COMPARING THE EFFECTS OF DRILL-BASED INTERVENTIONS ON MULTIPLICATION FACT ACQUISITION

A DISSERTATION

SUBMITTED TO THE GRADUATE SCHOOL

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

BY

SARAH RUTH ADAMS, M.A.

DR. KATHRIN E. MAKI, ADVISOR

DR. JANAY B. SANDER, CO-ADVISOR

BALL STATE UNIVERSITY

MUNCIE, INDIANA

JULY 2021

ACKNOWLEDGMENTS

Graduate school has been one of the most challenging but rewarding educational opportunities of my academic career. Needless to say, my success throughout this experience would not have been possible without the support of my family, friends and faculty to whom I owe many thanks.

To my advisor, Katie Maki, who nurtured my appreciation and passion for research. Much of my success and growth not only in research but throughout my graduate work is attributed to your constant guidance and encouragement. Thank you, Dr. Maki, for your continuous support and dedication to my learning and growth, even from afar. Also, thank you to my co-adviser, Janay Sander, as well as my other committee members, Theresa Kruczek and Angela Stefanski for your unwavering support as part of my committee but also in your many other roles as my professors and/or supervisors.

To my cohort and program friends, thank you for all the laughs and support throughout this journey. I can’t imagine my graduate school experience without each and every one of you. Caty, thank you for being my “work husband” these past four years and for being my biggest cheerleader. Jamie, thank you for your unwavering loyalty and unconditional friendship. Eric, I am thankful for our long phone calls and to have you as the papa bear to my mama bear. Asia, thank you for taking on new paths with me and being my partner in crime in the clinic. Amanda, thank you for always being a smiling face and source of support. Dominik, thank you for your friendship and encouragement. An additional thanks to Jamie and Jenna, without whom data collection would not have been possible. Each one of you along with many others have made a lasting impression on me.

Thank you to my family and friends who supported and encouraged me throughout graduate school. Mom, Dad, Emily, and Kevin, this would not have been possible without your support and unconditional love. The four of you have played a pivotal role not only these last four years but throughout my life journey. Kathy and Roger, thank you for the endless supply of leftovers as well as all your endless encouragement and love. Maureen, Connie, Annie, Hayley, and Ellie, thank you for sticking by me while I pursued this dream of mine. To all of my many other friends and family, which there are too many to name, thank you for all your support and love.

Finally, thank you to Bryce (and Zeke) for all your love and support. Bryce, I’m not sure if you quite knew what you were getting into when you joined me on this journey, but you have always believed in me even when I didn’t believe in myself. You truly have been my warrior in the face of the “grad school scaries.” I love you.

DEDICATION

To my parents, who somehow managed to survive my strong-willed, stubborn, and sassy nature to shape me into the woman I am today. Thank you for your unconditional love, support, and encouragement throughout all my life endeavors. All of my work is for you and I hope to one day make you proud. To my goddaughter, Addison, never be afraid to chase your passions and dreams, no matter how big they may be.

ABSTRACT

Math facts are foundational for later math skill development but have been identified as a common skill deficit among students in early schooling (Gersten, Jordan, & Flojo, 2005). Drill-based interventions effectively target and improve foundational academic skills (Burns, 2005), however, there is limited research examining drill-based interventions within the context of math facts and also a gap in the use of theoretical frameworks to guide such research. The purpose of this study was to compare the effects of incremental rehearsal (IR), traditional drill (TD), and strategic incremental rehearsal (SIR) on multiplication fact retention, maintenance, and fluency outcomes with 36 fourth and fifth graders. The study also examined intervention efficiency and treatment acceptability. Additionally, the instructional hierarchy was used to inform intervention procedures as well as discuss outcomes. Results showed no differences in retention ($p = .710$; $W = .01$), maintenance ($p = .231$; $W = .04$), and treatment acceptability ($p = .348$; $W = .03$) across the interventions. There were significant differences between pre-test and post-test fluency for IR ($p < .0125$; $\eta_p^2 = .14$) but no differences in fluency among IR and TD ($p = 1.000$), IR and SIR ($p = .966$), as well as TD and SIR ($p = 1.000$). Furthermore, TD was shown to be most efficient intervention overall ($p < .05$; $W = .39$). The findings from the present study were discussed within the context of previous research as well as implications for practice, future research, limitations and conclusions.

Table of Contents

ACKNOWLEDGMENTS i

DEDICATION iii

ABSTRACT iv

TABLE OF CONTENTS v

CHAPTER 1: INTRODUCTION 1

INTRODUCTION 1

STATEMENT OF PROBLEM 5

SIGNIFICANCE OF THE STUDY 6

DEFINITIONS 8

CHAPTER 2: LITERATURE REVIEW 12

LITERATURE REVIEW INTRODUCTION 12

ORGANIZATION OF THE CHAPTER 12

SYSTEMS-LEVEL FRAMEWORKS 13

Multi-tiered Systems of Support (MTSS) 13

Response to Intervention (RtI) 14

Tier 1 16

Tier 2 20

Tier 3 23

Outcomes and Challenges of RtI 27

Summary 28

THEORETICAL FRAMEWORKS 29

Instructional Hierarchy (IH) 32
Acquisition stage 33
Proficiency stage 35
Generalization stage 38
Adaption stage 41
Instructional Hierarchy (IH) and Academic Interventions 42
Summary 43
ACADEMIC INTERVENTIONS 44
Acquisition Interventions 45
Acquisition Interventions for Math 46
Drill-based Interventions and Math 50
Incremental Rehearsal (IR) 52
Traditional Drill (TD) 56
Strategic Incremental Rehearsal (SIR) 59
Treatment Acceptability 61
Summary 64
LITERATURE REVIEW SUMMARY 64
RESEARCH QUESTIONS 66
CHAPTER 3: METHOD 68
PARTICIPANTS AND SETTING 68
MEASURES 69
Known and unknown multiplication facts 69
Fluency 70

Retention 70
Maintenance 71
Efficiency 71
Treatment acceptability 71
MATERIALS 72
INTERVENTIONS 73
Incremental rehearsal (IR) 74
Traditional drill (TD) 75
Strategic incremental rehearsal (SIR) 76
PROCEDURES 77
Implementation Fidelity 79
RESEARCH DESIGN AND ANALYSES 80
CHAPTER 4: RESULTS 83
PRELIMINARY ANALYSES 83
ASSUMPTIONS OF ANALYSES 83
INTERVENTION EFFECTIVENESS 84
Multiplication fact retention 84
Multiplication fact fluency (DCPM) 85
Multiplication fact maintenance 87
INTERVENTION EFFICIENCY 88
TREATMENT ACCEPTABILITY 89
CHAPTER 5: DISCUSSION 94
INTERVENTION EFFECTIVENESS 94

Multiplication fact retention 94
Multiplication fact fluency (DCPM)....................96
Multiplication fact maintenance 97
Research procedures and effectiveness....................99
INSTRUCTIONAL HIERARCHY (IH) AND INTERVENTION EFFECTIVENESS....................99
INTERVENTION EFFICIENCY 101
TREATMENT ACCEPTABILITY 104
IMPLICATIONS FOR PRACTICE 105
LIMITATIONS AND IMPLICATIONS FOR FUTURE RESEARCH....................107
CONCLUSION 109
REFERENCES 112

List of Tables

Table 1: *Definitions* ...8

Table 2: *Study Procedure* ...79

Table 3: *Study Analyses* ...82

Table 4: *Descriptive and Inferential Statistics for Whole Sample* ...91

Table 5: *Descriptive and Inferential Statistics by Racial Groups* ...92

List of Figures

Figure 1: Instructional Hierarchy (IH) Overview 33

Figure 2: Incremental Rehearsal presentation sequence 75

Figure 3: Traditional Drill presentation sequence 76

CHAPTER 1

Introduction

According to the National Assessment of Educational Progress (NAEP), in 2017, approximately 40% of fourth grade students demonstrated proficiency in mathematics (Nations Report Card, 2017). Furthermore, the number of students demonstrating even basic level performance (i.e., demonstrating some understanding of NAEP's five areas of math, but not full proficiency) decreased from 42% to 39% (Nations Report Card, 2017). The large proportion of students in the United States who did not demonstrate proficiency in math (60%) underscores the need to systematically implement interventions targeting math skills in schools. Targeting foundational math skills through such interventions will also promote over math skill development throughout schooling.

Proficiency in early math skills (e.g., math facts) has been shown to predict later math skills (e.g., algebra; Duncan et al., 2007), and yet, math fact knowledge has been identified as a common skill weakness among students who have math difficulties during early schooling (Gersten, Jordan, & Flojo, 2005). Students who struggle to master math facts, specifically, have been shown to demonstrate difficulty understanding more complex math concepts, like algebra, later in formal schooling (Gersten et al., 2005). Such foundational math skills like math facts, therefore, represent an important target for evidence-based math interventions. However, the majority of the early intervention research has focused on early reading interventions as opposed to early math interventions (Methe et al., 2011). Given that math fact acquisition is a common area of math difficulty, targeting these skills through math intervention holds promise in addressing the low national proficiency levels. Such interventions are commonly implemented within the context of system-level frameworks.

Multi-tiered frameworks, such as Response to Intervention (RtI), are a means of providing academic supports to students demonstrating academic needs. RtI serves as a universal framework to provide instructional support to all students and intervention to students experiencing academic difficulties. It is a systems-level framework that uses school-wide screening measures to identify struggling students, and in response, implement evidence-based interventions in the students' area of academic difficulty. The goal of these academic supports is to increase students' academic skills, and consequently, decrease the need for educational supports. Thus, providing early, evidence-based, intervention not only promotes positive student academic outcomes but also reduces the number of students with academic difficulties. RtI may also be a cost-effective strategy given the support for effective resource allocation within the framework (Burns & Gibbons, 2012). In fact, implementation of RtI led to a reduction in time spent conducting special education evaluations and a reduction in costs to the school district (VanDerHeyden, Witt, & Gilbertson, 2007). Early intervention within an RtI context is of primary importance to support students' academic outcomes as well as broader life outcomes.

Early academic interventions play an important role in promoting positive outcomes for students both in and outside of school. For example, students experiencing academic difficulty are more likely to engage in problem behaviors such as drug and alcohol use (Wang & Fredricks, 2014) as well as dropping out of school altogether (Archambault, Janosz, Fallu & Pag, 2009; Janosz, Archambault, Morizot & Pagani, 2008; Fan & Wolters, 2014). In contrast, when students experienced academic success, specifically in math, there was a reduction in negative school outcomes (i.e., school dropout), which resulted in better vocational preparation, higher rates of employment and social contributions (e.g., volunteering within the community; Campbell & Ramey, 1995), fewer law-breaking instances (Berrueta-Clement, Schweinhart, Barnett, Epstein,

& Weikart, 1984), and increased independent living as adults (Tolar, Lederberg, & Fletcher, 2009). Given these positive outcomes associated with academic success and math proficiency, it is important to target students' math difficulties early through empirically supported interventions.

Despite the evident need for early math intervention, there is relatively little research examining such interventions (Lembke, Hampton, & Beyers, 2012; Hughes & Dexter, 2011; Methe et al., 2011). Moreover, the majority of academic intervention research has been related to reading (e.g., phonics, sight words). The research that has been conducted on acquisition-based math interventions has largely focused on broad instructional strategies (e.g., concrete-representational-abstract; Agrawal & Morin, 2016). Additionally, there is limited research examining interventions addressing students' math fact acquisition, specifically (e.g., drill-based interventions; Burns, 2005; Burns, Zaslofsky, Maki, & Kwong, 2016; McVancel, Missall, & Bruhn, 2018).

Drill-based interventions are one intervention approach that have demonstrated effectiveness for increasing basic skill acquisition (Klingbeil, Moeyaert, Archer, Chimboza, & Zwolski Jr., 2017; Volpe, Mulé, Briesch, Joseph, & Burns, 2011), and yet, most of the research examining such drill-based interventions have targeted reading skills (Burns et al., 2010). The relatively heavy focus on reading research is a significant problem given that acquisition of math facts is linked to proficiency in advanced math skills in later schooling. Additionally, research examining drill-based math interventions has largely used single-case research designs, which limits the generalizability of these study findings.

Currently, two drill-based interventions which have demonstrated effectiveness in improving students' math fact acquisition include incremental rehearsal (IR) and traditional drill

(TD; Burns, 2005). Although there is preliminary evidence supporting IR and TD in effectively improving students' math facts skills, these interventions have largely been examined within the context of reading skills (Burns, 2005). Recently, another drill-based intervention, strategic incremental rehearsal (SIR), has demonstrated effectiveness for the acquisition of basic reading skills (i.e., sight words) with two studies finding slightly greater effectiveness and efficiency of SIR when compared to IR (Kupzyk, Daly III, & Andersen, 2011; January, Lovelace, Foster, & Ardoin, 2017). As stated, SIR has solely been examined within the context of reading (Volpe, Mulé et al., 2011; January et al., 2017); however, given the preliminary evidence for IR and TD related to math fact acquisition, SIR may also hold promise for application to math facts because it combines components of these two effective instructional methods. Thus, the present study aims to further expand the current math intervention research base by comparing novel and familiar interventions as well as address other gaps in the math intervention literature.

Another limitation of the math intervention research base is the limited use of theoretical frameworks to guide implementation and interpret results of empirical studies (Burns, 2011). Theoretical frameworks are theory-based ideas or concepts that attempt to explain phenomena or guide understanding of a concept (Ellis, 2014) such as why certain academic interventions demonstrate effectiveness. Using theory to guide research is important because it increases the internal validity of the intervention (Volpe & Suldo, 2014) and allows for adaptation of interventions to specific settings (Burns, 2011) such as schools. Some of the theoretical frameworks identified within the intervention literature broadly include ecological, prevention, social cognitive, cognitive-behavioral, cognitive and behavioral frameworks (Mercer, Idler, Bartfai, 2014). Of these frameworks, cognitive and behavioral frameworks have been most commonly used to guide interventions targeting academic skills (Mercer et al., 2014; Axtell, McCallum, Bell, & Poncy, 2009). The Instructional Hierarchy (IH) is one framework which has been applied within the context of academic interventions, predominately related to reading interventions (Codding & Martin, 2016). Even though the IH has been applied within reading intervention research, the use of such frameworks within the context of math interventions is inconsistent and scarce (Burns, 2011). When theoretical frameworks, such as the IH, have been incorporated within math intervention research, they have largely either been referenced to broadly discuss instructional features of an intervention or in explaining research findings (Codding & Martin, 2016; Burns, 2011) but not to both guide intervention implementation and interpret research findings. The absence of such theoretical frameworks limits the generalizability of research findings as well as limits understanding of why interventions are effective. Therefore, research related to the application of theoretical frameworks, such as the IH, to guide and explain math interventions is also warranted.

Statement of problem

There are a limited number of students who demonstrate proficiency in math (Nations Report Card, 2017), and yet, much of the academic intervention research has focused on targeting students' reading difficulties. Math fact acquisition is a common area of difficulty among students experiencing math difficulties in elementary school (Gersten et al., 2005), but research examining interventions targeting foundational math skills, specifically math facts, is limited (Burns, 2005). Therefore, additional research examining math fact acquisition interventions is needed in an effort to improve such students' math skills and in turn increase the number of students demonstrating proficiency in math. Drill-based intervention research for math facts is lacking compared to the application of such interventions to other academic areas, such as reading (Burns, 2011). Additionally, drill-based math interventions for math facts have most commonly been examined through single-case research designs, which limits the generalizability of previous study findings. Therefore, the use of new research designs (i.e., group designs) for drill-based interventions related to math facts is needed to not only generalize findings from such studies but also to expand the math fact intervention research base. Moreover, math intervention research has inconsistently used theory to guide intervention implementation and explain research findings. In instances when theory has been incorporated it has either been used to explain the effectiveness of specific instructional features (e.g., modeling; Benner, Stage, Nelson, & Ralston, 2011; Burns, 2005), or has been used to retroactively study the results (Burns et al., 2011). Thus, future math intervention research should use theoretical frameworks to guide intervention implementation and interpret findings in an effort to better understand *why* interventions are effective. The present study aims to address these limitations and will be discussed in further detail.

Significance of the study

Given that only 40% of fourth-grade students demonstrated proficiency in math (Nations Report Card, 2017), there is a need for greater implementation of evidence-based math interventions in schools. However, in order to ensure implementation of appropriately targeted math interventions, there is also a need for additional research examining early math interventions. Currently, there is a relative dearth of research examining early math interventions compared to early reading interventions, particularly within the context of prevention models such as RtI (Bouck & Cosby, 2017). The present study, therefore, will examine the effectiveness of math fact interventions with elementary aged students demonstrating math difficulties and a need for a tiered math intervention. Additionally, the Instructional Hierarchy (IH; Haring & Eaton, 1978) will serve as the theoretical framework to inform instructional features of the interventions examined as well as to ensure the student demonstrates need for an acquisition base intervention targeting multiplication facts. The present study will expand the use of a theoretically based approach to guide the implementation and evaluation of math fact interventions.

