of the application of the treatment. Additionally, EST should not be a quick fix for psychological problems but should offer relevant and guided recommendations of interventions that have been deemed as effective towards populations with similar excess or deficiencies (Doll, 2000; Hughes, 2000). What results are ESTs that are carefully operationalized allowing for evaluations of effectiveness, verification of integrity, and replication of intervention procedures (Kratochwill & Stoiber, 2000).

A second purpose for intervention research is to advance the application of ESTs in both practice and research. This is accomplished by applying these ESTs to the knowledge base in such a way that is meaningful to practitioners in applied settings as well as researchers in more controlled research settings (Kratochwill & Stoiber, 2000). Essentially, knowledge that a treatment is empirically supported is not sufficient to contribute to practice or research in a meaningful way. Along with the knowledge that a treatment is empirically supported, it is important to know for what problem types the treatment is likely effective and what severity is an appropriate match for particular treatment strength.

When evaluating effectiveness of interventions Noell and Gresham (1993) propose the assessment of both the costs and benefits of treatments. In evaluating benefit or effectiveness, Doll (2002) also poses the question 'what psychological services offer the best chance of restoring psychological health'? Essentially, knowledge about what treatment will result in the most change for the client is crucial for research on intervention effectiveness. However, to determine what treatment will likely result in the most change requires a direct comparison of interventions. Direct comparisons can provide us with information on the relative effectiveness of interventions. This is important not only for the purposes of having empirical evidence of intervention effectiveness, but provides practitioners a means of matching intervention effectiveness to intervention need. Currently there are multiple ways to evaluate intervention effectiveness which include treatment acceptability, treatment integrity, convergence evidence scaling, evaluating statistical effect sizes, and direct observation of outcomes (Albers et al., 2005). All have strengths and weaknesses when evaluating and comparing the outcomes of specific interventions which assist in evaluating intervention strength.

Treatment acceptability is one method used to assess outcomes for interventions. Nastasi and Truscott (2000) define treatment acceptability as the "consideration of the responsiveness of key caretakers and social environments to the assessment and intervention procedures" (p. 118). Many interventions are seen as more effective based upon the utility of the intervention and how easy or feasible it is for a teacher to implement. All of the acceptability measures include self-report and paper-pencil questionnaires that delineate a rating of acceptability. Finn and Sladeczek (2001) reviewed many of the most common treatment acceptability measures that have specifically been used in behavioral intervention research and state that there is deep caution to be warranted when outcomes are based upon inferences made from subjective reports of acceptability. The authors also warn that the different types of measures have varying degrees of intended outcomes (Finn & Sladeczek,2001). While an intervention might be rated as acceptable, there may be discrepancies in actual implementation (integrity). Another weakness of the treatment acceptability literature and use of the construct as an evaluation tool is that it is generally a one-time measure and is usually not carried throughout the intervention and consultation process (Finn & Sladeczek, 2001). A limitation to a single assessment of acceptability is that there is a greater likelihood of variability in the response and may not be a true score of the consultee's acceptability or a direct measure of effectiveness.

A second method utilized to evaluate treatment effectiveness is treatment integrity. Noell and Gansle (2006) identify treatment integrity as the percent that consultee adheres to the treatment plan. Treatment integrity comprised of not only the 'accuracy', but also 'consistency' in which an intervention is executed (Wood, Umbriet, Liaupsin, & Gresham, 2007). When treatment gains or lack of treatment gains are observed it is necessary to first evaluate whether the intervention was carried out with integrity. If integrity of the treatment plan is not intact then the effort, resources, and time put forth to developing an intervention is a moot point. Effectiveness of an intervention is often impacted by the amount of integrity displayed. Treatment integrity becomes a large concern for consultants and school psychologists but beyond training teachers and other professionals to adhere to the treatment plan there are other obstacles. Noell and Gansle (2006) address a complex issue of how much treatment integrity is enough for the treatment to be deemed effective? If one intervention component is omitted but all others are carried out with integrity, is integrity intact? Some would argue that it depends upon the step omitted. If the intervention component is crucial to the interventions effectiveness, such as a reinforcement component or an instruction component this may be of greater detriment to an intervention than another component. For instance if a behavioral intervention is constructed so that the student earns points through a token economy system and all procedures are followed except the student never has the opportunity to "cash-in" the token, then intervention is missing a crucial component in implementation. The omission of this component would likely impact the overall effectiveness of the intervention significantly. Another problem associated with treatment integrity is the differences in how it is measured across studies (Noell& Gansle, 2006; Noell, Witt, Slider, Connell, Gatti, Williams, Koenig, Resetar, & Duhon, 2005). A variety of measures are taken when attempting to evaluate integrity. Some might include self-report measures, permanent products, or direct observation of integrity (Noell, Duhon, Gatti & Connell, 2002; Noell & Gansle, 2006; Noell et al. 2005; Wood et al., 2007). Overall, when treatment integrity has been investigated it is generally low, and while there are measures to improve integrity it is still a critical concern when evaluating effectiveness of interventions (Noell, Duhon et al., 2002; Noell et al., 2005, Wood et al., 2007).

Statistical methods of evaluation have also been established to assess treatment effectiveness. Convergence Evidence Scaling uses rating forms to assess how well the goals of the intervention were met, as opposed to acceptability of the treatment (Albers et al., 2005). The ability to meet the specified goal of the intervention is rated on a five point likert scale with +2 goal was fully met, +1 goal partially met, 0 no progress toward goal, -1 the behavior is somewhat worse, and -2 behavior is significantly worse (Albers et al., 2005).

Similar to this metric is the RCI (Reliability Change Index). RCI is calculated through a standard score which is determined by discovering the difference between the outcome score and the predicted score divided by the standard error measurement for the difference between the two scores. Then the ratio could be converted to a z-score and any z-score above +1.96 could be deemed as 95% or more effective than chance (Nunnally & Kotsch, 1983). Statistical configurations of effect sizes are another common way that has been used to evaluate interventions (Albers et al., 2005). Effect sizes are found by taking the difference between the post-test and the pre-test and then dividing by the standard deviation. However this method is often criticized for using group averages rather than the sole improvement of one individual (Gresham, 1991). Since often the individual client is the concern of a practitioner, it would be warranted to investigate individual effects. Additionally, effect sizes utilize statistical test based upon probability, which often does not link with educational outcomes or needs (Gresham, 1991).

Direct observation is one of the least subjective methods of evaluating effectiveness. Direct observation can be employed to evaluate effectiveness through means of graphical interpretations of the data. Within the realm of data based decision making, graphs of the participants' gains or lack of gains can be informative to the teacher as well as others who may work with the student. There is great benefit to graphic analysis in that it is easily interpreted and makes the data more visually apparent (Albers et al., 2005). However, direct observation methods for one intervention may be measured by a different 'metric' making it difficult to compare to other interventions. When investigating the differences of 3 different spelling procedures and two evaluation procedures (cumulative learning versus learning rate) Cates, Skinner, Watson, Meadows, Weaver, and Jackson (2003) found that evaluating effectiveness was more apparent when using learning rate than the cumulative growth. Even when direct observation has been used effectiveness still can be unclear depending on the method to evaluate.

All of these methods for evaluating treatment effectiveness suffer from weaknesses. Even direct observations, which are most closely related to treatment outcome, make direct comparisons between interventions difficult mainly because growth or even learning rate alone are not sufficient to allow for direct comparisons between interventions. What is needed is not only a measure of growth, but a means of equating one intervention to another. To accomplish this, the second aspect of treatment effectiveness described as cost (Noell & Gresham, 1993) must be incorporated into the evaluation. Cost can essentially be separated into two components: the effort required from those involved in carrying out the intervention and effort exhibited by the student when engaging in the intervention. Noell and Gresham (1993) define effort as the cost of the time, energy, and overall hassle that consultants, consultees, and the clients encounter. A barrier to integrity and acceptability of interventions is the amount of effort needed to implement the intervention. While it is important to investigate effort on the part of the implementing agent there are often limitations to measuring effort. Effort can include the amount of time used, amount of money put forth to the preparation and implementation, and amount of personnel used. A concrete measurement of consultant effort would be difficult to convert to a common metric and keep constant. Additionally, the measure of effort could change based upon time, setting and even the specific personnel involved making consistency in evaluation impossible. While consultant effort is an important piece to understanding intervention outcomes, participant effort may be a more reliable concept to measure.

Participant effort can be characterized in a variety of ways. Several definitions include the amount of time allocated for learning, the scheduled time for instruction, or the opportunities to respond to instruction (Heward, 1984). While all valid avenues for measuring student effort, Heward (1984) argues that there is even a more intricate level of student effort: Active Student Responses (ASR). ASR is defined as "an observable response made to an instructional antecedent" (Heward, 1984, p. 286). Response could be defined as an actual response emitted by a student (words read, problems answered, supportive comments spoken, basketballs shot). This measure of effort can be used to investigate the actual instances a student actively produces a behavior rather than simply the opportunities a student may have to respond. Heward (1984) goes further to illustrate an active response as "a response that produces movement or change in the environment that can be detected by someone who can provide feedback to the student, be it a teacher, peer, or the student her/himself" (p. 286). At the core of ASR is not only how much instruction has been administered but also a measurement of how much learning has occurred. ASR depicts a more specific rate of learning and provides a conduit for equating interventions and when combined with the outcome of growth makes direct comparison of interventions possible. Strength is often defined as the magnitude of positive change produced by the treatment (Gresham, 1991; Lentz et al., 1996). The amount of growth achieved as a result of the intervention can also be characterized as strength or the overall benefit. Strength, as opposed to the other methods may provide an objective measure of effectiveness. For the purposes of direct comparison a strength ratio can be created by utilizing the benefit (strength/growth) as the numerator and cost (participant effort as measured by ASR) as the denominator. Gresham (1991) discusses that intervention strength [growth] may vary with the particular situation or environment. However the amount of cost incurred through the intervention process could potentially be held constant and therefore allow for direct comparisons to be made across intervention components. Potentially a strength ratio can remove much of the subjectivity encountered within the consultation processes and allow for objective behaviors to be compared.

Several factors are involved when evaluating interventions. While all suggested evaluation tools have validity in their own respects, there are still weaknesses. Direct observation grants clear and measureable outcomes, but how are outcomes compared? This study seeks to resolve the problem by evaluating a method of comparing interventions, by using a ratio of the benefit over cost. The purpose of this study will be to evaluate treatments using a strength ratio designed to compare growth of the intervention (benefit/strength) over the effort exerted by the participant (cost). Research Questions:

1) Does the addition of supplementary fluency building intervention components result in increased gains in fluency?
2) Which combination of fluency building intervention components, when active student responding (ASR) is held constant, results in the best strength ratio?
3) Using the strength ratio are the gains additive or synergistic?

CHAPTER II

REVIEW OF LITERATUE

A portion of the role of the school psychologist to advance already stable methods for enhancing learning but specifically identify effective solutions for student's academic problems (Daly, Hintze, & Hamler, 2000). In remediating student's academic problems some interventions may entail short, small group procedures others may require more involved explicit instruction (Bryant, 2005; Shaprio, 2004). A common skill targeted for intervention is fluency. The National Council for Teachers of Mathematics (2000) stresses the importance of building specifically math computational fluency in American schools (U.S. Department of Education, 2008). Fluency is being able to "perform the skill accurate and rapidly" (Daly, Lentz, & Boyer, 1996, p. 376). Once fluency of a mathematic skill is achieved it can be linked to generalizability or being able to apply the skills in multiple situations (Daly et al., 1996). Established in research are several methods of fluency building activities from explicit timing, (Rhymer, Henington, Skinner, & Looby, 1999; Van Houten, Morrison, Javis, &McDonald, 1974; Van Houten & Thompson, 1976) to performance feedback, (Brosovic, Dihoff, Epstein, & Cook, 2006;

Coding, Eckert, Fanning, Shiyko, & Solomon, 2006) to goal setting with contingent reward, (Fuchs, Bahr, & Rieth, 1989; Schunk, 1985) and various other procedures (Skinner, McLaughlin, & Logan, 1997). Research conducted using academic intervention should assist educators “to understand how variables like [opportunities to respond and prior learning] interact with one another to affect student performance” (Daly et al., 2000, p. 65).

