

A COMPARATIVE ANALYSIS OF MATH FACTS
FLUENCY GAINS MADE THROUGH MASSED AND
DISTRIBUTED PRACTICE WITH VARIED INTER-
SESSION INTERVALS

By

STEVEN L. POWELL

Bachelor of Arts
Hood College
Frederick, Maryland
2014

Master of Arts
James Madison University
Harrisonburg, Virginia
2015

Educational Specialist
James Madison University
Harrisonburg, Virginia
2017

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
May, 2021

A COMPARATIVE ANALYSIS OF MATH FACTS
FLUENCY GAINS MADE THROUGH MASSED AND
DISTRIBUTED PRACTICE WITH VARIED INTER-
SESSION INTERVALS

Dissertation Approved:

Dr. Gary Duhon

Dissertation Adviser

Dr. Brian Poncy

Dr. Terry Stinnett

Dr. Mwarumba Mwavita

ACKNOWLEDGEMENTS

I would like to thank my adviser, Dr. Duhon, for his guidance and availability throughout the process of conducting and writing this dissertation. I'd like to thank Dr. Poncy for the inspiration for this research project and invaluable input on the procedural specifics of the project. I'd further like to thank Dr. Stinnett for serving on my committee and Dr. Mwavita for patiently teaching me the intricacies of HLM.

I'd like to thank Dr. Ashton Trice for taking me under his wing during my time in the specialist program at James Madison University.

I would like to express my unending gratitude to my peers for the many hours spent pursuing this research. Specifically, I'd like to thank Ryan Puccetti and Alex Englén for their unwavering support and availability to help at any minute. Without their hard work, completing this project would not have been possible.

I'd like to acknowledge my best friend, Grant Ostrander, for our frequent Friday-night discussions over research and the application of behavioral science. Those conversations helped mold me as a scientist and make me into the confident and theoretically- and ethically-grounded school psychologist I am today.

My family, who have patiently tolerated my apparent desire to live in as many different regions of the country as possible, deserve recognition as well.

Other people who have significantly influenced my career trajectory include Dr. Michael Hall, who guided me through my first teaching experiences as a graduate student, and Renee Sexton, who served as my supervisor during my specialist internship.

Name: STEVEN L. POWELL

Date of Degree: MAY, 2021

Title of Study: A COMPARATIVE ANALYSIS OF MATH FACTS FLUENCY GAINS
MADE THROUGH MASSED AND DISTRIBUTED PRACTICE WITH VARIED
INTER-SESSION INTERVALS

Major Field: SCHOOL PSYCHOLOGY

Abstract: We examined the math facts fluency growth of students receiving explicit timing practice in either a massed, short-distributed, or long-distributed fashion. In a repeated-measures group design, 50 mid-western 3rd grade students completed four one-minute math probes each day for 20 days. Hierarchical linear modeling revealed that students who received practice separated by a three-hour interval grew significantly more than students who practiced with a ten-minute distribution or a massed presentation.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION.....	1
Improving Total Learning.....	1
More Efficient Intervention	2
Altering Response Topography	4
Distributed Practice.....	4
Distributed Practice and Math Skills	6
Math Facts Fluency.....	6
Current Study	8
II. REVIEW OF THE LITERATURE.....	9
The Instructional Hierarchy	10
Cover, Copy, Compare	11
Taped Problems	12
Flash Card Drill.....	13
Explicit Timing	15
Proof of Concept.....	15
Explicit Timing as an Intervention	19
Instructional Efficiency.....	21
Opportunities to Respond	21
Timing.....	24
Set Size.....	26
Reinforcement.....	27
Saliency of Prompts and Cues	27
Feedback	29
Self-Monitoring and Goal Setting.....	30
Distributed Practice.....	32
Contemporary Study of Massed and Distributed Practice	34
Theories of Distributed Practice	36
Applied Studies of the Distributed Practice Effect	40
Distributed Practice and Vocabulary Learning.....	41
Distributed Practice and Reading.....	41
Distributed Practice and Science Knowledge	42
Distributed Practice and Math Skills	43
Current Study	45

Chapter	Page
III. METHOD	46
Participants.....	46
Materials	47
Experimental Design and Analysis.....	47
Independent Variable	48
Dependent Variable	48
Procedure	48
IV. FINDINGS.....	50
Procedural Fidelity.....	50
Baseline Analysis.....	52
Hierarchical Linear Model.....	52
V. CONCLUSIONS.....	56
Discussion	56
Limitations	58
Future Directions	59
REFERENCES	61
APPENDICES	76
APPENDIX A.....	77
APPENDIX B	78
APPENDIX C	79
APPENDIX D.....	80
APPENDIX E	81
APPENDIX F.....	82

LIST OF TABLES

Table	Page
1. Tasks for each group during daily intervention sessions.....	49
2. Fidelity results by session.....	51
3. Fidelity results by classroom.....	51
4. Means and SDs of groups in DCPM.....	52

LIST OF FIGURES

Figure	Page
1. Mean DCPM growth between the three conditions	54
2. Mean DCPM growth in the three-hour distributed condition and collapsed fifteen minute and massed conditions	55

CHAPTER I

INTRODUCTION

Students who fall behind their peers with regard to academic skillsets are frequently given targeted intervention in order to catch them up to their peers. Numerous empirically validated interventions exist which are designed to improve a student's accuracy or fluency, and many can be applied in more than one academic context. Cover, Copy, Compare (Skinner, Turco, Beatty, & Rasavage, 1989), Taped Problems (McCallum, Skinner, & Hutchins, 2004), Traditional Flash-Card Drill, and Explicit Timing (Van Houten & Thompson, 1976) are examples of such interventions. The purpose of these interventions is to increase a student's level of ability with specific academic skillsets in an efficient manner, since students who lack requisite skills to access instruction fall further and further behind their peers each day they are not caught up. As such, finding ways to improve total learning in an efficient manner is imperative to 'catching up' students who have fallen behind.

Improving Total Learning

Generally speaking, learning problems are not due to an inability to learn, but rather a problem of learning rate. Assuming individuals have a set learning rate, total learning can only be increased through more instruction, which can be achieved by either adding instructional time or making instruction more efficient (Skinner, Fletcher, & Henington, 1996). One way to increase instructional time is to lengthen the school day; however, lengthening a school day is not likely an option due to logistics and resources (e.g. re-routing school busses, parent pick-up times, etc.). If the school day cannot be extended, time must be taken from other subjects. It is not uncommon for a child to be held during recess or specials classes in order to participate in remedial activity. While this is an effective technique, it may not be a particularly palatable one for teachers and parents, who may prefer remediation efforts that keep recess and specials time intact.

If additional instructional time is a last resort, interventionists must increase the efficiency of their work. Skinner, Belfiore, & Watson (2002) operationally defined instructional efficiency as a ratio of the relative learning rate to instructional time used to administer the intervention.

More Efficient Intervention

One way to increase efficiency is to reduce inter-trial intervals (ITI), which is the amount of time between the end of one learning trial and the beginning of another. Skinner, Fletcher, and Henington (1996) speculated that fast-paced instruction may increase learning rates due to the increased quantity of potential learning trials. Sometimes, the authors posited, quantity is more important than quality. Skinner et al (2002) conducted a single-case study in which a student learned two sets of sightwords. The authors manipulated the ITI such that the student either moved to the next trial immediately or had a five-second interval time between trials. The authors showed that the student learned a similar number of words over the course of a similar number of days in both conditions. However, when the authors instead graphed words learned per instructional second, they

found large differences in favor of the immediate-ITI condition. While the authors point out that this is an example of the need for more precise measures of instructional time when graphing, their study also demonstrated that students can learn at higher rates when ITI is kept low.

Another way to improve the relative outcomes of interventions is to alter wait time. Wait time is the time between the delivery of a stimulus or cue and feedback from the instructor or interventionist. Wait time differs fundamentally from ITI in that wait times are imbedded within learning trials, while ITIs are between learning trials. In order to maximize efficiency, wait times must be balanced. If too much time elapses between stimulus and feedback delivery, time is wasted. If too little time elapses between the two, the student may not have enough time to produce a response (Riley, 1986; Row, 1974; Tobin, 1983). In a study that perhaps refutes this point, Poncy, Jaspers, Hansmann, Bui, and Matthew (2015) utilized a Taped Problems intervention and controlled for total number of trials but altered the wait time between the presentation of the intervention cue and the answer. While a two-second delay and no-delay caused relatively equal learning increases, the no-delay condition took 33% less time than the two-second delay, which indicates that a very short wait time between stimulus presentation and response can in some instances produce similar effects to longer wait times and can have the added benefit of taking less time, which allows for more of the intervention to be delivered in the same amount of instructional time.

Similar to altering the wait time, it has been demonstrated that the use of time limits tends to increase the rate at which children complete work. Derr and Shaprio (1989) found that children read significantly more words per minute when they knew they were timed than when they did not. Houten, Hill, and Parsons (1975) had two fourth grade classrooms complete a writing task. One room was told they had ten minutes to write as much as they can, while the other room had twenty minutes to write as much as they can, but were not told they would be timed. The classroom in the timed condition wrote, on average, twice as much as the untimed room.

A final method of increasing the efficiency of instruction is to utilize choral responding. Choral responding increases the number of learning trials in which each student may participate, which essentially introduces more intervention.

Altering Response Topography

Several studies have examined changing the topography of responses required of students. In this case, topography refers to the method in which a child responds to a presented stimulus in a learning trial. Gardner, Heward, and Grossi (1994) had children respond to instruction using reusable signs, called response cards, to indicate their response. Compared to traditional instruction, children who used response cards responded more frequently to instructor questions and performed more highly on quizzes. In traditional instruction, a question asked by the teacher is generally answered only by one student at a time. Using response cards, every student has the opportunity to respond to every question asked by the teacher. Gardner et al (1994) reported that students also preferred the use of the response cards to traditional instruction.

Two studies that examined Cover, Copy, Compare while requiring students to respond either in writing or verbally found similar learning rates for both topographies when the number of learning trials was held constant. However, when holding overall instructional time constant, verbal responding produced higher learning increases because verbal took less time (Skinner, Ford, & Yunker, 1991). This effect was replicated in two children with behavioral disorders by Skinner, Belfiore, Mace, Wililams-Wilson, and Johns (1997).

Distributed Practice

Distributed practice is a well-known and well-studied phenomenon first described by Ebbinghaus (1885/1913), who taught himself nonsense words using various inter-trial intervals. Distributed practice is the phenomenon in which longer inter-trial intervals tend to produce higher levels of overall learning compared to shorter inter-session intervals. Inter-trial intervals differ from

inter-session intervals in that multiple trials are generally nested within a single session. This effect was replicated numerous times in the late 1800's and early 1900's (Cook, 1934; Cook, 1944; Jost, 1897; Reed, 1924; Ruch, 1928; Thorndike, 1912). More recent studies have demonstrated that spaced, or distributed trials can produce up to twice the learning as the same number of trials presented at once (Bahrck & Phelps, 1987; Dempster, 1987; Underwood, 1970).

Many recent studies have examined the distributed practice effect in a laboratory setting or settings that utilize undergraduate students as participants (Cepeda, Pashler, Vul, Wixted, & Rohrer, 2006; Donovan & Radosevich 1999; Lee & Genovese (1988); Moss 1996; Pavlik & Anderson, 2003). Dempster (1987) found that undergraduate students more accurately recalled unfamiliar vocabulary words when they were learned using distributed practice than when learned using massed practice. Sobel, Cepeda, and Kapler (2011) taught 39 fifth grade students two roughly equivalent lists of Graduate Record Examination (GRE) words. Students practiced one list using distributed practice, and learned the other list using massed practice. The authors found that at follow up, their participants recalled approximately 20% and 7.5% of the distributed practice words and massed practice words, respectively.

In a somewhat unorthodox nine-year longitudinal study, Bahrck, Bahrck, Bahrck, & Bahrck (1993) utilized themselves as participants. The authors studied 300 pairs of English and foreign language words in six training conditions which manipulated the interval between learning sessions as well as the number of total sessions. Unsurprisingly, they found that recall increased as a function of the number of total learning sessions. They also found, however, that recall increased as a function of the interval between learning sessions. The authors recalled 20% more words in their longest inter-session interval (56 days) than in their shorted inter-session interval (14 days).

Distributed Practice and Math Skills

Studies which examine the utility of distributed practice in a laboratory setting can lack external and practical validity, especially when used to inform the implementation of distributed practice in a classroom setting. To date, several studies have examined the effects of distributed practice on the accuracy and fluency, or speed, with which students complete basic math fact problems (e.g. Rea & Modigliani, 1985; Rohrer & Taylor, 2006; Rohrer & Taylor, 2007; Schutte et al, 2015). Rohrer and Taylor (2006) taught 216 college students to complete a specific kind of math problem and had them practice using one of two practice schedules. All students practiced 10 problems. Roughly half of the participants split the practice into two sessions, which were one week apart. The other group of students practiced all 10 problems at once. Although both groups had comparable performance one week after their practice sessions were complete, the authors' distributed practice group was twice as accurate as the massed practice group (64% accuracy and 32% accuracy, respectively) after at a four-week follow up. In a similar experiment, Rohrer and Taylor (2007) had three groups of undergraduates practice math problems in three sessions, each one week apart. In this study, participants assigned to the distributed practice condition were called "spacers," those in the massed practice condition were called "massers." The authors included a third condition in which participants received half of the practice as the spacers and massers. These participants were referred to as "light massers." Rohrer and Taylor (2007) found that the "spacers" were 74% accurate at follow up, compared to 49% and 46% for "massers" and "light massers," respectively.

Math Facts Fluency

An intervention that is frequently utilized to increase math facts fluency in children is called Explicit Timing (ET). ET is a simple intervention that is easy to administer to large groups of students simultaneously. In most ET administrations, children are presented with a one-page sheet with a large number of math problems. These sheets, called 'probes,' usually have similar problems with respect

to their operation and difficulty, but variable operation and difficulty probes. During ET administrations, children are instructed to complete as many of the math problems on their probe as possible within some span of time, usually one minute. After the time has expired, the ET administrator instructs the children to stop completing problems and the intervention session is complete (van Houten & Thompson, 1976). The interventionist is then able to calculate the rate of accurate problem solving, as measured by digits correct per minute (DCPM). Probes can be scored after each intervention session, which provides the additional benefit of a built-in progress monitoring tool. The efficacy of ET has been demonstrated numerous times regarding its impact on math facts fluency (Duhon, House, & Stinnett, 2012; Poncy, Duhon, Lee, & Key, 2010; Rhymer, Skinner, Henington, D'Reaux, & Sims, 1998; Schutte et al 2015).

In the seminal study of distributed practice in ET interventions, Schutte et al (2015) split 53 third-grade students from three classrooms into three groups in a longitudinal stratified design. Students either completed their four one-minute timings back-to-back in one session, two timings in the morning and two in the afternoon (approximately four hours later), or one timing in each of four sessions throughout the day, which were approximately two hours apart. Since times in which some students were working on math problems and some were not, the authors had those who were not engaged in math problems complete reading worksheets. The authors found that students demonstrated more math facts fluency growth when intervention sessions were distributed, and the most growth when sessions were split into four one-minute timings per day. This study demonstrated the practical utility of distributed practice in the elementary classroom as well as provided further evidence of the efficacy of explicit timing interventions.

Current Study

Schutte et al (2015), allowed several hours to pass between sessions for students who received practice that was distributed. While this interval resulted in enhanced growth over the

massed practice, elementary school math periods are generally not this long nor occur more than once per day. Distributing practice could be more practical for implementation in schools if its effects were found over shorter inter-session intervals.