Flashcard or drill interventions are frequently implemented to target difficulty with math facts (Burns, 2005). TD is one technique that has been used to teach students math facts (Burns, 2004). This method involves repeated practice of facts, all of which are unknown (i.e., the student has not demonstrated mastery of the math facts). IR is another flashcard intervention used to rehearse basic skills such as math facts or letter names (Burns, 2005; Joseph, 2006; Codding, Archer, & Connell, 2010; Burns, Zaslofsky, Kanive, & Parker, 2012; Zaslofsky, Scholin, Burns, & Varma, 2016). This procedure uses a spaced practice approach in that one unknown fact is interspersed with eight known facts, which allows for increased opportunities to respond to (or practice) the unknown fact (OTR; McVancel et al., 2018). Like TD, IR has been shown to increase math fact retention (Burns, 2005; Burns et al., 2016). Although there is some research supporting the effectiveness of TD and IR to increase math fact acquisition and retention, the majority of the research has examined their use for basic reading skills (e.g., sight words; Burns, 2005; Daly & McCurdy, 2002). Finally, SIR (January et al., 2017; Kupzyk et al., 2011) is a modified drill rehearsal approach that uses a procedure similar to traditional IR but differs in that only unknown items are practiced, like in TD. SIR was shown to be slightly more effective for sight word acquisition and retention (January et al., 2017; Kupzyk et al., 2011); however, there is no known research that has compared SIR with other drill methods to target math fact acquisition. Research examining the effectiveness of all three drill interventions for math facts is therefore needed.

Given that a large number of students do not demonstrate proficiency in math (Nations Report Card, 2017), implementation of empirically supported math interventions is needed. An important consideration, however, is how to appropriately target math interventions to students' needs based on where students' skills fall within the math skill hierarchy. Given that early math difficulties have been linked to inadequate math fact acquisition, math facts represent an important skill to target through intervention. As discussed previously, drill-based interventions have demonstrated effectiveness in improving students' math fact acquisition (Burns et al., 2010). However, research examining drill-based intervention has largely examined basic reading skills, leaving a significant gap in the research base for drill-based math fact interventions. This discrepancy highlights the need for further research examining drill-based interventions for math. The goal of the present study, therefore, is to examine drill-based interventions for math facts. The present study aims to examine the comparative effects of IR, TD, and SIR on students' multiplication fact retention, multiplication fact fluency, and maintenance of multiplication facts taught. Additionally, intervention efficiency and differences across students' acceptability and perception of intervention effectiveness will be examined. A detailed outline of the research questions can be found at the end of chapter 2.

Definitions

Table 1 includes definitions of important terms used throughout the current literature review and present study.

Table 1.

Definitions of Terms.

Term	Definition
Acquisition	Acquisition is the first stage of the Instructional Hierarchy (IH). Students who fall in the acquisition stage are described as slow and inaccurate in using the skill (Haring & Eaton, 1978; Cates &

	Rhymer, 2003). The instructional focus of this stage is to improve the students' accuracy in responding.
Adaption	Adaption is the final stage of the Instructional Hierarchy (IH). At this stage, the instructional focus is to improve the students' ability to adapt or modify the skill they have learned to novel situations and to solve problems using the skill (Haring & Eaton, 1978).
Generalization	Generalization is the third stage of the Instructional Hierarchy (IH). At this stage, the student is able to proficiently perform the skill but is not yet able to apply the skill to a novel situation. During this stage, the instructional focus is improving the use of the skill across other situations outside of the practice procedures (Haring & Eaton, 1978).
Efficiency	Intervention efficiency is calculated by dividing the number of facts retained by total intervention time, in seconds. This number is then multiplied by 60 to determine the rate of retention per minute (i.e., $\frac{\text{Session Facts Retained}}{\text{Intervention Session Time}} X\ 60$; Cates, Skinner, Watson, Meadows, Weaver, & Jackson, 2003). Efficiency is used to identify and compare how quickly students acquire academic skills with the rate being per instructional minute.
Incremental Rehearsal (IR)	Incremental rehearsal is a drill-based intervention that intersperses unknown material with known material (Burns, 2005; Tucker, 1989). The known and unknown items are presented in a ratio of 10% unknown and 90% known (Burns, 2005).
Individuals with Disabilities Education Act (IDEA)	IDEA is a federal law which protects the educational rights of students with disabilities and their parents. This law requires all schools to provide a free and appropriate education (FAPE) to students with disabilities. Schools must evaluate students who are suspected of having a disability and provide services for students who qualify for special education to meet their individual educational needs.
Instructional Hierarchy (IH)	The Instructional Hierarchy (IH) is a theoretical framework that describes the four stages of academic skill progression as well as the instructional methods which support student progress through each of the learning stages. The four stages include acquisition, proficiency, generalization, and adaption (Haring & Eaton, 1978).
Multi-Tiered Systems of Support (MTSS)	MTSS is the overarching framework for Response to Intervention (RtI) and Positive Behavior and Intervention Supports (PBIS). This framework aims to integrate the resources, strategies, and practices from RtI and PBIS to enhance the core instruction,

	intervention, and behavioral supports as well as address barriers or difficulties related to student learning.
Positive Behavior Intervention and Support (PBIS)	PBIS is a specific MTSS framework which aims to create a positive school environment and prevent student behavior problems. PBIS utilizes a multi-tiered, data-based approach to service delivery. The first tier includes teaching and reinforcing a set of appropriate behaviors within the whole school. Tier two services include behavioral and social/emotional interventions for students who do not respond to the supports at the tier 1 level. Tier 3 services increase in intensity and individualization when students continue to demonstrate behavioral/social/emotional difficulties after previously receiving tier 1 and tier 2 supports (Bradshaw, Pas, Debnam, & Lindstrom Johnson, 2015).
Proficiency	Proficiency is the second stage of the Instructional Hierarchy (IH). At this stage, the learner is able to accurately perform the skills but is not able to use the skill fluently. The instructional focus of this stage is the improve the student's ability to quickly and accurately preform the skill (Haring & Eaton, 1978).
Response-to-Intervention (RtI)	RtI is an MTSS framework focusing on providing empirically based, high-quality academic instruction and interventions that match students' academic needs. This multi-tiered model typically has three or four tiers. At each tier level, instructional dosage and measurement intensity increases. For example, group sizes become smaller, frequency of progress monitoring increases, and time spent receiving supplemental interventions increases. Students' progress monitoring data are used to determine if the instructional supports in place are effective or if more intense instructional supports are necessary.
Specific Learning Disability (SLD)	A specific learning disability (SLD) is one of the thirteen disability categories under which students may receive special education. Fundamentally, an SLD is evident when a student demonstrates underachievement in academic performance.
Strategic Incremental Rehearsal (SIR)	Strategic incremental rehearsal (SIR) is a drill-based intervention that combines features of incremental rehearsal (IR) and traditional drill (TD). SIR uses unknown items only within the intervention sequence. Unique to SIR is the use of learner performance as the criteria for changing the stimulus items presented during intervention implementation (Kupzyk et al., 2011).

Traditional Drill (TD)	A drill-based intervention that uses all unknown stimuli and repeated presentation to increase students skill acquisition.
Treatment Acceptability	Whether the client perceived the treatment as effective and demonstrated a preference for the treatment when asked following treatment implementation (Arra & Bahr, 2005).

CHAPTER 2

Literature Review Introduction

Only 40% of fourth grade students demonstrated proficiency in math according to results from the National Assessment of Educational Progress in 2017 (NAEP; Nations Report Card, 2017). In order to increase the number of students who demonstrate proficiency in math, comprehensive instructional and intervention support in math are needed. In an effort to appropriately intervene with students' math difficulties, it is also important to consider the context and frameworks in which math interventions are delivered. Systems-level frameworks, such as Response-to-Intervention (RtI), are one approach to providing evidence-based academic supports in the schools. However, research in RtI for math related to interventions in the schools is relatively limited (Lembke et al., 2012; Bouck & Cosby, 2017). Given the limited number of students who demonstrated proficiency in math and the relative dearth of research examining math interventions, further research related to early intervention for students with math difficulties is warranted.

Organization of the Chapter

This chapter offers an overview of the research related to evidence-based math interventions as well as the service-delivery and theoretical frameworks that serve as a foundation for the development and implementation of such interventions. First, a systems-level prevention framework used to provide instructional supports is discussed. Second, the use of theoretical frameworks to identify appropriate interventions with a focus on the Instructional Hierarchy (IH; Haring & Eaton, 1978) will be examined. Third, math interventions specifically related to math facts and drill-based interventions are reviewed. Finally, the research questions for the present study will be outlined.

Systems-Level Frameworks

Math proficiency was shown to increase students' ability to solve complex problems and increase the likelihood of graduating college (Zaslofsky et al., 2016; Kilpatrick, Swafford, & Findell, 2001; National Mathematics Advisory Panel [NMAP], 2008). Given the importance of math proficiency on long-term student outcomes, early identification of math difficulties and early intervention through academic support systems is essential. Systems-level frameworks provide support to all students with the goal of identifying students demonstrating academic difficulties and consequently intervening early with those academic difficulties (Higgins-Averill & Rinaldi, 2013). These frameworks are an organized and data-driven means of providing struggling students with intervention supports and facilitate a decision-making process regarding student progress. The primary frameworks used in schools are Multi-tiered Systems of Support (MTSS), which includes Response to Intervention (RtI).

Multi-tiered Systems of Support (MTSS)

Multi-tiered systems of support (MTSS) are a systematic approach to providing academic and behavioral supports through "increasingly differentiated and intensified assessment and instruction" to support all students' academic and behavioral needs (Wixson, 2011; pg. 503; Bouck & Cosby, 2017). As educational reform and legislation (Every Student Succeeds Act, 2015) continues to push the use of data-based, outcome-driven models to scaffold instruction for all students, the implementation of MTSS has increased (Cusumano, Algozzine, & Algozzine, 2014). Therefore, schools may adopt MTSS frameworks to organize and evaluate both student needs and the resources available within the school to support students. Some of the key features of MTSS frameworks include using data to drive instructional decisions and using a problem-solving framework to determine how to appropriately provide intervention supports (Cusumano et al., 2014). Research suggested that providing educational services (i.e., general or special education) within the context of an MTSS model can optimize student outcomes (Griffiths, Parson, Burns, VanDerHeyden, & Tilly, 2007; Cusumano et al., 2014).

MTSS serves as an umbrella framework housing other systems models such as RtI and Positive Behavior Interventions and Supports (PBIS; Cusumano et al., 2014). RtI and PBIS are data-driven frameworks that support academic (RtI) and behavioral (PBIS) difficulties (Higgins-Averill & Rinaldi, 2013). Within MTSS, RtI systematically uses student data to drive academic instructional decisions across general and special education and can result in the modification of instruction (Wixson, 2011). Given the focus on math interventions in the present study and that such interventions are often implemented within the context of such frameworks, RtI will be outlined along with the relevant literature related to each component of the framework.

Response to Intervention Framework (RtI)

RtI is a systematic framework used to support students with academic difficulties by aligning appropriate instruction with student need (Cusumano et al., 2014) that grew out of prevention models such as the public health model developed within the fields of medicine and psychology (Bineham, Shelby, Pazey, & Yates, 2014; Mellard, McKnight, & Jordan, 2010). The public health prevention model assesses three levels: 1) whole population level of risk, 2) individual response to treatment, and 3) individual response to intensified intervention treatment (Bineham et al., 2014; Mellard et al., 2010). Based on this public health model, the National Research Council (1982) discussed the quality of general and special education as well as identification methods for SLD in their seminal report in an effort to develop a similar system in education (Bineham et al., 2014). Consequently, RtI was developed as a systematic data-driven framework to provide quality instruction and to implement evidence-based interventions.

Therefore, RtI is an important framework to consider when discussing academic interventions and the monitoring of student responding to such interventions.

There are two broad RtI approaches: the standard treatment protocol and the problem-solving approach (Cusumano et al., 2014). The standard treatment protocol approach largely uses pre-packaged, or standardized, interventions related to students' most common set of academic needs (e.g., reading fluency; Fuchs & Vaughn, 2012). In contrast, the problem-solving approach makes instructional decisions based on individual student's specific difficulties and selects interventions in accordance with each student's academic needs based on a systematic decision-making process (Fuchs & Vaughn, 2012). The problem-solving approach appeared to be more commonly implemented in the schools than the standard treatment protocol approach (Bouck & Cosby, 2017). The problem-solving approach typically includes four steps. First the problem, such as an academic difficulty, is identified and defined. Next, an appropriate intervention for the problem is selected and then implemented. After the selected intervention is implemented, the identified problem is re-evaluated post-intervention (Brown-Chidsey & Steege, 2010). Regardless of the approach used in the application of RtI, the tiered structure and implementation components remain broadly the same.

RtI is traditionally made up of three or four tiers, but there is no evidence supporting three versus four tiers being more effective (Mellard et al., 2010). Regardless of the number of tiers, the key difference between each tier of RtI is increased intensity related to instruction, curriculum, dosage, group size, and assessment frequency. In some models, the fourth tier is designated for special education, while other models may either integrate special education services throughout RtI, separate special education completely from the framework, or include special education at the tier 3 level (Brown-Chidsey & Steege, 2011).

Across the three, or four tiers, there are many important components required in order to effectively implement an RtI framework. One important overarching aspect is implementation fidelity. Although fidelity represents a significant challenge for RtI implementation (Burns, Maki, Warmbold-Brann, & Preast, 2018), it is essential to promote positive outcomes for students. Therefore, it is important for schools to be unified in their desire to implement such a framework (Mellard et al., 2010). Along with educators' buy-in and implementation fidelity, other key components of RtI also include quality core instruction, data-driven assessment practices, and evidence-based interventions to address academic difficulties (Bineham et al., 2014). Additionally, it is crucial for the RtI framework to be cohesive with other existing frameworks (i.e., PBIS) and be sustainable through the use of the existing resources within the school (Mellard et al., 2010). These features will be discussed in greater detail next as the three RtI tiers are reviewed.

Tier 1. Tier 1 of an RtI framework involves class wide core instruction in academic skills for all students (i.e., reading, math, writing; Stoiber & Gettinger, 2016). Given that both general education and special education students spend a large portion, if not the majority or all, of their instructional time in the general education classroom, it is essential for tier 1 instruction to be evidence-based and implemented with fidelity (Stoiber & Gettinger, 2016). Evidence-based instruction is instruction that is "grounded in scientifically based research" (Hughes & Dexter, 2011; pg. 5) and includes empirically supported instructional methods such as demonstration, frequent practice, and feedback (Doabler et al., 2012). Evidence-based core instruction prevents student academic difficulties and allows educators to rule out inadequate instruction as the cause of lack of student academic progress (Hughes & Dexter, 2011). If core instruction is inadequate and a large number of students therefore do not demonstrate academic proficiency, there will likely be a large number of students who appear to require more individualized and intensive supports. The RtI model is not designed to serve a majority of students at the tier 2 and tier 3 levels and will not be effective when a majority of students demonstrate academic difficulties (Brown-Chidsey & Steege, 2011). Thus, having strong and effective tier 1 core instruction is vital to an effective RtI framework.

Another important aspect of evidence-based instruction includes the use of differentiated instruction within the general education classroom to meet students' educational needs (Fuchs & Vaughn, 2012). The use of differentiated instruction allows for the majority of students to receive instruction solely at the tier 1 level. Evidence-based core instruction, including differentiated instruction, has been shown to lead to positive academic outcomes (Stoiber & Gettinger, 2016; Algozzine, Wang, & Violette, 2011), and when implemented effectively with fidelity, has led to a decrease in the number of students in need of academic intervention (Vaughn et al., 2009; Fuchs & Vaughn, 2012).

Another important component of tier 1 is universal screening, or benchmark assessments. These measures are typically administered three times per year to track all students' progress over the course of the school year (Gersten et al., 2009). The goals of universal screening measures are to ensure tier 1 instruction is effective for the majority of students, identify students who are at risk for academic difficulties, and broadly inform the academic skills in which students may need additional, targeted intervention support (Gersten et al., 2009). Universal screeners are effective tools in identifying students at-risk for academic difficulties (Hughes & Dexter, 2011); however, there is a lack of consistency related to the criteria to identify at-risk students when comparing various universal measures (e.g., cut-scores, percentile ranks; Hughes & Dexter, 2011; McMaster & Wagner, 2007). Although the inconsistencies in criteria make comparison across screening measures difficult, research broadly supports the use of a benchmark criteria for identifying at-risk students (Hughes & Dexter, 2011). Thus, early identification through universal screeners can prevent more significant academic difficulties later in schooling (Hughes & Dexter, 2011). For math, CBM universal screening includes measures for skills such as early numeracy (pre-k to 1st grade), computation, concepts and applications (elementary), as well as estimation and algebra (secondary; Lembke et al., 2012). Unfortunately, the research is relatively scarce for universal screening measures in areas other than reading, such as math (Glover & DiPerna, 2007; Bouck & Cosby, 2017).

When a student has been identified as at-risk and receives either tier 2 intervention or additional supports at tier 1, progress monitoring measures are used to track students' progress on a more frequent basis than universal screening measures are implemented (e.g., biweekly). Progress monitoring is another type of formative assessment which is important not only in tier 1, but across all tiers of the RtI model. One example of a commonly used screening tool at the elementary level includes Curriculum-based Measures (CBM; Bouck & Cosby, 2017) such as Dynamic Indicators of Basic Early Literacy Skills (DIBELS; Good, Kaminski, Smith, Simmons, Kame'enui, & Wallin, 2003; Good, Kaminski, & Moats, 2003).