A common tool used to assess student performance and instructional needs are Curriculum Based Assessment and Measurement (CBM/A) (VanDerHeyden & Burns, 2005). Curriculum based measurements (CBM) which are considered as ‘narrow band’ measures that usually measure one year’s worth of curriculum on one test (Shinn, 2002). The measurements are derived from the curriculum in place at that particular school or district within the general education classroom (Shinn, 2002). CBM is seen as a more treatment validated approach because of the direct link between what is taught and tested. The scores on CBM create a true score of what the student has gained from the instruction of the particular general education curriculum. A main purpose of CBM is to provide teachers with a simple series of assessments that they could use to evaluate their instruction and identify students that may need more or differential instruction in order to gain the skill (Deno, 2005). CBM is beneficial as an assessment tool is because it is sensitive enough to detect growth within short periods of time, used to measure the effectiveness of pre-referral interventions, and is a reliable source of data for screening or instructional planning (Burns, VanDerHeyden, Jiban, 2006; Deno, 2005). CBM cannot only serve as a progress monitoring tool but also a preventative tool to catch students early in their educational career that might be falling behind their peers (Shinn, 2002). CBM probes can also be utilized as a practice tool to increase fluency (Christ & Schanding, 2007).

Explicit Timing (ET)

Explicit Timing is a common intervention utilized to increase fluency among various academic tasks (Cates & Rhymer, 2006; Van Houten & Thompson, 1976). The premise behind the procedure of Explicit Timing is the idea of simple timed practice in order to improve students' ability to perform an academic task more quickly. Van Houten and Thompson (1976) implemented a timed procedure with a sample size of 20 in a second grade classroom. The students were given a packet 100 addition and subtraction worksheets and thirty minutes to work on the math facts. The skills were alternating within the packets. At the conclusion of the thirty minutes the teachers then graded the packets and then returned the packets the following day. At this point during the study the students were then asked to make the corrections on the problems. Van Houten and Thompson (1976) used a reversal design to evaluate results. The baseline condition was the procedures described above where the treatment condition involved the directions of the teacher explicitly stating that the students would have a thirty minute math period but at the end of each minute the students would be stopped and asked to draw a line where they completed. This went on for thirty minutes. The dependent measures the author used were the overall correct (number of problems completed accurately divided by 30 min) and local correct rate (number of problems completed accurately divided by the actual time that the children had available to work). Local correct rate was used due the transitions used for the treatment phase. The authors found that the students performed at higher fluency rates when the explicit timing condition was implemented for one minute intervals while keeping accuracy rates intact. The treatment procedure (ET) or also defined as a fixed ratio contingency was able to improve fluency upon mathematic facts.

Van Houten & Thompson (1976) was the beginning of utilizing explicit timing to improve learner's rates of fluency.

In a replication study, Rhymer, Henington, Skinner, and Looby (1999) examined the effects of ET on different populations compared a Caucasian sample to an African American sample of second grade students. The math skill selected for this investigation was a mix of single digit addition and subtraction problems. The students participated in the experimental procedures for two consecutive days. Students were administered both the timing condition and the non timing condition. In order to control for fatigue and practice effects classrooms were randomly assigned to the order in which they would receive the timing and the non timing conditions. During the untimed procedures students were given three sheets of problems and then were instructed to work on the packet, the students were not told they were being timed, yet the experimenters told them to stop after four minutes of practice. The timed procedure differed only in that the students had four, one minute intervals where they were told to circle the problem they were working on when the interval expired. Median digits correct per minute (DCPM) scores were used as the dependent measure for the conditions. A repeated measures ANOVA was conducted and showed a significant main effect for ET procedures, meaning students completed more problems during the timed procedures rather than the non-timed. When Race was added to the equation African American students completed more problems than the Caucasian students; however, there was not a significant main effect for Race. There was also no significant interaction between race and the treatment condition. Rhymer et al. (1999) also supports that ET is an effective method to increase mathematical fluency, but also an effective technique to increase fluency for various populations of students.

The use of explicit timing procedures has also been determined effective for other skills besides mathematics. Van Houten, Morrison, Jarvis, and McDonald (1974) used explicit timing procedures to increase writing composition skills. In an ABAB design to compare baseline simple mean of words written per minute to a explicit timing and feedback procedure as treatment. Feedback consisted of the students counting how many words they had written correctly after the timing procedure had completed. The student scoring was to serve as an immediate feedback for the student on how many words they had written. However, the following day the students returned to the baseline conditions were no explicit timing or feedback was presented. The final phase was a return back to the treatment. Results indicated that for the three different grades that participated in this study overall timing and feedback did impact the rate at which the students were writing sound stories.

Various studies have replicated the explicit timing procedures and compared the effects to other established fluency interventions. Clark and Rhymer (2003) sought to compare the effects of an explicit timing procedure versus an interspersal procedure. An interspersal procedure is where easier/known problems are intermixed with more difficult math problems. These authors were not only interested in the outcomes of the two interventions but also in the acceptability of the treatments as rated by the participants. The participants were 19 college students that were given a packet of a demographic survey and two 3 digit by 3 digit subtraction probes. These were to serve as control assignments prior to treatment. Trial one was comprised of an explicit timing sheet, followed by an interspersal sheet, and then a preference survey. Trial 2 was the counterbalanced order of the trial 1. During the trials the students had three minute time limit, however timing would occur in one minute intervals. The measures used for this study were the percent of total problems completed, the total target problems completed, and percent of target problems completed accurately. Results indicated that students completed more problems in the interspersal condition but completed more of the target problems in the explicit timing situation suggesting that both methods increased mathematic fluency.

Similar to Clark and Rhymer (2003), Codding, Shiyko, Russo, Birch, Fanning and Jaspen (2007) compared the explicit timing math fluency intervention to a Cover, Copy, and Compare procedure. The purpose of Codding et al. (2007) study was also to compare two empirically supported treatments with also a control condition on math fluency and to examine if these treatments would have different effects on students who were in a frustrational level of math fluency (<10 DCPM) versus students who were in an Instructional range of 10-19 DCPM . Cover, Copy, Compare (CCC) is also a fluency building procedure established in the literature (Skinner, McLaughlin, & Logan, 1997). In this procedure the student is presented with a set number of academic stimuli, then the student covers the stimuli and makes an academic response, then compares- which incorporates the student checking their work by uncovering the original stimuli to determine if accurate (Skinner et al., 1997). If the student was not accurate then the student participates in an error correction procedure. Codding et al. (2007) sought to answer which procedure (ET or CCC) when compared to a control group would produce the greater increases in slope over time. Procedures included screening all students who were solicited to participate by administering the Woodcock Johnson-III Tests of Achievement math calculation and fluency subtests to comprise the Math Calculation cluster. If students earned a standard score of 80 or above they were included in the study. After subjects were determined then the targeted math skill was chosen by using Curriculum Based Measurement Probes (single digit addition, single digit subtraction, and single digit multiplication). The majority of the sample was mastery at single digit addition, instructional at subtraction, and within the frustrational range for multiplication; therefore, subtraction was selected as the skill of choice. Codding et al. (2007) also chose to use CBM probes as the progress monitoring tool to measure growth twice a week for six weeks during the treatment phase. Similar to previous research, DCPM was used as the dependent variable measure. The students were then assigned to the three treatment groups control, ET, and CCC. Treatment occurred two times a week for 6 weeks. Procedures began with the students being in one large room where all participants were administered a two minute subtraction probe. No feedback was given to any of the treatment groups. The control group then left the room while the ET and CCC groups stayed. The ET and CCC groups participated in their treatments for five minutes. ET group took part in five one minute probes where as the CCC group had five minutes to solve as many problems in the packet as they could. In conjunction with the treatment the authors also administered a treatment acceptability measure after an explanation of each treatment condition in the study. The ET group rated their procedures as more favorable than the other two procedures where as the CCC and control group rated their respective treatments less positively. As discussed later, acceptability may contribute to effectiveness. A multivariate analysis of variance was conducted at the beginning of the study to indicate that there were no significant differences among the groups prior to treatment. Hierarchical Linear Model software was used to examine the effects of the treatment effects. Results indicated that slopes of the students over time determined that the average DCPM gain from session was .40 ($p<.001$), which indicated that there was “sufficient variation in DCPM over time” (Codding et al., 2007, p. 611). The results of this investigation indicated that neither ET or CCC was significant. However, the initial level of fluency that the student exhibited had significant effects on the overall performance as measured by DCPM. When an interaction analysis was conducted using initial fluency levels by ET treatment group significant effects were also discovered. This also translates initial fluency levels can serve as a valid predictor of the student's overall performance and the rate of the student's performance over time. All conditions led to increases in computational fluency- but differences emerged when initial CBM scores were considered. ET was considered the least and the most effective once initial fluency level was considered. If the student was in the instructional range of fluency (at or above 10 DCPM) the student improved in math fluency scores in single digit subtraction. However, for the student that was frustrational at single digit subtraction, these students had the lowest performance and progressed slower over time when compared to the CCC and control condition. Therefore for the student who is in the frustrational range CCC would be a better option to improve fluency. This study suggested that depending on level of achievement (frustrational, instructional, or mastery) should impact the selection of intervention to increase mathematical calculation fluency. Undoubtedly the student's need greatly impacts what intervention ideas should be considered.

Performance Feedback

Another manipulation that has been shown to increase mathematic fluency is the addition of performance feedback (Brosovic et al., 2006; Coding, Eckert, Fanning, Shiyko, and Solomon, 2006; Codding, Lewandowski, Eckert, 2005b; Rhymer, Dittmer, Skinner, & Jackson, 2000; Rhymer, Skinner, Hennington, D'Reaux, & Sims, 1998; Van Houten & Thompson, 1976). Performance feedback is the granting feedback on the student's performance evaluating their response (Van Houten & Thompson, 1976) . Performance feedback has also been used for various other skills to produce greater performance (Codding, Feinberg, Dunn, and Pace, 2005a; Gilbertson, Will, Singletary, and VanDerHeyden, 2007; Van Houten, 1979). While beyond the scope of this study feedback can also promote acquisition of new skills (Brosvic et al., 2006), however as discussed within the instructional hierarchy acquisition of new skills is the base for building fluency.

Brosvic et al. (2006) conducted three relatively linked studies investigating the differences between feedback methodologies. Study one sought to examine the differences between different techniques of delivering feedback upon acquisition of math facts for elementary students. The four treatment groups included delayed feedback, immediate feedback via visual analysis, immediate feedback via the educator, and a control group. The two immediate feedback conditions significantly impacted the participant's acquisition levels of the math facts but fluency as well, while the delayed and control groups were not significant. Study two however tried to remediate the non-effectiveness of the delayed and control group that they assigned those students to either the immediate feedback via the visual cue or the verbal feedback given by the educator. Significant growth was displayed in respect to fluency for those groups was observed as well. This evidence corroborated that performance feedback increased the acquisition skills. Study three conducted by Brosovic et al. (2006) continued to explore performance feedback on math acquisition but incorporating an error correction procedure condition. This "write-say" condition replaced the delayed feedback condition used in the first study. Results indicated that the "write-say" condition was just as effective as the previously investigated immediate performance feedback manipulations. Conclusions can be gained from Brosovic et al. (2006) in that immediate performance feedback can assist students in developing acquisition and fluency.

Performance feedback in combination with Cover, Copy, and Compare procedures was investigated to determine if feedback about the digits correct versus digits incorrect would affect mathematical fluency levels (Codding et al., 2006). The design chosen for this study was an alternating treatments design with counterbalancing the three conditions. The three conditions were CCC, CCC + performance feedback on DCPM, and CCC + performance feedback on digits incorrect per minute (DIPM). The findings indicated that there were no differences between any of the three conditions. All three increased the rates of math skill fluency. It is hypothesized that the design may have actually interfered with the treatment effects in that it could have created practice effects. Therefore a group investigation where treatments do not overlap could potentially provide more insight. A similar investigation was conducted with reading fluency specifically exploring if performance feedback on the words read correct versus words read incorrect would result in differentiated results (Eckert, Dunn, and Ardoin, 2006). Conducted much like the previously describe study the conditions were 1) oral reading fluency but no feedback, 2) feedback regarding the oral reading fluency with the words read correct per minute, and 3) feedback in reference to the student's words read incorrect per minute. Feedback for this study consisted of the amount of WCPM/WIPM the student earned and the student was instructed to try to beat the previous score. While previous research would suggest that feedback regarding the students words read correct per minute would influence the oral reading fluency, when in fact Eckert et al. (2006) displayed the performance feedback regarding words read incorrect per minute influence fluency greater. These findings are interesting in that they contradicted what was expected from previous research. While errors appeared to stay relatively constant, performance feedback upon WRICPM increased fluency levels as measured by WRCPM.