The purpose of the current study is to help determine a minimum effective temporal distribution of explicit timing intervention sessions in elementary students. We have two hypotheses: 1) Consistent with Schutte et al (2015), children who participate in distributed practice will show more growth in DCPM during intervention than children who participate in massed practice, and 2) Children who participate in distributed practice with short inter-session intervals will show similar growth to children who participate in distributed practice with long inter-session intervals.

CHAPTER II

REVIEW OF THE LITERATURE

Students who fall behind with regard to academic skillsets are frequently given targeted intervention in order to catch them up to their peers. Many of these targeted interventions are designed to improve a student's accuracy or fluency and can be applied in more than one academic context. Examples of such interventions are Cover, Copy, Compare (Skinner, Turco, Beatty, & Rasavage, 1989), Taped Problems (McCallum, Skinner, & Hutchins, 2004), Traditional Flash-Card Drill, and Explicit Timing (Van Houten & Thompson, 1976). Interventions are designed to increase a student's level of ability with specific academic skillsets in an efficient manner, since students who lack requisite skills to access instruction fall further and further behind their peers each day they are not caught up. As such, finding ways to improve total learning in an efficient manner is imperative to catching up students who have fallen behind.

Empirical Support for Common Academic Interventions

Barring severe developmental or intellectual disability, learning problems are generally not due to an inability to learn, but rather a problem with the rate at which a student learns. It is generally not the case that students who are referred for intervention or special education services have not learned at all, but rather the rate of their learning has not been determined to be adequate based on some set of normative criteria. The following section will discuss several commonly-used academic interventions. While many of these can be used to help students demonstrate growth in more than one academic subject, this discussion will direct its focus on the utility of these interventions in helping students improve their mathematics skills.

The Instructional Hierarchy

Academic interventions are not all created equally; different interventions are useful for different situations. When selecting appropriate intervention for a student who needs additional support, the interventionist must first consider the student's current level of skill. The instructional hierarchy (Haring, Lovitt, Eaton, & Hansen, 1978) is a framework which allows educators to conceptualize student level of skills into one of four broad, progressive categories.

Within the framework of the instructional hierarchy, progression to the next stage is usually dependent upon completion of the previous stage. To achieve the first stage of the instructional hierarchy, called acquisition, teachers must first promote accurate responding to prompts and cues through techniques like error correction and modeling. Once a student is able to accurately respond to prompts from his or her teacher, the focus shifts from simple accurate responding to rapid accurate responding, known in the instructional hierarchy as the fluency stage. Traditionally, fluent responding is built through instructional techniques like drill and practice. Ardoin and Daly (2007) define drill as the "repetition of responses to items being learned in order to specifically increase fluency" and practice as "the utilization of a learned

response in combination with other previously learned responses” (p. 4). In essence, the two differ in that drill presents known items repeatedly while practice utilizes known as well as unknown items which rely on the skill being taught.

After the student is able to respond accurately and fluently to academic stimuli, instruction should be catered toward the generalization of skills from one prompt to a similar – but not explicitly taught – prompt. For example, a student who has learned to successfully complete single digit addition problems may be able to complete some multiple digit addition problems which do not require carrying. The final stage of the instructional hierarchy is adaptation, in which the learner uses taught skills and generalizes them to completely novel situations.

Empirical support for an intervention does not necessarily mean that it will be appropriate in all situations with all students. Utilizing the framework set forth in the instructional hierarchy to choose interventions can help teachers, interventionists, and researchers be sure that they are choosing an appropriate intervention.

Cover, Copy, Compare

Cover, Copy Compare (CCC; Skinner, Turco, Beatty, & Rasavage, 1989) is an intervention that was originally developed to help students who struggled with multiplication. CCC as an intervention is targeted toward the first stage of the instructional hierarchy, which is acquisition.

The original procedure, which is still commonly used, consists of sets of five steps. The first two steps make up the “cover” portion of the intervention, and requires the student to look at a correctly completed multiplication problem and then cover it with a piece of paper. The next step is the “copy” portion of CCC, and requires the student to copy the problem and its solution while keeping the correctly completed problem covered. The last two steps make up the

“compare” portion of CCC, in which the student uncovers the correctly completed problem and compares their copied version to the original. In the original study, the authors utilized a multiple baseline design with four students recruited from a school for behaviorally-disordered children. Two of the students were in fourth grade, and the other two were in tenth grade. All four students improved their accuracy and digits correct per minute (DCPM) in multiplication problems. Skinner et al (1989) made the point that CCC is not only effective at increasing accurate rates of responding to multiplication problems, but is also cost effective in that it does not require expensive equipment.

A teacher of one of the students who participated in the original study reported that she observed the student using CCC to study for a science test, which implied the utility of CCC in subject areas other than mathematics. This implication has been supported in the literature for sight word acquisition (Kaufman, McLaughlin, Derby, & Waco, 2011), science (Smith, Dittmer, & Skinner, 2002), spelling (Darrow, McLaughlin, Derby, & Johnson, 2012; Hochstetler, McLaughlin, Derby, & Kinney, 2013), foreign language acquisition (Carter, Wong, & Michael, 2013), and geography (Skinner & Belfiore, 1992).

Taped Problems

Taped problems is a self-managing intervention originally designed to help students perform division facts fluently. McCallum, Skinner, and Hutchins (2004) based their taped problems intervention on the taped words intervention (Freeman & McLaughlin, 1984), which was designed to aid in fluent sight-word oral reading. Freeman and McLaughlin (1984) had six high school students identified by their district as students with a specific learning disability (SLD) listen to an audio recording of a list of sight-words and attempt to read along with the audio recording. After reading along with the recording, the students would attempt to read the list to their classroom teacher. The authors found that the taped words intervention caused a

significant increase in level and trend for correct responding and a significant decrease in errors for the practiced lists.

McCallum, Skinner, and Hutchins (2004) added a progressive time-delay component to Freeman and McLaughlin's (1984) taped words intervention. The literature supports the use of time-delay procedures in fluency development (for a review, see Aust, Wolery, Doyle, & Gast, 1989). The authors of taped problems (McCallum, Skinner, & Hutchins, 2004) examined its utility for the development of division facts fluency in a single ten-year-old student. The procedure for taped problems is simple; the student was presented with a division problem and tasked with answering the problem correctly before the recording listed the correct answer. If the student was unable to answer the problem or failed to provide a correct response, he was required to repeat the correct response before moving on to the next problem. The authors reported that they reinforced correct responding with verbal praise.

Using a multiple-probe-across-tasks design, the authors had their student learn three equivalent lists of division problems. The student completed each list four times, with the time delay between the presentation of the problem and the recording providing the correct answer reducing after each completed list. The authors found that the student maintained the high accuracy he had prior to intervention, but made rapid gains in fluency after implementing the intervention. The student moved from, on average, 12 DCPM on each list to 31 DCPM on each list over the course of 15 sessions, where one session occurred each school day.

Flash Card Drill

Flash card drill has been used with undeniable success for helping children learn their letters, sight words, and hone fluency of their mathematics skills. While flash card drill procedures can vary depending upon the goals of the intervention, the basic procedure is as follows: the teacher presents a flash card on which the target stimulus is written and prompts the

student to repeat what is on the flash card. The teacher then either provides reinforcement for correct responding or corrective feedback. This procedure is generally repeated several times until a “test” trial, in which the student must read what is on the cards without feedback to assess mastery of the material (Nist & Joseph, 2008). The flash card procedure has also been altered to be useful for practicing multiplication facts fluency (Burns, 2005).

A useful technique for enhancing the efficacy of flash card drill procedures is known as interspersal. Interspersal procedures involve the inclusion of known items alongside unknown items. The ratios at which students receive known to unknown items varies based upon the goals of the intervention, but a common ratio is three knowns to one unknown (Nist & Joseph, 2008). Incremental rehearsal is a type of interspersal technique in which unknown items are presented repeatedly between presentations of known words. Unknown items usually comprise 10% of the flash card set (MacQuarrie-Klender, Tucker, Burns, & Hartman, 2002). Incremental rehearsal has been supported by several empirical inquiries (e.g. Nist & Joseph, 2008; Skinner, Hurst, Teeple, & Meadows, 2002).

Nist and Joseph (2008) examined six first grade students who were referred by their teacher as having difficulties with reading sight words. They taught the students three sets of 36 words using traditional flash-card drill, an interspersal procedure, and an incremental rehearsal procedure. Once students exhibited mastery of a word, evidenced by a correct response within three seconds, the words were removed from the list. In the interspersal condition, the authors presented three unknown words, followed by one known word. In the incremental rehearsal procedure, the authors used a “folding-in” procedure, in which the student was presented one unknown word, followed by one known, then the unknown again, and then two known words, and so on until the student reached one unknown to nine known words. After that, the authors replaced the first known word with the previously unknown word and shuffled in the next unknown word to be learned. The authors found interspersal to be the most effective at teaching

the students new words, then incremental rehearsal, and then traditional flash-card drill. The authors point out that while interspersal seems to be the most efficient way to teach new words, there are situations in which incremental rehearsal is useful, such as when a student doesn't respond to traditional flash card drill or interspersal or for students who struggle to keep skills over time.

Explicit Timing

Explicit timing as stand-alone intervention came about from several studies conducted by Van Houten and various colleagues in the 1970s. Early examinations into its effectiveness were limited to written composition and included other variables, such as performance feedback, reinforcement, and public posting simultaneously.

Explicit timing proof of concept. Van Houten, Morrison, Jarvis, and McDonald (1974) examined the effects of explicit timing, immediate feedback, and the posting of high scores on the written composition rate of one second-grade and two fifth-grade classrooms. In the second grade classroom, the authors utilized an ABAB withdrawal design. During baseline, the students were told to write as much as they could, but were not informed of the ten-minute time limit, given feedback on their writing, nor were their scores posted publicly. During the treatment phases, students were told that they had ten minutes to write as much as they could about the chosen topic for that day. After each writing session, the children counted up their total words written as a form of immediate feedback. The authors also counted up the students' total words written and put them on a board in the classroom. Each student was told to "try to beat their score" before each subsequent treatment session. The authors found an increase in the level and trend of the average words written per minute during treatment and a return to baseline during the withdrawal phase.

In the fifth grade classrooms, the authors utilized a multiple-baseline across classrooms design and a similar procedure to that of the second grade classroom. The authors found a clear increase in level and trend for the total written words per minute for both classrooms.

In a follow-up study, Van Houten, Hill, and Parsons (1975) more closely examined the effects of timing on the writing rate and reading comprehension of elementary students in two experiments. In the first experiment, the authors used a similar procedure to that of Van Houten et al (1974), where children were told to write as much as they could on a pre-selected topic. The authors examined two classrooms of children. One classroom wrote for ten minutes and the other wrote for 20 minutes. The authors explained this was due to teacher perceptions of the ability of the first classroom to write for longer than ten minutes. Both classrooms received the same reversal design. After the baseline phase, the children were told they would be timed and were given feedback in the form of counting up their own total words written. The authors then alternated phases in which they posted scores publicly. They found increases in total words written during all treatment phases compared to baseline and slight increases during phases which included public posting of scores. In the second experiment, the authors examined the reading comprehension of 19 fifth grade children. During each session, the students worked independently to read a story and answer ten questions about what they had read. After the baseline phase, the authors introduced feedback, explicit timing, and public posting of scores. The authors found these variables combined to increase correct responding on comprehension measures as well as word meaning measures during the treatment phase and a return to baseline during the withdrawal phase.

While the Van Houten et al (1974) and Van Houten, Hill, and Parsons (1975) studies implied that explicitly timing students would increase their academic performance, studies at that time had not yet examined the effects of explicit timing alone. Van Houten and Thompson (1976) sought to demonstrate the effectiveness of explicit timing on math performance. The authors

implemented a reversal design and utilized a classroom of 20 second grade students. Students in this study completed worksheets of simple addition and subtraction to and from ten, respectfully. The worksheets were arranged into packets, whose problems alternated between addition and subtraction. During baseline, the authors gave the students 30 minutes to complete as many problems as they could, but did not inform the students that they were being timed. During the treatment phases, the authors had students complete similar worksheets to those completed during baseline. The authors set a timer for 30 minutes, but also broke the session into one-minute sessions. Using a separate timer, the classroom teacher told the students to work as quickly as they can for one minute and to draw a line after their last completed problem. The teacher repeated this one-minute timing procedure until the 30 minute timer ran out. Once the 30 minute timer ran out, all students stopped working, even if they were currently working a one-minute timing.

Overall, the authors found that during the one-minute timing condition, students completed two to three times more mathematics problems than during baseline. These results were replicated in the second baseline and treatment phases. During all phases, the students were highly accurate; all sessions had a mean accuracy of 90% or greater.

Van Houten and Little (1982) examined the impact of timing on the response rate of three adolescent students who received special education services in a self-contained classroom for part of the day. All three students had slow rates of responding to math problems, which ranged from 0.5 to 2 correct problems per minute in baseline. Using an across-subjects multiple baseline with reversal design, students alternated between the baseline phase, where they had 20 minutes to work mathematics problems on a worksheet, and the intervention phase, where they were told they were being timed and to work as fast as they could for five minutes. The authors found increases in the rates of correct responding for all three students during the explicit timing phases as well as a return to baseline during non-treatment phases. These results are commensurate with

the results from Van Houten and Thompson (1976), and show that explicitly timing students can increase their rates of responding.

This finding was demonstrated again in Rhymer and Morgan (2002). Rhymer and Morgan (2005) had 45 third grade students complete sheets of subtraction problems. Using a within subjects design, the authors had each student complete problems using an explicit timing procedure and an interspersal procedure. In the interspersal condition, the probes had a 1-1 digit subtraction problem every third item, while all other items on the page were 2-2 digit subtraction and required borrowing in the ones column. The explicit timing and control sheets had only 2-2 digit subtraction with borrowing in the ones column. In both the explicit timing and the interspersal condition, the authors told participants to work as quickly as they can for the three-minute trial. However, in the explicit timing condition, participants worked in timed one-minute segments. While the authors found that in this instance, the interspersal procedure produced the highest rates of responding, both interspersal and explicit timing produced higher rates of responding than the control condition. One limitation the authors noted was that every third problem in the interspersal condition was easier than the others as it did not require borrowing. Similar results were found in Clark and Rhymer (2003), who used the same procedures but had college students for participants.

Rhymer et al (2002) examined the effects of explicit timing against a timed (but not explicitly timed) control condition using a within subjects design of 54 sixth grade students. In addition, the authors sought to examine the impact of item difficulty in the aforementioned conditions. The authors compared a one-step addition task, a three-step subtraction task, and a difficult multiplication task in their explicit and covertly timed conditions. The study was conducted over the course of six days, with the first three days serving as the covert condition and the last three days serving as the explicit condition. The authors found similarly accurate responding from students regardless of the timing condition. Unsurprisingly, students performed

significantly worse on the difficult multiplication problems than the one step addition and three step subtraction problems. The authors also found that students completed significantly more problems in the timing condition for both addition and subtraction, but not for multiplication, which implies that accurate responding is necessary for the explicit timing effect to occur. The importance of using explicit timing with children who are already able to produce accurate responses is further explained in Rhymer, Skinner, Henington, D'Reaux, and Sims (1998).