Overall, the research examining tier 1 supports to address math difficulties is limited (Bouck & Cosby, 2017). One study by Gresham and Little (2012) reviewed the application of RtI for math within an elementary classroom. Many of the components used for classroom math instruction included common components used in RtI for reading such as universal screening, supplemental instruction for specific skills as well as progress monitoring using CBM probes. However, this study did not examine whether such components were more effective compared to when the components were not implemented; therefore, it is unclear if these components actually led to improved student outcomes. Another study conducted by Fuchs and colleagues (2004), examined the effects of supplemental tier 1 instruction on three groups of students (i.e., students at-risk for reading and math disabilities, students at-risk for math disability only, and students not at-risk for disabilities). Results showed that supplemental tier 1 instruction was generally ineffective in improving story problem math skills for students at-risk for math and/or reading disabilities (Fuchs, Fuchs, & Prentice, 2004), suggesting more intensive or alternative instruction beyond tier 1 may be required for supporting students with math difficulties. However, participants assigned to the experimental condition outperformed participants who did not receive classroom instruction on math problem solving, suggesting classroom supplemental instruction is better than no supplemental instruction, even for students at-risk for a learning disability.

Another shortcoming of math instruction broadly relates to the limited incorporation of research-based components of effective core instruction in math. Instructional components such as modeling of content, structured practice, and performance feedback have been inconsistently implemented within math instruction (Doabler et al., 2012; Bryant, Bryant, Gersten, Scammacca, & Chavez, 2008). There is some evidence supporting class wide curricula or strategies such as Mathematics Peer-Assisted Learning Strategies (PALS; Fuchs, Fuchs, Yazdian, & Powell, 2002), Mastering Math Facts (Otter Creek Institution; Lembke et al., 2012), and multitiered math curricula such as Transmath and Vmath (Voyager Learning; Lembke et al., 2012). Although there is some support for such curricula, overall, they are narrow in terms of the academic skills of focus and are not necessarily identifiable as whole class, tier 1 core instruction curricula but rather serve as supplemental supports at the tier 1 level. In sum, the available research examining tier 1 math instruction suggested that core-curriculum at the tier 1 level has limited empirical evidence, but there are supplemental supports which have demonstrated promising outcomes for class wide performance. However, these supports have not resulted in improved performance across all students and therefore students demonstrating continued math difficulties may receive math support through more intensive tiered interventions rather than in tier 1.

Tier 2. Tier 2 within an RtI framework involves the implementation of evidence-based intervention as supplemental supports to students who do not demonstrate adequate progress in response to tier 1 core instruction. When RtI is implemented with fidelity, it is estimated that 15% to 20% of students should receive tier 2 intervention supports in combination with tier 1 instruction (Shapiro, 2015). Gersten and colleagues (2009) asserted that instruction at the tier 2 level should be explicit, systematic, and supplemental to core instruction in tier 1. This instruction includes modeling of the skills, guided practice with corrective feedback, and repeated practice and review (Gersten et al., 2009; Bouck & Cosby, 2017). Instructional intensity and dosage increase at the tier 2 level compared to tier 1.

Dosage increases at the tier 2 level, however, dosage amount appears to vary between intervention specific and broad RtI focused research. Intervention specific research examined academic interventions not implemented within the context of an RtI framework. In contrast, RtI-focused research examined academic interventions within the context of an RtI framework. Dosage can include an increase in instructional time, smaller group size, and alternative instructional methods based on student needs (Wixson, 2011; Mellard et al., 2010). In a review of school-based RtI implementation, average group sizes at the tier 2 level ranged from 2 to 5 or 6 to 10 students (Mellard et al., 2010). When interventions were examined outside the context of an RtI model but were identified as being similar to interventions at the tier 2 level, 6 or fewer students were identified as the appropriate group size (Fuchs & Vaughn, 2012), indicating that broad RtI research supported the inclusion of larger intervention group sizes. Such differences in group size suggested that tier 2 implementation may vary in dosage between research examining interventions in isolation and research examining such interventions within the context of an RtI model. Instructional time, however, was fairly consistent across intervention-focused and RtI-focused research at the tier 2 level which may be due to fixed-schedules with limited time within the school day (Mellard et al., 2010). Tier 2 math intervention sessions should, ideally, be implemented four to five times per week lasting 20 to 40 minutes each session (Lembke et al., 2012). Length of intervention time often depends on student needs and time set aside for tier 2 interventions during the school day. Implementation of evidence-based interventions should typically last for 10 to 15 weeks and include teacher led explicit instruction and high implementation fidelity (National Center for RtI, 2010). Regardless of group size and instructional time, however, evidence-based intervention is of greatest importance to ensure high quality supplemental supports to students with academic difficulties.

Four types of evidence-based instructional components for math have been identified for tier 2 interventions with elementary students (Bouck & Cosby, 2017). These four instructional components are explicit instruction, the concrete-representational-abstract (CRA) instructional sequence, schema-based instruction, and number talks (i.e., talking through solutions to math problems). In a review of the literature, Gersten and colleagues (2009) found that all four of these instructional strategies demonstrated statistically significant outcomes for enhancing students' math proficiency. These findings supported the importance of evidence-based instructional strategies which serve as the foundation for evidence-based interventions across standard treatment intervention packages and individual skill-based interventions (i.e., those determined through problem-solving approaches).

One type of evidence-based intervention is standard treatment, or pre-packaged, interventions which include an instructional approach in which the same intervention is used across students who demonstrate academic difficulties. In other words, they target a wide set of academic skills for a specific academic area (e.g., decoding, sight words, and fluency for reading). Standard treatment intervention packages have resulted in improved student math skills such as arithmetic, complex calculations, and integrative number knowledge (e.g., Galaxy Math; Fuchs, Fuchs, & Gilbert, 2019). Other standard treatment math interventions demonstrating effectiveness included The Numeracy Booster (Bryan et al., 2008), Hot Math (Fuchs, Fuchs, & Hollenbeck, 2007), Pirate Math (Fuchs, Fuchs, Powell, Seethaler, Cirino, & Fletcher, 2008), Number Rockets (Fuchs et al., 2005), and Number Worlds (Griffin, 2004). These interventions targeted skills such as early number sense, solving world problems, conceptual learning, and other skills aligning with the common core state standards for math. However, standard treatment intervention packages are not the only instructional strategies used to address student difficulties. Specific skill-based, or non-standard treatment, interventions have also been shown to improve students' math skills (e.g., drill-based intervention; Burns, 2005).

Overall, evidence-based tier 2 interventions, standard treatment and specific skill-based, have addressed a wide variety of math skills and have been examined more often than math interventions at the tier 1 or tier 3 levels. Thus, there is a significant gap in the research base regarding tier 3 interventions. Although some of the standard treatment interventions have also been examined within the context of a tier 3 intervention (e.g., one-on-one or smaller group sizes), these standardized intervention packages have also been researched more than skill-based individualized interventions (i.e., those determined though the problem-solving approaches) even though these interventions are more commonly implemented within RtI frameworks (Bouck &

Cosby, 2017). Therefore, additional research regarding interventions determined through individualized, problem-solving methods is needed. Evidence-based math interventions will be further reviewed as part of the tier 3 section and math intervention sections of this literature review.

As in tier 1, progress monitoring tools remain an important component at the tier 2 level. At the tier 2 level, progress monitoring tools are administered frequently to evaluate student progress in response to interventions (e.g., weekly; Silberglitt, Parker, & Muyskens, 2016). Progress monitoring tools are used to identify any instructional changes or supplemental instruction needed based upon student performance. Additionally, progress monitoring measures allow teachers to provide performance feedback to students (Cusumano, et al., 2014). One of the most common progress monitoring tools is CBM. CBM are versatile in that they can also be used as screening tools, as previously described in the tier 1 section of this review. Advantages of using CBM as a progress monitoring tool include their sensitivity to student growth (or lack thereof) in response to intervention, allowance for frequent administration, and measurement of various skills related to math (Lembke et al., 2012). It is important for progress monitoring measures to be sensitive to changes in skill growth because they are used to make instructional decisions in response to students' progress (or lack thereof). Students who continue to demonstrate difficulties in math will likely receive more intensive supports beyond tier 2, such as tier 3.

Tier 3. Tier 3 interventions are implemented when students demonstrate inadequate growth in response to tier 1 instruction and tier 2 intervention (Brown-Chidsey & Steege, 2011). Within tier 3, students receive the most intensive academic supports (Hosp, Huddle, Ford, & Hensley, 2016). Dosage within tier 3 is intensified in terms of group size, with typical groups being 2 to 5 students (Mellard et al., 2010), and sometimes one-on-one instruction (Brown-Chidsey & Steege, 2011). Intervention is also typically implemented more frequently and for longer periods of time during intervention sessions compared to the tier 2 level (e.g., 30 - 45 minutes; Gersten et al., 2009). Consistent with the other tiers of an RtI framework, as evidence-based intervention intensity increases, progress monitoring measures are administered more frequently. At the tier 3 level, frequency of progress monitoring typically increases to at least once per week so that students' responding can be closely monitored allowing for immediate intervention modification when needed (Klingbeil, Bradley, & McComas, 2016). Apart from increased intensity in intervention implementation and progress monitoring, interventions adapted across tiers 2 and 3 maintain similar instructional features, such as modeling and corrective feedback, sustaining the efficacious features of evidence-based interventions (Cusumano, et al., 2014). Because students receiving tier 3 interventions receive more intense intervention compared to tier 2, particularly in terms of time, it is of even greater importance that the interventions demonstrate effectiveness in addressing students' academic difficulties.

At the tier 3 level, there are some evidence-based interventions that are flexible for implementation across small groups and individual formats, thus being appropriate for tier 2 or 3. Cover-Copy-Compare (CCC) is an example of an evidence-based intervention which can be implemented either in a small group or individually. CCC is an effective self-management intervention in which students guide their own practice of academic skills (e.g., spelling, math problems; Joseph, Konrad, Cates, Vajcner, Eveleigh, & Fishley, 2012). CCC encompasses effective instructional techniques such as modeling and corrective feedback as well as multiple opportunities to respond accurately (Joseph, Konrad et al., 2012). Although CCC has not been prescribed as a tier 3 intervention, it has often been examined within the context of tier 3 based upon intervention intensity (e.g., individual instruction and frequency of intervention implementation). However, CCC is more likely an effective intervention for continued self-monitored practice of a skill to improve accuracy opposed to acquisition of an academic skill because modeling and corrective feedback are not provided directly by the instructor.

Some research has suggested that students' academic outcomes are more contingent upon the quality and consistency of intervention instruction rather than group sizes in tier 3 because differences in students' academic progress in small group versus individual intervention formats were negligible when qualified teachers consistently implemented the evidence-based interventions (Elbaum, Vaughn, Tejero Hughes, & Watson Moody, 2000). These findings highlighted the importance of high quality and consistent instruction not only at the tier 3 level but across tiers as well. Thus, as long as the intervention is evidence-based and implemented with fidelity, group size is not necessarily the most important variable on which to focus within tier 3 interventions. Perhaps of most importance is focusing on establishing an adequate evidence base for targeted math interventions generally, particularly in light of the relative dearth of such research.

Although evidence-based instruction is a key component to intervention success, there are some academic interventions that are designed for tier 3 individual implementation. For example, drill-based interventions are most commonly implemented individually (Burns, 2005; Volpe, Mulé et al., 2011). Incremental rehearsal (IR) is an effective drill technique to increase acquisition of foundational academic skills, such as math facts (Burns, 2005). IR teaches unknown academic material (i.e., math facts or sight words) to the student while interspersing known academic material. Therefore, IR may be more suitable for individual implementation in order to include each student's specific known and unknown material. Traditional drill flashcard interventions are also commonly implemented in a one-on-one format like IR (Volpe, Mulé et al., 2011). These interventions serve an important role in addressing students' individual academic needs across subject areas (i.e., reading and math). For example, students may differ in the sight words they are struggling to learn and therefore will require individualized intervention time to practice their specific sight words. Later in this chapter, research examining math fact drill interventions will be reviewed in further detail.

Given that a majority of the interventions, both standard treatment protocol and skill-based individualized interventions, discussed thus far are fairly adaptable across tier 2 and tier 3, there appears to be a significant gap in the research base regarding tier 3 specific interventions. Although adapting traditionally tier 2 interventions at the tier 3 level can be beneficial when a group of students all share a specific academic difficulty, this approach may not necessarily support students with more individualized skill difficulties or students who show inadequate responding to tier 2 interventions. One such intervention previously discussed as best individually implemented at the tier 3 level are drill-based interventions. Even though drill-based interventions have been examined specific to tier 3, the research base has largely examined such interventions within the context of reading skills (i.e., sight words; Burns, 2005). Similarly, the RtI research base includes some evidence supporting interventions for math difficulties within an RtI framework; however, the effectiveness of math interventions is largely under-studied (Lembke et al., 2012; Hughes & Dexter, 2011) specifically at the tier 3 level. Therefore, the goal of the present study is to add to the math intervention research base, particularly for tier 3 interventions.

Outcomes and Challenges of RtI

There are many beneficial features of RtI frameworks such as the use of evidence-based instruction and frequent monitoring of student progress. Close monitoring of student progress allows for early intervention to prevent student failure (Silberglitt et al., 2016). The implementation of RtI also led to a reduction in inappropriate special education referrals and disproportionality in special education identification (Kauffman, Bruce, & Lloyd, 2012). RtI promotes early intervention which has resulted in improved student academic outcomes and a decrease in the likelihood of negative long-term outcomes for students such as risky behavior (e.g., drug and alcohol use; Wang & Fredricks, 2014), and dropping out of school (Fan & Wolters, 2014). Thus, when implemented with fidelity, RtI demonstrated overall positive academic and broader life outcomes for the students served within the framework. Despite these positive outcomes associated with RtI implementation, Sarason (1990) argued that system changes often fail when the individuals within the system (e.g., school) do not implement the model cohesively. Therefore, it is important to evaluate implementation fidelity and correct any discrepancies to promote optimal student outcomes (Higgins-Averill & Rinaldi, 2013). Even though there are positive academic outcomes associated with RtI implementation, it is also important to consider barriers to implementing a system-level framework.

Systems-level changes are often difficult for a variety of reasons. Castro-Villarreal and colleagues (2014) examined teachers' perceptions of the most common barriers to the implementation of RtI. In their study, 97 teachers were surveyed through qualitative methods. Five themes emerged based on the responses of the participants and included: training, time, resources, the RtI process itself, and paperwork. It is important to consider such barriers related to implementation of systems-level frameworks and also broadly when looking to implement any new resource, such as evidence-based interventions. Interventions are a key feature of an RtI framework and directly address and support student academic skill growth. Thus, it is important to identify the most effective instruction and interventions to address students' academic difficulties while considering the implementation barriers identified. As mentioned, however, research related to RtI, and math interventions specifically, is lacking, particularly at the tier 3-level (Bouck & Crosby, 2017), and is therefore an important focus for future research to promote high quality implementation.

Summary

MTSS is a systematic, data-based decision-making model which provides increasingly intensive supports through specific frameworks (e.g., RtI and PBIS) to address student difficulties (e.g., academic or behavioral). MTSS frameworks seek to provide early identification and intervention to struggling students in hopes of decreasing long-term student difficulties and increase overall student success in school. These goals are accomplished through the use of evidence-based interventions throughout the three, or four, tiers of the framework. At each tier, instruction intensity typically increases in terms of intervention dosage, group size, and assessment frequency. RtI, as part of MTSS, addresses the academic needs of students universally, in small groups, and individually. Academic interventions used within RtI serve as the key mechanism to improving students' academic skills. Therefore, it is essential to have a thorough understanding of empirically supported academic interventions in order to promote optimal student outcomes. RtI serves as an effective means to deliver interventions for students demonstrating academic difficulties and thus represents an important consideration when intervening in response to student academic difficulties. However, RtI is a broad service delivery framework in which academic interventions are implemented. In contrast, theoretical frameworks can be used to drive specific intervention delivery decisions based on individual student needs and intervention components. Next, the importance and use of theoretical frameworks when identifying appropriate interventions will be discussed.

Theoretical Frameworks

In this section, the importance of using theoretical frameworks to guide intervention research will first be reviewed. Then, Haring and Eaton's Instructional Hierarchy (1978) will be discussed along with how it can be used to drive development and implementation of academic interventions. Although the research methodology of studies examining evidence-based intervention has improved, such research is still lacking in terms of reliance on theoretical frameworks to guide and evaluate interventions (Burns, 2011). However, Burns (2011) asserted that it is important to focus on theory when conducting educational research because research findings cannot be adequately interpreted without being grounded in theory.

Theoretical frameworks serve as a foundation to guide research, practice, and problem solving (Burns, 2011; Tilly III, 2008). According to Ellis (2014), theory is defined as a provisional prediction of a phenomenon, including a set of ideas that have been cautiously considered. In educational research and practice, theoretical frameworks inform development of instructional methods or activities (Ellis, 2014). One advantage of testing the underlying mechanism of change, or theory, is doing so increases the internal validity of empirical research (Volpe & Suldo, 2014). Another advantage of using theoretical frameworks is that they allow interventions to be adapted for specific settings or circumstances "without sacrificing efficacy" by ensuring that intervention modifications do not disrupt the underlying efficacious intervention components (Burns, 2011; pg. 133). Along with these advantages, theory also aids in the identification of "common mechanisms of change across" various interventions (Volpe & Suldo, 2014; pg. 116), which may be important to help develop additional effective intervention practices that capitalize on such mechanisms of change. Research regarding academic interventions in the schools has typically lacked an emphasis on theory and the mechanisms leading to change but instead, such interventions have generally emphasized effectiveness rather than *why* an intervention may be effective (Burns, 2011; Mercer et al., 2014). However, it is important to examine how theoretical frameworks can be used to develop and implement academic interventions to help inform future research, such as the current study.