Using a multi-component treatment to increase math fluency Rhymer et al. (2000) evaluated a peer-tutoring dyad using three sets of problems and supplementary performance feedback. An alternative treatments single subject design was used when evaluating the performance of these students. The students were paired into dyads for the tutoring portion of the intervention. Baseline conditions consisted of the students answering the multiplication problems for the three sets which were assigned to the tutee, tutor, or control problems for a minute. The next phase was the intervention package where they students participated in their assigned set of problems. For a two minute interval the tutee had to answer as many flash cards that the tutor presented. The tutor provided feedback as to whether the answer was correct or incorrect, and if the answer was incorrect then the tutor said "incorrect, the answer is____". If incorrect the tutee then participated in an overcorrection procedure where they rewrote the problem and answer three times. Then the next flash card was presented. As the tutor was presenting the flash cards to the tutee the tutor would place them in the correct or incorrect piles so that the experimenter could document how many the participant answered correct and incorrect. Then the dyad completed an assessment sheet that contained the tutee's set of problems also identical to the baseline probes. Then the tutee and tutor switched roles however the participants used a different set of the flash cards (the now tutee's set). At the conclusion of the two minutes the students took another paper assessment of the tutee's problems followed by the control set. By the conclusion of the procedures the tutee and tutor were answering all three sets of problems. The intervention package treatment proceeded for eight sessions then the addition of the assessment performance feedback was added. The feedback involved that prior to the administration of the assessment probe the student was told their highest number of digits correct and directed to beat their score. While the tutoring dyad provided feedback on the session this phase bestowed feedback on the assessment session. After the conclusion of the study, the examiners took a maintenance point. Findings indicated that three of four students responded with slight increases in problems completed per minute after the implementation of the intervention peer-tutoring package and over correction. Two students displayed more increases on the tutee condition problems as compared to the 'tutor' condition. None of the students exhibited gains in the control problems. After the initiation of the assessment probe, performance feedback for three of the students demonstrated even greater increased levels of fluency. Maintenance data also displayed that the students were able to sustain relatively the same level of fluency two weeks after the completion of the study. Rhymer et al. (2000) displayed that addition of fluency building components influenced growth for three out of the four students. This study provides a starting point to understand how fluency can be built upon. Therefore continued research is needed to investigate the addition of already supported treatments to increase math fluency.

Finally, Van Houten (1979) used a delayed performance feedback when trying to increase writing fluency. While not as explicit as the performance feedback administered in the previously discussed treatments, the classes used public posting of the amount of words they wrote per minute. Students not only were able to see their overall performance they were able to engage feedback when they graded their own passages first for words written as well as action words. All of these investigations provide a base of literature to support the use of performance feedback to increase mathematical fluency.

Goal Setting

Goal setting procedures have been used to increase fluency levels on academics but also behavioral goals such as work completion (Konrad, Fowler, Walker, Test, & Wood, 2007). Codding et al. (2005b) investigated the effects of goal-setting and performance feedback upon two students diagnosed with ADHD. Specifically Codding et al. (2005b) was looking at the goal setting procedures on math fluency. One of the criteria for inclusion to the study was that the students had to display accuracy on the specified math skill however, slow with the rate of completion. The students fluency levels were determined by curriculum based measures of basic math skills. Digits correct per minute from the CBM probes were used as the dependent measure for the purposes of the study. Experimental design utilized was an alternating treatments design to compare the performance feedback and experimenter selected goal setting (PFEG) versus the performance feedback and the student selected goal setting procedures (PFSG). The differences between the two treatment groups consisted of either the teacher setting a overall target that the student was attempting to beat (mean from the baseline performance of the previous sessions performance) or the student themselves were to select a goal that they strove to beat. Performance feedback consisted of the student graphing their score from that session on a bar graph that had their previous data points. There was also a performance feedback component imbedded in grading their probes with the experimenter and by circling correct answers. The treatment conditions were randomized as to minimize sequence effects. At the final implementation of the treatment the experimenter chose the most successful treatment for that student to be implemented. Treatment sessions lasted for five weeks (20 sessions) for the two participants. Results indicated that for both participants, students were more successful when they utilized the student selected goals.

Schunk (1985) wanted to evaluate goal setting procedures with students who had been identified as learning disabled. Specifically the population displayed not only had an achievement math score 1-1.5 standard deviations lower than their WISC-R intelligence scores but was identified by their teachers as not being able to complete 25% or less of a subtraction skill test. Similar to Codding et al. (2005b), researchers found that goal setting procedures specifically the student selected goals presented greater growth when compared to an examiner selected goal. Although both groups displayed growth, greater change was attributed to the self-selected goals for students identified as learning disabled. It should be noted that within the procedures of the Schunk (1985) study there was specific instruction in setting appropriate goals rather than unrealistic goal setting. The accuracy of the subtraction skills was not affected by either goal setting procedure. A large limitation to this study is that the treatment was only carried out for a week. Fuchs, Bahr, and Rieth (1989) sought to remediated that limitation by replicating the Schunk (1985) procedure.

The purpose of the Fuchs et al. (1989) was to first review the differential effects of self-selected versus assigned goals on math performance and investigate the effects of contingencies on participant outcomes when goals are set for the student. Similar to the Schunk (1985) all students were identified as having a learning disability. The 20 participants were assigned to four treatment groups by a stratified random assignment. The four treatment groups: assigned goal with non contingent reinforcement, self-selected goals with non contingent reinforcement, assigned goals with contingent reinforcement, and self-selected goals with contingent reinforcement. This study required that the students engage in two minute randomly generated probes in four different skills (additions, subtraction, multiplication, and division). The dependent measure for the purpose of this study was the average correct digits per minute across all four problem types. Treatment consisted of a computer assisted practice program that consisted of four one minute practice trials during each computer session. The computer program was specified to move through successive skills once the criterion of 15 correct digits per minute was met during the practice session. There were a total of 16 objectives or skills to complete in the computer program. If criterion was not met than more problems that were targeted to meet the current objective were given until the student met the criterion. The goal for the student was then generated before four one minute sessions began. The goal would appear on the screen accompanied by the potential reinforcement the student could earn. The reinforcement was an amount of time to play a specified computer game. After each one-minute probe the students had the opportunity to play the computer game. Treatment differed for the groups by the modality of goals. For the assigned goal groups the goal was always to beat 15 DCPM which was determined as an ambitious but obtainable goal by the average performance of the sample population. Self-selected goals were selected from a group of three choices. The choices were paired with a specific duration that the student could participate in the game play that was chosen for the reinforcement. Authors chose various durations of the reinforcement according to how ambitious the goal was. For example, a goal of 10, which was deemed as un-ambitious when evaluating the pre-treatment means, was paired with 30 seconds of game play. A goal of 15 DCPM was deemed as a moderately ambitious the student had the opportunity to engage in one minute of the game play. The most ambitious goal was 20 DCPM, which when examining the pre-treatment means the students would actually be doubling the pretreatment means if the student's beat this goal. The goal of 20 was paired with 90 seconds of game play. During treatment conditions the self-selected goal was selected prior to the four one-minute sessions. Game play (reinforcement) was also manipulated by treatment group. Some groups received non contingent reinforcement in which they were able to participate in the game play associated with their goal (assigned or selected). Other groups were assigned to the contingent reward system where they were only allowed to participate in the reinforcement if they met or exceeded their goal (assigned or self-selected). For the purposes of analyses the authors took a pre, mid, and post-treatment measures. Growth from pre to mid treatment was greater for the students a part of the self-selected goal treatments versus the assigned goal treatment group. Growth from mid to post treatment was not significant. Whether the treatment group received non contingent reward vs. contingent reward contributed to the goal selection. The self-selected goal condition participants on average chose the 15 DCPM goal. The non-contingent reward group self-selected goals were somewhat higher than the contingent reward group. This study supports the impact that goal setting and reinforcement can have on math fluency. Fuchs et al. (1989) state that “these findings support the notion that participation during goal selection (a) improves performance outcomes (b) may affect performance by mediating the individual’s sense of goal commitment, and (c) may enhance the sense of potential accomplishment with which youngsters with learning disabilities approach learning tasks” (p. 558).

Goal setting can be effective in that the student can have the opportunity to set and reach their own goals, but teachers can feel satisfied that the students are accomplishing educational goals (Martens & Witt, 2004). When rewards are efficiently used such as with setting ambitious yet realistic goals for student success: “can they achieve on one of the ultimate outcomes of education”: for the child to become a ‘self-motivated life-long learner (Martens & Witt, 2004, p. 28). Bandura and Schunk (1981) investigated the differential effects of how goal specifications are presented affects student performance and subtraction knowledge. The treatment conditions were comprised of first a proximal goals, distal goals, no goals, and finally no treatment. The proximal goal treatment condition, the students were told by the experimenter that they “should consider setting themselves a goal of completing at least six pages” of the items each session (p.

589). This suggestion was made at the beginning of the treatment phase as well as the second session. Within the distal goals condition the children were told that they should consider setting a goal of finishing the 42 page packet by the end of the seventh session. Next, the no goals condition were instructed without any goals specified but the participants were told to finish as many problems as they could. Finally, the no treatment (control) group was presented the same packet of problems without any exposure to the material prior. For the purposes of the study students participated in a pretreatment and post treatment measure which consisted of a 25 subtraction problems ranging in difficulty. The findings indicated that the proximal goals setting procedure produced the most gains in the subtraction problems. Distal goals delivered moderate gains, followed by the no goals, while the no treatment made minimal (non significant gains). Even when goals are vaguely stated there appears to be a response in performance.

Reward

Also used in the literature to increase fluency of an academic skill that is already in the student's repertoire is the use of reinforcement. Rewards have been shown to increase interest and enjoyment in academic tasks; except for one sceanario- when non-contingent rewards are given independent performance often referred to as "non-contingent reinforcement" (Martens & Witt, 2004; Vollmer, Ringdahl, Roane, & Marcus, 1997). Martens and Witt (2004) argue that reinforcement can provide some incentive for practice of different academic skills, which int turn could promote practice in general among students. The fact that the students are able to progress through facts more quickly with less errors could generalize to be reinforcing itself (Martens & Witt, 2004). Noell, Gansle, Witt, Whitmarsh, Freeland, LaFleur, Gilbertson, and Northup (1998) utilized a multiple baseline design with 3 male students to demonstrate the utility of using reinforcement to increase oral reading fluency. Each of the three students worked on 3 different grade level passages. After baseline, contingent reward was applied, followed by reward, modeling and practice if words read correct per minute was not in the mastery range. If the student reached mastery range before the following phase change took place then there was no need to continue to the phase change. Contingent reward procedures consisted of the participant receiving a token coupon if the WCPM exceeded the median WCPM and if errors did not surpass the previous session. Modeling and practice procedures encompassed the experimenter reading the passage to the participant 20% more words than the median score achieved on the participant's previous session. Then the student would practice the reading the passage with performance feedback given on errors from the experimenter, then the time was recorded and repeated for all three passages per session. The Reward, Modeling, Practice phase was a combination of the contingent reward phase and the modeling and practice phase. Translating into trying to beat the previous sessions scores after the listening preview procedures. The first student was able to reach mastery on the 2nd grade passages with the contingent reward procedures however the third and fourth grade passages took until the final reward, modeling and practice phase to reach mastery. The second student showed some gains in the contingent reward phase but did not reach mastery until reward, modeling and practice phase was implemented for all three grades of material. The third student was at the mastery range for the 2nd grade material during baseline but consistently displayed mastery level when contingent reward phase was implemented. The third student never achieved to the mastery range on the third grade passages in any of the phases but was considerably closer to mastery on fourth grade passages during the modeling and practice phase. Overall the three participants were able to increase their reading rate by 59% or greater than baseline. While this particular study had some limitations in respect to the different phase changes it did display that the addition of reinforcement did increase reading fluency. The addition of reinforcement, especially with specific contingencies can improve fluency levels.