Explicit timing as an intervention. As a result of the work by Van Houten and Thompson (1976), explicit timing as a standalone intervention for improving math facts fluency became relatively common. Since that time, several more researchers have examined its utility with somewhat mixed results. While explicit timing procedures do tend to increase the rate of responding, they may not necessarily increase the rate of *correct* responding. Rhymer et al. (1998) attempted to partially replicate the findings of Van Houten and Thompson (1976) by using a multiple baseline across group design. The authors split 44 third grade students into three groups and conducted three explicit timing sessions each day for three days and measured the students' rate of completed problems as well as their rate of correctly completed problems. While the treatment phase showed clear increases in the rate of completed mathematics problems, it did not show any increases in the rate of correctly completed problems. In fact, the students' accuracy actually decreased by approximately 10% for each of the three groups during the explicit timing phase. The authors used their results as evidence that while explicit timing is useful in increasing responding, it may not by itself be enough to increase correct responding. The authors argued that some sort of correct feedback procedure is necessary in order to avoid the reinforcement of potentially incorrect but rapid responding over correct responding. Further, the authors used their results as a warning to practitioners to ensure that their students are accurate before beginning an explicit timing intervention. These findings are further supported by the findings of Rhymer et al (2002).

While the implications of Rhymer et al (1998) are practically sound, their study was limited in part by its short duration and lack of accurate responding from its participants. Van Houten and Thompson (1976) conducted daily sessions over the course of 26 school days and examined students who were already highly accurate, whereas Rhymer et al (1998) conducted nine sessions over three school days and had students who were able to complete the mathematics problems with approximately 75% accuracy.

Poncy, Duhon, Lee, and Key (2010) sought to examine the extent to which addition practice with various intervention methods lead to the generalization of correct responding to subtraction problems. The authors utilized a multiple baseline design with three fourth-grade students and had their students complete phases of explicit timing, conceptual lessons, and cloze procedures. Conceptual lessons in this context refers to direct instruction of part-part whole relationships (e.g. $2 + 3 = 5$ and $5 - 3 = 2$, etc.), while cloze procedures were horizontally presented math problems which did not include one of the two addends (e.g. $_ + 5 = 11$, $5 + _ = 11$). While the authors did not find that any of the three examined interventions promoted generalization to subtraction problems, each intervention phase promoted more fluent responding in all three participants.

In another study which examined the effectiveness of explicit timing on math facts fluency, Duhon, House, and Stinnett (2012) assigned 32 children to two groups. While each group completed a two-minute explicit timing procedure once per day, one group completed theirs using traditional paper-and-pencil while the other completed theirs using a computer. After 20 consecutive school days, the children completed a post-test using both the computer and paper-and-pencil. The authors found moderate to large gains in fluency over the 20-day course of the study for both paper-and-pencil and computer, but did not find that students who participated in one modality generalized their responding to the other modality.

Instructional Efficiency

As much of the literature discussed so far implies, there are numerous ways to increase the efficiency of academic instruction and intervention. Making instruction and intervention as efficient as possible is critical to catching up students who have fallen behind. Generally speaking, learning problems are not due to an inability to learn, but rather a problem of learning rate. If students are not learning at the desired rate, it can be increased through more instruction, which can be only achieved by either adding instructional time or making instruction more efficient (Duhon, House, Hastings, Poncy, & Solomon, 2015; Skinner, Belfiore, & Watson, 2002; Skinner, Fletcher, & Henington, 1996; Skinner, McCleary, Skolitis, Poncy, & Cates, 2013). One way to increase instructional time is to lengthen the school day; however, lengthening a school day is not likely an option due to logistics and resources (e.g. re-routing school busses, parent pick-up times, etc.). If the school day cannot be extended, time must be taken from other subjects. It is not uncommon for a child to be held during recess or specials classes in order to participate in remedial activity. While this is an effective technique, it may not be a particularly palatable one for teachers and parents, who may prefer remediation efforts that keep recess and specials time intact.

If additional instructional time is not an option, interventionists must increase the efficiency of their interventions. Intervention efficiency is referred to as a ratio of the relative rate of learning and the time it takes to administer the intervention (Skinner et al, 2002). The following section will discuss the empirical support for various methods of increasing instructional efficiency.

Opportunities to Respond

Much of traditional instruction is lecture-based, where the teacher stands at the front of the classroom and instructs the entire class at once. While this method is likely the most practical

way to instruct large numbers of students simultaneously, more individualized methods, like one-on-one tutoring, tend to be more effective (Bloom, 1984; Brophy, 1988). One theory as to why this is the case has to do with the number of opportunities that students have to respond to academic content. In whole-class instruction, the teacher often lectures, then calls on a single student or poses a question to the class to which only one student may respond at a time. As a result of this system, students who are high achieving are more likely to respond than those that are low achieving, which can then further increase the skill disparity between high and low achievers (Brophy & Evertson, 1976; Maheady, Mallette, Harper, & Sacca, 1991).

Several alternative instructional methods have been posed and studied in the last 30 years which attempt to remedy this problem. One such instructional method was studied by Maheady et al (1991). Heads Together (HT) was created by Kagan (1989), and is described by Maheady et al (1991) as a combination of teacher-directed and peer-mediated instruction. This combination, the authors argue, allows for more opportunities for student responding and as a result will increase total learning. In HT, students are placed into groups of four which are matched by achievement. Groups contain one high, one low, and two average achieving students. Students in each group are then randomly assigned a number from one to four, irrespective of their achievement status, and the teacher begins instruction. During whole-class instruction, the teacher will ask a question and give each group time to discuss the possible answer (i.e. to 'put their heads together'). Then, instead of calling on a single student, she calls a number and asks how many students with that number know the correct answer, and then reinforces the students and their groups who had the correct response. Since students do not know which number will be called, they are likely to make sure all students in their group can produce the correct response. The authors utilized an alternating treatment design which alternated between traditional (whole-class) instruction and HT with a classroom of 20 third-grade students during their social studies period. The authors

found that in all sessions, students performed more highly on quizzes when instructed using HT than when instructed using whole-class instruction.

Another way to increase student opportunities to respond is to alter the response topography to include more students per instructional prompt. One easy, cheap, and low-tech way to do this is through choral responding. In choral responding, a teacher poses a question to the class, who then all respond aloud simultaneously. This method allows for large percentages of the class to respond at once instead of one at a time and is recommended in numerous publications (e.g. Heward, 1994; Godfrey, Grisham-Brown, Schuster, & Hemmeter, 2003; Skinner et al 1996). Haydon et al (2013) reviewed six studies which examined the effects of choral responding. All six studies showed improvement in their dependent variables, which generally included the number of active student responses, on-task behavior, and correct responding on quizzes or tests. For example, Kamps, Dugan, Leonard, & Daoust (1994) utilized a reversal design with 24 students with developmental disabilities and measured active student responding as well as correct responding on weekly quizzes. They found that 75% of their sample had more responses per session as well as higher scores on weekly quizzes. Haydon et al (2010) examined six students and compared the efficacy of choral responding, individual responding, and mixed responding. They found that five out of six students had higher on-task percentages and lower numbers of disruptive behaviors. All six of their students had higher rates of responding when using choral responding than when using individual responding.

While choral responding has been found to effectively increase student rates of accurate responding, there may be times when using choral responding is not appropriate, and a quieter approach is necessary. In those instances, teachers can use response cards, which, like choral responding, allow for all students in the room to respond at once. Response cards are generally dry-erase boards or paper sheets on which students will write their response following an instructional prompt by their teacher. This technique was first examined by Narayan, Heward,

Gardner, Courson, & Omness (1990). Using a reversal design, the authors compared hand raising and response cards in a fourth-grade classroom of 20 students, and measured the performance of six students chosen by their teacher to represent the range of skill in the class. In their study, the teacher alternated between having the class respond using response cards and via hand raising. Unsurprisingly, the authors found significant increases in response rates, likely due to the vastly increased opportunities to respond in the response card condition. In the hand raising condition, the authors pointed out that an average of only 40 learning trials (i.e. teacher-posed questions) occurred while an average of 480 learning trials occurred in the response card condition. In a replication of Narayan et al (1990), Gardner, Heward, and Grossi (1994) examined the efficacy of hand raising against response cards in a fifth grade classroom. Similar to their predecessors, the authors measured the performance of five students from the class who were selected by their teacher to represent the range of skill in the room. The authors measured student responses, accuracy of responses, and quiz scores. They found that the percentage of student responding (hand-raising in the hand-raising condition or using the response card in the response card condition) increased from 4% to 68% of teacher-posed questions. Quiz scores were higher for almost all of the students in the class – not just the target students – for content taught using response cards as well as for review tests. Gardner et al (1994) replicated the findings of Narayan (1990) and demonstrated that the use of response cards is an empirically-supported method of increasing student opportunities to respond.

Timing

One way to increase efficiency is to reduce inter-trial intervals (ITI), which is the amount of time between the end of one learning trial and the beginning of another. Inter-trial intervals are distinct from inter-session intervals in that inter-*session* intervals are the amount of time between the end of one learning *session* and the beginning of another. For example, one could consider the

time between math classes, which are learning sessions, to be 24 hours during the normal school year and week. There are usually numerous learning trials in one learning session.

Skinner et al (1996) speculated that fast-paced instruction may increase learning rates due to the increased quantity of potential learning trials. Sometimes, the authors posited, quantity is more important than quality. Skinner et al (2002) conducted a single-case study in which a student learned two sets of sightwords. The authors manipulated the ITI such that the student either moved to the next trial immediately or had a five-second interval time between trials. The authors showed that the student learned a similar number of words over the course of a similar number of days in both conditions. However, when the authors instead graphed words learned per instructional second, they found large differences in favor of the immediate-ITI condition. While the authors point out that this is an example of the need for more precise measures of instructional time when graphing, their study also demonstrates that students can learn at higher rates when ITI is kept low.

Another way to improve the relative outcomes of interventions is to alter wait time. Wait time is the time between the delivery of a stimulus or cue and feedback from the instructor or interventionist. Wait time differs fundamentally from ITI in that wait times are imbedded within learning trials, while ITIs are between learning trials. In order to maximize efficiency, wait times must be balanced. If too much time elapses between stimulus and feedback delivery, time is wasted. If too little time elapses between the two, the student may not have enough time to produce a response (Riley, 1986; Row, 1974; Tobin, 1983). In a study that perhaps refutes this point, Poncy, Jaspers, Hansmann, Bui, and Matthew (2015) utilized a Taped Problems intervention and controlled for total number of trials but altered the wait time between the presentation of the intervention cue and the answer. While a two-second delay and no-delay caused relatively equal learning increases, the no-delay condition took 33% less time than the two-second delay, which indicates that a very short wait time between stimulus presentation and

response can in some instances produce similar effects to longer wait times and can have the added benefit of taking less time, which allows for more of the intervention to be delivered in the same amount of instructional time.

Similar to altering the wait time, it has been demonstrated that the use of time limits tends to increase the rate at which children complete work. Derr and Shaprio (1989) found that children read significantly more words per minute when they knew they were timed than when they did not. Van Houten, Hill, and Parsons (1975) had two fourth grade classrooms complete a writing task. One room was told they had ten minutes to write as much as they can, while the other room had twenty minutes to write as much as they can, but were not told they would be timed. The classroom in the timed condition wrote, on average, twice as much as the untimed room.

Set Size

Set size refers to the number of items per set in a learning session. For example, someone teaching a student the alphabet may not choose to teach them all 26 letters at once and instead may opt for a smaller set size, depending on the age and existing skill of the child. Purdum (2014) described several studies which demonstrated differential learning efficiency after altering set size. Duhon, Poncy, Hubbard, Purdum, & Kubina (2012) conducted a group design study in which a control group received 20 mathematics problems at once, experimental group one received five of the 20 problems for three days before moving onto the next set of five, and experimental group three received ten of the 20 problems for three days before moving on. The authors found experimental group one to make superior gains in accurate responding compared to control or experimental group two, which indicates that in this instance, a small set size was most efficient. While the literature seems to have supported small set sizes over larger ones, reducing the set size can act as a double-edged sword in that it frequently increases the rate of student learning at the cost of efficiency since more time is necessary to focus on a smaller set of items.

In the previous example, if a teacher chooses to focus on two letters, the student is likely to learn them rapidly, however, if the teacher had focused on six letters, the rate of student learning may have been lower, but the student may have mastered more letters in the same period of time.

Reinforcement

While they may not realize it, all educators use reinforcement in their instructional practices. A reinforcer is any stimuli that increases the probability of a behavior occurring in the future (Mazure, 2017) and can include peer and adult attention, access to tangibles (stickers, candies, etc.) or removal of unpleasant stimuli. The distinction between the application and removal of stimuli to increase behavioral probability is known as positive and negative reinforcement, respectively. The use of reinforcement is ubiquitous in intervention research (e.g. Daly, Witt, Martens, & Dool, 1997; Hofstadter-Duke & Daly, 2015; O’Conner & Daly, 2018; Van Houten et al, 1975). As Pipkin, Winters, and Diller (2007) point out, reinforcement is frequently nested within or alongside other intervention components. As such, several of the following discussions of other components discuss the impact of reinforcement.

Saliency of Prompts and Cues

It is not uncommon for children to struggle to discriminate between relevant academic stimuli. Children who add when they should subtract or read the letter ‘d’ when they should read ‘b’ are commonplace in elementary schools. These children tend to benefit from instruction which makes relevant academic stimuli more salient, or noticeable. For example, a teacher may circle the addition sign in a math problem or actively prompt a student to notice the sign by pointing at it. Since educators do not want children to be permanently dependent upon a circled addition sign or teacher prompts, they use a procedure called stimulus fading. In their applied behavior analysis text, Cooper, Heron, and Heward (2007) describe various procedures in which students are taught to respond to academic stimuli which has some dimension of it highlighted to

prompt correct responding. Over time, the highlighted dimension is slowly removed, or faded, to the natural stimulus.

Egeland and Winer (1974) demonstrated the efficacy of this approach. To teach letter discrimination, the authors compared a technique called errorless discrimination training (EDT) against the traditional flash-card drill with a group of 64 preschoolers. In the EDT condition, the authors highlighted in red characteristics that make letters distinct from one another. The authors used four pairs of words: R-P, C-G, Y-V, and K-X. Children in both conditions were trained using a matching-to-sample task. Both groups of children completed 10 trials for each pair of letters in the first half of the training phase. In the second half of the training phase, children in the EDT condition had the red highlights faded. By the tenth trial in the second half of the training phase, the antecedent stimuli was identical to that of the traditional flash-card drill condition. After the training phase, the children completed a post-test in which they distinguished between unaltered versions of the pairs of letters. The authors found that children exposed to the EDT condition made significantly fewer errors than those in the flash-card condition.

Van Houten and Rolider (1990) further demonstrated the efficacy of stimulus fading on the number identification of three elementary students with disabilities. Using a color mediation technique embedded in a multiple baseline design, the authors taught students to respond to colored numbers on flashcards by their color and their name (e.g. Green 1, Red 2, etc.). After four sessions with 100% accuracy on target items, the colored ink was removed and students were not expected to respond with the color and then the correct answer. The authors found that all three students rapidly increased their accuracy with identification of target numbers after implementation of the color mediation condition. After the authors withdrew the color mediation, the students maintained 100% accurate responding. This effect maintained at one, two, three, and four months after the intervention terminated.

Feedback

The provision of feedback is ubiquitous in school buildings, and is to many an obviously efficacious component of good instruction. Feedback can take the form of corrective feedback or performance feedback. Corrective feedback is given when a child makes an incorrect response, and serves to correct the child's responding. Performance feedback is usually given following a set of responses and serves to alert the child to their performance on the task (i.e. "You got eight problems right that time!").