Mercer and colleagues (2014) reviewed the literature in four school psychology research journals and identified studies which used theory to both guide their research regarding interventions broadly (e.g., social-emotional, behavioral, and/or academic) and discuss the study findings. The authors differentiated between theoretical frameworks and causal program theories. Theoretical frameworks are broad theoretical approaches to research that are not linked to a specific intervention. Causal program theories provide explanation for specific intervention outcomes. Overall, Mercer and colleagues (2014) found that out of the 94 studies reviewed, 48% identified a causal program theory, 37% assessed constructs or intervention components which had been identified as effective within the context of a causal program theory, 7% used mediation analyses, 20% used moderation analyses, and 37% considered theoretical implications when discussing their results. Thus, the authors did not identify a single study that used a theoretical framework, not a causal program theory, to guide the intervention implementation and discuss the study findings. Although these are meaningful findings regarding the incorporation of theory into research of interventions across domains (i.e., academic, behavioral, etc.), the authors only examined research within a six-year span and they also limited their sources to four journals. It is possible that there may be other school-based intervention research published in other journals during this time frame, or outside of this time frame, that were not included in the review. However, these findings are important because they demonstrated the underutilization of theoretical frameworks to guide and explain the findings within school-based intervention research.

Although theoretical frameworks were not used to guide and explain results, Mercer and colleagues (2014) identified seven categories of theoretical frameworks used to discussed components of intervention within the intervention literature in their review. These theoretical frameworks were also identified as broadly important and relevant within intervention research. The theoretical frameworks discussed within the reviewed literature included the ecological, prevention, social cognitive, cognitive-behavioral, cognitive, and behavioral frameworks (Mercer et al., 2014). Ecological frameworks pose that interventions are most effective when grounded in the larger system (e.g., RtI, MTSS; Burns, 2011; Apter & Conoley, 1984). Similarly, prevention frameworks aim to increase the system's ability to meet student needs and prevent student difficulties while also increasing the student's skills or competencies (Burns, 2011; Ysseldyke et al., 2006). Ecological and prevention frameworks appeared to be most commonly used in tandem with other theoretical frameworks (e.g., social-cognitive; Mercer et al., 2014). Likewise, cognitive and behavioral frameworks have been conjointly used to guide research on interventions such as those targeting academic skills (e.g., math fluency; Mercer et al., 2014; Axtell et al., 2009). Cognitive theory frameworks (e.g., spaced practice) have been used to explain the effective mechanisms of interventions such as incremental rehearsal, a drill-based intervention used to improve reading and math skills (Varma & Schleisman, 2014). Behavioral frameworks have also been used to guide the implementation of academic interventions. The Instructional Hierarchy (IH; Haring and Eaton, 1978) is one behavioral theoretical framework which can be used to guide research on the implementation of academic interventions in the schools (Ardoin & Daly, 2007). However, the IH has been inconsistently applied within intervention research (Burns et al., 2010; Codding & Martin, 2016).

These theoretical frameworks may represent appropriate models to guide research on school-based interventions, including academic interventions (Burns, 2011). Examination of interventions without the use of a guiding theoretical framework can result in trial and error rather than a true evaluation of the instructional features of the intervention and their impact on student skill growth (Burns, 2011). This trial and error approach may also impact intervention generalization to school-based practice because the findings, in the absence of theory, may only apply to the sample and setting in which the study took place. In spite of the importance of theory-driven research, there has been inconsistent use of such frameworks to evaluate and explain the findings of academic interventions. A goal of the present study, therefore, is to use a theoretical framework, the IH, to guide intervention implementation and to help contextualize the findings.

Instructional Hierarchy (IH)

Learning can be defined as being able to perform a specific skill in increasingly complex opportunities or activities (Haring & Eaton, 1978). In order to understand how learning occurs, consideration of the theoretical underpinnings of the learning process is important. The Instructional Hierarchy is a framework used to match instructional features to students' skill levels and can consequently be used to guide intervention implementation. The four stages of the IH in order of skill progression include acquisition, proficiency, generalization, and adaption (Haring & Eaton, 1978). Students' skills are evaluated through empirically supported assessments (e.g., CBM) prior to instruction to determine the stage at which a student's skills fall within the hierarchy. This evaluation helps to identify the type of instruction the student requires to further develop their academic skills. For example, if a student falls in the acquisition stage of learning math facts, it is likely that they will need instruction involving repeated practice of the facts before moving on to the proficiency stage or generalization to higher-order math skills. Haring and Eaton (1978) outlined instructional techniques to develop students' academic skills from initially being slow and inaccurate to eventually being fluent, accurate, and generalized across academic tasks. Each stage of the IH along with relevant literature supporting the instructional components will be reviewed. Below, Figure 1 outlines each stage of IH broadly.

Figure 1. Instructional Hierarchy (IH) Overview

	Acquisition	**Proficiency**	**Generalization**	**Adaption**
Learning Hierarchy	Slow and inaccurate responding	Accurate but slow responding	Can apply academic skill to new setting	Can use information to problem solve
Instructional Hierarchy	• Modeling • Explicit Instruction • Immediate Corrective Feedback	• Novel practice opportunities • Independent practice • Timings • Immediate Feedback	• Discrimination training • Differentiation training	• Problem solving • Simulations

Acquisition stage. Students who are identified as falling in the acquisition stage exhibit slow and inaccurate responding (Haring & Eaton, 1978). The goal for students in this stage is to increase accurate responding. The key instructional strategies to promote skill acquisition are demonstration, or modeling, explicit instruction, and immediate feedback (Haring & Eaton, 1978). Explicit instruction, or direct instruction, has been defined as teacher led instruction which uses evidence-based instruction procedures such as "cueing, modeling, verbal rehearsal, and feedback" (Wendling & Mather, 2009; pg. 204). Explicit instruction also includes well-organized lessons which are clearly communicated and include student engagement (Wendling & Mather, 2009). This instructional strategy has been identified as effective when teaching foundational academic skills in reading (e.g., phonics; Ryder, Tunmer, & Greaney, 2008) and math (e.g., arithmetic and problem-solving; Kroesbergen & Van Luit, 2003). Explicit instruction is not only an individual instructional strategy, but also encompasses some of the other instructional strategies identified within the IH framework, like modeling and feedback.

Modeling is defined as the explicit and accurate demonstration of the academic skill of focus (Malouf, Reisener, Gadke, Wimbish, & Frankel, 2014). Modeling has been used to instruct students on basic reading skills such as phonics (Benner et al., 2010) and sight words (Skinner, Logan, Robinson, & Robinson, 1997). Modeling through strategies such as listening passage preview (LPP) have also demonstrated effectiveness for increasing words read correctly per minute (WRCM) when combined with repeated reading strategies (i.e., repeatedly reading the same passage of text; Rogers & Ardoin, 2018). Additionally, math fact interventions which include modeling demonstrated greater gains in digits correct per minute (DCPM) compared to math fact interventions without modeling (Codding, Burns, & Lukito, 2011). Thus, the use of modeling during skill acquisition interventions has been supported across subject areas.

Immediate feedback is another key instructional component during the acquisition stage. Immediate feedback occurs when the interventionist provides feedback to the student about the accuracy of their responding. One type of feedback used is error correction. Error correction is defined as correcting a student when they make an error immediately after the error occurs (Malouf et al., 2014). Error correction improved students' accuracy on skill-based tasks such as math computation (Ardoin, Witt, Connell, and Koenig, 2005), math facts (Reynolds, Drevon, Schafer, & Schwartz, 2016) and sight words (Marvin et al., 2010). Another type of error correction commonly examined in the intervention literature is response repetition or direct rehearsal. This corrective feedback strategy requires the instructor to model the correct response and then the student immediately practices the correct response multiple times (Drevon & Reynolds, 2018). Response repetition has been shown to increase academic skill acquisition related to math facts (Reynolds et al., 2016) and sight words (Marvin et al., 2010).

Prompting is another feedback strategy to promote skill acquisition. Some prompting strategies demonstrating effectiveness included no-no-prompting (NNP; Leaf, Sheldon, & Sherman, 2010), simultaneous prompting (Drevon & Reynolds, 2018; Rao & Mallow, 2009), and most-to-least prompting (MTL; Fentress & Lerman, 2012). No-no prompting first requires the instructor to provide positive reinforcement for correct responses and corrective feedback such as ‘try again’ when incorrect responses are provided (Leaf et al., 2010). If the student has two errors in a row, the error is corrected by the instructor followed by a ‘remedial trial’ or additional opportunity to practice the same stimuli (Leaf et al., 2010). Simultaneous prompting occurs when prompts are provided immediately after instruction to ensure 100% response accuracy from the student (Leaf et al., 2010). Most-to-least prompting is a prompt fading technique which includes significant guidance and feedback during initial practice and less guidance and feedback as practice trials continue (Libby, Weiss, Bancroft, & Ahearn, 2008). Preliminary research suggested that error correction was more effective than prompting strategies (Drevon & Reynolds, 2018). These findings are important to consider when implementing acquisition interventions to ensure that effective feedback procedures are used.

Proficiency stage. Students who fall in the proficiency, or fluency, stage are accurate in their responding to the task demand; however, their responding is slow (Haring & Eaton, 1978).

Thus, the instructional goal within this stage is to continue to strengthen accurate responding while also increasing response rate. The instructional strategies to promote proficiency are immediate feedback, timing, novel practice opportunities, and independent practice (Haring & Eaton, 1978). Immediate feedback can take the form of corrective feedback strategies or simple response accuracy feedback, as discussed in the acquisition section. However, immediate feedback may not always be appropriate during students' practicing of a skill, such as reading, because it interferes with the student's fluent practice of the skill (Burns, Riley-Tillman, & VanDerHeyden, 2012). Moreover, delayed feedback is more effective than no feedback at all. For example, during written expression fluency interventions, providing performance feedback following writing practice completion led to greater growth in writing fluency compared to receiving no performance feedback (Truckenmiller, Eckert, Codding, & Petscher, 2014). In some interventions, providing immediate corrective feedback may be appropriate such as in math fact interventions (Burns, 2005). During math fact interventions, there are natural breaks in between stimuli practice, thus providing an opportunity for immediate feedback.

Another component to promote skill proficiency is timing. Timing, or explicit timing, occurs when the student is given a specific amount of time to complete a set of problems or items (Poncy & Skinner, 2011; Van Houten & Little, 1982). Providing a set amount of time in which the student provides multiple responses has been identified as a key component to increase student math computation fluency (Codding et al., 2011) and basic reading skills such as word reading (Hughes, Beverley, & Whitehead, 2007). Although explicit timing has been shown to enhance proficiency (Evans-Hampton, Skinner, Henington, Sims, & McDaniel, 2002), this instructional strategy is more effective when the student can engage in the skill with adequate accuracy (Codding, Shiyko, Russo, Birch, Fanning, & Jaspen, 2007). If the student cannot accurately practice the skill (i.e., the material is too hard), skill acquisition may not have been obtained and further modeling, practice, and corrective feedback is required before adding instructional components such as explicit timing to promote fluency.

Independent practice and novel practice opportunities are the last two instructional components identified as important during the proficiency stage. Repeated independent practice has been shown to effectively improve academic performance in math (Mong & Mong, 2010; Harber et al., 2004) and reading (Therrien, 2004; Rogers & Ardoin, 2018). In reading, repeated reading of the same passage or multiple novel passages (i.e., continuous reading) has led to increased oral reading fluency (Hammerschmidt-Snidarich, Maki, & Adams, 2019). For math facts, the amount of repeated practice trials required to obtain fluency may vary depending on students' pre-intervention math skills as well as the math fact family being practiced (Burns, Ysseldyke, Nelson, & Kanive, 2015). For example, students with limited initial known math facts typically required more repeated practice to reach fluency compared to students who initially demonstrated more known math facts (Burns et al., 2015). Additionally, certain math fact families (i.e., 6s, 7s, 8s) required more repetition to reach fluency compared to other math fact families (e.g., 2s and 3s; Burns et al., 2015). However, repeated practice remains an important instructional component for increasing students' computational fluency (Codding et al., 2011) regardless of pre-intervention skill level. When students demonstrated computational fluency, they were better able to perform more advanced math skills (e.g., multidigit multiplication and word problems; Singer-Dudek & Greer, 2005) likely because they could then devote more cognitive resources to higher-order computation skills (Burns et al., 2015). Thus, computational fluency is important for students' more advanced math skill development and ideally leads to skill generalization.

Generalization stage. Students in the generalization stage demonstrate accurate and quick skill responding, but need practice applying their skill to a novel setting or situation (Haring & Eaton, 1978). For example, generalization occurs when a student is able to accurately respond to a math fact presented through a different format other than the traditional math fact form (e.g., 2 x 2), such as within a word problem. Although it is crucial for students to first learn accurate and proficient responding with an academic skill, there are limited practical advantages to skill acquisition if they are unable to generalize the skills to new settings or task demands. Additionally, it is not possible for educators to teach every type of math computation or how to read every word to students. Therefore, students must be able to generalize skills from direct instruction to novel situations. However, there is limited research examining interventions targeting students' skill generalization (Codding & Poncy, 2010).

The use of discrimination and differentiation training are key instructional components to enhance students' generalization of skills within the IH framework (Haring & Eaton, 1978). Differentiation training is defined as "programmed manipulations of stimulus conditions that increase the probability of correct and fluent responding under different but similar (and appropriate) stimulus conditions to those that were used during initial training" (Skinner & Daly, 2010; pg. 108). For example, if a student can fluently count with buttons, the student's skill generalization could be tested by having them count other objects such as pencils or coins. Discrimination training occurs when the learner provides a specific response to one stimulus, but not in the presence of another (Haring & Eaton, 1978). For example, addition is taught so that it occurs when there is a "+" sign in the math problem but not when there is the "x" sign. Generalization can be tested by presenting math equations with both "+" and "x" signs and tracking the student's responding.

As mentioned, there is limited research examining interventions that target generalization of academic skills compared to research on acquisition and proficiency interventions. However, some preliminary findings of instructional techniques that target skill generalization have emerged. For example, initial training in phoneme blending followed by practice and reinforcement on untrained word lists (i.e., differentiation training) showed increases in oral reading accuracy of the target words across word lists and passages (Martens, Werder, Hier, & Koenig, 2013). However, generalization of phoneme blending skills to oral reading fluency gains was variable across the participants in the study (Martens et al., 2013).

Similar to phoneme interventions, the generalizing effects of a repeated reading intervention also resulted in mixed findings with some studies supporting generalizing effects (Wagner & Espin, 2015; Therrien, 2004). When repeated reading (RR) of the same passage and reading of multiple passages was compared, equal practice across multiple passages resulted in better performance on outcome measures with medium word overlap (i.e., passages that share a moderate amount of the same words). However, there were no differences in reading performance between conditions on generalized passages with high word overlap (i.e., passages that share a significant portion of the same words; Ardoin et al., 2008). Alternatively, the multiple exemplar strategy (i.e., reading three different passages one time each) has shown greater gains in the context of learning rates indicating it may be more generalizable than repeated reading (Silber & Martens, 2010). Ardoin and colleagues (2018) suggested combining features of RR and multiple exemplar interventions to explore new avenues in the generalization intervention research. For example, students could practice repeatedly reading the same passage in addition to engaging in additional reading practice across multiple passages to help promote both fluency and generalization.

Similarly, research examining generalizability of math interventions has demonstrated conflicting results across various strategies (Skinner & Daly, 2010). Some research has shown that the implementation of drill-based math fact interventions (e.g., incremental rehearsal) in conjunction with administration of math computation probes (e.g., multiplication, fractions, word problem) during each session led to generalization of the math facts presented in inverse form (i.e., 2 x 6 and 6 x 2; Codding et al., 2010). Other interventions such as taped problems demonstrated similar results to Codding and colleagues' (2010) findings (i.e., generalization to inverse presentation of math facts); however, the components of the intervention which led to generalization was unclear (Miller, Skinner, Gibby, Galyon, & Meadows-Allen, 2011). Still, other studies identified a lack of skill generalization when students moved from computer-based math practice to paper-pencil performance measures (Duhon, House, & Stinnett, 2012). However, when students received multiple stimulus exemplars with both computer-based and paper-pencil practice, they demonstrated greater skill generalization compared to those who received computer only or paper-pencil only practice (Rich, Duhon, & Reynolds, 2017). Thus, differentiation in instruction across various stimuli appears to be an important mechanism to enhance students' generalization of learned skills across various stimuli.

Overall, research examining skill generalization within academic interventions is lacking for academic subjects other than reading and math, with the majority of research examining reading interventions. Although not the focus of the current study, skill generalization from intervention to different settings, times, and materials needs further examination. Advances in this research could inform how to best implement academic interventions so that students are able to generalize their skills within the classroom or on assessments.

Adaption stage. The last stage within the IH is adaption. Adaption occurs when the learner is able to modify and apply their learned skills in response to novel stimuli and to solve problems (Daly III, Lentz, & Boyer, 1996; Haring & Eaton, 1978). This stage can only occur after the learner has mastered accurate, fluent, and generalized responding. Adaption is promoted through problem solving practice and engaging in various simulations, or instances with novel stimuli in which students must apply their skills (Haring & Eaton, 1978; Daly III et al., 1996). Given that educators cannot teach students the accurate response across all situations and contexts, students must learn how to identify and apply their learned responses to novel situations to solve problems.