Literature on Computer Based Probes

Curriculum based measurement has been seen as predictive and informative when developing treatment for students who have difficulty in varying tasks (VanDerHeyden & Burns, 2005). Literature supporting the comparability of paper-pencil to computer based probes is in the beginning stages. Several studies have investigated tests such as the GRE, an exam in a business class, and the Nelson-Denny Reading Comprehension test (Clariana & Wallace, 2002; Mead & Drasgow, 1993; Pomplun, Frey, & Becker, 2002) . Differences that have existed when comparing paper-pencil and computer tests have been small. One has investigated the validity and reliability when comparing the paper-pencil versus computer based CBM (Duhon, Wong, & Mesmer, 2007).

Clariana and Wallace (2002) investigated the differences between a computer based test or a paper-pencil test with 105 freshman business students. All students that participated in the study were all receiving lectures in a face-to-face class, the only variable that differed was the mode the participants took their examinations. Researches sought to also account for individual learner characteristics such as gender, content familiarity, computer familiarity, and finally competitiveness. Clariana and Wallace (2002) used a post-test only with the dependent variable being the participant's score on a 100 item multiple choice test as well as a self-report survey on distance learning characteristics. The paper-pencil condition consisted of six or seven questions written on a page, where as the computer based test the student's received one question at a time on the screen where they were to click their answers. On both forms student were allowed to review previously answered questions. Item order also differed for the versions. The results indicated that the students who were in the computer based examination condition preformed significantly higher than the paper pencil students. When individual learner characteristics were analyzed, students with higher familiarity of the content preformed better than the lower familiarity group. Authors also investigated the interaction effects of test mode x content familiarity. The interaction was also significant, displaying that the computer based examination assisted the higher familiarity group when compared to the paper-pencil group or otherwise the paper-pencil based examination impeded the success of the high familiarity group. The other learner characteristics did not have a significant findings displayed through the examination score. This particular study provides evidence that computer methods of evaluation may actually allow for greater success for the student as far as this type of examination. This is a limited sample but it is unknown to what effect computer based examinations may have on children.

Computerized versions of other tests have also been developed. Pomplun, Frey, and Becker (2002) compared two forms of a computerized Nelson-Denny reading assessment to the same forms on paper-pencil. The Nelson-Denny Reading test is a standardized test to evaluate students vocabulary and comprehension skills. The test delineates a vocabulary, comprehension, and a total score. A unique feature of the Neslon-Denny test is that it is structured so that most students taking the test do not finish the reading assessment in the allotted amount of time. The authors were interested in the comparability of the two different modes of the Nelson-Denny test. The schools (two four year colleges, two community colleges, and one high school) depending on what type of institution were assigned to the treatment groups. However, some of the schools that dropped out which altered with some of the percentages of the students within the conditions. Using the stratified sample by institution the students were then randomly assigned. The mode (computer vs. paper-pencil) of the test was counterbalanced. Each student took the Nelson-Denny both on paper pencil but also on the computer. Dependent measures used were the scores of the students on the two sub scores (vocabulary and comprehension) as well as the total composite score. Research control methods were used to first eliminate outliers and then investigate if there were practice or fatigue effects. Results indicated that there were no practice or fatigue effects found. Also it was determined that there were minute differences between modality of the test. The largest difference was displayed when the vocabulary scores were compared. The computer based test generated higher scores specifically on the vocabulary sub score than the paper-pencil test. The authors attribute the differences to the different response procedures (bubbling in an answer vs. clicking an answer). Analyzing the amount of time took to bubble an answer versus click could add seconds and potentially even minutes to answer more questions on the test. By only allowing a limited amount of time, that is generally not sufficient to finish the test, could allow more seconds to proceed through more of the test. With more opportunities to earn more points this could contribute to higher scores. Ultimately the tests were not seen as completely comparable.

Other research has been done with computers and mathematics (Ku, Harter, Liu, Thompson, & Cheng, 2007; Reimer and Moyer, 2005; Wilson & Majsterek, 1996). Ku et al. (2007) found that personal individualized computer math instruction influenced attitudes and achievement in mathematics. Results indicated that there was a significant two way interaction in that the students with the lower level math achievement performed better at post test when they received the personalized computer instruction than the non-personalized lower math achievement. When investigating the overall treatment effect of the personalization of the computer program there was not a significant outcome; however, for the lower math achievement students the treatment did show significance. Another two way interaction was indicated for the higher-level math achievement students when examining the scores and problem type. Both the higher and lower level students performed significantly higher on the computational problems versus the word problems.

Cates (2005) sought to compare two math fluency interventions: peer drill vs. computer drill. This study was not designed to increase fluency but to investigate which method produce greater accurate responses. The study used two dyads of students (N=4). The peer tutoring method used was a technique in which the students tutor each other through a flash card program providing immediate feedback to their partner on the accuracy of the other student's ability to answer the math facts orally correct from the flash cards. The peer drill was timed for 3 minutes, the students were to answer as many facts as they could. The immediate feedback consisted of verbal recognition if the answer was right or wrong. If the student answered incorrectly then the corresponding peer would say try again. The computer drill was a flash card program where the flash cards were constructed to be similar to the peer tutoring flash cards. Analogous to the peer flash cards the students had 3 minutes to complete as many as they could. If the student responded correctly then a "ding" would sound, if incorrect then the same "ding" sounded only the same flash card stayed on the screen until the student answered it correctly. Results of Cates (2005) varied by dyad. One dyad responded at to the computer flash card system with higher rates of responding while the other dyad preformed the opposite. The second dyad attained greater levels of fluency during the peer-tutoring phase. The author also noted that the variability in responses could be linked to the fact that the second dyad was twin sisters. Overall, this computer assisted fluency program did provide gains for all students, just not as much for one dyad. Further research is warranted when using computer based programs to increase mathematical fluency.

Evaluating Effectiveness

Daly et al. (2000) argue that if assessment does not lead to effective treatment, then the utility of treatments is obsolete. Often school psychologists are asked and expected to be problem solvers from behavioral problems, to crisis plans, to academic remediation for general education. Without a base and knowledge of the effective treatments and literature the consultant has an empty bag of tricks. Systematically delivering treatment with methods to measure the individual effects is the core of the school psychologist's role. Currently there are multiple ways of evaluating effectiveness used, such as measures of treatment acceptability, treatment integrity, specific metric tools, and direct observation.

Treatment Acceptability

Treatment acceptability is assessed by rating forms addressing forms of social validity (Finn and Sladeczek, 2001; Graham, 1999). Generally these rating forms ask for the teacher to evaluate "does the treatment seem appropriate for the target behavior?" As discussed earlier, there are limitations in relying on an acceptability measure to determine whether the treatment is effective or not. While social validity is important for the purposes of buy in, acceptability measures are not the litmus test for effectiveness.

Finn and Sladeczek (2001) conducted a literature review of many of the treatment acceptability measures. These ranged from consultant, teacher, and even child acceptability measures. While each form displayed various degrees of usefulness the author's argued that the concept of "treatment acceptability" is more complex than at first believed. The technical and psychometric property of each of these acceptability measures does not provide evidence for one form being superior over another. The majority of these measures display adequate reliability and validity. However, as mentioned acceptability while supplementary to sound consultation practices and social validity concerns does not answer the question of treatment effectiveness in its entirety. Self-report, which is often so variable, without other forms of evaluating effectiveness has the potential to be biased (Finn & Sladeczek, 2001). While often multiple informants of acceptability could provide constructive advice in how to be perceived as an more effective consultant. Finn and Sladeczek (2001) advise that when measuring treatment acceptability that the practitioner will need to decide what purpose an acceptability measure will serve their intervention and treatment in order to decide on which rating from to use.

When evaluating treatment acceptability there is another component that influences the perceptions of the treatment: consultant effectiveness. Graham (1998) investigated the ratings of treatment acceptability and consultant effectiveness while there were differences in the approach (expert vs. collaborative) the consultant took during the consultation meetings. The authors also manipulated the mode the teacher explained the problem behavior (vague or clear terms) to the consultant. Results indicated that when the problem was defined in a vague manner by the teacher, raters preferred a more collaborative approach in solving the target problem. However, if the problem was defined more clearly ratings indicated that the consultant should exhibit an expert style.

A treatment acceptability measure was administered during this study to assess the alleged treatment plan that was developed for the purposes of this study. The results on the treatment acceptability measure were also influenced by the style of the consultant (expert or collaborative), but no significant main effect for the treatment acceptability measure and the request type (vague or clear). Graham (1998) displays that acceptability measures can even be predisposed by variables that specifically involve the methods of the consultant when the treatment plan is held constant. This is a giant limitation to basing intervention effectiveness on ratings of treatment acceptability.

Treatment acceptability is often measured by having the teacher rate the treatment, but what about the students that are receiving the treatment? These ratings of acceptability may reflect the amount of the treatment that they found to be unpleasant or impractical (Noell & Gresham, 1993). Waas and Anderson (1991) conducted an investigation of student's acceptability of common classroom interventions. The students were grouped by grades (second, fifth, and college students) and were presented the same scenario. The scenario included a child experiencing some behavioral problems and then presented three different common intervention ideas. After the treatment descriptions the participants rated the treatment using a Treatment Expectancy Scale as well as a treatment acceptability measure. Following statistical analyses the results indicated that the treatment acceptability and treatment expectancy measures are correlated. This is a finding that supports the idea that these two constructs are in fact related when students are evaluating treatments. Across each group of subjects the intervention that involved the target student being placed within a special education classroom (the more stringent intervention component) was rated most negatively. However, the behavioral contingency program with the most inclusion components was rated as the most positive. Wass and Anderson (1991), hypothesized that there would be a relation between the expected outcome, view of acceptability, and cooperation of the students. Once again these studies, support that there exists more components influencing the overall measures of treatment acceptability and may not be a true measure of treatment effectiveness.

Treatment Integrity

Consistency but also accuracy of the treatment protocol often defines treatment integrity (Wood et al., 2007). Practitioners are consistently presented with constructing an effective intervention but often outcomes may not progress in the path expected. Wood et al. (2007) investigated whether intervention outcomes could be attributed to the treatment or to the poor implementation of the intervention protocol. This is an important question to investigate for many reasons specifically for the efficiency of the consultant's time and resources (Noell & Gansle, 2006). When investigating data it may not be evident to the parent, teacher, or administration why the student appears as though they are not progressing when in fact it may be due to the intervention not consistently being implemented. Gilbertson et al. (2007) found that at after verbal instruction of the intervention that the integrity levels were between 50-25%. However when continued performance feedback is administered there is gains in intervention integrity (Codding et al., 2005a; Gilbertson et al., 2007)

Integrity procedures are not only important for academic interventions but for behavioral as well (Codding et al., 2005a). Wood et al (2007) provided behavioral consultation services for a student who was referred for disruptive behavior such as yelling out, pushing peers, or not complying with requests of the teacher. Through proper problem identification the consultant developed an intervention specifically targeting antecedents to increase the likelihood of the student displaying the appropriate replacement behavior. On-task behavior and treatment integrity data were collected at the end of each interval; therefore permanent products of implementation were collected through observation of the teacher going through the specific protocol. Simply evaluating the intervention data from might be concerning drops in the effectiveness specifically since there were some data points that returned to baseline levels of behavior. However, once integrity data was presented over the treatment data it is evident that treatment integrity is highly related to the intervention success. In fact the level of the student's success paralleled the treatment integrity data, thus indicating that the intervention was not being conducted with integrity. This study provides evidence that when evaluating interventions it is vital to calculate integrity, because treatment effects may be disguised by the integrity data. As displayed by Wood et al. (2007) integrity is essential, but often to get the levels of integrity that are acceptable it requires continued monitoring and direct assessment of the consultee.

Another investigation of treatment integrity was conducted using four different teachers who had all referred students for mathematic difficulties. Gilbertson et al. (2007) utilized a performance feedback procedure to increase treatment integrity. Baseline procedures was the verbal instruction of how to carry out the intervention. Only permanent products were collected during this phase and integrity rates ranged from 25-50% of the protocol. The next phase, faded three criteria classroom training, was implemented once integrity levels were stable and below 100%. The three criteria classroom training consisted of three teaching stages that included step by step, immediate feedback following the session, and then delayed feedback. Each stage was required to have at least one session of 100% integrity before progressing to the next stage. The third and final stage was the Response-dependent performance feedback which consisted of the removal of the consultant's feedback when the intervention was run with 100% integrity determined by the permanent products collected. If the integrity was not at 100% the consultant met with the teacher prior to the next scheduled time of implementation and provided a graph that displayed the student's success but also the percentage of the teacher's integrity. Positive feedback was given for the steps the teacher followed but for the steps omitted the consultant reviewed them. After the integrity performance feedback sessions the experiment advanced to the maintenance phase. Results obtained from Gilbertson et al. (2007) indicate that performance feedback from the consultants was salient to produce adequate levels of integrity. It should be noted the amount of feedback it took to gain the appropriate integrity performance from the teachers. Another key result gained from Gilbertson et al. (2007) is the trend of the student's achievement data in comparison to the integrity data. It is evident from the data of the student's achievement that their performance was correlated. In fact student's made gains that ranged from 25-150 digits correct per minute.