The literature supports the view of feedback as an effective instructional component. In a previously discussed study, Van Houten et al (1975) found feedback to increase the amount of written work in two fourth grade classrooms. Krohn, Skinner, Fuller, and Greear, (2012) taught four children to identify their numbers 0-9 using a taped problems intervention. The authors found that three of their four students quickly improved to 100% accuracy in their number identification, and that the fourth student improved to 100% accuracy after the introduction of performance feedback and reinforcement components. Hier and Eckert (2016) conducted a randomized-controlled trial of 118 elementary school students and studied the impact of performance feedback on student writing performance. Students were broken into three groups, one which received no feedback, one which received feedback, and one which received writing probes designed to program generalization. In the feedback condition, students had their total words written from last session written in a box at the top of their worksheet alongside an upward-facing or downward facing arrow to indicate if last session's performance was higher or lower than the session before it. In the no-feedback condition, students simply received blank writing probes. In the generalization condition, students received probes designed by the authors to program generalization. The authors found that students in the performance, measured by total words written, was significantly higher for the performance feedback group than the no feedback and generalization groups.

Several studies have included feedback as a component in explicit timing interventions. Poncy, Duhon, Lee, and Key (2010) included a performance feedback element in their explicit timing study which compared ET with conceptual teaching and cloze procedures and found that explicit timing with a feedback component increase math facts fluency in their sample. Duhon et al (2015) studied the impact of immediate feedback on ET procedures in second grade students against no immediate feedback ET procedures and no ET procedures using a computerized ET program described in Gross and Duhon (2013). The results, as one might guess, showed no improvement in math fluency in the control group, which was given no intervention, and significant pretest-posttest differences for both ET groups. With immediate feedback, participants more than doubled their digits correct per minute (from 21 DCPM to 44 DCPM), while participants who did not receive immediate feedback improved their performance from 20 DCPM to 32 DCPM. All of these comparisons were found statistically significant.

Self-Monitoring and Goal Setting

Self-monitoring procedures generally entail the child keeping track of their performance themselves, and as such is usually a form of performance feedback. Carr and Punzo (1993) found self-monitoring of academic performance to improve accuracy and on-task behavior in reading, math, and spelling. Kern-Dunlap and Dunlap (1989) examined a self-monitoring procedure in three elementary-aged children with learning disabilities. Using a multiple-baseline design, they compared three instructional techniques for subtraction: traditional instruction, which consisted of verbal teacher direction on how to complete subtraction problems, teacher direction plus a point system for accurate responding, and a self-monitoring procedure. In the self-monitoring condition, students were given a checklist of subtraction steps, and the child checked off each step as they completed a problem. The authors found improvements in level of accurate responding to subtraction problem as well as reduction in variability in the self-monitoring condition than in either of the other two conditions. The improvements in level and variation

maintained for all three students in follow-up sessions, which occurred each day following removal of the self-monitoring phase and varied by student. One student's performance was measured for four days following treatment, one for 12, and one for 16 days.

In a similar study, Maag, Reid, and DiGangi (1993) used a multiple baseline design to examine the effects of attention, accuracy, and productivity self-monitoring procedures on the mathematics performance of six sixth-graders with learning disabilities. Following a tone, students would mark whether they were on-task, count up the total number of problems they'd completed since the previous tone sounded, and tally the number of correct problems since the last tone to measure attention, productivity (fluency), and accuracy, respectively. The authors found that all three forms of self-monitoring improved productivity and accuracy for all six students, while on-task behavior improved for four of six students.

Another form of self-monitoring is called self-graphing. Self-graphing involves the student keeping track of their performance on some task and graphing the data themselves. Self-graphing procedures have been supported in the literature several times (e.g. Coddling, Hilt-Panahon, Panahon, & Benson, 2009; Fuchs, Fuchs, Hamlett, & Whinnery, 1991; Schutte et al, 2015). Coddling, Chan-Iannetta, George, Ferreira, and Volpe (2011) examined this technique on a mathematics intervention called Kindergarten Peer-Assisted Learning Strategies in Mathematics (KPALS; Fuchs, Fuchs, & Karns, 2001), which utilizes peer-tutoring embedded within typical instruction to teach early numeracy skills. The authors compared KPALS with and without self-graphing in a sample of 96 kindergarten students. The authors measured progress using probes from AIMSweb Tests of Early Numeracy (Clark & Shinn, 2002) and found that students who received KPALS with or without self-graphing had significantly higher post-test performance on measures of early numeracy than did students in the control group who did not receive KPALS, and that students who had self-graphing in addition to KPALS performed higher than those who received KPALS alone.

One way to increase the efficacy of self-monitoring procedures is to include goal setting. With goal setting, students are informed of their performance goal and are told to try to meet or exceed it before each instructional setting. Many interventions, both academic and behavior, include a goal-setting component, and this practice has considerable empirical support (e.g. Graham, MacArthur, & Schwartz, 1995; Koenig, Echert, & Hier, 2016; Schunk & Schwartz, 1993). Despite this support, some researchers have found it lacking when compared to other interventions. Koenig et al (2016) conducted a randomized controlled trial of 118 third grade and compared writing fluency growth of third graders who received performance feedback and goal setting, performance feedback alone, and no additional intervention components. Progress was measured using curriculum-based measurement (CBM) writing probes. Students in the performance feedback with goal setting condition received their writing probes which had their previous performance alongside a goal for that session, while students in the performance feedback only condition did not have a goal on their probe. In the control condition, students simply received a blank writing probe. The authors conducted the study over eight sessions which occurred at a rate of one per week. While the authors found no differences between the feedback only and combined conditions, both were significantly superior to the control conditions with regards to both total words written and correct word sequence.

Distributed Practice

The distributed practice effect has been studied for well over 100 years. Distributed practice is the phenomenon in which longer inter-session intervals tend to produce higher levels of overall learning compared to shorter inter-session intervals. Inter-trial intervals differ from inter-session intervals in that multiple trials are generally nested within a single session; that is, a session is usually composed of several trials. The temporal placement of practice was first explored by Ebbinghaus (1885/1913), who taught himself nonsense words, which consisted of all three-letter consonant-vowel-consonant (CVC) combinations of letters. After creating the

nonsense syllables, Ebbinghaus removed those which had semantic meaning (e.g. “cat”) and was left with approximately 2,300 nonsense syllables. In order to study nonsense words of varying length, Ebbinghaus would place syllables together to create nonsense words. For example, the nonsense syllables “bim,” “ner,” and “tep” could be put together to make the three-syllable nonsense word, “bimnertep.”

Though the following study examined forgetting, it has implications for distributed practice. Published in the same book as the previous study, Ebbinghaus completed 163 tests over the course of a year. In each test, he learned eight nonsense words which each had 13 syllables. After learning each list, he would conduct a “relearning session” after some period of time. Ebbinghaus conducted relearning sessions after 20 minutes, one hour, nine hours, one day, two days, six days and 31 days. In this study, relearning sessions occurred after the completion of the first learning session in which Ebbinghaus learned the nonsense words to mastery. The dependent measure was the total time it took to learn each list of words, and Ebbinghaus described the “saving of work” (p. 67) as the difference between the amount of time it took for him to learn the list of words initially and the amount of time it took for him to relearn the list to mastery. Ebbinghaus found that the “work saved” decreased after each successive increase in time between initial learning and relearning sessions, which implies that he forgot more information when sessions were more spread out. This finding demonstrated that distributing practice too far apart can be antithetical to the facilitation of efficient learning. There is clearly a point at which distributing practice no longer provides benefits.

Ebbinghaus’ (1885/1913) work is credited with inspiring other works at the time which examined the effects of various temporal placements of practice (e.g. Jost, 1897; Reed, 1924; Ruch, 1928; Thorndike, 1912). Thorndike (1912) examined the efficiency of work of five graduate students, who were made to work continuously on mathematics problems for varying periods of time and discussed various characteristics of their work. Among the characteristics,

Thorndike described the practice effect, which was defined by him as “a measure of difference between two tests separated by an interval of time” (p. 186). He goes on to describe the implied phenomenon in which efficiency increases as people practice something, but then lose this gained efficiency over periods of time in which they are not practicing.

Contemporary Study of Massed and Distributed Practice

Much of the early work on distributed practice focused more on aspects of forgetting than improving the efficiency of learning over time, which is generally the focus of more contemporary study on the phenomenon. This review will begin with studies that took place in cognitive labs and transition to studies of the distributed practice effect in applied settings.

Generally speaking, *massed* practice refers to practicing material to be learned all at once and has an inter-session interval of zero, while *distributed* practice refers to practicing material over several sessions and as such has inter-session intervals of greater than zero. Practice can be distributed in two ways: by time or by items. While the current study focuses on distributing practice by varying the inter-session time intervals of practice, others have examined the impact of varying the distance (in items) between repeated presentations of the same target item. For example, one condition may have a target item, three other items, and then the target item again, while another condition may have the target item, then ten other items, and then the target item again.

Cepeda, Pashler, Vul, Wixted, and Rohrer (2006) conducted a meta-analysis of 317 experiments that studied the distributed practice effect on verbal memory tasks, most of which were paired, list, or fact recall. The authors examined the impact of spacing, or distributing, practice sessions and grouped inter-session intervals into the following categories: 1-59 seconds, one minute to 10 minutes, 10 minutes to less than one day, one day, two to seven days, eight to 30 days, and 31 or more days. The authors found significant retention improvements over massed

practice when practice sessions were distributed 1-59 seconds, 1 minute to ten minutes, and eight to 30 days. The authors also found significant overall improvement in retention in distributed practice conditions when comparing all studies, which had a total N of 14,811.

Janiszewski, Noel, and Sawyer (2003) conducted a meta-analysis of 97 verbal learning studies comprised of 284 experiments in an attempt to apply the results to consumer verbal learning based on exposure to advertisements. While the authors' argument for the application of the findings to advertising may be on shaky ground, the evidence supporting distributed practice is not. All 284 experiments the authors examined concluded that spaced, or distributed, practice was superior in learning verbal material ($r = .339$). The authors calculated a fail-safe N , which is an estimate of the number of studies with a zero effect size that would need to exist to negate their significant finding. The authors reported that nearly 150,000 unpublished studies would need to exist with a zero effect size to render their finding insignificant. Lee and Genovese (1988) conducted a meta-analysis of 47 articles which examined the distributed practice effect on the acquisition of motor skills and found similar results to the literature on verbal learning: distributed practice consistently produces superior results to massed practice.

Cepeda, Vul, Rohrer, Wited, and Pashler (2008) had over 1,300 participants memorize obscure facts and measured their retention after a varying period of practice. The authors split the participants into several groups, which completed two learning sessions and several measures of retention. The learning sessions varied in temporal distance and had gaps of 0 days (i.e. session two occurred immediately following session one), one day, two days, seven days, 21 days, and 107 days. Testing occurred after seven days, 35 days, 70 days, and 350 days. Simply put, participants had two practice sessions in which feedback was delivered and four testing sessions in which their retention of information was tested. The authors found that participants who had more spread out retention intervals retained the information for longer than those who had shorter

retention intervals. However, those who had short retention intervals tended to perform more highly on tests that occurred more closely to the beginning of the study.

While the consensus in much of the literature is that distributed practice trumps massed practice, perfect agreement does not exist. In her doctoral dissertation, Moss (1996) reviewed 120 studies which examined the effects of distributed practice across learning of motor skill, intellectual skill, and verbal information. The author defined intellectual skill as the completion of mathematics problems. She found distributed practice to be superior in 83% of the studies she examined, which left approximately 17% of studies that found either no differences between the two conditions or found massed practice to be superior. Donovan and Radosevich (1999) conducted a meta-analysis of 63 studies with 112 experiments which demonstrated again that overall, distributed practice is superior to massed practice. However, additional analyses found that this effect was mediated by two other variables: the type of practiced task and inter-trial time interval (i.e. the time between learning trials within a session). Simpler tasks which had low inter-trial intervals – that is, were quickly-paced within the learning session – demonstrated more powerful distributed practice effects than more complex tasks. Further, the authors point out that higher effect sizes were found in studies that had low methodological rigor, measured through a nine-item checklist. It appears that the literature generally supports the use of distributed practice, but suggests that it is pertinent to keep the task type in mind when planning interventions which will utilize the distributed practice effect.

Theories of Distributed Practice

In any phenomenon that carries with it over 100 years of research, theory surrounding its cause will abound. Unsurprisingly, there is a noticeable schism between cognitive and behavioral theories of distributed practice. From a constructivist point of view, Glenberg (1979) presented a comprehensive cognitive theory on distributed practice. Glenberg's component levels theory

assumes internal processes which encode various visual features of stimuli, which are then used to retrieve information from long-term memory. According to Glenberg's theory, these visual features can be split into three "components:" contextual components, structural components, and descriptive components. Contextual components refer to things surrounding the item to be learned – physical environment, time, and the cognitive and emotional state of the learner. Structural components are similar to Piaget's (1957) schemas, and represent the way in which items are chunked or categorized. Descriptive components include physical descriptors of the item to be learned, such as its color, shape, texture, etc. In distributed practice, Glenberg (1979) asserted that distribution, or *repetition lag*, leads to the storage of different components of the item, and that more stored components of the learning item, the higher the chance of recall of the item by the learner. Others have suggested similar theories which describe distributed practice as an effect of contextual fluctuation (see Estes, 1955; Raaijmakers, 2003).

To support his theory, Glenberg (1979) utilized 64 college students as participants. Instead of using time to measure the lag between item presentations, the author used two to six intervening items to space presentation of the target items. In this study, items were pairs of words which were practiced through exposure to the pairs and the instruction to memorize the pairs. Items were practiced in either a massed or distributed fashion, and later tested in both cued and free-recall conditions. The author found distributed practice of items to improve memory in free recall, but not cued recall, where there were no differences between massed and distributed practice conditions. The author suggested that this finding is due to the possibility of cues activating descriptive components in the participants' memory.

Of course, behavioral theorists would likely use quite different language to describe Glenberg's theory. What he referred to as contextual components would likely be called discriminative stimuli by behaviorists, and his descriptive and structural components would be talked about in terms of stimulus control and concept formation. The key difference between

these competing theories is *where* the distinction is taking place. Behavioral scientists study the impact of the environment on the behavior of the individual, while cognitive scientists attempt to study the impact of the individual's various mental constructs on the behavior of the individual. While worthy of research, "internal" variables are difficult to measure. For example, it is possible, and even likely, that some biological process controls human memory. However, we have historically struggled to pinpoint an exact brain region which controls memory. There is evidence to suggest the hippocampus is associated with memory, but in case studies where the human hippocampus is destroyed or damaged, memory does not cease to exist. In several cases, damage to the hippocampus only resulted in anterograde amnesia, which is the inability to form new memories after the brain damage occurred – memories from before the damage are usually relatively intact (for a review, see Carlson & Birkett, 2017, p. 445).

Thios and D'Agostino (1976) suggested that the distributed practice effect in the recall of verbal information is due to retrieval during the practice phases. Their idea, called the study-phase hypothesis, states that the act of recalling information during the practice reduces the changes of forgetting in the test phases. Put simply, this hypothesis posits that recalling information acts as a 'reminder' of the information during training, and that reminder is why participants tend to recall more information in the final phase. While an interesting theory, it does not explain the increased performance of distributed practice on tasks which require skill fluency and not rote recall. Further, Küpper-Tetzel (2014) pointed out both of the previously discussed theories, component-level and study-phase retrieval, are weak because they often fail to produce strong and consistently accurate predictions.