The basis of problem-solving skills is continual practice of those skills through increasingly challenging, novel scenarios (Ross & Maynes, 1983). The research related to problem-solving in math largely focused on word problems with emerging evidence for proportional problem-solving (Jitendra, Harwell, Dupuis, & Karl, 2017). Additionally, instruction in problem-solving strategies such as schema-based instruction (SBI) has been examined for math (Flores, Hinton, & Burton, 2016; Jitendra et al., 2017). When math instruction with SBI was compared to a control group, students in the SBI group had better performance on math problem-solving tasks (Jitendra et al., 2015). Additionally, the use of concrete-representational-abstract (CRA) instructional approaches and SBI in combination led to improvements in students' math word problem-solving (Flores et al., 2016). Although there has been some research evaluating problem-solving to enhance math computation (i.e., word problems), direct examination of problem-solving interventions and simulation of novel situations through the lens of the IH is lacking. This gap in the research base could be due to the instructional strategies identified for acquisition, fluency, and generalization interventions being viewed as treatment components which make an intervention effective (Daly III et al., 1996). This perspective is in contrast to viewing problem-solving as an instructional feature of adaption interventions but rather a skill the student must master independently.

Instructional Hierarchy (IH) and Academic Interventions

As discussed above, there has been relatively extensive research examining the instructional components targeting acquisition of and proficiency in academic skills while the research supporting the instructional components targeting skill generalization and adaption is limited (Cates et al., 2007). There is some empirical support for problem-solving interventions (e.g., schema-based instruction) but it is limited and not directly discussed within the context of the IH. Although more research has examined acquisition interventions, the majority of the research targeted foundational reading skills with limited evidence for math acquisition interventions. Additionally, the research examining acquisition interventions often lacked theoretical guidance (Burns, 2011), which undermines the understanding of why an intervention may or may not be effective. Given the long-term outcomes related to early difficulties with foundational math skills along with the lack of math proficiency students have demonstrated, acquisition interventions for math are vital to improve students' math performance. Therefore, the present study seeks to fill this gap in the research base related to math acquisition interventions within the context of the IH.

The IH has been applied to the investigation of effective treatment components for reading difficulties (Daly et al., 1996; Parker & Burns, 2014; Ardoin, Binder, Zawoyski, & Foster, 2018). Such research has examined the use of acquisition, proficiency, and generalization strategies to improve reading skills of learners. Reading research grounded in the IH supported its use to intervene in skills such as phonics (Benner et al., 2010; Martens et al., 2013), sight words (Marvin et al., 2010), word list reading (Skinner, Johnson, Larkin, Lessley, & Glowacki, 1995), oral reading fluency (Ardoin et al., 2018; Ardoin, McCall, & Klubnik, 2007) and comprehension (Therrien, 2004) as well as spelling (Cates et al., 2007). Thus, there has been broad application of the IH within the context of reading, however, such application of the IH within math intervention research is scarce.

Although there are some math intervention studies that have discussed features of an intervention in reference to the IH or that use the IH in their explanation of research findings, math intervention research has very infrequently used the IH to guide its implementation (Codding & Martin, 2016). For example, Burns and colleagues (2010) retroactively applied the IH in their review of math fluency interventions. They discussed the IH to explain intervention outcomes in their meta-analysis after identifying studies and examining student outcome data and effect sizes. Thus, it is important to expand math intervention research to include the use of theoretical frameworks (such as the IH) both to inform intervention implementation and to explain student outcomes following intervention.

Summary

Theoretical frameworks are important to guide and inform academic intervention research because it increases the internal validity of the study and allows for interventions to be adapted across implementation settings (Volpe & Suldo, 2014; Burns, 2011). However, the use of theoretical frameworks to guide intervention implementation and to explain research findings is limited (Burns, 2011). The IH has been used to inform and guide the instructional features of academic interventions but has been largely applied to research related to reading interventions (Codding & Martin, 2016). Therefore, research regarding use of the IH to inform and explain effective math interventions is needed. The present study will use the IH to inform the implementation of math fact interventions and interpret the research findings. Of the four stages within the IH, the acquisition stage will be the focus for the present study. Given that early math skills predict the acquisition of higher-order math skills (Gersten et al., 2005), it is important to examine the effectiveness of math interventions which address early math skills (e.g., math facts). Like theoretical frameworks, evidence-based interventions are a key component to providing academic supports to students with academic difficulties. Given that the present study's focus on early math skill intervention related to math facts, the next section will review the literature related to the efficacy of evidence-based math fact interventions.

Academic Interventions

Evidence-based, high-quality instruction is one of the essential components of any RtI model (Stoiber & Gettinger, 2016). Evidence-based instruction has been defined as instructional strategies that have empirical support related to the use, efficacy, and generalization of such strategies across settings and subjects (e.g., modeling; Stoiber & DeSmet, 2010). Evidence-based interventions (EBI), specifically, are prevention or intervention practices that have "strong scientific support or research evidence" (Stoiber & Gettinger, 2016; pg. 121). Research regarding academic interventions is generally more developed than other areas of prevention and intervention research (e.g., social-emotional and behavioral; Stoiber & Gettinger, 2016). Although the focus on academic EBIs is a strength within intervention research broadly, within academic intervention research, reading has been examined more extensively than other academic areas (e.g., math), thus underscoring the need for additional research examining the effectiveness of interventions in other academic areas.

For all academic interventions, it is important to consider the instructional techniques demonstrating efficacy to ensure the effectiveness of such interventions. In a review of empirically supported academic interventions, there were four common intervention features across academic subjects. These four components included practice or drill, modeling with error correction, intervention intensity, and reward or reinforcement (Bramlett, Cates, Savina, & Lauinger, 2010). These features of empirically supported academic interventions align with instructional components previously discussed for the acquisition stage of the IH thus supporting the utility of the IH in guiding acquisition-based academic interventions. Another consideration for academic interventions is intervention efficiency. Efficiency has been a topic of focus within the intervention research due to the time restrictions for implementing interventions during the school day (Nist & Joseph, 2008). Efficiency within drill-based interventions is defined as the number of facts or words retained per instructional minute (Burns & Sterling, 2010). This section will first review acquisition-based academic interventions and student outcomes associated with those interventions. Next, acquisition-based interventions targeting broad math skills will be discussed. Then, the research related to three-types of drill-based interventions will be reviewed. Last, treatment acceptability research for academic interventions and specifically math interventions will be presented. Intervention efficiency will be discussed throughout this review of academic intervention research.

Acquisition Interventions

Students fall in the acquisition stage of the IH when they demonstrate slow and inaccurate responding to academic stimuli (Haring & Eaton, 1978). Acquisition interventions target students' accuracy in responding by explicitly teaching the skill through modeling, error correction, and repeated practice (Burns et al., 2010). According to the IH, students must first demonstrate accurate responding (i.e., they have acquired the skill) before being proficient in a skill. The acquisition stage serves as the foundation for the hierarchy and is of greatest importance for the initial stage of the learning process (Daly III et al., 1996). If students do not acquire an academic skill, they will not become proficient or be able to generalize and adapt the academic skills across subjects and activities in school. Thus, when students demonstrate difficulty acquiring academic skills, it is imperative to implement interventions directly targeting those skill deficits.

Acquisition interventions target foundational academic skills such as phonics (Benner et al., 2010), letter sounds (Rahn et al., 2015), sight words (Skinner et al., 1997; Kupzyk et al., 2011), word reading (Joseph, 2006; Volpe, Mulé et al., 2011), spelling (Cates et al., 2007) and math facts (Burns, 2005; Drevon & Reynolds, 2018). Early acquisition of basic reading skills (e.g., sight words) increases the acquisition of more advanced skills such as reading fluency and reading comprehension (Burns, 2004; Volpe, Mule et al., 2011). Similarly, acquisition of spelling has been linked to improvements in reading performance (Cates et al., 2007). Acquisition interventions targeting math facts have resulted in improved student performance on computation measures as well as generalization to other math skills such as word problem solving (Codding et al., 2010). Broadly, there is significant support showing that acquisition-based academic interventions positively impact later math skills (Kroesbergen & Van Luit, 2003). However, the majority of acquisition intervention research has examined reading skills underscoring the need for further research examining acquisition interventions for math.

Acquisition Interventions for Math

The acquisition of foundational math skills in early elementary school has been shown to be one of the strongest predictors of later math achievement (Duncan et al., 2007). However, the foundational skill of math fact knowledge is a common skill weakness among students who have math difficulties during elementary school (Gersten et al., 2005). Given the frequency of student math fact difficulties, identifying effective early math fact interventions to target students' math fact skills is important.

Some commonly implemented interventions targeting foundational math skills include using manipulatives in the concrete-representational-abstract approach (CRA; Burns et al., 2010; Agrawal & Morin, 2016), detect, practice, and repair (DPR; Poncy et al., 2015), comprehensive instruction (Poncy et al., 2015), taped problems (Poncy et al., 2015), cover-copy-compare (CCC; Burns et al., 2010; Cates et al., 2007), and drill-based interventions (Szadokierski & Burns, 2008; Mule et al., 2018). CRA is an intervention framework which first uses concrete or physical objects to teach math concepts (concreate phase). Next, instruction moves to the representational phase where students use drawings to represent and solve math problems. Finally, in the abstract stage, students use traditional notations (e.g., 2 + 2) to solve math problems (Bouck, Satsangi, & Park, 2018). CRA has been shown to be effective (Bouck et al., 2018) but has largely been examined within single-case design studies and within the context of teaching math computation with regrouping (Bouck et al., 2018). Additionally, within the school context, there are no clear criteria regarding when to move from one phase of the CRA sequence to the next (Bouck et al., 2018), which may impact intervention efficacy. Although CRA has empirical support for math computation, it is likely not the most efficient approach to aiding students in the acquisition of math facts but is more appropriate for introducing the concepts of computation (e.g., addition, subtraction), number sense, place value, and fractions (Agrawal & Morin, 2016). This is likely due to CRA's focus on conceptual understanding of computation procedures.

CCC is another intervention which has been examined across academic subjects and has been identified as an effective acquisition-based intervention for academic skills (i.e., reading, math, and writing/spelling). CCC provides modeling of the item (e.g., math computation or spelling of a word) and allows the student to independently respond and then evaluate the accuracy of their response. CCC has demonstrated effectiveness in increasing math computation accuracy (Poncy, Skinner, & Jaspers, 2007) as well as spelling (Cates et al., 2007). However, CCC does not include repeated practice of the correct response following inaccurate responding but rather focuses on error correction only. Although error correction is important within acquisition interventions, repeated practice represents an important component of acquisition interventions and has been shown to increase academic skill acquisition, specifically related to math facts (Reynolds et al., 2016). Moreover, when additional repeated responding was incorporated within CCC, students' performance further improved (Joseph, Konrad et al., 2012), but most CCC interventions do not include repeated practice. Even though there is empirical support for CCC, it may not be the most effective acquisition-based intervention to target math facts. This is likely due to the absence of important instructional features such as repeated practice following feedback which have been demonstrated to increase math fact retention and in turn intervention effectiveness (Burns et al., 2010). The exclusion of these important instructional features may then impact intervention efficiency because valuable intervention time is used on tasks that may not promote optimal skill development.

Similar to CRA, DPR includes multiple components to teach math computation (Musti-Rao & Plati, 2015). DPR has been shown to be an effective small group intervention for improving students' computational accuracy for subtraction (Poncy, Skinner, & O'Mara, 2006), multiplication (Poncy, Skinner, & Axtell, 2010), and division (Poncy, Fontenelle, & Skinner, 2013). First, during the detect stage, students respond to math problems presented through PowerPoint slides at a fixed interval of 1.5 seconds for each math fact. Next, students identify the problems they answered incorrectly via an answer key and practice those problems using

CCC. CCC serves as an opportunity to re-practice the facts answered incorrectly but it occurs for limited amount of time (i.e., 5 minutes) and each fact is typically practiced once (Musti-Rao & Plati, 2015). The third stage requires students to respond to the original math problems presented during the detect stage through a timed computation assessment. Last, students score their computation assessment and graph their results (Musti-Rao & Plati, 2015). Similar to CCC, the opportunity for repeated practice is limited with DPR and lacks direct feedback from the instructor related to their performance (i.e., using an answer key), which are both crucial instructional features identified in the acquisition-based intervention literature (Burns et al., 2010).

Another acquisition-based intervention which has demonstrated effectiveness in enhancing student's math fact accuracy is taped problems (TP) (Kleinert, Codding, Minami, & Gould, 2018). During a TP intervention, math facts are presented via a tape recording and students must provide the correct answer before the recording gives the answer. If the student provides an incorrect response or does not provide a response prior to an answer being given, they must write the answer to the math fact (Poncy et al., 2015). Similar to CCC, the response feedback students receive during TP and DPR is either self-monitored (Codding, Hilt-Panahon, Panahon, & Benson, 2009) or briefly provided by the interventionist (Miller et al., 2011). As previously discussed, corrective feedback is an important instructional strategy for the acquisition phase of the IH. When feedback is provided in multiple formats (e.g., self-monitored, peer, or adult), it is typically inconsistent and non-systematic. When instructional strategies like corrective feedback are not implemented as outlined within the IH model and in empirically supported ways, students' skill acquisition may be negatively impacted. The current study will fill such gaps in the research base by examining tier 3 drill-based interventions which incorporate systematic and immediate feedback.

Drill-Based Interventions and Math

Another common type of acquisition intervention is drill-based interventions. Drill-based tasks have been used to teach foundational academic skills and have been linked to improved performance on higher order skills across academic areas (Burns, 2004). Drill is the repeated rehearsal or practice of a skill (Burns, 2004) and has been examined across reading and math skills. One important component of drill interventions is the ratio of known to unknown stimuli presented during the intervention (Cooke, Guzaukas, Pressley, & Kerr, 1993). The "optimal ratio" of known to unknown items during drill interventions was important for providing an appropriate level of challenge during instruction (Burns, 2004; pg. 167). However, findings regarding the optimal ratio during drill interventions have been mixed, with ratios including 90% known and 10% unknown (i.e., IR; Tucker, 1989), 70% to 80% known and 15% to 30% unknown (Gickling & Thompson, 1985), 50% known and unknown (Neef, Iwata, & Page, 1980), and less than 50% known items (Robinson & Skinner, 2002). Meta-analytic research suggested that drill interventions with less than 50% known items was the least effective ratio for acquisition of a skill (Burns, 2004); however, this study did not identify the optimal ratio for known versus unknown items. One possible explanation for a lack of universal definition of optimal ratio is that the drill ratio may be specific to the academic skill area taught (Cooke et al., 1993; Burns, 2004). Therefore, it is important to understand the evidence regarding drill-based research across academic skills, but within the context of the present study, drill-based interventions for math are of particular importance.

Drill rehearsal has been shown to be one of the most effective approaches to target acquisition of math facts because they incorporate multiple components of effective acquisition interventions (Burns, 2004). The acquisition of new information (i.e., math facts) has been linked to the number of practice trials (Daly, Hintze, & Hamler, 2000) with a greater number of practice trials, or opportunities to respond (OTR), leading to increased skill retention (Volpe, Mulé et al., 2011). OTR has been defined as the number of opportunities a student has to practice an unknown stimulus and receive feedback from an instructor regarding their performance on that stimulus (Zaslofsky et al., 2016). The high number of OTR within drill interventions are theorized to be an underlying causal mechanism for the effectiveness of drill-based interventions specifically and acquisition interventions generally because students are provided with repeated practice of the skill (Burns, 2005; Volpe, Burns, DuBois, & Zaslofsky, 2011). OTR are not the only mechanism researchers have examined while developing drill-based interventions.

Drill interventions have also been developed by manipulating some of the features of IR and TD. For example, interventions that incorporate features of drill-based methods (i.e., IR and TD) include WordSheets (Mulé et al., 2018), Strategic Incremental Rehearsal (SIR; Kupzyk et al., 2011; January et al., 2017), and IR with exponentially increased spacing of unknown material (Swehla, Burns, Zaslofsky, Hall, Varma, & Volpe, 2016). Having peers serve as the interventionist also represents a promising practice for drill-based interventions (e.g., IR; Klingbeil et al., 2017). Although there is empirically support for math drill-based interventions (Burns, 2005), the findings lack generalizability because single-case designs have been most commonly used to examine their effectiveness. Thus, it is important to expand the research for acquisition interventions, specifically drill-based interventions, to other research models (i.e., group designs) as well as within the context of basic math skills. The present study aims to do both.

Incremental Rehearsal (IR). IR is a drill intervention that intersperses unknown material with known material (Burns, 2005; Tucker, 1989). The interspersing of known, easier, items has been linked to more on-task behavior as well as task preference (Cates & Skinner, 2002; McCurdy, Skinner, Grantham, Watson, & Hindman, 2011). The known and unknown items are presented at a ratio of 10% unknown and 90% known material (Burns, 2005). The first unknown item is presented to the student followed by the presentation of the first known item. When the unknown item is presented this first time, the interventionist models the correct answer before eliciting a response from the student. Next, the first unknown item is presented again followed by the first known and second known items. Then the first unknown item is presented a third time followed by the first, second, and third known items. This presentation order continues until all the items are presented (see Figure 2 for IR sequence). Following traditional IR procedures (Burns, 2005), after all nine items are presented, a new unknown item is introduced through modeling of the correct response. When the new unknown item is introduced, the original unknown item becomes the first known item in the sequence, and the ninth item (i.e., eighth known item) from the original sequence is eliminated. Traditional IR has been shown to be one of the most effective interventions for academic skill acquisition (Burns, 2005; Burns et al., 2010; Codding et al., 2011; Codding et al., 2009; Klingbeil et al., 2017; Varma & Schleisman, 2014). It embodies the instructional features described at the acquisition stage of the IH such as modeling, feedback, repeated practice, as well as high OTRs.