Codding et al. (2005a) also employed performance feedback procedures to increase treatment integrity for behavioral support plans. The six student's that participated for this study all exhibited problem behaviors such as yelling out, aggression, non-compliance, and instigation. The integrity was measured by how the teachers paired with these students implemented the procedures specified for the antecedent behaviors as well as the consequences. Overall performance feedback assisted the integrity levels but also allowed for continued communication between the consultants and the teachers. Codding et al. (2005a) provides more evidence that integrity can be improved with specifically performance feedback but that often it is imperative for the integrity of treatments to be implemented for the effectiveness of the treatment.

Direct Observations

Direct observation of the target behavior can provide insight to the level of growth as well as rate. "Direct assessment techniques that are being used widely in both practice and research are example of an outgrowth of methods from behavioral assessment research for which absolute units of measurement are the standard" (Daly et al., 2000, p. 64). While growth of the dependent variable is important to understand student's learning there are alternative method beyond the obvious. Cates et al. (2003) conducted a exploration of three different spelling techniques (traditional drill, interspersal, and a high-p procedure) when evaluating the data of the alternate treatment design the discovered there was a clearer indication of the learning when they evaluated the treatment by using learning rate. Cates et al. (2003) used two methods: cumulative learning and learning rate. Cumulative learning was defined as the cumulative amount of mastered spelling words across the conditions, where as learning rate was the number of words mastered per minute of instruction. When evaluating the graphical displays if Cates et al. (2003) for each participant the cumulative learning for the most part is very similar across the interventions however when learning rate is calculated there is a distinct difference for the method of treatment. This specific evaluation of learning creates some more insight in to how these particular methods of increasing spelling mastery operate. An important consideration is that while all three treatment procedures were effective, utilizing the learning rate procedure allowed for the distillation of which treatment could potentially provide greater growth in the same amount of time. There is a great need for more research regarding direct assessment to address many of the methodological issues that truly affect outcomes of academic interventions (Daly et al., 2000).

Strength and Intensity

Intervention strength is often compared to medical definitions of dosage, what level strength or potency can deliver an effective outcome? (Noel & Gansle, 2006). How much of the dosage is necessary to produce adequate effects- is often the concern of the school psychologist as well as the teacher. Lentz et al. (1996) provides simple points when considering the strength of interventions. These characteristics are as follows:

1) *An intervention is a set of events and relationship within problem ecology.*
2) *Strong interventions have socially significant, lasting outcomes that extend beyond the term that any short term intervention is actually in place.*
3) *Strong interventions disrupt the existing classroom/school ecology as little as possible to achieve significant results.*
4) *Outcomes of strong interventions satisfy those who presented the problem.*
5) *Strong interventions result from a structured problem solving sequence that allows accurate hypotheses about problem maintenance and the type of intervention matches that hypotheses. Intervention planning in reference to the existing knowledge base is most likely to produce this relationship.*

6) *Strong interventions contain elements to promote and insure treatment integrity (usually involves both efforts to promote acceptability as well as monitoring of and feedback about treatment process.*

The second point is vital, in that the outcomes achieved by the intervention must continue to provide the advantage for the student and the teacher. Results of interventions would be considered obsolete if the effects were not carried throughout. The third point discussed by Lentz et al. (1996) is important not only for the acceptability considerations from the teacher but services for the targeted student. Intervention procedures should not be so distracting and demanding that it is impossible for it to be executed within the classroom. Utilizing the naturalistic environment will ensure the fit and more acceptable procedures to the teacher (Lentz et al., 1996). Strong interventions do not translate into multi-faceted, rigorous individual instruction. The intent behind intervention/consultation services is to promote parsimonious individualized treatment to promote achievement. As discussed earlier acceptability of the intervention is important for the “buy in” and integrity concerns but also that the teacher might be willing to use the intervention procedures when she/he faces a similar problem behavior in the future. One of the most important qualities of strong interventions is that they must be “correctly matched to the reason underlying the problem” (Lentz et al., 1996, p. 121). Without the match the intervention would not be influencing the target problem. Intervention strength is “multidimensional in nature and involves consideration of treatment duration, specificity, and relevance for a particular type of problem, as well as intensity (Daly, Martens, Barnett, Witt, & Olson, 2007).

Intensity more refers to the frequency and/or complexity of the intervention (Daly et al., 2007). While strong interventions might show conceptually sound interventions but intensity refers to how much of the intervention will of the strong intervention will be needed for a student to gain the necessary skills. The strength and intensity are linked in that when a robust intervention is used often if growth is not occurring at the rate to be expected intensity of the intervention may need to be adjusted. Duhon, Atkins, Mesmer, Greguson, and Olinger (2009) utilized a math intervention to increase frequency to evaluate if this increase would in fact produce the effects desired. Using a group of 35 students participated in the initial phase of the experiment where they used CBM and a goal setting procedure with reward to increase digits correct per minute. All but three students appeared to respond to the phase one of implementation which lasted for 20 sessions. These three students continued on to phase two of the procedures where frequency of the intervention was increased until all students achieved the benchmark goal. Intensity for phase two was to increase the frequency of the intervention to five sessions daily and for the students who were not at benchmark goals intensity increased to ten times per day. One month after data collection a maintenance data point was collected. Results indicated that for the majority of students the base (phase 1) intervention was enough to meet benchmark, however for three students more intensive (phase 2) was needed. Two of the students who responded at 5x the frequency maintained on benchmark level, the student who required 10x the initial intervention did not maintain the previously gained levels of fluency. Duhon et al. (in submission) argue that using this metric of intensity could be informative to understanding strength and intensity, indicating how much support is needed for academic success for all students.

Effort

Frequently when consultation services are warranted there is an amount of effort required in order to serve the student. Implementation effort could be delineated to the amount of resources, time, and energy expended to train, implement, and monitor the effects (Noell & Gresham, 1993). Unfortunately when an intervention is developed that requires too much *effort* from the teacher, it is either rejected or implemented with poor integrity (Noell & Gansle, 2006; Noell & Gresham, 1993). In some instances more treatment is effective, but a balance must be reached in how much effort is put forth to solicit appropriate outcomes (Noell & Gansle, 2006). As discussed in Lentz et al. (1996) interventions should be naturalistic of the classroom not only for the purposes of implementation integrity but also so that the effort on the part of the teacher is minimal. It is a common occurrence to hear teachers state that there is not enough time to execute the prescribed intervention. Noell & Gresham (1993) state the need for evaluation techniques that evaluate the cost relative to the benefit of the intervention. This could be attempted by the functional outcome analysis proposed by Noell & Gresham (1993). Gains in achievement or increases of the desired behavior over the cost (time, resources, and or monetary cost) could delineate the overall 'cost-benefit' analysis of implementing a certain intervention. For the practitioner this could be linked to the strength of the intervention. For instance, if an intervention is conceptually strong and is at fairly low intensity the Functional Outcome Analysis (FOA) ratio would portray the benefit over the cost. Essentially, the more intensive the intervention it would be expected that the denominator of cost (aka effort) would increase therefore impacting the FOA ratio. While FOA would provide indispensable information it in fact may add more effort for the stakeholders implementing the interventions (Sheridan, 1993). Effort is arduous to measure without adding more resources and time. This is not an argument to distil the importance to evaluating the cost benefit analysis, but that there could potentially be other methods to evaluate various interventions when dissecting the power behind them.

Participant Effort and ASR (Active Student Response)

Participant or student effort can amount of time engaged in the learning task or the opportunities to respond to the learning task. While each of shown growth in student performance (Burns, 2001; Swing, Stoiber, & Peterson, 1988). At the basis of student effort is the concept of Active Student Responses (ASR). ASR as discussed earlier is the most core absolute form of measurement for learning (Heward, 1984). The only way to essentially measure if a student understands a concept is if the student can demonstrate the performance of the skill. By utilizing this concise measurement of learning can be the first step to contributing to the advancement of school psychology (Daly et al., 2000). ASR can provide a link to evaluating intervention strength to effectiveness. Standard measurement of absolute variables has the potential to greatly increase the literature about the functional relationships of learning (Daly et al., 2000). "When a student produces fluid, accurate responses to academic tasks in the sequence desired by the teacher, the teacher infers that the student "knows" or has mastered the content"- however this display of knowledge is just a behavioral chain that has come to fruition

Student learning unfortunately can be a passive act (Heward, 1984). Increasing ASR can not only increase learning in students but the maintenance effects sustain (Barbetta, Heron, & Heward, 1993; Gardner, Heward, & Grossi, 1994). Barbetta et al. (1993) differentiated an error correction procedure while teaching student's identified as having a developmental disability sight words. The two procedures involved a no-response component and the ASR error correction. If the participant missed the sight word on the flash card, then the experimenter indicated to the student that the answer was wrong and then modeled the correct word, the student then was asked to repeat the word. Where as in the no-response error correction the student received feedback that they had said the word wrong but just asked to look at the word as the teacher told them specified sight word. The dependent measure for was the correct responses. The ASR phase without a doubt increased the student's performance when compared to the no response manipulation on the same day tests as well as the next day tests of the sight words. Maintenance data was also collected two weeks later also indicated that ASR error correction was superior. Barbetta et al. (1993) argue that requiring the participants to emit a response contributed to the overall effectiveness of this intervention.

Building upon previous study, Gardner et al. (1994) applied the ASR principles to a classwide intervention to increase learning in a science lecture. Manipulations included Hand-Raising (HR) versus Response Cards (RC) to increase student's opportunities to respond but also the effects of the manipulations upon maintenance of the material being instructed. Five students from the class were selected by the teacher and the experimenter to be representative of the class as a whole. Using an ABAB design the experimental manipulation alternated between HR and RC. Hand raising would ask questions during the lesson and students were to raise their hand if they knew the answer. The teacher would wait 3 seconds before calling on an individual student. If the student was correct she would acknowledge so; however, if incorrect she would give feedback as to the correct answer. For the Response Cards condition the five students were given dry erase boards to write their responses to the teacher's questions. The teacher would ask questions during the lecture identical to the HR procedures. When the teacher asked questions the students would write their answers and then hold up their response cards. At this point the teacher would scan the answers and give feedback accordingly. The RC condition required that the students emitted a response to every question and receive feedback accordingly, while the HR procedures allowed for only one student to truly emit an answer and receive feedback. Gardner et al. (1994) used four dependent variables: teacher presentation rate of questions, the number of student responses, accurate student responses, next day score upon a quiz, and maintenance by a bi-weekly quiz. The teacher asked a mean of 1.54 questions per minute, while the student responded during the HR sessions 9.9 times throughout the session. During the RC sessions the student averaged 21.8 responses throughout the session. When the students were responding to the teacher's questions they were above 90% accurate (HR: 92% and RC: 93%). The next-day quiz scores corresponded with the treatment phase. Higher scores were observed during the RC condition as compared to the HR, the same was true for the maintenance points too. Overall, results indicated that promoting more active student responding increased learning for the same-day quizzes as well as the maintenance points. The link between requiring more active responding for students and learning is substantial.

As displayed by various investigations, Active Student Responding is pivotal to enhancing greater learning for students (Barbetta et al., 1993; Gardner et al., 1994; Heward, 1984). A more important question to be inquiring within research is how much Active Student Responding are these interventions requiring of our students, and is the shortcomings due to the lack of the amount of responses the student is executing.

CHAPTER III

METHODOLOGY

Participants and Setting

The participants were solicited from five 3rd grade classrooms in a Midwest rural school district. Teacher and parent permission was obtained in order for the participants to be involved in the study. The study took place within the computer labs located within the respective school buildings. A total of 97 students were recruited. Four students were dismissed from the study due to a low accuracy rates (below 80%), four more students' data were dismissed due to an inaccurate post test, and five students were removed from the study due to the lack of a post test score. After attrition and exclusion of participants who did not meet criteria, the population was a total of 86 third grade students (44 male, 42 female).