Pavlik and Anderson (2003) put forth a more parsimonious solution. Distributed practice is the result of convergence between two basic learning principles: Jost's (1897) law of forgetting and the power law of learning. Jost's (1897) law of forgetting states that if two associations are of equal strength but different ages, the older will lose strength more slowly than the newer. Jost's

law of forgetting would later become the power law of forgetting, which in the context of distributed practice simply states that the rate of forgetting slows as time passes. Put onto a graph, the law would demonstrate that forgetting is fast at first, but slows as time goes on. The power law of learning in this context states that learning rate is steep at first and then slows over time. In most laboratory studies of distributed practice, there is an initial training, an intersession interval, a second training, a retention interval, and then the test. The authors posit that the number of learning sessions participants are exposed to represents the power law of learning while the retention interval (i.e. the time between the last practice session and the posttest) represents the power law of forgetting. In their study, the authors recruited forty college students and assigned them to either a one day or seven day retention-interval condition. Participants were exposed to pairs of English and Japanese words and told to remember the pairs. Within the two retention-interval groups, participants studied pairs of words one, two, four, or eight times. Material was spaced by items, and participants were exposed to two, 14, or 98 intervening items. After one or seven days, participants returned for the posttest session. Perhaps unsurprisingly, the authors found learning of the material to be consistent with the power law of learning and that more exposures to the material facilitated more accurate responding in a logarithmic fashion. The authors also found that accuracy on the final session depended upon the spacing condition the participant was exposed to. Those who had material spaced between two intervening items performed the highest at the end of the training session, but the lowest on the second day, regardless of their retention interval. Subsequent spacing conditions were similar. Those who had 14 intervening items retained the second-most at the end of day one and the second-most at the end of day two. Those who had 98 intervening items retained the least on day one and the most on day two.

These findings support the theory of distributed practice as a convergence of the power laws of learning and forgetting. Participants demonstrated low but stable learning of the material

when it was spaced at very long intervals because the long intervals meant fewer presentations of the target material, and fewer presentations of material reduce the rate of learning of the material. Further, they retained the material precisely because of the long intervals. The power law of forgetting posits that forgetting is steep at first but slows over long periods of time. It is possible that just as participants were forgetting material, they were exposed to it again. The long-interval repeated exposures could have reasonably maintained the low but stable responding to the learned material.

While the study of the theoretical causes of distributed practice effects is interesting, it is possible that science will never know the true underpinnings of the phenomena. With that reality in mind, the study of the utility of distributed practice in applied settings should be the more immediate focus of scientists.

Applied Studies of the Distributed Practice Effect

The distributed practice effect has been shown effective numerous times in the literature for a wide variety of academic outcomes. Distributed practice is superior to massed practice in vocabulary learning (e.g. Bahrick, Bahrick, Bahrick & Bahrick, 1993; Dempster, 1987; Seabrook, Brown, & Solity, 2005), reading (Dempster, 1989; English, Wellborn, & Killian, 1934), science knowledge (Reynolds & Glaser, 1964; Vlach & Sandholfer, 2012), and mathematics facts fluency (e.g. Rea & Modigliani, 1985; Rohrer & Taylor, 2006; Schutte et al, 2015) and has been demonstrated effective in all ages of children and adults (Cornell, 1980; Cahill & Toppino, 1993; Rea & Modigliani, 1987). For example, Rea and Modigliani (1987) conducted a study of 96 children from kindergarten through third grade and had them practice identifying words and pictures in a massed or distributed fashion. The authors found the distributed practice effect in all ages of children they studied. The following section will review the available literature on the use of the distributed practice effect for applied purposes.

Distributed practice with vocabulary learning. Dempster (1987) found that undergraduate students more accurately recalled unfamiliar vocabulary words when they were learned using distributed practice than when learned using massed practice. Sobel, Cepeda, and Kapler (2011) taught 39 fifth grade students two roughly equivalent lists of Graduate Record Examination (GRE; Educational Testing Service, 2019) words. Students practiced one list using distributed practice, and learned the other list using massed practice. The authors found that at follow up, their participants recalled approximately 20% and 7.5% of the distributed practice words and massed practice words, respectively.

In a somewhat unorthodox nine-year longitudinal study, Bahrick, Bahrick, Bahrick, & Bahrick (1993) utilized themselves as participants. The authors studied 300 pairs of English and foreign language words in six training conditions which manipulated the interval between learning sessions as well as the number of total sessions. Unsurprisingly, they found that recall increased as a function of the number of total learning sessions. They also found, however, that recall increased as a function of the interval between learning sessions. The authors recalled 20% more words in their longest inter-session interval (56 days) than in their shortest inter-session interval (14 days).

Distributed practice and reading. In the area of reading, the literature tends to support the view that distributed practice is beneficial, but the rate of return diminishes rapidly after a few hours. English et al (1934) had their participants read texts four times. Participants read the text four times consecutively, three hours apart, one day apart, or three days apart. The authors found that those who practiced reading the text in a distributed fashion performed more highly on measures of comprehension than those who practiced them in a massed fashion. However, the author found no group differences in comprehension between the three distributed conditions. Other researchers at the time found similar results in distributed reading practice (Lyon, 1914; Peterson, Ellis, Toohill, & Kloess, 1935; Sones & Shroud, 1940).

Distributed practice and science knowledge. Reynolds and Glaser (1964) sought to use distributed practice to teach biology facts to 75 middle school students matched for intellectual ability. After breaking the curriculum down into ten topics, the students were split into five groups that either reviewed the material all at once (i.e. massed) or reviewed it in a distributed fashion. In the massed condition, students reviewed the material as intended, half as much, or one and a half times as much as the original curriculum, which varied in length by topic (i.e. some topics were longer than others). In the distributed condition, students either studied the as intended or one and a half times as much as intended. The authors found that for all topics, participants in the distributed practice condition performed more highly on post-test measures of the material than those who studied in a massed fashion. Interestingly, the authors found no effect of the amount of review.

Vlach and Sandhofer (2012) gave four science lessons to 36 first and second grade children in a massed, clumped, or spaced fashion. Students in the massed condition received all four lessons consecutively, students in the clumped condition received two lessons spaced one day apart, and students in the spaced condition received one less per day for four days. Pretests for particular lessons occurred directly before the presentation of that lesson. Posttests occurred one week after the last lesson the student had. For example, students in the massed condition who received all four lessons on a Monday would take the posttest the following Monday. Children in the spaced condition who received their last lesson on a Thursday would take the posttest the following Thursday. The authors found that both clumped and spaced conditions increased posttest performance over the massed condition. No differences between the clumped and spaced conditions were found for simple generalizations, which were essentially rote recall of lesson material (e.g. “Which of these does a snake eat?”). Spaced practice was, however, found superior to clumped practice in what the authors refer to as complex generalization. An example of a complex generalization item would be: “the grass gets sprayed with a poison that makes animals

die when they eat it. What happens to the number of crickets? Does it go up, down, or stay the same?" (Vlach & Sandhofer, 2012, p. 1140).

Distributed practice and math skills. To date, several studies have examined the effects of distributed practice on the accuracy and fluency, or speed, with which students complete basic math fact problems (e.g. Rohrer & Taylor, 2006; Roher & Taylor, 2007; Schutte et al, 2015). Rohrer and Taylor (2006) taught 216 college students to complete a specific kind of math problem and had them practice using one of two practice schedules. All students practiced 10 problems. Roughly half of the participants split the practice into two sessions, which were one week apart. The other group of students practiced all 10 problems at once. Although both groups had comparable performance one week after their practice sessions were complete, the authors' distributed practice group was twice as accurate as the massed practice group (64% accuracy and 32% accuracy, respectfully) after at a four-week follow up. In a similar experiment, Rohrer and Taylor (2007) had three groups of undergraduates practice math problems in three sessions, each one week apart. In this study, participants assigned to the distributed practice condition were called "spacers," those in the massed practice condition were called "massers." The authors included a third condition in which participants received half of the practice as the spacers and massers. These participants were referred to as "light massers." Rohrer and Taylor (2007) found that the "spacers" were 74% accurate at follow up, compared to 49% and 46% for "massers" and "light massers," respectfully.

Gay (1973) conducted two experiments which sought to examine the effects of distributed practice in middle school students who struggled in math. In the first experiment, 53 students were split into four groups and given computerized instruction on several rules for solving algebraic problems over two sessions. The program began with a pre-test and followed with the instruction. Instruction consisted of a presentation of the mathematics rule, an example problem, and two practice problems. If the student accurately responded to the two practice

problems, they moved on to the next rule. If the student did not respond accurately, they were shown their error and given two more problems. Though the procedure for all four groups was identical, the time between the two sessions varied by group. One group returned for the second session after one day, another after one week, and a third after two weeks. The fourth group completed a second session immediately following the first session. Consistent with other literature (English et al, 1934; Lyon, 1914; Peterson, Ellis, Toohill, & Kloess, 1935; Sones & Shroud, 1940), the authors found that post-test performance was higher for the students who received their instruction in a distributed fashion. However, no differences were found between the three distributed conditions.

Schutte et al (2015) split 53 third-grade students from three classrooms into three groups in a longitudinal stratified randomized design. This design was chosen to most effectively minimize error associated with nonrandom sampling. The authors rank ordered each student on their math facts fluency ability, and then split into clusters of three. Students in each cluster were assigned a random number from one to three and assigned to groups based on their number. To measure math facts fluency, the authors used curriculum-based measurement probes that measured fluency for addition problems which have correct answers of up to 18. Each probe takes one minute to administer, during which time students were told to complete as many items as they could. These probes can be used to determine the student's digits correct per minute, which has been demonstrated to be a reliable and valid method of measuring performance in basic math (Deno, 2003).

To control for the amount of practice students received, each student completed four probes per day for the duration of the study. Students either completed their four one-minute timings back-to-back in one session, two timings in the morning and two in the afternoon (approximately three hours later), or one timing in each of four sessions throughout the day, each of which were approximately two hours apart. Since there were times in which some students

were working on math problems and some were not, the authors had those who were not engaged in math problems complete reading worksheets. The authors found that students demonstrated more math facts fluency growth when intervention sessions were distributed, and the most growth when sessions were split into four one-minute timings per day. This study demonstrated the practical utility of distributed practice in the elementary classroom as well as provided further evidence of the efficacy of explicit timing interventions.

Current Study

Schutte et al (2015) allowed several hours to pass between sessions for students who received practice that was distributed. While this interval resulted in enhanced growth over the massed practice condition, elementary school math periods are generally not this long nor occur more than once per day. Distributing practice could be more practical for implementation in schools if its effects were found over shorter inter-session intervals. This more practical use of distributed practice would be cost effective (Roediger & Pyc, 2012) and easy to implement (Carpenter, Cepeda, Rohrer, Kang, & Pashler, 2012).

The purpose of the current study is to determine if a shorter temporal distribution of explicit timing intervention sessions in elementary students can result in a more effective fluency gain over massed practice. We have two hypotheses: 1) Consistent with Schutte et al (2015), children who participate in distributed practice will show more growth in DCPM during intervention than children who participate in massed practice, and 2) Children who participate in distributed practice with short inter-session intervals will show similar growth to children who participate in distributed practice with long inter-session intervals.

CHAPTER III

METHOD

Participants

We utilized three 3rd grade classrooms from a small school district in a rural Midwestern town. Participants were 50 children in a general education setting. However, seven students were excluded from analysis, bringing the total participant pool to 43. Students who were absent for three or more days of intervention were still given intervention on days they were present for the administration of intervention sessions, but were excluded from analysis. Six students met criteria for exclusion from analysis based on absences. Further, one student was excluded from analysis due to an extreme decrease in level from baseline which remained stable and was likely attributable to a performance deficit rather than a skill deficit. During fidelity observations, this student was observed to not participate in the explicit timings. To maintain experimental control, reinforcement procedures were not put in place for this student but were recommended to classroom teachers at the conclusion of the study.

Materials

Materials and procedures used in this study were based upon Schutte et al (2015). We used paper-and-pencil mathematics probes which were modeled from those in the Measures & Interventions for Numeracy Development (MIND, Poncy & Duhon, 2017). All problems were presented in a vertical format using Arial font size 16, and addition to ten problems were randomly generated between days using Microsoft Excel. That is, all students completed the same four math probes in the same order each day, but math probes between days differed. A sample math probe can be found in Appendix A.

Due to our procedure, all students were not always engaged in completing mathematics probes. During times when a student did not have a mathematics probe to complete, they instead worked on a word search in which target words were composed of the class' spelling words for that week. An example of a word search can be found in Appendix B. The reason for having students complete non-mathematics related timed work is twofold. First, it avoids wasting instructional time for those not working on mathematics probes. Second, keeping all students engaged in some task throughout intervention administration reduces chances of distraction to those who are working.

Experimental Design and analysis

In keeping with a partial replication of the work of Schutte et al (2015), we utilized a longitudinal stratified sample (Shadish, Cook, & Campbell, 2002). Using this technique, we rank ordered students based upon their baseline performance. Then, we assigned students to groups in triads by taking the three highest performers and randomly assigning them to treatment groups, then the next three, and so on. This design increases the chances of group equivalency, especially in group design experiments with a relatively small sample size.

Independent Variable

The independent variable in this study was the time between intervention sessions. The massed practice received all four one-minute mathematics probes back-to-back. The three-hour received two one-minute mathematics probes in the morning and two in the afternoon with approximately three hours between administrations. The ten-minute group received two one-minute mathematics probes concurrently with the other two groups, and then received the second set of two mathematics probes approximately ten minutes later. These sessions took place at approximately 9:10am, 9:25 am, and 12:50pm during each day of intervention. While there was 15 minutes between the first and second administration, the first administration took approximately five minutes to complete, keeping the actual intersession interval to ten minutes.

Dependent Variable

The dependent variable in this study was be the mean number of digits correct per minute (DCPM) students achieved each day across their four one-minute math probes. In scoring DCPM, digits are correct if the student places the correct number in its corresponding column. For example, a student would be scored as having two digits correct if they provided an answer of “12” for “8 + 4.” They would be scored one digit correct if they provided an answer of “11,” “22,” or “14,” because only one column contains a correct response. We would score the student has having zero digits correct for a response of “9” or “33.”

Procedure

Due to the following procedures falling within the realm of standard educational practice, individual parent consent was not required (Protection of Human Subjects, 2018). However, a mailer created to match standard letter formatting at our participating school site was sent home with each student outlining the study and allowing parents to contact the researchers with

questions or comments. No parents contacted the researchers during the study. The de-identified letter can be found in Appendix C.

We had students practice their math for a total of four minutes each day for 20 consecutive school days. After 6 calendar days with no treatment, we conducted a maintenance assessment to check for the retention of fluency skills. Though Schutte et al (2015) used a ten-day maintenance period, a six day period was chosen due to the spring break scheduled in the school calendar.

All three classroom teachers were trained to conduct the intervention by the experimenter, and the classroom teachers carried out the entirety of the intervention sessions using a script (Appendix D). Procedural fidelity was monitored by the experimenter and graduate research assistants for 30% of intervention sessions using a form that can be found in Appendix E. Intervention fidelity was calculated as a percentage of steps completed by the administrator of the intervention session.

Table 1

<i>Tasks for each group during daily intervention sessions.</i>				
	9:10 AM	9:25 AM	12:50 PM	
Massed	Four Math Probes	2 Word Searches	2 Word Searches	
10 Minute Distributed	2 Math Probes 2 Word Searches	2 Math Probes	2 Word Searches	
3 Hour Distributed	2 Math Probes 2 Word Searches	2 Word Searches	2 Math Probes	

Prior to the first day of intervention, each student completed three addition to ten fluency probes, with the median score representing their baseline. During each day of intervention, teachers conducted three administration sessions. The first administration began as close to 9:10am each day as the teachers were able. The second administration took place at 9:25am, and the third took place at 12:50pm. A summary of the tasks each group completed can be found in Table 1.