Repeated presentation of the unknown items within the traditional IR sequence allows for repeated practice, or high OTR, and multiple opportunities to receive corrective feedback

(McVancel et al., 2018). As discussed previously, high OTR has been linked to increased skill retention (Volpe, Mulé et al., 2011). Szadokierski and Burns (2008) examined the impact of OTR on retention of unknown words when using IR and found that the number of OTR had a significant and large effect on word retention 1 week after intervention. However, OTR is not the only feature that has been examined as a causal mechanism underlying IR's effectiveness. OTR has been compared to other variables such as response generation (Zaslofsky et al., 2016). Response generation is a concept within cognitive psychology referring to the production a response as opposed to simply reading or being supplied with the correct response (Volpe, Mulé et al., 2011). When a high number of OTR and high response generation were present, IR demonstrated the largest learning gains for students compared to low OTR and low response generation (Zaslofsky et al., 2016). Therefore, high response generation is likely another important instructional component to enhance students' skill acquisition.

Another important instructional component of IR is modeling. Drill-based interventions with modeling have been found to produce the largest effect sizes for treatment outcomes compared to other math fact interventions (Codding et al., 2011). IR uses modeling within the intervention procedures which contributed to intervention effectiveness (Codding et al., 2010). Modeling occurs both when the unknown item is first taught (e.g., two times one equals two; Codding et al., 2010) as well as during corrective feedback (i.e., providing the correct response and having the student repeat it; Joseph, Eveleigh, Konrad, Neef, & Volpe, 2012).

Another advantage of IR is that it can be tailored to each student's unique needs. The unknown and known items used within the intervention session are based on the items the student has and has not learned (Burns, et al., 2016). However, like all interventions, there are limitations to IR. For instance, when IR is implemented in the traditional sequence, the instructional time needed to implement the intervention is greater than other drill-based interventions (e.g., TD; Volpe, Mulé et al., 2011). The relatively longer amount of time IR requires has been a common critique of the intervention.

Efficiency has been a topic of focus within the intervention research and a major critique of traditional IR (Volpe, Mulé et al., 2011). There has been some research which has tried to address this critique; however, such studies have demonstrated mixed findings related to IR efficiency (Burns, Zaslofsky, et al., 2012; McVancel et al., 2018). These mixed findings could be attributed to the ambiguity and differences in the definition of efficiency across intervention studies (Burns, Zaslofsky et al., 2012). For example, efficiency has been defined based upon next day retention as well as maintenance assessments occurring one or two weeks after intervention (i.e., number of facts correctly stated one or two weeks after intervention per instructional minute; McVancel et al., 2018). Given such differences in the definition of efficiency, it is difficult to compare intervention efficiency across studies. In studies examining efficiency within the context of next day retention, IR has typically been less efficient compared to TD (Burns, Zaslofsky et al., 2012). However, when efficiency uses maintenance checks as part of the definition (i.e., facts retained 1 week after intervention per instructional minute), IR was more efficient than other drill-based methods (e.g., TD; Burns, Zaslofsky et al., 2012). Additionally, most of the intervention research examining efficiency has been within the context of reading (e.g., letter sounds and sight words; Burns, Zaslofsky et al., 2012). While the definition of efficiency contains inconsistencies and flaws, there have still been attempts to address such criticism of IR through manipulation of the IR stimuli presentation.

Although instructional time has been identified as a potential weakness of IR, taking into consideration students' individual acquisition rate has been shown to increase intervention effectiveness and intervention efficiency. Acquisition rate (AR) is defined as the amount of information a student can practice and later recall after one intervention session (Gravois & Gickling, 2002; Burns et al., 2016). It has been argued that AR is an important consideration when matching the intervention to students' individual needs (Burns & Parker, 2014; Burns et al., 2016). AR is determined by practicing all items until a student makes three errors during the practice of either known or unknown items (Burns, 2001). Using students' AR is important in determining when intervention instruction should end as well as how many unknown stimuli should be taught per intervention session. For example, if a student is only able to retain three new math facts per intervention session, it would be most efficient to discontinue the intervention session after the third fact has been taught. The use of students' AR led to more efficient use of instructional time as well as long-term retention compared to using predetermined set amounts (i.e., 2 or 8 stimuli; Haegele & Burns, 2015). Matching the number of unknown items taught to students' AR could address the instructional time weakness as well as enhance the individualization of IR to meet each student's learning needs. The present study will use AR research to inform the number of facts taught to students across intervention conditions.

In an effort to address the relatively long instructional time of IR, Volpe, Mulé and colleagues (2011) modified the IR presentation sequence such that the unknown item taught was removed after the presentation of all nine items, standardizing OTR across various intervention conditions. When OTR were controlled across IR and other drill interventions, such as TD, TD was more efficient than IR (Joseph & Schisler, 2007; Nist & Joseph, 2008; Volpe, Mulé et al., 2011). These results may be due to IR effectiveness being contingent on the number of times an unknown stimuli is presented throughout the intervention, which results in longer intervention time. When the unknown stimuli are not folded-in as known stimuli, the OTR to the stimuli is reduced, which likely compromises the instructional effectiveness of the IR approach. Instructional time was also held constant in an attempt to address the issue of IR taking a longer amount of time to implement. Minimal differences in effectiveness and efficiency between IR and TD were evident when instructional time was held constant (Volpe, Mulé et al., 2011). Given OTR have been shown to be an important mechanism underlying IR's effectiveness, it is likely that controlling, or limiting, OTR negatively impacted the intervention's effectiveness in Volpe, Mulé and colleagues' (2011) study. Even though holding OTR constant shortened the instructional time, removing the fold-in component of IR compromised a causal and theoretically based component of the intervention (i.e., repeated practice).

Manipulation of the spacing of the unknown facts in IR has also been examined in an effort to minimize instructional time (Swehla et al. 2016). In comparing traditional IR to IR with spacing increasing exponentially (IR-Exp; i.e., one unknown fact, one known fact, one unknown fact, two known facts, one unknown fact, four known facts, one unknown fact, eight known facts), traditional IR resulted in superior retention rates compared to IR-Exp, even though IR-Exp took half the amount of instructional time to implement. Thus, although traditional IR took longer to implement, the high OTRs within traditional IR, again, appeared to be an important component underlying its effectiveness. Given the evidence which supports the high OTR within traditional IR, future research should continue to examine new implementation procedures which maintain the key components of IR (i.e., OTR) while attempting to address shortcomings of the intervention (i.e., instructional time). The current study hopes to expand upon this point by comparing SIR to IR.

Traditional drill (TD). TD is a drill-based intervention targeting basic skill acquisition and presents only unknown stimuli to a student in a repeated fashion (Volpe, Mulé et al., 2011;

Burns, 2004). During the first round of practice, each unknown item is modeled by the instructor followed by prompting the student to provide the correct answer. On remaining trials, students are prompted to provide a response to the stimuli and corrective feedback is provided when needed (January et al., 2017). Practice of the items continues until all of the items can be answered correctly or a predetermined criterion has been met (e.g., predetermined amount of instructional time). TD has been shown to increase students' acquisition of sight words (Joseph, Eveleigh et al., 2012; MacQuarrie, Tucker, Burns, & Hartman, 2002; Nist & Joseph, 2008; Volpe, Mulé et al., 2011) reading comprehension (Joseph & Schisler, 2007), and math facts (Codding et al., 2011).

TD was shown to be a more efficient intervention compared to IR because, historically, participants have demonstrated a higher rate of next day fact or word retention per instructional minute (Burns & Sterling, 2010; Volpe, Mulé et al., 2011). Additionally, TD was shown to take less time to implement overall leading to increased efficiency. TD has also demonstrated greater efficiency than IR when instructional time is controlled (i.e., words learned per instructional minute; Burns & Sterling-Turner, 2010; Cates et al. 2003; Joseph & Nist, 2006; Nist & Joseph, 2008). However, efficiency when examined over the long-term yielded conflicting findings, as discussed previously. TD demonstrated greater efficiency in the short term (e.g., immediate retention), but was equally or less efficient compared to other drill-based interventions (e.g., IR) when calculated based on maintenance assessments (e.g., 1 week after intervention; Burns & Sterling, 2010).

Another limitation of TD is that it resulted in less generalization of learned words to other reading conditions compared to other drill-based interventions (e.g., IR; Joseph, Eveleigh et al., 2012). Thus, TD may be more effective for larger, immediate increases in unknown stimuli, but not the best intervention to promote generalization of learned information to other academic contexts. However, the time at which generalization measures have been administered has been inconsistent (e.g., immediately after intervention versus one week after the intervention; Haegele & Burns, 2015). As a result, additional research examining generalization of facts taught during drill interventions is needed. Moreover, much of the research examining TD has focused on word reading with minimal research related to the use of TD for math fact acquisition.

Despite these mixed findings, TD has also been compared to various other drill-based interventions of similar and different stimuli ratios such as sandwich drill, IR, and a novel drill procedure, WordSheets. Coulter and Coulter (1989) adapted the traditional drill technique to create the sandwich drill (SD) approach. SD uses a ratio of 30% unknown and 70% known stimuli, with the unknown stimuli mixed in with the known stimuli, and each stimulus is presented three times (MacQuarrie et al., 2002). SD resulted in greater retention of the unknown stimuli than the TD approach (MacQuarrie et al., 2002; Browder & Shear, 1996; Cooke et al., 1993). However, when comparing SD to TD and IR, IR led to overall greater retention of sight words (MacQuarrie et al., 2002).

TD has also been compared to other drill-based methods such as WordSheets (Mulé et al., 2018). WordSheets were derived from the TD approach with two key differences: (1) increased OTR and (2) presenting the stimuli in a context similar to passage reading (Mulé et al., 2018). In WordSheets, all of the words presented on a worksheet are unknown and are first modeled by the instructor and then practiced by the student. Next, the words are presented in random order through a WordSheet generated by a computer. The student practices reading the words presented on the WordSheet and the instructor provides corrective feedback throughout the intervention. When comparing the effectiveness of WordSheets and TD on retention of words practiced, TD led to greater overall retention. There were no differences between TD and WordSheets related to generalization and maintenance of words learned (Mulé et al., 2018). These findings may be reflective of the similarities in intervention procedures between TD and WordSheets and therefore, the difference in presentation mode (i.e., flashcard vs. worksheet) did not impact the underlying mechanisms consistent across TD and WordSheets (i.e., modeling, all unknown stimuli).

Although findings regarding the relative effectiveness of TD compared to other drill-based interventions have been mixed, TD has been shown to increase students' acquisition of basic academic skills (Volpe, Mulé et al., 2011). Despite the support for TD approaches to teaching basic skills, there is relatively little research examining TD for math fact interventions. One way to address these shortcomings could be to examine alternative drill-based interventions in the context of math facts. Additionally, a majority of the research base has used single-case designs to compare TD to IR and other drill methods. The current study will use a within-subjects design to compare TD and IR, thus improving upon the external validity of findings.

Strategic incremental rehearsal (SIR). SIR is a novel drill intervention which modified traditional IR by incorporating features of TD procedures (Kupzyk et al., 2011; January et al., 2017). When implementing SIR, all of the stimuli presented are unknown to the student, like TD, but the items are introduced in an incremental fashion, like IR, to maintain spaced practice (January et al., 2017). Similar to non-traditional IR (i.e., without fold-in of unknown items), in SIR, target stimuli are removed from the instructional sequence; however, SIR bases the removal and introduction of stimuli on the student's correct and incorrect responding to the stimuli and does not follow a preset sequence, as in modified IR (January et al., 2017). The number of unknown stimuli practiced during IR is limited during an intervention session, but with SIR, the number of stimuli practiced is dependent upon student's acquisition rate within the intervention session (Kupzyk et al., 2011). During a traditional IR session, three to five unknown stimuli can be presented within the intervention session with known items being presented potentially three times more than unknown items (Kupzyk et al., 2011; Burns, 2007; MacQuarrie et al., 2002; Nist & Joseph, 2008), which results in more time spent practicing the known stimuli than unknown stimuli. Additionally, IR often requires more instructional time compared to other drill procedures (i.e., TD; Cates et al., 2003; Nist & Joseph, 2008). SIR aims to increase intervention efficiency by introducing new unknown stimuli only after the student has practiced the previous unknown item without any errors (January et al., 2017). Preliminary findings suggested SIR is slightly more effective and efficient than IR when teaching sight words (January et al., 2017). However, SIR has not been compared to TD nor has it been used to examine math fact instruction. Although SIR includes the use of AR in the implementation procedures, it is a relatively novel intervention and therefore has limited research (Kupzyk et al., 2011; January et al., 2017). The current study aims to add to the SIR intervention research as well as expand the research to the subject of math facts.

Kupzyk and colleagues (2011) compared the effectiveness of IR and SIR for sight word acquisition through an A-B-A-B design across four participants. Their findings showed SIR resulted in greater gains in sight word acquisition compared to IR. SIR also resulted in greater maintenance of the sight words two weeks after intervention completion. January and colleagues (2017) also compared SIR and IR for sight word acquisition through an alternating treatments single-case design, but procedures for removal of learned unknown words differed from the original SIR procedures used by Kupzyk and colleagues (2011) which were described earlier in this section. In January and colleagues' (2017) study, words were considered acquired when a student responded correctly to the word during three consecutive intervention sessions across three days. The researchers then tested the words the student had acquired on two separate days to check for maintenance of the word. If the student did not respond correctly to the stimuli, the word was added back to the student's target words for practice, but it was not modeled when re-introduced to the student. Like January and colleagues (2017), Kupzyk and colleagues (2011) also did not include modeling when stimuli were reintroduced. This is a shortcoming of the implementation of SIR because, as discussed previously, modeling is a key instructional component when students are acquiring a new skill. In the present study, modeling will be used when first introducing a new fact and when providing corrective feedback; however, given the current study design, facts will not be reintroduced across days. The Kupzyk and colleagues (2011) and January and colleagues (2017) studies are the only known studies examining SIR. Therefore, researching examining SIR targeting other academic areas (i.e., math) and compared to interventions other than IR (i.e., TD) is warranted. Additionally, SIR has only been examined through single-case design research which has limited generalizability beyond the students who participated. Additional research examining SIR through group designs in an effort to increase generalizability is therefore needed. The present study aims to expand the SIR research base to include application to math facts which will be examined through a group-based research design.

Treatment Acceptability

Treatment acceptability is clients' (i.e., students) and other stakeholders' (i.e., teachers) judgment regarding whether or not they perceived treatment procedures as "appropriate, fair, and reasonable for a problem" (Arra & Bahr, 2005; Kazdin, 1981). In other words, treatment acceptability evaluates whether a client perceived the treatment as effective and demonstrated a preference for the treatment. In the case of academic interventions, the client is typically the student receiving the intervention or the teacher or interventionist implementing the intervention. Treatment acceptability is an important consideration because it can inform whether the intervention is targeting the individual skill with which a student has difficulties (Mautone et al., 2009) as well as evaluate student engagement in instruction.

Teacher acceptability has been of greater focus compared to student acceptability within the intervention research base (Arra & Bahr, 2005). Research on teacher acceptability has largely focused on rating treatments for behavioral interventions (McCurdy & Cole, 2014; Cowan & Sheridan, 2003; Mautone et al., 2009), interventions developed in consultation (Cowan & Sheridan, 2003; Mautone et al., 2009), or assessment procedures (Eckert & Shapiro, 1999; Mautone et al., 2009). Teacher acceptability is important because it can impact implementation fidelity when they are implementing the intervention. Similarly, taking teachers' opinions into consideration during intervention evaluation and development can positively impact implementation fidelity because teachers may feel their opinion and expectations are valued (Mautone et al., 2006). Although there has been some research examining teacher intervention acceptability, there is a significant gap in the literature regarding student acceptability of academic interventions and very little research examining student acceptability of math interventions specifically (McVancel et al., 2018; Mong & Mong, 2010; McCallum, Skinner, Turner, & Saecker, 2006; Arra & Bahr, 2015).

There has been some recent research which has examined student acceptability of math interventions including CCC and Math to Mastery (MTM; Mong & Mong, 2010), taped problems (McCallum et al., 2006), IR (McVancel et al., 2018; Volpe, Mulé et al., 2011), and TD (Volpe, Mulé et al., 2011) as well as broad types of math intervention frameworks (e.g., cognitive, behavioral; Arra & Bahr, 2015). Across broad types of math interventions (e.g., cognitive, behavioral, and traditional), students reported high levels of acceptability across conditions (Arra & Bahr, 2015). However, the students received only one of the three intervention conditions and therefore these results may not be comparable. When MTM and CCC were compared, CCC was identified by students as the preferred intervention approach (Mong & Mong, 2010). One possible explanation for the difference in student acceptability could be due to MTM taking twice as long as CCC to implement. Student participants also rated TP highly (McCallum et al., 2006). However, there was not a comparative intervention condition so it is unclear whether TP would still receive high ratings when another intervention condition is present. When comparing two drill-based interventions, IR and TD, student acceptability was split in that half of the students preferred IR while the other half preferred TD (Volpe, Mulé et al., 2011). Students indicated they preferred one intervention over the other because they liked having known words as part of their drill practice (IR) while conversely others liked having only unknown words as part of their drill practice (TD; Volpe, Mulé et al., 2011). There is currently no student acceptability research which examines the preference between IR and TD within the context of math facts. The present study aims to expand the student acceptability research for drill interventions to include math facts.

Although some recent research has incorporated examination of treatment acceptability of math interventions for students and teachers broadly, there is still a gap in the research base regarding student acceptability when comparing multiple math interventions, specifically drill-based interventions. Given this gap in the research base and the importance of student engagement during intervention (Mautone, et al., 2009), treatment acceptability is an important area for future research to examine, particularly in the context of math fact interventions.