Materials

Materials consisted of a web based mathematic flashcard program. The flash card program contained randomized electronic flash cards that post to the computer screen. The participants responded to the flash card by typing their answer using either set of generated flash card is presented on the screen. Once a predetermined number of responses have been emitted the program would terminate.

Dependent Measure

The dependent measure is the amount of work completed within a minute or digits correct per minute (DCPM). When scoring using DCPM the individual digits if the number are within the correct column (Shapiro, 2004). For example if the problem is "7 x 3 =" equaling an answer of "21". An answer of 41, would result in one digit correct because the one in the ones column is correct. In order to compute DCPM since the responses are being held constant the amount of seconds for the student to complete the fifty problems was recorded. The digits correct was then be multiplied by 60 then divided by the seconds taken to complete the 50 ASR.

Procedures

Target Skill Identification.

Participants were given single skill addition, subtraction, and multiplication math probes to determine accuracy and fluency levels for each skill administered (Deno, 2005). Each single skill math probes were administered via the web based program. After initial CBM measures were administered a targeted skill of addition was identified after evaluating fluency levels observed within the population. The simple skill of addition resulted in the majority of participants with performance at or above 80% accuracy.

Pre-Test

The pre-test consisted of one administration of target skill probes. The probe was administered using the computer based program until the student completed 50 problems. Digits correct per minute were calculated for and served as the pre-test measure. The participants were then rank ordered and assigned to one of the four treatment conditions or control group.

Treatment Conditions

There were four treatment conditions which will consisted of individual and combined treatments. The treatments utilized will employ the components of explicit timing, goal setting, and reward and will be combined as described in the four conditions below.

Condition 1: Explicit Timing (ET). Students engaged in an explicit timing procedure where the problems are held constant rather than the time. While the students were completing the flash card program, a timer positioned on the screen calculated the time had elapsed since initiating the session.

Condition 2: ET combined with Performance Feedback (PFB). This condition was combined the previous condition (ET) with performance feedback occurring throughout the session. The feedback given was an audible tone and a visual cue indicating that the student had answered the question correctly or that they made an error.

Condition 3: ET combined with Goal Setting and Reward (GS^+). This condition combined ET and the addition of a goal setting procedure with a reward for meeting or exceeding an assigned goal. Participants were timed while completing the single skill math probe, however, prior to initiating the probe students were given a goal based on the previous session's score. The students were instructed to work to exceed the goal, and cued to the amount of rewards they had earned for beating their score throughout the week. If the goal was met or exceeded then the student received a reward. The reward was delivered at the conclusion of the week. The rewards consisted of various tangibles approved by the teachers (mechanical pencils, stickers, pencil holders, etc.).

Condition 4: ET + PFB + GS^+. This condition will combine all previous intervention components into one intervention.

Explanation of treatment components:

Explicit Timing administered on the computer (ET).

Explicit Timing (ET) is the administration of timed practice at fixed ratio contingencies to improve fluency of an academic task (Cates & Rhymer, 2006; Van Houten, Morrison, Javis, & McDonald, 1974; Van Houten & Thompson, 1976). In previous research, participants are usually asked to practice for an extended period of time. For the purpose of this study the participants practiced the specified math skill not for an explicit amount of time but rather for an explicit number of active student responses. Active Student Responses (ASR) is defined as the "observable response made to instructional antecedent" (Heward- book chapter, pg. 286). The participant will then practice the specified math skill until they have reached 50 active student responses (problems).

Performance Feedback.

Performance feedback has determined as an effective manipulation to increase student's fluency levels in various academic tasks (Brosovic et al., 2006; Codding et. al., 2004; Codding et al., 2006; Rhymer et al., 2000). Performance feedback provides the student with immediate response to their progress (Van Houten & Thomspson, 1976). Specifically, studies have utilized performance feedback in conjunction with other procedures such as goal setting (Codding et al., 2004), Cover, Copy, Compare (Codding et al., 2006), and peer tutoring (Rhymer et al., 2000) to increase growth in mathematical computation. Based upon the current efficacy of performance feedback interventions, continued research is warranted to understand the additive of component performance feedback to other fluency based interventions.

Goal Setting.

Several studies have investigated the efficacy of adding a goal setting component to increasing fluency, work completion, and other academic tasks (Codding, Lewandowski, & Eckert, 2005; Fuchs, Bahr, & Rieth, 1989; Schunk, 1986). In a goal setting condition the student is encouraged to try to meet a specified goal either self-selected goals or examiner selected goals (Codding et al., 2005; Fuchs et al., 1989; Schunk, 1986). Often goal setting (GS) is an effective method for increasing student behavior (Codding et al., 2005; Fuchs et al., 1989; Schunk, 1986). For the purposes of this study the participants engaged in the experimenter selected goal setting which consisted of the participant using the previous session's time as the goal. At the beginning of the session the participant was given their goal upon the computer screen. The participant then began the computer program where they were required to work continuously on the specified math fluency skill until they meet 50 ASR. At the conclusion of the session that particular student's time was presented on the screen.

Reward.

Contingent reward systems have been effective for also increasing fluency in academic tasks (Noell, Gansle, Witt, Whitmarsh, Freeland, LeFleur, Gibertson, & Northup, 1998; Smith & Lovitt, 1976). The participant was asked to complete 50 problems on the computer. Once the student had completed their 50 problems of active student responding their time from that day was displayed on the screen. On the following day, when the student logged in there was a running total of their rewards upon the screen. The rewards were earned throughout the week then on Fridays the students could choose their rewards from the goodie box.

Post-Test

Three post-test administrations were completed at the conclusion of the treatment phase (following 21 consecutive school days). The median DCPM score was used as the post-test measure.

Design

The design utilized for this study is a pre-test post-test control group experimental design. The pre-test measurement consisted of the participants initial time to complete 50 ASR, which was then converted to DCPM. Using the pre-test data students were then rank ordered by their (DCPM) and assigned to one of the four treatment groups or control.

Determining growth and the strength ratio

In order to evaluate the comparisons of different interventions this study suggests the idea of utilizing a strength ratio. The strength ratio of an intervention was defined as the response (change from baseline) divided by the effort. For the purposes of this study response will be defined as the growth from pre-test and post-test assessments, while effort will be defined as the active student responses emitted by the participant. However, if response is actually the difference from baseline the ratio could be utilized to compare interventions that decrease behavior in conjunction with those that increase behavior.

Denominator: Effort.

Heward (1994) argued that the most intricate part of learning is not the opportunities to respond but the active student responses (ASR) that a student engages in to demonstrate proficiency and receive feedback. In order to keep a common denominator for each participant included in the study, the amount of ASR per session was held constant at 50 ASR per session for twenty one sessions for a total of 1050 ASR. Translating each participant per session engaged in 50 active responses per math skill probe. ASR being held constant allows for rates of comparison among participants and treatment groups. Daily the students were not able to complete treatment each day until the 50 responses were complete. Integrity of this data was investigated daily through the database that contained each student's answers for each problem.

Numerator: Response.

Response for the purpose of the strength ratio was defined as the change from baseline observed in the various treatment groups. The change from baseline for each treatment is the difference between the pre-test score of digits correct per minutes (DCPM) and the post-test score of DCPM. Due to holding the effort constant the response was converted to DCPM. This will be done by multiplying the digits correct (constant ASR) x 60 seconds divided by the participant's time to complete the session, yielding the participant's DCPM. Growth will then be determined by subtracting the pre-test DCPM group mean from the post-test DCPM group mean.

Integrity

Integrity data was measured on 77% of the treatment sessions throughout the study. Measurement was conducted through a check sheet that one clinician managed while the other assistants carried out the procedures. Treatment procedures were followed with 99.8% integrity.

Analysis

Initially an ANOVA was conducted upon the pre-test scores to determine that there were no significant differences between the treatment groups prior to the initiation of treatment. In order to investigate group differences after treatment an Analysis of Covariance was selected due to the slight differences between the groups at the post test. By utilizing a covariate, differences will then be adjusted for when investigating the treatment effects upon the posttest. Once it is determined if there is a significant effect upon the dependent variable as a function of the covariate, pairwise comparisons were examined to determine differences between the five treatment groups.

.

CHAPTER IV

RESULTS

Target Skill Identification

Students were administered addition, subtraction and multiplication probes via the web based program to determine accuracy and fluency levels for each skill. Addition was selected as the skill for study because this skill provided the highest accuracy rates for the participants with 94% of the participants with accuracy rates above 80%. For the subtraction skill probe, only 65% of the participants had accuracy levels above 80%.

Participants were then rank ordered by fluency levels and using a randomized block design assigned to the five treatment groups. The assignment of participants was selected to control for the varying fluency levels.

Pre-Test

Pre-test means and standard deviations for each treatment group are as follows: ET= 21.32 (6.93), ET + PFB = 17.95 (5.22), ET + GS^+ = 17.65 (6.85), ET + PFB + GS^+ = 18.25 (7.89), and the Control = 21.07 (8.76) (See Table 1).

Post Test

The posttest was administered following the 21st day of treatment. Four students were dismissed from the study due to an inaccurate (below 80% accuracy) posttest score and five additional students were dismissed due to the lack of a posttest score. Posttest mean scores as a function of the treatment group for ET= 33.23 (13.04), ET + PFB= 34.02 (7.66), ET + GS^+ = 32.60 (10.81), ET + PFB + GS+ = 34.93 (14.51), and Control 24.44 (8.97) (See Table 1).

Analysis of Group Differences

An analysis of covariance was used to assess which math fluency intervention provided the greatest gains when the pre-test was utilized as a covariate. Results indicate that overall the treatments did have a significant effect upon math fluency, $F(4,80) = 10.68$, $p < .001$ (See Table 2). An examination of the pairwise comparisons reveals that all treatment groups significantly differed from the control group, p< .001. Pairwise comparisons also indicated that ET versus the ET+PFB+GS^+ also significantly differed from each other, $p = 0.032$, while ET vs. ET + PFB was approaching significance at $p = 0.054$ (See Table 3). The first research question this study sought to answer was "does the addition of supplementary fluency building intervention components result in increased gains in fluency"? By the pairwise comparisons revealing that the combination intervention (ET + PFB + GS+) was statistically significant from the base intervention (ET), the first research question can be answered that supplemental fluency building interventions does result in increased gains in fluency.

Determining growth and the strength ratio

The second question this study proposed to address was to utilize a strength ratio to evaluate which treatment offered the best strength ratio, when active student responding (ASR) was held constant. The strength ratio for each intervention was calculated by taking the difference from baseline or growth (post test mean - pre test mean), then dividing the growth by the effort (ASR). The strength ratio for each intervention are as follows: ET = 0.0113, ET + PFB = 0.0153, ET + GS^+ = 0.0142, and ET + PFB + GS^+= 0.0159 (See Table 4). By examining the strength ratio alone the combination intervention (ET + PFB + GS^+) offered the largest ratio.

Finally, the strength ratio was employed to further investigate if the additive components of the interventions would provide additive, synergistic, or diminishing returns effects. Additive results would produce findings that each treatment would sum to equal to the strength ratio of ET + PFB + GS^+ (0.0159). If ET + PFB + GS^+ intervention's strength ratio (0.0159) was greater than the sum of strength ratios from each treatment, then the effects would be considered synergistic. Finally, if the sum of the treatments was actually less than ratio of the ET + PFB + GS^+ treatment, then the result would be considered to be a diminishing returns effect. By using the ET strength ratio as the base intervention (0.0113) the individual treatment's strength can be ascertained. ET's strength can then be removed from the ET+PFB intervention (0.0153): [ET + PFB: 0.0153]- [ET: 0.0113]= 0.0040. ET's strength can also be taken away from the ET+GS^+ (0.0029) strength: [ET+$GS^{+:}$ 0.0153] - [ET: 0.0113] = 0.0029 (See Table 4). If each additional component's ratio is added, then the sum is 0.0182 [(ET + PFB + GS^+) 0.0113 + 0.0040 + 0.0029 = 0.0182]. The strength ratio yielded by the actual combination intervention (ET+PFB+GS^+) was 0.0159. The individual components sum was greater than the actual strength ratio yielded by the (ET+PFB+GS^+) indicating that effects were not additive, nor synergistic but actually described as a diminishing return.