CHAPTER IV

FINDINGS

Procedural Fidelity

We trained several school psychology graduate students to observe intervention sessions to check for procedural fidelity. Overall, there were 60 intervention sessions per classroom (20 days of intervention times 3 sessions per day) for a grand total of 180 sessions between the three classrooms and a total of 54 observed sessions (30%) between the three classrooms. Of the 54 observed sessions, 18 observations took place in each of the three participating classrooms. Due to research assistant availability, observations of each of the three sessions were skewed toward the first morning session ($N = 22$) and the second morning session ($N = 22$) over the afternoon session ($N = 10$).

The procedural fidelity form can be found in Appendix E. The form consisted of a nine-point scale which listed each of the steps a participating teacher would need to complete in order to conduct the session with fidelity. In order to obtain an average deviation score, we also asked observers to record the start time of each session, defined by the time in which the teacher started the timer for the first one-minute timing. Finally, we asked each of the three participating teachers to record their start time in the same manner as our research assistants for each session and each day of intervention. This form can be found in Appendix F. The results of these measures of procedural fidelity can be found in Table 2.

Table 2

Fidelity results by session.

	Observed Fidelity	Observed Deviation	Reported Deviation
9:10am Session	98%	+/- 57s	+/- 56.8s
9:25am Session	92.8%	+/- 42s	+/- 45s
12:50pm Session	98.8%	+/- 40	+/- 81s*
Overall	96.5%	+/- 48s	+/- 16.5s

** On day 5, the 12:50pm session had an eleven-minute delay for one classroom. No other outliers were present. When removing that session from analysis, the average reported absolute deviation in session three was +/- 44.2s.*

We also calculated fidelity by classroom to ensure there were no systematic effects within any of our three classrooms. Our measures of procedural fidelity by classroom are available in Table 3. Note that reported deviations by classroom are lower than reported deviations by session because the former takes all 180 sessions into account, whereas the latter only accounts for 54 sessions.

Table 3

Fidelity results by classroom.

	Observed Fidelity	Observed Deviation	Reported Deviation
Mrs. C	96.3%	+/- 14s	+/- 16s
Mrs. F	98.1%	+/- 68s	+/- 21s
Mrs. O	95%	+/- 18s	+/- 13s

Baseline Analysis

To demonstrate that our three classrooms and groups were equivalent at baseline, we conducted two one-way ANOVAs. We found no significant mean DCPM differences between classrooms, $F(2, 43) = 1.411, p = .255$ or between practice conditions, $F(2, 43) = .529, p = .593$. Descriptive statistics for baseline, end of intervention, and maintenance are listed in Table 4.

Table 4
Means and SDs of groups in DCPM.

	10 Minute ISI		3 Hour ISI		Massed Practice	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Baseline	33.34	9.25	34.31	9.91	31.56	9.92
End of Intervention (20 Days)	40.19	14.97	52.55	16.84	45.31	19.43
Six-Day Maintenance	44.83	16.04	49.8	20.52	47.13	20.39

Hierarchical Linear Model

The data were analyzed using hierarchical linear modeling (HLM). In accordance with a partial replication of Schutte et al (2015), each observation point served as a level one variable nested within individual students, who were then nested within distributed practice conditions. HLM can also be referred to as multi-level modeling (MLM), and its procedures are described in detail in Greenberg and Phillips (2013) as well as in Raudenbush and Bryk (2002).

Our unconditional model, which is the multi-level model with no predictors added, can be represented by the following set of equations:

$$\text{Level 1: } \text{Score}_{i} = \pi_{0i} + \pi_{1i} * (\text{Occasion}_{i}) + e_{i}$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + r_{0i}$$

$$\pi_{1i} = \beta_{10} + r_{1i}$$

Where $Score_{ti}$ represents a student's average DCPM across their four math probes on each day and $Occasion_{ti}$ represents each of the 20 days of intervention. The intra-class correlation (ICC) in our unconditional model equaled 0.873, which indicates that 87.3% of the variance in the model is attributable to level two clustering. Across groups, the average participant started at 34 DCPM and grew .7 DCPM during each day of intervention. Both the intercept $t(42) = 20.619, p < .001$ and slope $t(42) = 8.119, p < .001$ were significantly different from zero. These significant findings indicate that there is sufficient variance in the model to justify the introduction of our level two predictors. The next iteration of our model compared our ten-minute and three-hour conditions to our massed condition using a two-vector dummy coding scheme. This model can be represented by the following equations:

$$\text{Level 1: } Score_{ti} = \pi_{0i} + \pi_{1i} * (Occasion_{ti}) + e_{ti}$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01} * (Vector1_i) + \beta_{02} * (Vector2_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} * (Vector1_i) + \beta_{12} * (Vector2_i) + r_{1i}$$

In this iteration of the model, we compared the ten-minute (Vector 1) and three-hour (Vector 2) distributed groups to the massed condition, which served as our reference group. We did not find a statistically significant slope for either our fifteen-minute or our three-hour groups, $t(40) = 1.936, p = .06$, and $t(40) = -.52, p = .606$, respectively. Figure 1 shows average group DCPM scores for each day of intervention. Visual analysis of these data indicate a clear level and trend difference between the three-hour condition and the other two conditions.

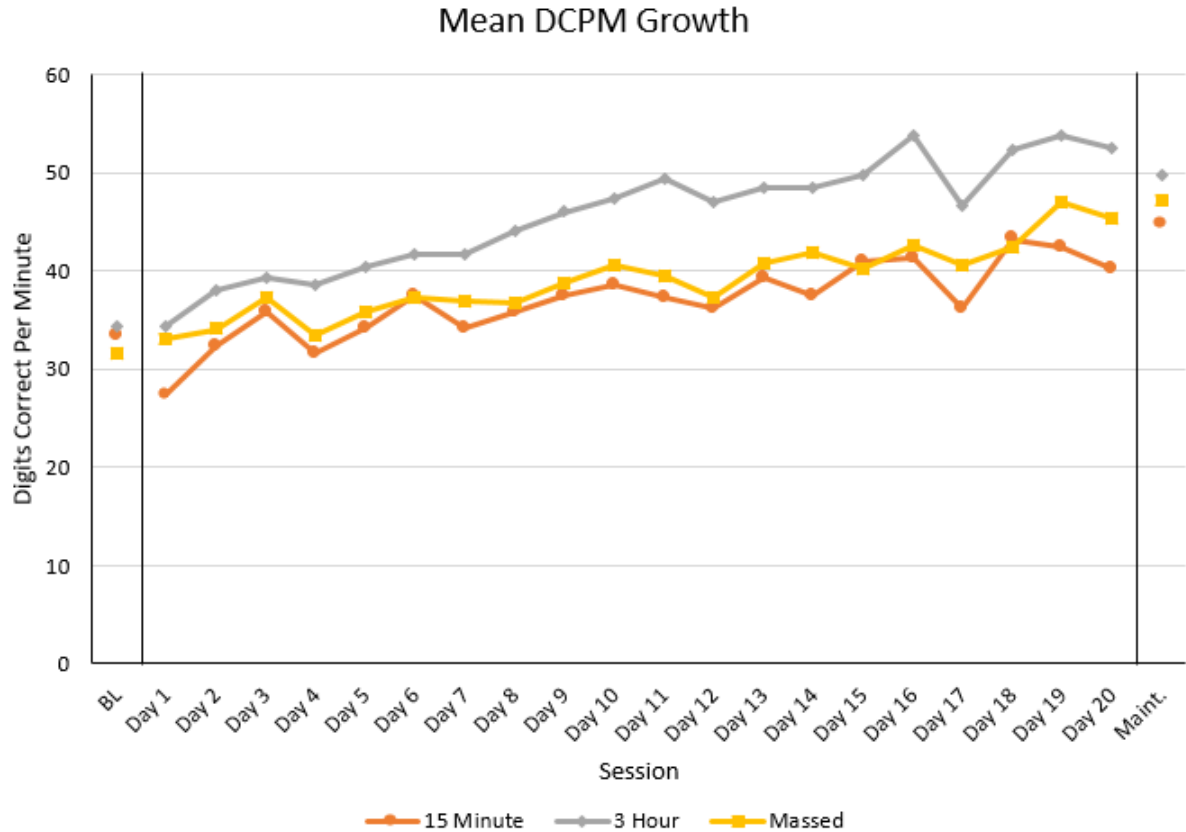


Figure 1. Mean DCPM growth between the three distribution conditions.

Due to these differences and lack thereof, we collapsed the fifteen minute and massed conditions and used a simple one-vector dummy coding scheme to compare growth between the three hour condition and the rest of the participants. Our final model can be summarized by the following equations:

$$\text{Level 1: } \text{Score}_{i_t} = \pi_{0i} + \pi_{1i} * (\text{Occasion}_{i_t}) + e_{i_t}$$

$$\text{Level 2: } \pi_{0i} = \beta_{00} + \beta_{01} * (\text{Vector3}_i) + r_{0i}$$

$$\pi_{1i} = \beta_{10} + \beta_{11} * (\text{Vector3}_i) + r_{1i}$$

In the final iteration of the model, we compared the three-hour distribution to the newly-collapsed fifteen-minute and massed conditions. On average, a student in the three-hour distributed condition grew .433 DCPM more between each day of intervention than a student in either the

fifteen-minute or massed conditions. This slope was statistically significant, $t(41) = 2.42, p = .02$.

Figure 2 shows average group DCPM scores with the collapsed groups.

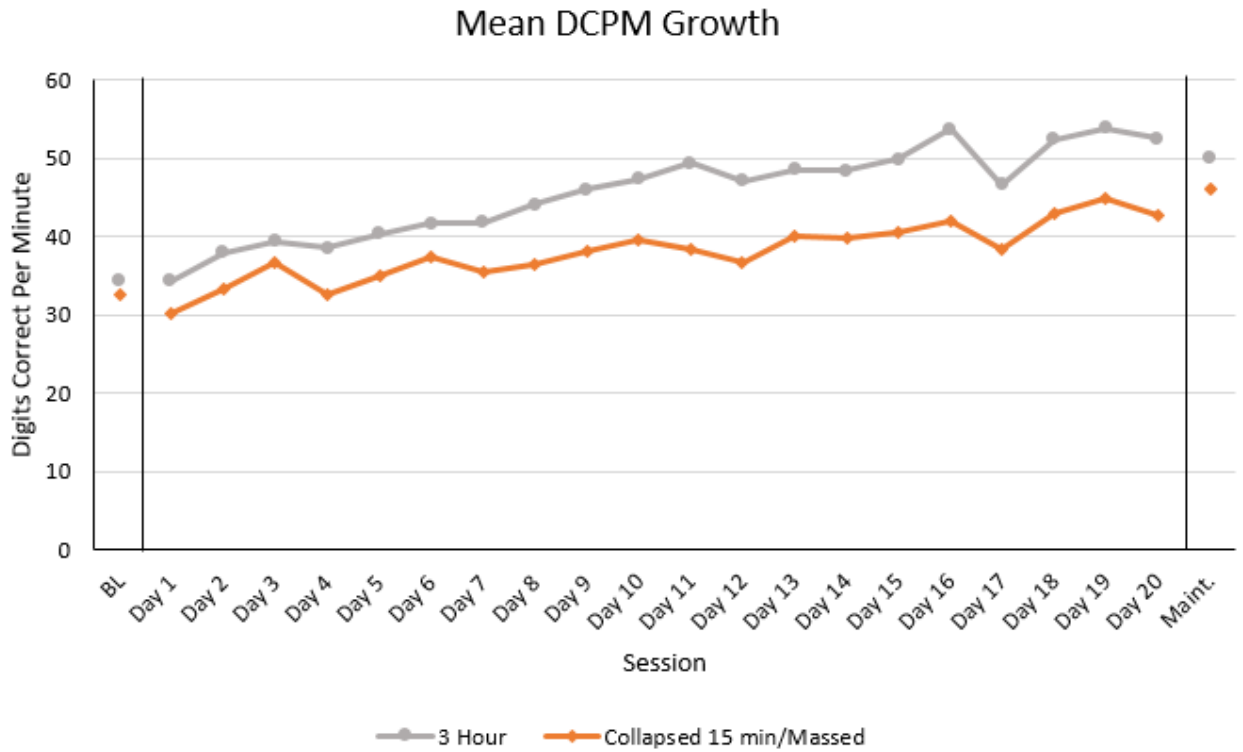


Figure 2. Mean DCPM growth in the three-hour distributed condition and collapsed fifteen minute and massed conditions.

CHAPTER V

CONCLUSIONS

Our results indicate that we partially replicated Schutte et al (2015). The group of students who received the three-hour intersession interval grew at a significantly higher rate than did the students who practiced in a massed or a short-distributed fashion. In our original hypothesis, we indicated that both of our distributed groups would grow at a higher rate than the massed group. However, while all three groups grew at statistically significant rates, the short-distributed group and massed group grew at similar rates. This finding implies that a ten-minute distribution is insufficient to trigger the distributed practice effect, though a three-hour distribution continues to be supported, consistent with Schutte et al (2015). Our second hypothesis, that the two distributed practice conditions would be superior in growth rate to a massed condition, was not supported by our data.

Though we were unsuccessful in finding support for the brief intersession interval, we were able to add to the body of literature (e.g. Duhon, House, & Stinnett, 2012; Poncy, Duhon, Lee, & Key, 2010; Rhymer & Morgan, 2002; Rhymer, Skinner, Henington, D’Reaux, & Sims, 1998; Schutte et al, 2015; Van Houten, Hill, & Parsons, 1975; Van Houten & Little, 1982; Van Houten, Morrison, Jarvis, & McDonald, 1974) that brief, intensive practice with high opportunities to respond can cause rapid improvement in skill fluency. On average across all three conditions, children grew by approximately 13 digits correct per minute from baseline to the end of the intervention. When considered in terms of instructional time, our class-wide intervention took 80 instructional minutes (20 sessions x 4 minutes of practice) over the course of the entire month – about the length of a single math period in some school buildings. Considering that recommendations for skill mastery in sums to ten fluency range from 40 digits correct per minute (Poncy & Duhon, 2017) to 100 digits correct per minute (Kubina & Yurich, 2013), an average improvement of 13 digits correct per minute is substantial even if it did not also include a low time commitment for teachers and interventionists.

The intervention described in this study was implemented at a class-wide level for an entire grade. Explicit timing can also be used to improve the learning rate of small groups of children who have fallen behind their peers in academic skills. The purpose of this study was to find support for a practical distribution of practice which could show significant student growth beyond that which is attainable through massed-style intervention. To our knowledge, no other study has sought to find the shortest effective intersession interval for math facts fluency growth rates using explicit timing. While our data did not support a ten-minute distribution of practice, it did reinforce the effectiveness of a three hour distribution. It is important to note that all three groups grew significantly from baseline – that is, using explicit timing in a massed fashion is not detrimental to the learning rate of students. However, those completing interventions in schools would be well advised to distribute the practice of their explicit timing interventions for math

facts fluency by at least three hours, and future research may demonstrate shorter effective distributions.