Summary

Drill-based interventions are some of the most commonly implemented interventions to target students' acquisition of basic academic skills and include IR (Burns, 2005; Burns et al., 2010; Codding et al., 2011; Codding et al., 2009; Klingbeil et al., 2017; Varma & Schleismann, 2014), TD (Joseph, Eveleigh et al., 2012; MacQuarrie et al., 2002; Nist & Joseph, 2008; Volpe, Mulé et al., 2011), and SIR (Kupzyk et al., 2011; January et al., 2017). However, drill-based intervention research has largely focused on the examination of drill-based interventions for early reading skills (e.g., sight words), with less research examining the impact of drill-based interventions for improving math skills. There have also been mixed findings regarding the comparative effectiveness of IR and TD for teaching math facts (Volpe, Mulé et al., 2011; Burns, 2005). However, SIR may be a promising alternative to these drill-based intervention methods based on preliminary research. The current study will therefore compare the effectiveness and efficiency of IR, TD, and SIR on the retention and maintenance of multiplication facts through a group research design. This study will add to the drill-based intervention literature for math facts as well as expand the acquisition-based intervention literature by using a within-subjects design study.

Literature Review Summary

This chapter reviewed the research related to service-delivery frameworks, theoretical frameworks, and evidence-based math interventions. Service-delivery frameworks, such as RtI, serve as a vehicle to implement evidence-based instruction, additional academic supports (e.g., interventions), and assessment for students experiencing academic difficulties (Cusumano et al., 2014). Currently, math interventions are largely understudied compared to other academic skills (i.e., reading) within the context of RtI (Lembke et al., 2012; Hughes & Dexter, 2011).

Specifically, there is a gap in the research related to tier 3 math interventions which target acquisition of foundational math skills, such as math facts. Similarly, the use of theoretical frameworks to guide the examination of such interventions is inconsistent and limited within math intervention research (Burns, 2011). Using theoretical frameworks to guide and inform academic interventions research not only increases internal validity but also allows for the adaptation of interventions across environments (e.g., schools; Volpe & Suldo, 2014; Burns, 2011).

The research base for math interventions, specifically drill-based interventions for math facts, is also limited. The studies that have examined drill-based interventions for math have largely used single-case research designs which limits the generalizability of the findings. There has also been limited research examining student acceptability related to math intervention (Arra & Bahr, 2015; Mong & Mong, 2010; McCallum et al., 2006; McVancel et al., 2018; Volpe, Mulé et al., 2011) and specifically related to drill-based interventions such as IR and TD (Volpe, Mulé et al., 2011). Student acceptability is important in evaluating both student engagement in intervention as well as students' perception of whether or not the intervention helped them (Mautone et al., 2009). Given that math fact acquisition has been identified as a common area of math difficulty for elementary students, targeting these skills through drill-based intervention holds promise in addressing the overall low national proficiency levels in math. Thus, the current study aims to address these gaps outlined in this literature review by using the IH to guide the examination of the effectiveness, efficiency, maintenance, and acceptability of three drill-based math fact interventions through a within-subjects research design.

Research Questions

The following research questions will guide the present study:

1. What are the comparative effects of traditional incremental rehearsal (IR), traditional drill (TD), and strategic incremental rehearsal (SIR) on students' multiplication fact retention?
    a. What are the differences in multiplication fact retention between IR and TD?
    b. What are the differences in multiplication fact retention between IR and SIR?
    c. What are the differences in multiplication fact retention between TD and SIR?
2. What are the comparative effects of traditional incremental rehearsal (IR), traditional drill (TD), and strategic incremental rehearsal (SIR) on students' multiplication fact fluency?
    a. What are the differences in multiplication fact fluency between IR and TD?
    b. What are the differences in multiplication fact fluency between IR and SIR?
    c. What are the differences in multiplication fact fluency between TD and SIR?
3. What are the comparative effects of traditional incremental rehearsal (IR), traditional drill (TD), and strategic incremental rehearsal (SIR) on students' multiplication fact maintenance?
    a. What are the differences in multiplication fact maintenance between IR and TD?
    b. What are the differences in multiplication fact maintenance between IR and SIR?
    c. What are the differences in multiplication fact maintenance between TD and SIR?
4. What are the differences in intervention efficiency among the three intervention conditions?
    a. What are the differences in efficiency between IR and TD?
    b. What are the differences in efficiency between IR and SIR?
    c. What are the differences in efficiency between TD and SIR?

5. To what extent are there differences in treatment acceptability across intervention conditions?

CHAPTER 3

Method

Participants and Setting

This study was conducted in partnership with a local community public intermediate school located around a mid-size Midwestern city. The participating school included 454 students in third through sixth grade. Out of the 454 students, approximately 251 (55.3%) identified as male. The majority of the students identified as Caucasian (88.5%), 7.3% identified as Latinx, 3.7% identified as two or more races, and less than 1% identified as African American. Additionally, 56.8% of the students were eligible for free or reduced-price lunch (National Center for Education Statistics, 2018).

A total of 268 consent forms were sent home with all fourth (*n* = 133) and fifth (*n* = 135) grade students at the participating school. Of the 268 consent forms, 67 consent forms were returned, and the guardians of 53 students provided consent to participate in the study. Of the 53 students whose guardians provided consent, 46 provided assent. The students who provided assent participated in pre-testing and 37 of those students met study inclusion criteria. The pre-test inclusion criteria required students to know less than 132 multiplication facts out of 144 multiplication facts presented (i.e., facts 1 through 12). One student who met study inclusion criteria did not participate because he was absent prior to a school break and therefore making up his data would have resulted in a significant deviation from study procedures.

A total of 36 students in fourth (*n* = 20) and fifth (*n* = 16) grade participated in this study. Of those 36 students, 18 (50%) were male. The majority of the participants identified as Caucasian (80.5%), five (13.9%) identified as Multiracial, and two (5.6%) identified as Latinx. Additionally, eight (22.2%) of the students were eligible for special education services in the areas of Speech (*n* = 3), Language (*n* = 2), Specific Learning disability (SLD; *n* = 2), Other Health Impairment (OHI; *n* = 1), and/or Autism Spectrum Disorder (ASD; *n* = 2). Two of the eight students were eligible under more than one category (i.e., ASD and Language Impairment, and ASD and OHI). Of the eight students eligible for special education services seven of the students identified as Caucasian (87.5%) and one identified as Latinx (12.5%).

Measures

Known and unknown multiplication facts. Participants' known and unknown multiplication facts were determined prior to the start of intervention implementation. All participants were administered 144 multiplication facts (facts 1 through 12) via flashcards following Burns' (2005) procedures. The multiplication facts were written horizontally on an index card (e.g., 3 x 4) with numbers that were approximately 2 inches in height. There was no equal sign written on the card and the numbers were written in black marker (Nist & Joseph, 2008). All multiplication facts 1 through 12 were presented including each commutative property form (e.g., 1 x 2 *and* 2 x 1). The primary researcher and graduate students presented one multiplication fact flashcard at a time and participants had 3 s to provide a response. The facts were presented in a random order. Participants' correct (known) and incorrect (unknown) responses to each multiplication fact were recorded. Known facts were defined as multiplication facts to which the participant provided a correct response within 3 s, including self-corrections. Unknown facts were defined as incorrect responses, no response, or a response (correct or incorrect) provided after 3 s to the multiplication fact. All 144 multiplication facts were administered across three separate days to ensure each student's known and unknown facts were accurate. In order for a fact to be considered known at the end of pre-testing, the student had to respond correctly to the multiplication fact across all three pre-test sessions. If a student responded incorrectly to a presented multiplication fact across all three pre-test sessions, the fact was considered unknown. Multiplication facts to which students did not respond correctly across all three days of pre-testing were not used as either knowns or unknowns.

Fluency. At pre-test and post-test, participants were administered three subskill mastery measures (SSMM) across three separate days. Standardized curriculum-based measures (CBM) administration and scoring procedures were followed (Shinn, 1989) for the SSMM. Students had two minutes to complete the SSSM multiplication fact probe. Students' digits correct per minute (DCPM) were calculated for each probe following administration. The DCPM were calculated by dividing the total digits correct on each probe by 2 (i.e., 30 digits correct total / 2 min = 15 DCPM). Participants' median DCPM score across the three probes served as the students' pre-test score. CBM, such as SSMM, have demonstrated adequate reliability for measuring student skill performance (Ketterlin-Geller & Yovanoff, 2009). SSMM provided meaningful data to monitor student progress (Fuchs, 2004) and for use within RtI (Ysseldyke, Burns, Scholin, & Parker, 2010). Additionally, SSMM probes generated through websites showed moderate test-retest (*M* - Fluency = 0.66; *M* - Accuracy = 0.68) and alternate form reliability (*M* - Fluency = 0.66; *M* - Accuracy = 0.65) for fluency and accuracy (Strait et al., 2015).

Retention. Participants' retention of multiplication facts taught during intervention sessions was assessed at three different post-test time points. The day immediately following an intervention session, retention of the multiplication facts taught during the previous day's intervention session was assessed. For example, multiplication facts taught during the IR intervention session implemented on Monday were assessed for retention on Tuesday prior to the next intervention condition being implemented. A fact was considered retained if the student provided the correct response to the fact within 3 s. These retention post-test assessments of the taught multiplication facts followed the same procedures used to determine known and unknown facts previously described.

Maintenance. Participants' maintenance of taught multiplication facts was assessed two weeks after each of the intervention sessions. At each maintenance assessment, the four facts evaluated for retention two weeks prior were assessed. For example, if a student received post-test retention assessment for the IR intervention session on Wednesday October 16, TD October 17, and SIR October 18, the maintenance assessment for the facts took place on Wednesday October 30 (IR), Thursday October 31 (TD), and Friday November 1 (SIR), respectively. Thus, after three days of maintenance assessments, all twelve math facts across the three intervention conditions were evaluated. Maintenance assessments of the multiplication facts followed the same procedures used to determine known and unknown facts and retention previously described.

Efficiency. Intervention efficiency was calculated for each of the three intervention conditions. Efficiency was calculated by computing the number of multiplication facts retained per instructional minute (Cates et al., 2003; Swehla et al., 2016). The number of facts retained in each condition was divided by total intervention session time in seconds and then multiplied by 60 to obtained the number of facts retained per minute (i.e., $\frac{\text{Session Facts Retained}}{\text{Intervention Session Time}} \text{x } 60$; Cates et al., 2003; Nist & Joseph, 2008).

Treatment acceptability. When identifying effective interventions, it is also important to consider students' perspectives related to the value and effectiveness of the instructional procedures being examined (Nist & Joseph, 2008). Therefore, the current study assessed students' treatment acceptability of the interventions. After each intervention session was completed, the participating students completed a treatment acceptability survey. Thus, students completed the survey three times, once for each intervention condition. The survey was adapted from the Children's Intervention Rating Profile (Elliott, 1986). Students rated five statements using a four-point Likert scale. The four response options within the Likert scale were: 1 = I do not agree, 2 = I kind of disagree, 3 = I kind of agree, and 4 = I Agree. The five statements that students rated included:

1. This way of practicing multiplication facts is a good one.
2. Other ways to practice multiplication facts are better than this one.
3. Working with a helper to practice multiplication facts is a good idea.
4. I like this way of practicing multiplication facts.
5. I think practicing multiplication facts this way will help me do better in school.

The interventionist read the statements and response options to the student. Students circled their responses after the interventionist read both the statements and response options. However, due to interventionist error, 5 treatment acceptability scores had to be excluded from data analyses because students circled a score in between Likert items (e.g., circling the space between 1 and 2) on the scale, so scores could not be calculated.

Materials

Pre-test. The materials that were required to complete pre-testing included flashcards of all 144 multiplication facts for facts 1 through 12, a record sheet to record students' known and unknown facts, and SSMM probes. Pre-testing occurred across three days, thus one unknown vs. known record sheet and three SSMM probes were needed during this phase of the study.

Intervention. During each intervention session, the materials required included multiplication fact flashcards, a timer, a pencil, and a record sheet to record intervention session time, fact retention, and fact maintenance two weeks later.

Post-test. During the three retention assessments (i.e., post-tests) following the implementation of each intervention condition, the materials required included the flashcards with the multiplication facts taught during the intervention session the day prior, a timer, a pencil as well as three SSMM probes. Other materials that were needed included small rewards (e.g., pencils, puzzles, bracelets, Rubik's cubes) students earned at the end of the study for giving their best effort during pre-test, intervention, and post-test.

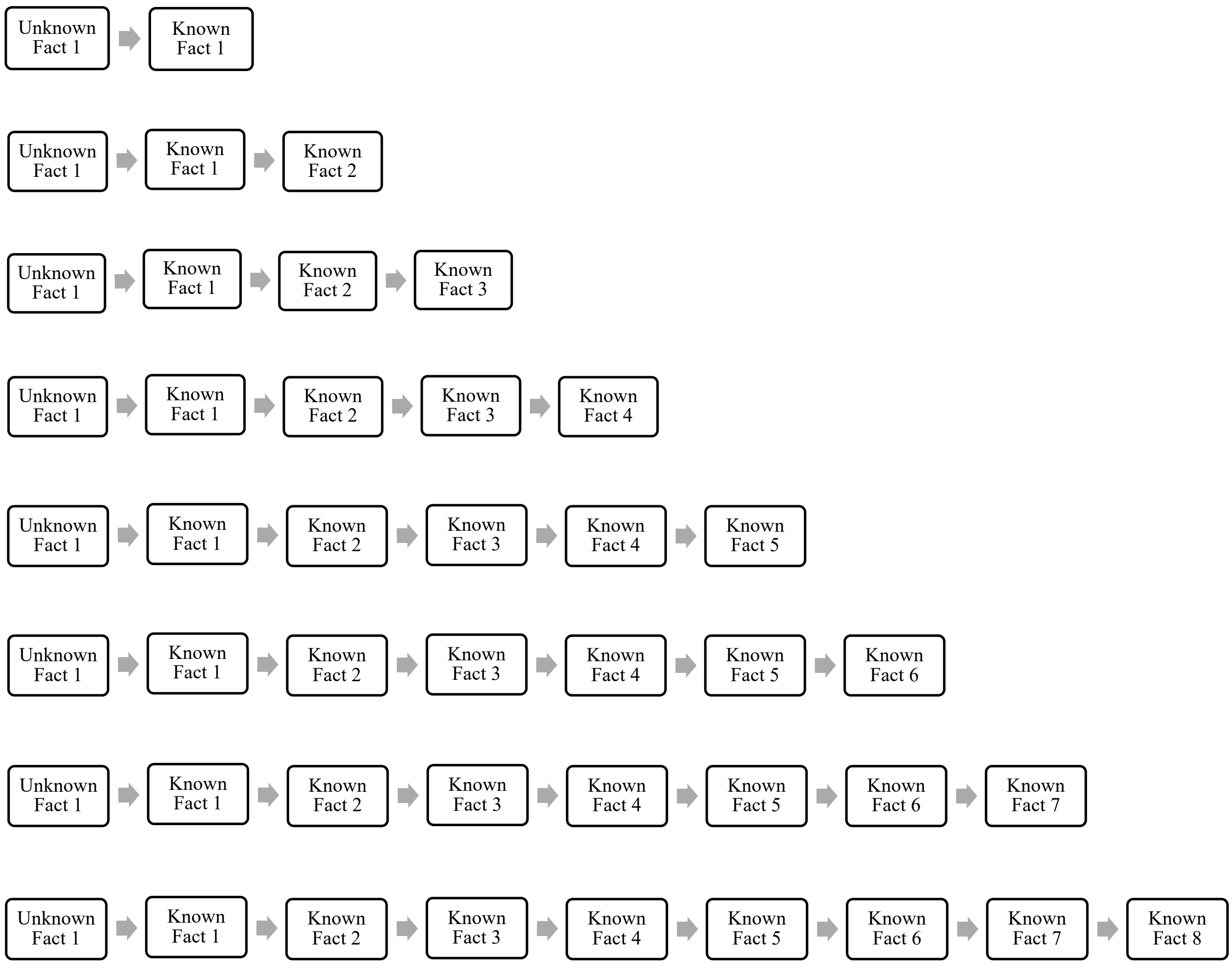
Interventions

All students participated in all three intervention conditions, Incremental Rehearsal (IR), Traditional Drill (TD), and Strategic Incremental Rehearsal (SIR). Across all students, the unknown facts determined during pre-testing were randomly assigned to each intervention condition. The known facts used for the IR condition were also randomly assigned from each student's pool of known facts. Additionally, the intervention condition implementation order was randomly assigned across participants.

The interventionists implemented scripted intervention protocols for each intervention condition to promote implementation fidelity across interventionists and conditions. At the start of each intervention session, the interventionist started a timer when the first unknown fact was presented. The interventionist stopped the timer after the student provided their response to the last multiplication fact presented. The total time of the intervention session was used to calculate efficiency for each of the intervention conditions. Students rehearsed four unknown multiplication facts for each of the three intervention conditions, based on previous intervention research examining acquisition rate (AR; Haegele & Burns, 2015), and therefore rehearsed twelve unknown multiplication facts total over the course of the study. AR is defined as the amount of information that can be taught during one intervention session and can be recalled at least one day after intervention (Haegele & Burns, 2015; Gravois & Gickling, 2002). Following a student error on either known or unknown facts, the interventionist provided corrective feedback using the following standardized language: “That’s not quite right. Let’s try again. BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?”.