CHAPTER V

DISCUSSION

This study proposed a method to compare the effects of mathematics fluency interventions. To accomplish this, active student responses (Heward, 1983) were held constant in contrast to previous studies and practice that have often held time constant (Cates & Rhymer, 2006, VanHouten, et al. 1974). By holding the active student responding constant, the strength ratio proposed made it possible to evaluate the intervention effect more precisely. Results indicate that the combination of all intervention components (ET + PFB + GS^+) offered the strongest treatment when compared to the other interventions; however by evaluating each component, this study was able to investigate the effects of each intervention component in isolation. For instance, when the simple effects of the base intervention (ET) were taken away from the ET+PFB treatment, the result yielded a remaining strength ratio of 0.0040 and when the (ET) intervention ratio was subtracted from the ET + GS^+ intervention it produced an isolated ratio of 0.0029. The sum of all of the intervention components in isolation was greater than the actual strength ratio of the combination intervention indicating that the ET + PFB + GS^+ intervention was weaker than each of the components added together, resulting in what might be described as a diminishing return. With that in mind, the intervention that could then be considered the strongest would be the ET +

PFB with a strength ratio of 0.0150.

For the purposes of this study, “effort” was defined as the participant effort and more specifically the ‘Active Student Responses’ that the students were required to produce. As discussed, the ASR was held constant across all participants. While this allowed for a simple metric of effort, it may have not provided the most comprehensive description of effort. Within the schools and practice there could be many definitions for “effort”, especially when implementing evidence based treatments within a school environment. Effort on the part of the implementers, the amount of time needed, monetary cost, as well as effort on the part of the student can all impact integrity and effectiveness (Noell & Gansle, 2006; Noell & Gresham, 1993). This study sought to investigate growth when student effort could be measured and held the same across all participants. If implementation effort could also be quantified, a more comprehensive cost/benefit ratio could be employed to evaluate treatments. Future studies should consider how to incorporate the implementation effort into the strength ratio.

Previous research and practitioners have often used growth or change from baseline as a measure of effectiveness (Codding et al., 2005). By evaluating growth over effort creating a “cost/benefit” metric (strength), this study has provided an innovative way of comparing interventions within the schools and practice. The results of this study suggest that ET + PFB offers the greater “strength”, which could essentially inform interventionists that in order to produce the most growth with the most efficient level of effort, ET + PFB should be selected. While each intervention resulted in growth, the ET + PFB is considered the most efficient.

While the original hypothesis predicted that the combination intervention (ET + PFB + GS^+) would offer the most strength this was not the case. As with all studies there are limitations that must be discussed. One limitation is that while the goal setting component was evident everyday through a prompt on the computer, the reward was not delivered every day. Each day the students knew how many rewards they would receive at the end of the week. Due to the group administration of the interventions, all treatment groups were intermixed throughout the class. To deliver rewards everyday to some of the children and not all may have caused disruption in the classroom. Delivering rewards immediately would have also required more effort on the part of the implementer. While considering what is feasible to implement in a classroom or small group instruction delivering rewards weekly rather than daily could offer an easier procedure for the teacher or interventionist. However, future studies may want to consider delivering rewards daily, immediately following the ET + GS^+ and the ET + PFB + GS^+ interventions.

This study is not suggesting that the delivery of a reward is obsolete. Results of this study indicate just the opposite. By investigating the statistical analysis, all treatments were considered significantly different from the control. Explicit Timing, Performance feedback, goal setting with rewards, and the combination of all of those components were all effective interventions, this study is proposing that each offer a different level of strength. It is important to evaluate strength of interventions because while some interventions may offer growth they may require the student to produce a more effort.

In considering the performance feedback condition versus the goal setting with reward, can one answer the question: is not goal setting with reward also a performance feedback condition? While GS^+ did not offer immediate feedback, it did offer an overall feedback by simply telling the student that they beat their score for that day. If GS^+ is another form of performance feedback, is there any way to delineate the two conditions? This may be a question for future research to consider.

Active student responding was utilized to hold the effort on the part of the student equal. This allowed for a direct comparison of interventions components. While the goal of this study was to investigate group effects, an investigation into individual growth and rates of change could provide insight within the context of an RtI framework. When working with individuals to increase fluency, knowing intervention strength can offer guidance in how to produce that growth in the most efficient manner. Holding the amount of ASR constant was a parsimonious way to compare interventions but in doing so the study limited the students who may have been able to actually achieve greater growth if time had been held constant. For instance, a student who is able to complete 50 problems within 3 minutes could potentially finish 65-70 problems if given a longer time frame of say 5 minutes; therefore, allowing for more active student responding and potentially becoming even more fluent and achieving more growth over the course of the treatment. Another potential limitation to consider is the duration of the treatment. The participants received the treatment conditions for 21 days. While previous studies have been implemented for approximately the same amount of time, ASR was not held constant. Students may have received the same amount of days but not the same amount of forced responses. So the same phenomenon described earlier occurs in this procedure. The higher fluency children are earning more active student responses because they can move more rapidly through the problems. So essentially, they are able to become more fluent. By holding the responses constant this study could have limited the growth. The question then remains, is it the goal to increase growth or to evaluate strength?

In conclusion, ASR did allow for comparisons but when conducting math fluency interventions targeted for a student or several students it might be more beneficial to hold time constant to produce greater growth, however, this research simply points to what intervention offers the best strength to utilize within the context of a Tier 2 or Tier 3 RtI framework.

Another potential limitation of this study is the consideration of the strength ratio as a whole and the generalizability of this idea to other intervention research. As discussed, previous research has utilized other methods of comparison such as z-scores, effect sizes, and the

Reliability Change Indices.   However effect size and z-scores are often criticized for using group averages rather than the sole improvement of one individual (Gresham, 1991). Additionally, effect sizes utilize statistical test based upon probability, which often does not link with educational outcomes or needs (Gresham, 1991). Is the strength ratio a comparable metric? This study argues that by including the amount of responses and holding them constant the strength ratio potentially could offer an innovative method of comparison and potentially the beginning of a new metric.

This study offered a proposed way to evaluate and compare interventions. Intervention strength can inform practitioners of what empirically supported treatment offers the most strength or “bang for the buck” and can be guided to implement accordingly. For instance, if it is not feasible to offer a reward every week, ET or ET + PFB may be a better fit for the needs of the student and the teacher. Utilizing the strength ratio, could also have future benefit for other intervention research as well. While previously, school psychologists have been aware of the need to implement Empirically Supported Treatments, we know have a method to investigate the strength of the treatments in order to provide the “best psychological health” for our clients (Doll, 2002).

REFERENCES

Albers, C.A., Elliot, S.N., Kettler, R.J., and Roach, A.T. (2005). Evaluating intervention outcomes. In R. Brown-Cidsey (Ed.), *Assessment for Intervention: A Problem Solving Approach* (pp. 329-351). New York: Guilford Press.

Bandura, A. and Schunk, D.H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586-598.

Barbetta, P.M., Heron, T.E., and Heward, W.L. (1993). Effects of active student response during error correction on the acquisition, maintenance, and generalization of sight words by students with developmental disabilities. *Journal of Applied Behavior Analysis, 26*(1), 111-119.

Burns, M.K. (2001). Comparison of opportunities to respond within a drill model when rehearsing sight words with a child with mental retardation. *School Psychology Quarterly, 22*(2), 250-263.

Burns, M.K., VanDerHeyden, A.M., Jiban, C.L. (2006). Assessing the instructional level for mathematics: A comparison of methods. *School Psychology Review, 35*(3), 401-418.

Bryant, D.P. (2005). Commmentary on early identification and identification for students with mathematics disabilities. *Journal of Learning Disabilities, 38*(4), 340-345.

Brosovic, G.M., Dihoff, R.E., Epstein, M.L., and Cook, M.L. (2006). Feedback facilitates the acquisition and retention of numerical facts series by elementary school students with mathematics learning disabilities. *The Psychological Record, 56*, 35-54.

Cates, G.L. (2005). Effects of peer versus computer-assisted drill on mathematics response rates. *Psychology in the Schools*, *42*(6), 637-646.

Cates, G.L. and Rhymer, K.N. (2006). Effects of explicit timing on elementary students' oral reading rates of word phrases. *Reading Improvement*, *43*(3), 148-156.

Cates, G.L., Skinner, C.H., Watson, T.S., Meadows, T.J., Weaver, A., and Jackson, B. (2003). Instructional effectiveness and instructional efficiency as considerations for data-based decision making: An evaluation of interspersing procedures. *School Psychology Review*, *32*(4), 610-616.

Clark, S.L. and Rhymer, K.N. (2003). Student preference for explicit timing and interspersal procedures as a function of math prolem completion rates: Testing the discrete task completion hypothesis. *Journal of Behavioral Education, 12*(4), 275-286.

Clariana, R. and Wallace, P. (2002). Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology, 33*(5), 593-602.

Codding, R.S., Eckert, T.L., Fanning, E., Shiyko, M., and Solomon, E. (2007). Comparing mathematics interventions: The effects of cover-copy-compare alone and combined with performance feedback on digits correct and incorrect. *Journal of Behavioral Education, 16*(2), 125-141.

Codding, R.S., Feinberg,A.B., Dunn, E.K., and Pace, G.M. (2005a). Effects of immediate performance feedback on implementation of behavior support plans. *Journal of Applied Behavior Analysis*, *38*(2), 205-219.

Codding, R.S., Lewandowski, L., and Eckert, T. (2005b). Examining the efficacy of performance feedback and goal-setting interventions in children with ADHD: A comparison of two methods. *Journal of Evidence-Based Practices for Schools, 6*(1), 42-58.

Codding, R.S., Shiyko, M., Russo, M., Birch, S., Fanning, E., and Jaspen, D. (2007). Comparing mathematics interventions: Does initial level of fluency predict intervention effectiveness? *Journal of School Psychology, 45*, 603-617.

Daly, E.J., Lentz Jr., F.E., & Boyer, J (1996). The Instructional Heirarchy: A conceptual model for understanding the effective components of reading interventions. *School Psychology Quarterly, 11*(4), 369-386.

Daly, E.J., Hintze, J.M., and Hamler, K.R. (2000). Improving practice by taking steps toward technological improvements in academic intervention in the new millennium. *Psychology in the Schools, 37*(1), 61-72.

Daly, E.J., Martens,B.K., Barnett, D., Witt, J.C., and Olson, S.C. (2007). Varying intervention delivery in response to intervention: Confronting and resolving challenges with measurement, instruction, and intensity. *School Psychology Review, 36*(4), 562-581.

Deno, S.L. (2005). Problem solving assessment. In R. Brown-Chidsey (Ed.), *Assessment for Intervnetion: A Problem Solving Approach* (pp. 10-40). New York: Guilford Press.

Doll, B. (2000). A commentary on Hughes: The tension between pragmatics and rigor. *Journal of School Psychology*, *38*(4), 331-335.

Duhon, G.J., Atkins, M., Mesmer, E.M., Gregurson, L., & Olinger, E (2009). Quantifying intervention intensity: A systematic approach to evaluating student response to increasing levels of intervention intensity. *Journal of Behavioral Education*, *18*(2), 101-118.

Duhon, G.J., Wong, C., and Mesmer, E.M. (2007). Evaluating reliability and validity of computerized version of CBM. Poster Presented at the National Association of School Psychologists conference, New York, NY.

Eckert, T.L., Dunn, E.K., and Ardoin, S.P. (2006). The effects of alternative forms of performance feedback for elementary-aged students' oral reading fluency. *Journal of Behavioral Education, 15,* 149-162.

Finn, C.A., and Sladeczek, I.E. (2001). Assessing the social validity of behavioral interventions: A review of treatment acceptability measures. *School Psychology Quarterly*, *16*(2), 176-206.

Fuchs, L.S., Bahr, C.M., and Rieth, H.J. (1989). Effects of goal structures and performance contingencies on math performance of adolescents with learning disabilities. *Journal of Learning Disabilities*, *22*(9), 554-560.

Fuchs, D., Mock, D., Morgan, P.L., & Young, C.L. (2003). Responsiveness-to-intervention: Definitions, evidence, and implications for learning disabilities construct. *Learning Disabilities Research & Practice*, *18*(3), 157-171.

Gardner III., R., Heward, W.L., and Grossi, T.A.(1994). Effects of response cards on student participation and academic achievement: A systematic replication with inner-city students during whole-class science instruction. *Journal of Behavior Analysis*, *27*(1), 63-71.