Limitations

Although the use of HLM to model student growth protects against low statistical power due to a low number of participants, this study could have been more sensitive to group differences between the ten-minute distributed condition and the massed condition. Further, we did not run an a priori power analysis and did not run a post-hoc power analysis due to the limited utility of post-hoc power analyses (Lenth, 2007). However, our study had a comparable number of participants to the only other study of this kind ($N = 50$ vs $N = 53$ in Schutte et al (2015)). Another limitation of this study was the failure to replicate Schutte et al (2015)'s maintenance-time of 10 calendar days. Due to student availability, we were only able to collect maintenance data six calendar days after the end of intervention, which may have made our maintenance data incomparable with Schutte et al (2015).

Another limitation that we encountered during this study was performance drift on the part on the participants. As observable in both Figures 1 and 2, student performance dropped dramatically on Day 17 and began recovering on the following days of intervention. While not assessed experimentally, it is possible that this performance drift was due to the lack of an explicit reinforcement paradigm in the intervention program. We opted to not use an overt reinforcement paradigm to maintain the purity of our independent variable in our study – we strove to determine the effects of distributed practice on explicit timing alone. The addition of additional components makes the determination of experimental effects difficult. Students completed their math work simply because they were told to – not to earn any overt tangible or praise from teachers or the experimenter. While it is interesting that students grew in the absence of an overt reinforcement paradigm, our results imply that this growth without overt

reinforcement may not last indefinitely. The literature supports the use of explicit reinforcement paradigms in all academic and behavioral interventions (e.g. Daly, Witt, Martens, & Dool, 1997; Hofstadter-Duke & Daly, 2015; O’Conner & Daly, 2018; Van Houten et al, 1975).

Finally, we noted a small, but potentially powerful average deviation from the prescribed start time observed by our research assistants. Our assistants noted an average observed start time deviation of +/- 48 seconds, while our teachers reported an average deviation of +/- 16.5 seconds. When trying to find group differences in which a distribution is as short as ten minutes, it is imperative to have sessions start exactly on time. While it is not possible to examine this effect after the fact, future studies seeking to answer questions similar to ours should be aware of the practical difficulties in implementing controlled experimental research in school settings.

Future Directions

Future studies in this area should continue to search for support of a minimally-effective intersession interval. Due to the “noise” introduced by the deviation in prescribed start times, it may be pertinent to first seek support for an intersession interval of one-hour compared to a three-hour distribution and a massed presentation. The deviations, while potentially large in effect when a distribution of practice is as short as ten minutes, may not be as potentially impactful when the short-distribution is slightly longer. Further, future replications would do well to either utilize an overt reinforcement schedule for participants or to shorten the length of the intervention from 20 days to 10 or 11 days. Our data demonstrated that all groups grew in DCPM by day ten and separation of group slopes was evidence via visual analysis by that point. The purpose of these options is to keep the task novel for participating students, who then may not experience performance drift as the students in the current study did.

Future researchers should also examine the possibility of an average rate of growth for students given this particular dose and distribution of practice in explicit timing. The average

DCPM growth of our participants for both massed (40% vs 43.8%) and distributed (62.5% vs 53%) practice was comparable to those same groups in Schutte et al (2015). This may imply the existence of an “average” student response to specific doses and distributions of explicit timing interventions.

Researchers conducting future replications may also consider the addition of a social validity measure of some sort. While explicit timing as an intervention is a relatively easy-to-implement intervention at the class-wide level, sticking to strictly-prescribed intervention start times is not practical in a school building, where schedules routinely run several minutes fast or slow.

REFERENCES

- Ardoin, S., & Daly, E. III. (2007). Introduction to the special series: Close encounters of the instructional kind – how the instructional hierarchy is shaping instructional research 30 years later. *Journal of Behavioral Education, 16*, 1-6.
- Ault, M., Wolery, M., Doyle, P., & Gast, D. (1989). Review of comparative studies in the instruction of students with moderate and severe handicaps. *Exceptional Children, 55*, 346-355.
- Bahrlick, H., Bahrlick, L., Bahrlick, A., & Bahrlick, P. (1993). Maintenance of foreign language vocabulary and the spacing effect. *Psychological Science, 4*(5), 316-321.
- Bloom, B. (1984). The search for methods of group instruction as effective as one-to-one tutoring. *Educational Leadership, 41*(8), 4-17.
- Brophy, J. (1988). Research linking teacher behavior to student achievement: Potential implications for instruction of chapter 1 students. *Educational Psychologist, 23*, 235-286.
- Brophy, J., & Evertson, C. (1976). *Learning from teaching: A developmental approach*. Boston, MA: Allyn & Bacon.

- Burns, M. (2005). Using incremental rehearsal to increase fluency of single-digit multiplication facts with children identified as learning disabled in mathematics computation. *Education and Treatment of Children, 28*, 237-249.
- Cahill, A., & Toppino, T. (1993). Young children's recognition as a function of the spacing of repetitions and the type of study and test stimuli. *Bulleting of the Psychonomic Society, 31*, 481-484.
- Carlson, N., & Birkett, M. (2017). *Physiology of Behavior*. Harlow, England: Pearson Education Ltd.
- Carpenter, S., Cepeda, N., Rohrer, D., Kang, S., & Pashler, H. (2012). Using spacing to enhance diverse forms of learning: Review of recent research and implications for instruction. *Educational Psychology Review, 24*, 369-378.
- Carr, S., & Punzo, R. (1993). The effects of self-monitoring of academic accuracy and productivity on the performance of students with behavioral disorders. *Behavioral Disorders, 18*(4), 241-250.
- Carter, S., Wong, C., & Michael, M. (2013). Enhancing foreign language competency using the cover, copy, compare technique: An exploratory evaluation. *Education & Treatment of Children, 36*(2), 105-116.
- Cepeda, N., Coburn, N., Rohrer, D., Wixted, J., Mozer, M., & Pashler, H. (2009). Optimizing distributed practice: Theoretical analysis and practical implications. *Experimental Psychology, 56*(4), 236-246.
- Cepeda, N. Pashler, H., Vul, E., Wixted, J., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin, 132*(3), 354-380.

- Cepeda, N., Vul, E., Rohrer, D., Wixted, J., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science, 19*(11), 1095-1102.
- Clark, S., & Rhymer, K. (2003). Student preference for explicit timing and interspersal procedures as a function of math problem completion rates: Testing the discrete task completion hypothesis. *Journal of Behavioral Education, 12*, 275-286.
- Clarke, B., & Shinn, M. (2002). *Tests of early numeracy measures (TEN): Administration and scoring of AIMSweb early numeracy measures for use with AIMSweb*. Eden Prairie, MN: Edformation.
- Codding, R., Chan-Lanetta, L., George, S., Ferreira, K., Volpe, R. (2011). Early number skills: Examining the effects of class-wide interventions on kindergarten performance. *School Psychology Quarterly, 26*(1), 85-96.
- Codding, R., Hilt-Panahon, A., Panahon, C., & Benson, J. (2009). Addressing mathematics computation problems: A review of simple and moderate intensity interventions. *Education and Treatment of Children, 32*(2), 279-312.
- Cooper, J., Heron, T., & Heward, W. (2007). *Applied behavior analysis*. Upper Saddle River, NJ: Pearson.
- Cornell, E. (1980). Distributed study facilitates infants' delayed recognition memory. *Memory and Cognition, 8*, 539-542.
- Daly, E., Witt, J., Martens, B., & Dool, E. (1997). A model for conducting a functional analysis of academic performance problems. *School Psychology Review, 26*, 554-574.

- Darrow, D., McLaughlin, T., Derby, M., & Johnson, K. (2012). Using cover, copy, compare spelling with and without timing for elementary students with behavior disorders. *International Electronic Journal of Elementary Education, 4*(2), 417-426.
- Dempster, F. (1987). Effects of variable encoding and spaced presentations on vocabulary learning. *Journal of Educational Psychology, 79*(2), 162-170.
- Dempster, F. (1989). Spacing effects and their implications for theory and practice. *Educational Psychology Review, 1*(4), 309-330.
- Deno, S. (2003). Curriculum-based measures: Development and perspectives. *Assessment for Effective Intervention, 28*(3), 3-12.
- Derr, T., & Shapiro, E. (1989). A behavioral evaluation of curriculum-based assessment of reading. *Journal of Psychoeducational Assessment, 7*(2), 148-160.
- Donovan, J., & Radosevich, D. (1999). A meta-analytic review of the distribution of practice effect: Now you see it, now you don't. *Journal of Applied Experimental Psychology, 84*, 795-805.
- Duhon, G., Poncy, B., Hubbard, M., Purdum, M., & Kubina, R. (2012). An examination of various intervention components with explicit timing procedures to increase math fact fluency. Symposium at the meeting of the Association of Behavior Analysis, Seattle, WA.
- Duhon, G., House, S., Hastings, K., Poncy, B., & Solomon, B. (2015). Adding immediate feedback to explicit timing: An option for enhancing treatment intensity to improve mathematics fluency. *Journal of Child and Family Studies, 24*(3), 74-87.

- Duhon, G., House, S., & Stinnett, T. (2012). Evaluating the generalization of math fact fluency gains across paper and computer performance modalities. *Journal of School Psychology, 50*, 335-345.
- Ebbinghaus, H. (1913). *Memory: A contribution to experimental psychology*. New York, NY: Teachers College, Columbia University (Original work published 1885).
- Educational Testing Service. (2019). *Graduate Record Examination*. Retrieved from: www.ets.org/gre.
- Egeland, B., & Winer, K. (1974). Teaching children to discriminate letters of the alphabet through errorless discrimination training. *Journal of Reading Behavior, 2*, 143-150.
- English, H., Wellborn, E., & Killian, C. (1934). Studies in substance memorization. *Journal of General Psychology, 11*, 233-260.
- Estes, W. (1955). Statistical theory of distributional phenomena in learning. *Psychological Review, 62*, 369-377.
- Freeman, T., & McLaughlin, T. (1984). Effects of a taped-word treatment procedure on learning disabled students' sight-word oral reading. *Learning Disability Quarterly, 7*(1), 49-54.
- Fuchs, L., Fuchs, D., Hamlett, C., & Whinnery, K. (1991). Effects of goal line feedback on level, slope, and stability of performance within curriculum-based measurement. *Learning Disabilities Research, 6*, 66-74.
- Fuchs, L., Fuchs, D., & Karns, K. (2001). Enhancing kindergarteners' mathematical development: Effects of peer assisted learning strategies. *The Elementary School Journal, 101*, 495-585.

- Gardner, R., Heward, W., & Grossi, T. (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 Applied Behavior Analysis, 27*(1), 63-71.
- Gay, L. (1973). Temporal position of reviews and its effect on the retention of mathematical rules. *Journal of Educational Psychology, 64*, 171-182.
- Glenberg, A. M. (1979). Component-levels theory of the effects of spacing of repetitions on recall and recognition. *Memory & Cognition, 7*(2), 95-112.
- Godfrey, S., Grisham-Brown, J., Schuster, J., & Hemmeter, M. (2003). The effects of three techniques on student participation with preschool children with attending problems. *Education and Treatment of Children, 26*, 255-272.
- Graham, S., MacArthur, C., & Schwartz, S. (1995). Effects of goal setting and procedural facilitation on the revising behavior and writing performance of students with writing and learning problems. *Journal of Educational Psychology, 87*, 230-240.
- Greenberg, D., & Phillips, J. (2013). Hierarchical linear modeling of growth curve trajectories using HLM. In D. Garson (Ed.), *Hierarchical Linear Modeling* (pp. 219-247). Raleigh, NC: SAGE Publications.
- Gross, T., & Duhon, G. (2013). Evaluation of computer-assisted instruction for math accuracy intervention. *Journal of Applied School Psychology, 29*(3), 246-261.
- Haring, N., Lovitt, T., Eaton, M., & Hansen, C. (1978). *The fourth R: research in the classroom*. Columbus, OH: Merrill.

- Haydon, T., Conroy, M., Sindelar, P., Scott, T., Barber, B., & Orlando, A. (2010). A comparison of three types of opportunities to respond on student academic and social behaviors. *Journal of Emotional and Behavioral Disorders, 18*, 27-40.
- Haydon, T., Marsicano, R., & Scott, T. (2013). A comparison of choral and individual responding: A review of the literature. *Preventing School Failure, 57*(4), 181-188.
- Heward, W. (1994). Three “low tech” strategies for increasing the frequency of active student response during group instruction. In R. Garder III, D. M. Sainato, J. O. Cooper, & T. E. Heron (Eds.), *Behavior analysis in education. Focus on measurably superior instruction* (p. 283-320). Monterey, CA: Brooks/Cole.
- Hier, B., & Eckert, T. (2016). Programming generality into a performance feedback writing intervention: A randomized controlled trial. *Journal of School Psychology, 56*, 111-131.
- Hochstetler, E., McLaughlin, T., Derby, M., & Kinney, M. (2013). The effects of cover, copy, compare to teach spelling to middle school students with learning disabilities and OHI. *Educational Research Quarterly, 36*(4), 25-48.
- Hoffstadter-Duke, K., & Daly, E. (2015). Identifying controlling variables for math computation fluency through experimental analysis: The interaction of stimulus control and reinforcing consequences. *Behavior Modification, 39*, 342-364.
- Janiszewski, C., Noel, H., & Sawyer, A. (2003). A meta-analysis of the spacing effect in verbal learning: Implications for research on advertising repetition and consumer memory. *Journal of Consumer Research, 30*, 138-149.
- Jost, A. (1897). The strength of associations in their dependence on the distribution of repetitions. *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane, 14*, 436-472.

- Kagan, S. (1989). The structural approach to cooperative learning. *Educational Leadership*, 47(4), 12-15.
- Kamps, D., Dugan, E., Leonard, B., & Daoust, P. (1994). Enhanced small group instruction using choral responding and student interaction for children with autism and developmental disabilities. *American Journal on Mental Retardation*, 99, 60-73.
- Kaufman, L., McLaughlin, T., Derby, M., & Waco, T. (2011). Employing reading racetracks and DI flashcards with and without covey, copy, compare and rewards to teach sight words to three students with learning disabilities in reading. *Educational Research Quarterly*, 34(4), 27-50.
- Kern-Dunlap, L., & Dunlap, G. (1989). A self-monitoring package for teaching subtraction with regrouping to students with learning disabilities. *Journal of Applied Behavior Analysis*, 22, 309-314.
- Koenig, E., Eckert, T., & Hier, B. (2016). Using performance feedback and goal setting to improve elementary students' writing fluency: A randomized controlled trial. *School Psychology Review*, 45(3), 275-295.
- Krohn, K., Skinner, C., Fuller, E., & Greear, C. (2012). Using a taped intervention to improve kindergarten students' number identification. *Journal of Applied Behavior Analysis*, 45(2), 437-441.
- Kubina, R., & Yurich, K. (2013). *The Precision Teaching Book*. Athens, Greece: Greatness Achieved Publishing.
- Küpper-Tetzl, C. (2014). Understanding the distributed practice effect: Strong effects on weak theoretical grounds. *Zeitschrift für Psychologie*, 222(2), 71-81.

- Lee, C., & Genovese, E. (1988). Distribution of practice in motor skill acquisition: Learning and performance effects considered. *Research Quarterly for Exercise and Sport*, 59, 277-287.
- Lenth, R. (2007). Post hoc power: Tables and commentary. University of Iowa: Department of Statistics and Actuarial Sciences.
- Lyon, D. (2014). The relation of length of material to time taken for learning and the optimum distribution of time. *Journal of Educational Psychology*, 5, 85-91.
- Maag, J., Reid, R., & DiGangi, S. (1993). Differential effects of self-monitoring attention, accuracy, and productivity. *Journal of Applied Behavior Analysis*, 26, 329-344.
- MacQuarrie-Klender, L., Tucker, J., Burns, M., & Hartmann, B. (2002). Comparison of retention rates using traditional, drill sandwich, and incremental rehearsal flash card methods. *School Psychology Review*, 30, 23-32.
- Maheady, G., Mallette, B., Harper, G., & Sacca, K. (1991). Heads together: A peer-mediated option for improving the academic achievement of heterogeneous learning groups. *Remedial and Special Education*, 12(2), 25-33.
- Mazure, J. (2017). *Learning and behavior*. New York, NY: Routledge.
- McCallum, E., Skinner, C., & Hutchins, H. (2004). The taped-problems intervention. *Journal of Applied School Psychology*, 20(2), 129-147.
- Miller, A., Hall, S., & Heward, W. (1995). Effects of sequential 1-minute time trials with and without inter-trial feedback and self-correction on general and special education students' fluency with math facts. *Journal of Behavioral Education*, 5(3), 319-345.

- Moss, V. (1996). The efficacy of massed versus distributed practice as a function of desired learning outcomes and grade level of the student. *Dissertation Abstracts International*, 56, 5204.
- Narayan, J., Heward, W., Gardner, R., Courson, F., & Omness, C. (1990). Using response cards to increase student participation in an elementary classroom. *Journal of Applied Behavior Analysis*, 23(4), 483-490.
- Nist, L., & Joseph, L. (2008). Effectiveness and efficiency of flashcard drill instructional methods on urban first-graders' word recognition, maintenance, and generalization. *School Psychology Review*, 37(3), 294-308.
- O'Connor, M., & Daly, E. (2018). Selecting effective intervention strategies for escape-maintained academic-performance problems: Consider giving 'em a break! *Journal of School Psychology*, 66, 41-53.
- Pavlik, P., & Anderson, J. (2005). Practice and forgetting effects on vocabulary memory: An activation-based model of the spacing effect. *Cognitive Science*, 29(4), 559-586.
- Peterson, H., Ellis, M., Toohill, N., & Kloess, P. (1935). Some measurements of the effects of reviews. *Journal of Educational Psychology*, 26, 65-72.
- Piaget, J. (1957). *Construction of reality in the child*. London: Routledge & Kegan Paul.
- Pipkin, C., Winters, S., & Diller, J. (2007). Effects of instruction, goals, and reinforcement on academic behavior: Assessing skill versus reinforcement deficits. *Journal of Early and Intensive Behavior Intervention*, 4(4), 648-657.
- Poncy, B. & Duhon, G. (2017). *Measures & Interventions for Numeracy Development*, Retrieved from <http://brianponcy.wixsite.com/mind>

- Poncy, B., Duhon, G., Lee, S., & Key, A. (2010). Evaluation of techniques to promote generalization with basic math skills. *Journal of Behavioral Education, 19*, 76-92.
- Poncy, B., Jaspers, K., Hansmann, P., Bui, L., & Matthew, W. (2015). A comparison of taped-problem interventions to increase math fact fluency: Does the length of time delay affect student learning rates? *Journal of Applied School Psychology, 31*(1), 63-82.
- Protection of Human Subjects, 45 CFR 46, §46.104 (2018).
- Purdum, M. (2014). Evaluating the effect of item set size on student learning rates with multiplication facts (Doctoral dissertation). Retrieved from UMI Dissertation Publishing. 3629921.
- Raaijmakers, J. (2003). Spacing and repetition effects in human memory: Application of the SAM model. *Cognitive Science, 27*, 431-452.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, CA: Sage Publications.
- Rea, C., & Modigliani, V. (1987). The spacing effect in 4- to 9-year-old children. *Memory and Cognition, 15*(5), 436-443.
- Reynolds, J., & Glaser, R. (1964). Effects of repetition and spaced review upon retention of a complex learning task. *Journal of Educational Psychology, 55*(5), 297-308.
- Rhymer, K., & Morgan, S. (2005). Comparison of explicit timing and interspersal interventions: Analysis of completion rates, student preference, and teacher acceptability. *Journal of Behavioral Education, 14*(4), 283-303.

- Rhymer, K., Skinner, C., Henington, C., D'Reaux, R., & Sims, S. (1998). Effects of explicit timing on mathematics completion rates in African-American third-grade elementary students. *Journal of Applied Behavior Analysis, 31*(4), 673-677.
- Rhymer, K., Skinner, C., Jackson, S., McNeill, S., Smith, T., & Jackson, B. (2002). The 1-minute explicit timing intervention: The influence of mathematics problem difficulty. *Journal of Instructional Psychology, 29*(4), 305-311.
- Roediger, H., & Pyc, M. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition, 1*, 242-248.
- Rohrer, D., & Taylor, K. (2006). The effects of overlearning and distributed practice on the retention of mathematics knowledge. *Applied Cognitive Psychology, 20*(9), 1209-1224.
- Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science, 35*(6), 481-498.
- Ruch, T. (1928). Factors influencing the relative economy of massed and distributed practice in learning. *Psychological Review, 35*, 19-45.
- Schutte, G., Duhon, G., Solomon, B., Poncy, B., Moore, K., & Story, B. (2015). A comparative analysis of massed vs. distributed practice on basic math fact fluency growth rates. *Journal of School Psychology, 53*(2), 149-159.
- Seabrook, R., Brown, G., & Solity, J. (2005). Distributed and massed practice: From laboratory to classroom. *Applied Cognitive Psychology, 19*(1), 107-122.
- Schunk, D., & Schwartz, C. (1993). Goals and progress feedback: Effects on self-efficacy and writing achievement. *Contemporary Educational Psychology, 18*, 337-354.

- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth, Cengage Learning.
- Shin, T. (2007). Comparison of three growth modeling techniques in the multilevel analysis of longitudinal academic achievement scores: Latent growth modeling, hierarchical linear modeling, and longitudinal profile analysis via multidimensional scaling. *Asia Pacific Education Review*, 8(2), 262-275.
- Skinner, C., & Belfiore, P. (1992). Cover, copy, compare: Increasing geography accuracy in students with behavior disorders. *School Psychology Review*, 21(1), 73-82.
- Skinner, C., Belfiore, P., Mace, H., Williams-Wilson, S., & Johns, G. (1997). Altering response topography to increase response efficiency and learning rates. *School Psychology Quarterly*, 12(1), 54-64.
- Skinner, C., Belfiore, P., & Watons, T. (2002). Assessing the relative effects of interventions in students with mild disabilities: Assessing instructional time. *Journal of Psychoeducational Assessment*, 20(4), 346-357.
- Skinner, C., Fletcher, P., & Henington, C. (1996). Increasing learning rates by increasing student response rates: A summary of research. *School Psychology Quarterly*, 11(4), 313-325.
- Skinner, C., Ford, J., & Yunker, B. (1991). An analysis of instructional response requirements on the multiplication performance of behaviorally disordered students. *Behavioral Disorders*, 17(1), 56-65.
- Skinner, C., Hurst, K., Teeple, D., & Meadows, S. (2002). Increasing on-task behavior during mathematics seatwork in students with emotional disturbance by interspersing additional brief problems. *Psychology in the Schools*, 39, 647-659.

- Skinner, C., McCleary, D., Skolits, G., Poncy, C., & Cates, G. (2013). Emerging opportunities for school psychologists to enhance our remediation procedure evidence base as we apply response to intervention. *Psychology in the Schools, 50*(3), 272-289.
- Smith, T., Dittmer, K., & Skinner, C. (2002). Enhancing science performance in students with learning disabilities using cover, copy, compare: A student shows the way. *Psychology in the Schools, 39*(4), 417-426.
- Sobel, H., Cepeda, N., & Kapler, I. (2011). Spacing effects in real-word classroom vocabulary learning. *Applied Cognitive Psychology, 25*(5), 763-767.
- Sones, A., & Shroud, J. (1940). Review with special references to temporal position. *Journal of Educational Psychology, 31*, 665-676.
- Thorndike, E. (1912). The curve of work. *The Psychological Review, 19*(2), 165-194.
- Van Houten, R., Hill, S., & Parsons, M. (1975). An analysis of a performance feedback system: The effects of timing and feedback, public posting, and praise upon academic performance and peer interaction. *Journal of Applied Behavior Analysis, 8*(4), 449-457.
- Van Houten, R., & Little, G. (1982). Increased response rate in special education children following an abrupt reduction in the time limit in the absence of a token economy. *Education and Treatment of Children, 5*(1), 23-32.
- Van Houten, R., Morrison, E., Jarvis, R., & McDonald, M. (1974). The effects of explicit timing and feedback on compositional response rate in elementary school children. *Journal of Applied Behavior Analysis, 7*, 547-555.

- Van Houten, R., & Rolider, A. (1990). The use of color mediation techniques to teach number identifications and single digit multiplication problems to children with learning problems. *Education and Treatment of Children, 13*(3), 216-225.
- Van Houten, R., & Thompson, C. (1976). The effects of explicit timing on math performance. *Journal of Applied Behavior Analysis, 9*(2), 227-230.
- Vlach, H., & Sandhofer, C. (2012). Distributed learning over time: The spacing effect in children's acquisition and generalization of science concepts. *Child Development, 83*(4), 1137-1144.
- Wahlheim, C., Maddox, G., & Jacoby, L. (2014). The role of reminding in the effects of spaced repetitions on cued recall: Sufficient but not necessary. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*, 94-105.

APPENDICES

APPENDIX B

NAME: _____ DATE: _____



www.AtoZTeacherStuff.com

A V H U S L F I D T M A P D P
C O A Q R T C X S K K S O I I
Q O U L F C M O Q I M A J N L
I A N T Q D C R M M Z M S S G
G N G T Q V J Q J P O P P P R
C D S L R P V F S P L L B E I
E G S T R A P H X Y W E U C M
S I U H A G S S S L M E T T N
C V R J U N D T D P O A A E T
B T P V R N T K W P N D W B W
J I R T P J D E P W S P X C P
M X I J Z L O R C Q T D R J A
T H S L N O R P E D E N J T R
S W E H D P J P T D R I Q W W
V Z U I P J X D T W H U B U I

MONSTER
SURPRISE
HUNDRED
COMPLETE
SAMPLE
INSTANT
INSPECT
PILGRIM
CONTRAST

APPENDIX C

[Date]

Dear Parent,

This spring, the 3rd grade class at [school] will be working on their math facts fluency. Fluency is the speed at which students can complete simple math problems, and high fluency is associated with better long-term outcomes in math. Mr. Steven Powell, a doctoral student in school psychology from Oklahoma State University is studying this topic for his dissertation and will be overseeing the implementation of the math facts fluency practice. He is being supervised by Dr. Gary Duhon, school psychology faculty at Oklahoma State University. As part of this study, your child will be assigned to one of three groups, and will practice their math at varying times of the day. Each student in [grade] will receive the same amount of math practice and the practice time will be built into normal instructional time and be in addition to any mathematics support your child is already receiving. After the study is complete, Mr. Powell and his research assistants will remove any identifying information from each student's data before using the results for his dissertation. If you have questions or concerns regarding the study, please contact Mr. Powell or Dr. Duhon at steve.powell@okstate.edu or gary.duhon@okstate.edu, respectively.

Sincerely,

[Principal/Administrator of School]

APPENDIX D

Explicit Timing Protocol

1. Pass out folders.
 2. *“This morning we’re going to do some math problems and word searches.”*
 3. *“Please open your folders and take out your packet of worksheets.”*
 4. *“When I say begin, start with the first problem and work across the row. When you finish a row go on to the next one. If you come to one you don’t know, you may skip it, but try to work all of the problems. You will have one minute to complete as many problems as you can. Please do your best and work as quickly as you can.”*
 5. *“Are there any questions?”*
 6. *“Ready? Begin.”* (Time for 1 minute)
 7. (After 1 minute) *“Stop. Please put your pencils down.”*
 8. *“Please turn to the second page of your packet.”* (wait for students to all have turned their pages)
 9. Repeat steps 6-7.
 10. *“Please turn to the third page of your packet. Some of you will have word searches, and some of you will have math problems. Remember, everyone will get the same number of word searches and math problems each day. If you have a word search, I want you to find as many words as you can in one minute. If you have a math worksheet, keep completing the problems as quickly as you can.”*
 11. Repeat steps 6-7.
 12. *“Please turn to the fourth page of your packet.”*
- Repeat steps 6-7.

APPENDIX E

Procedural Fidelity Form

Date: _____ Start Time: _____ End Time: _____

Observer: _____ Administrator: _____ Teacher: _____

#	Step	Complete?
1	Pass out folders.	
2	<i>“This morning we’re going to do some math problems and word searches.”</i>	
3	<i>“Please open your folders and take out your packet of worksheets.”</i>	
4	<i>“When I say begin, start with the first problem and work across the row. When you finish a row go on to the next one. If you come to one you don’t know, you may skip it, but try to work all of the problems. You will have one minute to complete as many problems as you can. Please do your best and work as quickly as you can.”</i>	
5	<i>“Are there any questions?”</i>	
6	<i>“Ready? Begin.”</i> (Time for 1 minute)	
7	(After 1 minute) <i>“Stop. Please put your pencils down.”</i>	
8	<i>“Please turn to the second page of your packet.”</i> (wait for students to all have turned their pages)	
9	Repeat steps 6-7.	
10	<i>“Please turn to the third page of your packet. Some of you will have word searches, and some of you will have math problems. Remember, everyone will get the same number of word searches and math problems each day. If you have a word search, I want you to find as many words as you can in one minute. If you have a math worksheet, keep completing the problems as quickly as you can.”</i>	
11	Repeat steps 6-7.	
12	<i>“Please turn to the fourth page of your packet.”</i>	
13	Repeat steps 6-7.	
		_____/13

Notes:

APPENDIX F

Teacher Fidelity Form

Please write the start time for each intervention session. Fidelity information from you as teachers is important to the integrity of the project. Because this project is examining the timing of intervention sessions, it is *vital*ly important that intervention sessions occur at the same times each day. Thank you so much for helping me with this project!

Week 1

	Monday	Tuesday	Wednesday	Thursday	Friday
9:10 Session					
9:25 Session					
12:50 Session					

Week 2

	Monday	Tuesday	Wednesday	Thursday	Friday
9:10 Session					
9:25 Session					
12:50 Session					

Week 3

	Monday	Tuesday	Wednesday	Thursday	Friday
9:10 Session					
9:25 Session					
12:50 Session					

Week 4

	Monday	Tuesday	Wednesday	Thursday	Friday
9:10 Session					
9:25 Session					
12:50 Session					

VITA

Steven L. Powell

Candidate for the Degree of

Doctor of Philosophy

Dissertation: A COMPARATIVE ANALYSIS OF MATH FACTS FLUENCY
GAINS MADE THROUGH MASSED AND DISTRIBUTED
PRACTICE WITH VARIED INTER-SESSION INTERVALS

Major Field: School Psychology

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in School Psychology at Oklahoma State University, Stillwater, Oklahoma in May, 2021.

Completed the requirements for the Educational Specialist in School Psychology at James Madison University, Harrisonburg, Virginia in 2017.

Completed the requirements for the Master of Arts in School Psychology at James Madison University, Harrisonburg, Virginia in 2015.

Completed the requirements for the Bachelor of Arts in Psychology at Hood College, Frederick, Maryland in 2014.

Experience:

Graduate Teaching Assistant – Oklahoma State University
Clinic Practicum – OSU School Psychology Clinic
MTSS Site Coach – Osage County Interlocal Cooperative
School Psychologist Intern, Halifax Co. Public Schools, Virginia
Teaching Assistant – James Madison University

Professional Memberships:

NASP, APA, OSPA, VASP, BACB