Gilbertson, D., Witt, J.C., Singletary, L.L., and VanDerHeyden, A. (2007). Supporting teacher use of interventions: effects of response dependent performance feedback on teacher implementation of a math intervention. *Journal of Behavioral Education, 16*, 311-326.

Graham, D.S. (1998). Consultant effectiveness and treatment acceptability: An examination of consultee requests and consultant responses. *School Psychology Quarterly*, *13*(2), 155-168.

Gresham, F.M. (1991). Conceptualizing behavior disorders in terms of resistance to intervention. *School Psychology Review 20*(1), 23-36.

Heward, W.L. (1984). Three "low-tech" strategies for increasing frequency of active student response during group instruction. In W.L. Heward, T.E. Heron, D.S. Hill, & J. Trapp-Porter (Eds.), Focus on behavior analysis in education (pp. 283-320). Columbus, OH: Merrill.

Hughes, J.N. (2000). The essential role of theory in the science of treating children: Beyond empirically supported treatments. *Journal of School Psychology, 38*(4), 301-330.

Konrad, M., Fowler, C.H., Walker, A.R., Test, D.V., and Wood, W.M. (2007). Effects of self-determination interventions on the academic skills of students with learning disabilities. *Learning Disability Quarterly, 30,* 89-113.

Kratochwill, T.R. and Stoiber, K.C. (2000). Empirically supported interventions and school psychology: conceptual and practice issues—part II. School Psychology Quarterly, 15(*2*), 223-253.

Ku, H., Harter, C.A., Lieu, P., Thompson, L., and Cheng, Y. (2007). The effects of individually personalized computer-based instructional program on solving mathematics problems. Computers and Human Behavior, 23, 1195-1210.

Lentz, F.E., Allen, S.J., and Ehrhardt, K.E. (1996). The conceptual elements of strong interventions in school settings. *School Psychology Quarterly, 11*(2), 118-136.

Martens,B.K. and Witt, J.C. (2004). Competence, persistence, and success: The positive psychology of behavioral skill instruction. *Psychology in the Schools, 41*(1), 19-30.

Mead, A.D. and Drasgow, F. (1993). Equivalence of computerized and paper-and-pencil cognitive abilities tests: A meta-analysis. *Psychological Bulletin, 114*(3), 449-458.

Nastasi, B.K. and Truscott, S.D. (2000). Acceptability research in school psychology: Current trends and future directions. *School Psychology Quarterly, 15*(2), 117-122.

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

Noell, G.H., Duhon, G.J., Gatti, S.L., and Connell, J.E. (2002). Consultation, follow-up, and implementation of behavior management interventions in general education, *School Psychology Review, 31*(2), 271-234.

Noell, G.H. and Gansle, K.A. (2006). Assuring the form has substance: Treatment plan implementation as foundation for assessing response to intervention. *Assessment for Effective Intervention, 32*(1), 32-39.

Noell, G.H., Gansle, K.A., Witt, J.C., Whitmarsh, E.L., Freeland, J.T., LaFleur, L.H., Gilbertson, D.N., and Northup, J. (1998). Effects of contingent reward and instruction in oral reading performance at differing levels of passage difficulty, *Journal of Applied Behavioral Analysis, 3*(4), 659-663.

Noell, G.H. and Gresham, F.M. (1993). Functional outcome analysis: Do the benefits of consultation and prereferral intervention justify the costs?. *School Psychology Quarterly, 8*(3), 200-226.

Noell, G.H., Witt, J.C., Slider, N.J., Connell, J.E., Gatti, S.L., Williams, K.L., Koenig, J.L., Resetar, J.L., and Duhon, G.J. (2005). Treatment implementation following behavioral consultation in the schools: A comparison of three follow-up strategies, *School Psychology Review, 34*(1), 87-106.

Nunnally, J.C., and Kotsch, W.E. (1983). Studies of individual subjects: Logic and methods of analysis. *British Journal of Clinical Psychology, 22*, 83-93.

Pomplun, M., Frey, S., and Becker, D.F. (2002). The score equivalence of paper-and-pencil and computerized versions of speeded test of reading comprehension. *Educational and Psychological Measurement, 62*(2), 337-354.

Reimer, K. and Moyer, P.S. (2005). Third-graders learn about fractions using virtual manipulatives: A classroom study. *Journal of Computers in Mathematics and Science Teaching, 24*(1), 5-25.

Rhymer, K.N., Dittmer, K.I., Skinner, C.H., and Jackson, B. (2000). Effectiveness of multi-component treatment for improving mathematics fluency. *School Psychology Quarterly, 15*(1), 40-51.

Rhymer, K.N., Henington, C.,Skinner, C.H., and Looby, E.J. (1999). The effects of explicit timing on mathematics performance in second-grade Caucasian and African American students. *School Psychology Quarterly, 14*(4), 397-407.

Rhymer, K.N., Skinner, C.H., Henington, C., D'Reaux, R.A., and Sims, S. (1998). Effects of explicit timing on mathematics problem completion rates in african-american third-grade elementary students. *Journal of Applied Behavior Analysis, 31*, 673-677.

Schunk, D.H. (1985). Participation in goals setting: Effects on self-efficacy and skills of learing-disabled children. *The Journal of Special Education,* 19(3), 307-317.

Shapiro, E.S. (2005). *Academic skills problems: Direct assessment and intervention, third edition.* New York, NY: Guilford Press.

Sheridan, S.M. (1993). Functional outcome analysis: Do the costs outweigh the benefits? *School Psychology Quarterly, 8*(3), 227-230.

Shinn, M.R. (2002). Best practices in using curriculum-based measurement in a problem-solving model. In A. Thomas & J. Grimes (Eds.), *Best Practices in School Psychology IV* (Vol. 1, pp. 671-697). Bethesda, MD: National Association of School Psychologists.

Smith, D.D. and Lovitt, T.C. (1976). The differential effects of reinforcement contingencies on arithmetic performance, *The Journal of Learning Disabilities, 9*(1), 32-40.

Skinner, C.H., McLaughin, T.F., and Logan, P. (1997). Cover, copy, and compare: A self-managed academic intervention effective across skills, students, and settings. *Journal of Behavioral Education, 7*(3), 295-306.

Swing, S.R., Stoiber, K.C., and Peterson, P.L. (1988). Thinking skills versus learning time: Effects of alternative classroom-based interventions on students' mathematics problem solving. *Cognition and Instruction, 5*(2), 123-191.

United States Department of Education (2008). *Foundations for success: The final report of the national mathematics advisory panel.*

Van Houten, R., Morrison, E., Jarvis, and McDonald, M. (1974). The effects of explicit timing and feedback on compositional response rate in elementary school children. *Journal of Applied Behavioral Analysis, 7*(4), 547-555.

Van Houten, R. and Thompson, C. (1976). The effects of explicit timing on math performance. *The Journal of Applied Behavioral Analysis, 9*(2), 227-230.

Vollmer, T.R., Ringdahl, J.E., Roane, H.S., and Marcus, B.A. (1997). Negative side effects of noncontingent reinforcement. *Journal of Behavior Analysis, 30*, 161-164.

Waas, G.A. and Anderson, G.P. (1991). Outcome expectancy and treatment acceptability: perceptions of school-based interventions. *Professional Psychology: Research and Practice, 22*(2),149-154.

Wilson, R., Majsterek, D., and Simmons, D. (1996). The effects of computer-assisted versus teacher-directed instruction on the multiplication performance of elementary students with learning disabilities. *Journal of Learning Disabilites, 29*(4), 382-390.

Wood, B.K., Umbreit, J., Liaupsin, C.L., and Gresham, F.M. (2007). A treatment integrity analysis of function-based intervention, *Education and Treatment of Children, 30*(4), 105-120.

APPPENDICES

Table 1

Pre- and Posttest Mean Scores and Standard Deviations as a Function of Treatment Group

Source	Pre-Test		Post-Test	
	M	SD	M	SD
Explicit Timing	21.32	6.93	33.23	13.04
Explicit Timing + Performance Feedback	17.95	5.22	34.02	7.66
Explicit Timing + Goal Setting	17.65	6.85	32.60	10.81
Explicit Timing + Performance Feedback + Goal Setting	18.25	7.89	34.93	14.51
Control	21.07	8.76	24.44	8.97

Table 2

Analysis of Covariance of Posttest DCPM Scores as a Function of the Treatment Conditions, With Pretest Scores as a Covariate

Source	df	SS	MS	F	ω^2
Covariate	1	6136.47	6136.47	117.27**	.594
Treatment Group	4	2235.69	558.92	10.68**	.348
Error	80	4186.21	52.33		
Total	86	99259.236			

*p < .05 **p < .01

Table 3

Pairwise comparisons between treatment groups

Group		Mean Difference	Standard Error	Significance
ET	ET + PFB	-4.856	2.483	0.054
	ET + GS^+	-3.798	2.632	0.153
	ET + PFB + GS^+	-5.403	2.478	0.032*
	Control	8.488	2.520	0.001**
ET + PFB	ET	4.856	2.483	0.054
	ET + GS^+	1.058	2.499	0.673
	ET + PFB + GS^+	-0.0547	2.347	0.816
	Control	13.345	2.440	0.000**
ET + GS^+	ET	3.798	2.632	0.153
	ET + PFB	-1.058	2.499	0.673
	ET + PFB + GS^+	-1.605	2.499	0.523
	Control	12.286	2.591	0.000*

ET + PFB + GS^+	ET	5.403	2.478	0.032*
	ET + PFB	0.547	2.347	0.816
	ET + GS^+	1.605	2.499	0.523
	Control	13.891	2.435	0.000**
Control	ET	-8.488	2.520	0.001**
	ET + PFB	-13.345	2.440	0.000**
	ET + GS^+	-12.286	2.591	0.000**
	ET + PFB + GS^+	-13891	2.435	0.000**

* $p < 0.05$, ** $p < 0.01$

Table 4

Evaluating the Strength Ratio

Intervention	Strength Ratio	Intervention - ET
ET	0.0113	--
ET + PFB	0.0153	0.0040
ET + GS^+	0.0142	0.0029
ET + PFB + GS^+	0.0159	

VITA

Kimberly A. Hastings

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE ISOLATED AND COMBINED EFFECTS OF INTERVENTION COMPONENTS ON MATHEMATIC FLUENCY

Major Field: Educational Psychology with an option in School Psychology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Educational Psychology with an option in School Psychology at Oklahoma State University, Stillwater, Oklahoma in July, 2010.

Completed the requirements for the Master of Science in Educational Psychology with an option in Applied Psychometrics at Oklahoma State University, Stillwater, OK in 2006.

Completed the requirements for the Bachelor of Science in Psychology at Abilene Christian University, Abilene, Texas in 2005.

Experience:

Pre-doctoral Intern for Cypress Fairbanks Independent School District
Clinic Practicum Student
School Psychology Practicum Student for Stillwater Public Schools

Professional Memberships:
American Psychological Association
National Association of School Psychologists

Name: Kimberly A. Hastings Date of Degree: July, 2010

Institution: Oklahoma State University Location: Stillwater, Oklahoma

Title of Study: THE ISOLATED AND COMBINED EFFECTS OF INTERVENTION COMPONENTS ON MATHEMATIC FLUENCY

Pages in Study: 81 Candidate for the Degree of Doctor of Philosophy

Major Field: Educational Psychology with an option in School Psychology

Scope and Method of Study: The scope of the study was to compare empirically supported math fluency interventions to a control group as well as each other. The intervention components were isolated to investigate their initial strength as well as adding them together to investigate if the combined intervention was actually stronger or more efficient than the isolated components added together. Design was a pre-test/post test control group design. Analysis chosen included an ANCOVA and the development and evaluation of a strength ratio.

Findings and Conclusions: Results indicated that all treatment interventions were considered significant when compared against the control group. ANCOVA pairwise comparisons also indicated that the ET (explicit timing) condition and the ET + PFB + GS^{+} were significantly different from each other. The evaluation of the strength ratio revealed that the ET + PFB + GS^{+} offered the largest strength ratio. When the isolated components were broken into their individual strength ratios and summed, the sum of the strength ratios was actually larger than the combination intervention's strength ratio indicating that the effects were considered an example of diminishing returns. Therefore, the ET + PFB offered the most cost efficient or "strong" intervention.

ADVISER'S APPROVAL: Dr. Gary J. Duhon