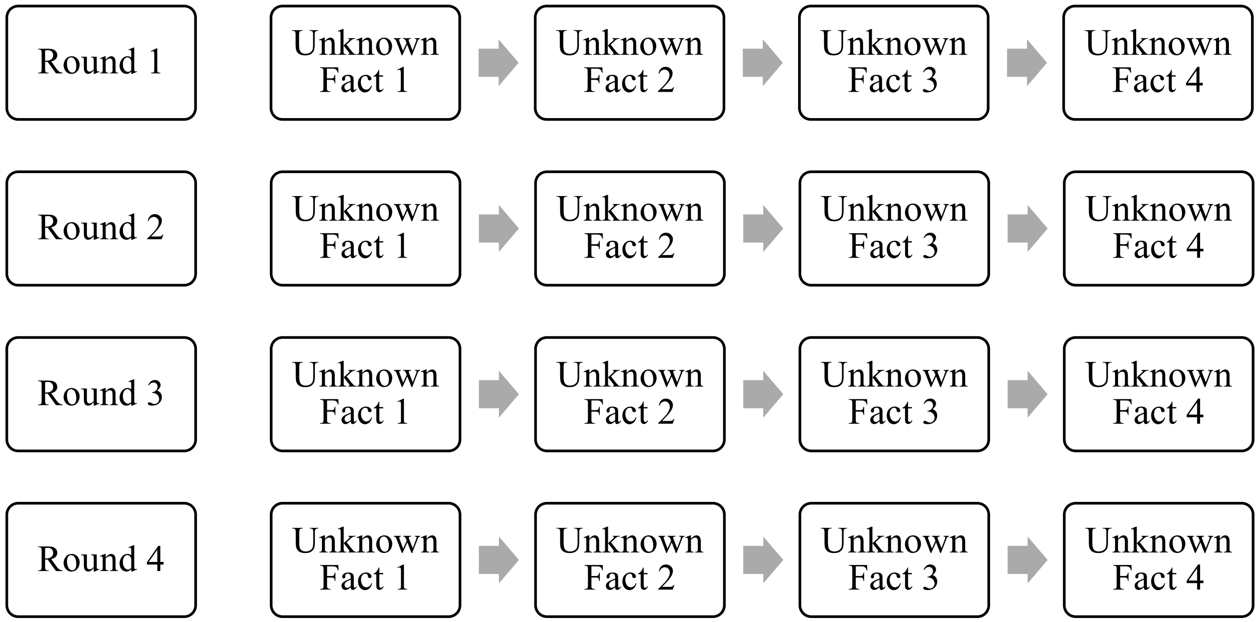
Incremental rehearsal. For the IR condition, the interventionist taught the unknown multiplication facts following the traditional standardized IR administration procedure (Burns, 2005). The interventionist modeled the correct response for the first unknown fact by stating, “BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?”. If the student provided a correct response, the interventionist responded with “That’s right!” and then presented the first known fact. If the student responded with an incorrect answer to any of the known or unknown multiplication facts presented, the interventionist responded by saying, “That’s not quite right. Let’s try again. BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?”. Following the teaching and presentation of the first unknown fact and then the first known fact, the first unknown fact was then presented again. When the unknown fact was presented a second time, the interventionist did not model the fact. Then, the interventionist presented the first and second known facts (i.e., unknown fact 1, known fact 1, known fact 2). New known facts were added to this sequence following the same procedure until all eight of the known facts were presented. After the initial sequence, the first unknown fact became the first known fact and a new unknown fact (unknown fact 2) was introduced. Additionally, the last known fact (known fact 8) was removed from the sequence. This sequence was repeated for the remaining three unknown facts. See Figure 2 for an outline of round one of the IR sequence.

Figure 2. Round one of Incremental Rehearsal presentation sequence.



Traditional drill. For the TD condition, interventionists taught four unknown multiplication facts through individual presentation of each unknown fact. The first time the interventionist presented each of the four unknown facts, they stated, “BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?”. There were no known facts presented as part of the TD intervention condition. In line with previous research, each unknown fact was presented in four rounds, resulting in each fact being presented four times, and facts were not shuffled in between each round (Nist & Joseph, 2008). See Figure 3 for an outline of the TD sequence.

Figure 3. Traditional Drill presentation sequence.



Strategic incremental rehearsal. For SIR, a total of four unknown multiplication facts were taught during the intervention session. January and colleagues' (2017) SIR procedures were followed. First, the interventionist modeled the first unknown fact by stating, "BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?". Positive praise such as "That's right!" was provided if the student provided a correct response and the interventionist provided corrective feedback if the student provided an incorrect response. Corrective feedback followed standardized procedures, "That's not quite right. Let's try again. BLANK times BLANK equals BLANK. What does BLANK times BLANK equal?". Next, the second unknown multiplication fact was presented using the same modeling script used for the first unknown fact. Again, either positive praise or corrective feedback was provided based on the student's response. The first and second unknown facts were then shuffled. Then, the first and second unknown facts were individually presented again. In order for a new unknown fact to be introduced, the student had to respond accurately to both unknown facts within the same round.

After the student responded accurately to both unknown facts, the facts were shuffled, and the third unknown fact was introduced following the same procedures as unknown facts 1 and 2. The same procedure was implemented for unknown fact four. In order for the intervention session to end, the student had to respond accurately to each of the four facts during two consecutive presentations (January et al., 2017; Kupzyk et al., 2011).

Procedures

The researcher obtained a letter of support from the participating school's principal. After the letter of support was obtained, an application was submitted to the university Institutional Review Board (IRB). Upon IRB approval, consent forms were sent home with all fourth and fifth grade students at the participating school. With approval from the school, and the discretion of the teachers, incentives were provided (e.g., pencil, piece of candy, school PBIS dollar) to students who brought their signed consent form back. Students received the incentive as long as their guardian signed the form, not based on whether their guardian provided consent for participation. The students whose guardians provided consented for participation met with the primary researcher or a graduate student to review assent. First, the researcher reviewed the assent form and explained the purpose of the study using age-appropriate language. During this time, the student was able to ask questions and indicated if they wished to participate in the study. If the student assented to participate in the study, they signed the assent form. If the student did not want to participate in the study, they were thanked for their time and sent back to class. If assent was received, pre-testing began on a day after assent was received.

Pre-tests were administered across three days. During each pre-testing session, the primary researcher or a graduate student presented students with all 144 multiplication facts 1 through 12 and one SSMM probe of multiplication facts. After pre-testing was completed, the primary researcher determined if the participant met the study inclusion criteria (i.e., at least 12 unknown multiplication facts). If the student met inclusion criteria, the primary researcher then randomly assigned the participant's unknown multiplication facts to one of the three intervention conditions. The first step of random assignment was to assign a number to each student's unknown facts (i.e., 1 – 12) and each intervention condition (i.e., 1 = IR, 2 = TD, 3 = SIR). Next, a random number generator was used to assign each student's unknown facts to each of the three conditions. A random number generator was then used to randomly assign intervention implementation order for each participant. Across students, the known facts used within the IR intervention were assigned a number and then a random number generator was used to select eight of the known facts to include in each students' IR known pool.

One to two weeks after pre-testing was completed, participants received their first intervention condition followed by post-test session one the next day. One the same day, immediately following post-test session one, students participated in their second intervention condition. Post-test session two was conducted the following day, after which participants received intervention condition three. Then, post-test session three occurred the day immediately after intervention condition three was implemented. Last, participants received maintenance assessments across three days which took place two weeks after each post-test session. Following each intervention session, students completed a treatment acceptability survey regarding each intervention condition. All pre-test, intervention, post-test, and maintenance assessment sessions took place during tier 2 intervention time or non-core instruction (e.g., specials) in an available classroom, sensory room, or conference area. Below, Table 2 outlines study procedures.

Table 2

Study procedures

Pretest	Intervention 1	Posttest 1	Intervention 2	Posttest 2	Intervention 3	Posttest 3
Known & unknow math fact identification* SSMM Probe*	IR *or* TD *or* SIR Acceptability Survey	Retention Assessment SSMM probe (3) Maintenance Assessment**	IR *or* TD *or* SIR Acceptability Survey	Retention Assessment SSMM probe (3) Maintenance Assessment**	IR *or* TD *or* SIR Acceptability Survey	Retention Assessment SSMM probe (3) Maintenance Assessment**

**Administer multiplication facts (1 – 12)/SSMM probe 3 times on 3 separate days.*
***Maintenance assessments will occur 2-weeks after the post-test session was implemented.*

Implementation fidelity. Prior to data collection, the author trained and assessed five interventionists (i.e., school psychology graduate students) in the intervention conditions and assessment procedures. Interventionists were not permitted to implement interventions with students until they successfully implemented each procedure with 100% accuracy on competency assessments during training. A procedural checklist was used both during interventionist competency assessments prior to beginning data collection and implementation fidelity assessments during data collection. Fidelity assessments were completed for 21% of the TD intervention sessions and 25% of the IR and SIR intervention sessions and across all interventionists. Implementation fidelity was calculated by dividing the number of correctly implemented intervention steps by the total number of intervention steps.

Most implementation fidelity assessments were completed in-person; however, when in person fidelity assessments could not occur due to problems with schedule logistics, the first author video recorded intervention sessions with IRB approval and parental consent. When video recordings were conducted, the video frame included only the interventionist's face with the students' body and face not in the video frame. The primary researcher stored the recordings securely in a locked filing cabinet in her locked on-campus office. The primary researcher and one of the school psychology graduate students reviewed the videos using the same fidelity checklists implemented during live observations. Of the 26 fidelity assessments across all three conditions (IR = 9; TD = 8; SIR = 9), 19.2% were completed through video recordings. Interrater observation agreement (IOA) for fidelity video reviews was 100% across both raters. After the videos were reviewed for fidelity purposes, they were deleted. Implementation fidelity for in-person observations and video observations ranged from 90.4% to 100% and averaged 99.44% across all conditions.

Research Design and Analyses

A within-subjects design was used to evaluate the comparative effectiveness and efficiency of the three intervention conditions. Students were pre-tested on their multiplication facts and participated in retention assessments (post-tests) following the implementation of each intervention condition. Each student participated in each of the intervention conditions, therefore serving as their own control. Prior to data collection, a power analysis for a within-subjects design was conducted to identify the minimum number of participants needed to detect statistically significant differences with estimated power of 0.80. A priori analyses indicated that a sample of approximately 36 participants was needed. For effect sizes the following metrics were used to interpret results: Kendall's W (.1 = small, .3 = medium, .5 = large; Tomczak & Tomczak, 2014), eta squared (.02 = small, .13 = medium, .26 = large; Bakeman, 2005), and partial eta squared (.01 = small, .09 = medium, .25 = large, Cohen, 1988).

The outcome variables (DV) for the present study included multiplication facts retained, median fluency scores (DCPM) on the SSMM probes, maintenance of multiplication facts, intervention efficiency (i.e., multiplication facts retained per instructional minute), and students' intervention acceptability. The independent variable (IV) for all six research questions was intervention condition (i.e., IR, TD, SIR).

A Friedman rank test was conducted to answer research questions one, three and four to examine differences in multiplication facts retained (Research Question 1), differences in maintenance of multiplication facts two-weeks after intervention (Research Question 3), and differences in intervention efficiency (i.e., facts retained per minute; Research Question 4) across intervention conditions. A nonparametric analysis was used given the outcome variables were ordinal in nature and therefore normality was violated for both outcome variables. For the sub-research questions under Research Question 4, a Wilcoxon signed-rank test was used to compare the efficiency of each of the three intervention conditions to one another (i.e., IR vs. TD, TD vs. SIR, & IR vs. SIR).

A one-way repeated measures analysis of variance (ANOVA) was conducted for research questions two to analyze differences in intervention effectiveness related to multiplication fact fluency (DCPM on SSMM probes; Research Question 2). No follow-up analyses were conducted.

A Friedman rank test was also conducted to evaluate treatment acceptability (Research Question 5) across intervention conditions. No follow-up analyses were conducted.

In addition to the examining overall outcomes, differences among racial groups was also examined during follow-up analyses. It is important to examine differences among racial groups in order to attend to the individual and cultural differences (e.g., race) which may impact intervention outcomes (Sander, in press). Thus, intervention effectiveness, efficiency, and treatment acceptability outcomes were examined among the three racial groups represented in the sample (i.e., White, Latinx, and Multiracial). For retention, maintenance, efficiency, and treatment acceptability, a Kruskal-Wallis was used to examine differences among racial groups (Howell, 2012). For the fluency outcome variable, an ANCOVA was used to examine differences, with pre-test fluency scores serving as the covariate. See Table 3 for an overview of the analyses.

Table 3

Study analyses

Research Question	Analyses
Initial Analyses for RQ2.	T-test for grade-level T-test for gender
RQ1. Multiplication fact retention	Friedman rank test Kruskal-Wallis for differences among racial groups
RQ2. Multiplication fact fluency	One-way repeated measures ANOVA ANCOVA for differences among racial groups
RQ3. Multiplication fact maintenance	Friedman rank test Kruskal-Wallis for differences among racial groups
RQ4. Intervention efficiency	Friedman rank test Wilcoxon signed-rank test (if significant results from Friedman rank test are present) Kruskal-Wallis for differences among racial groups
RQ5. Treatment acceptability	Friedman rank test Kruskal-Wallis for differences among racial groups

CHAPTER 4

Results

Preliminary Analyses

A t-test was conducted prior to examining the research questions to determine whether there were significant differences in pre-test SSMM DCPM scores (i.e., multiplication fact fluency) across grade level and gender. These tests did not result in statistically significant differences across grade (t = -.293, p = .771) or gender (t = -1.367, p = .181). Therefore, all additional analyses were conducted by collapsing the data across grade and gender.

Assumptions of Analyses

The assumptions of a one-way repeated measures ANOVA (i.e., independence, normality, sphericity) were initially evaluated for the outcome variables for all research questions (i.e., retention, fluency, maintenance, efficiency, treatment acceptability) to determine the appropriateness of using a one-way repeated measures of ANOVA to analyze the data. Across all outcome variables for research questions one through five, the assumption of independence was met because the students' unknown facts were randomly assigned to each condition and intervention implementation order was randomly assigned for each student. It was hypothesized that the data for the retention (RQ 1) and maintenance (RQ 3) assessments would not be normally distributed because these variables were not continuous. If the assumption of normality was violated for any of the research questions, then a Friedman rank test was used to examine differences in the outcome variables across the intervention conditions (van Gorp, Segers, & Verhoeven, 2014). In these instances, sphericity was no longer evaluated because it is not an assumption for the Friedman rank test. Normality was evaluated both visually with histograms of the dependent variables' residuals and statistically through the Shapiro-Wilk test, which is appropriate for analyzing normality for samples smaller than 50 using an alpha level of .01 (Tabachnick & Fidell, 2014). When the normality assumption was met, sphericity was examined through Mauchly's test of sphericity. Descriptive and inferential statistics for all outcome variables can be found in Table 4.

Intervention Effectiveness

Several outcome variables were examined regarding the differential effectiveness of the three interventions, IR, SIR, and TD. These results are described in detail below.

Multiplication fact retention. Before examining differences in retention across intervention conditions, the residuals of the retention data were examined to determine if the data met the assumption of normality. Because the variable was not continuous, it was hypothesized that the residuals may not be normally distributed. First, based on visual analysis of histograms of the retention data residuals, the distributions for the IR and TD residuals appeared to be platykurtic, which suggested that the assumption of normality was violated. Next, normality was also examined using the Shapiro-Wilks test, which resulted in a statistically significant test for all three intervention conditions (IR $p = .004$; TD & SIR $p = .003$), suggesting the assumption of normality was violated for the next day retention assessment data. Thus, the Friedman rank test was used to examine differences in multiplication fact retention across intervention conditions. The Friedman rank test was not statistically significant ($X^2 = .695$; $df = 2$; $p = .710$) with a small effect ($W = .010$; Tomczak & Tomczak, 2014). These results suggested that students demonstrated similar rates of multiplication fact retention across intervention conditions and there were no meaningful differences in retention among intervention conditions. Results are presented in Table 4.

A Kruskal-Wallis test was used to determine if there were differences in retention for each intervention condition across racial groups in the sample. Because of the normality violation in the residual data for retention, the Kruskal-Wallis test was identified as an appropriate analysis (Howell, 2012). The results of the Kruskal-Wallis test indicated no statistically significant differences among racial groups in next day retention for IR (p = .663), TD (p = .856), and SIR (p = .546) with effect sizes of η^2 = .04 (IR), η^2 = .05 (TD), η^2 = .03 (SIR) indicating a small effect (Bakeman, 2005). These results suggested similar rates of multiplication fact retention across racial groups. Results are presented in Table 5.

Multiplication fact fluency (DCPM). Prior to examining differences in multiplication fact fluency (DCPM) across intervention conditions, the residuals of the fluency data were examined to determine if they met the assumption of normality for a one-way repeated measures ANOVA. Based on visual analysis, the residuals appeared to be normally distributed for all three conditions. The Shapiro-Wilks test was not statistically significant (IR p = .096; TD p = .082; SIR p = .028), suggesting that the residuals for multiplication fact fluency (DCPM) were normally distributed for all three intervention conditions. The Mauchly's test of sphericity was then conducted to evaluate the assumption of sphericity. Mauchly's test of sphericity was not statistically significant ($x^2(5)$ = 1.815; p = .874), indicating the assumption of sphericity was met. Because the assumptions were met, a one-way repeated measures ANOVA was conducted to examine the differential impact of each intervention condition on multiplication fact fluency (DCPM). Because there were four levels of comparison (i.e., pre-test DCPM, IR post-test DCPM, TD post-test DCPM, SIR post-test DCPM), a Bonferroni correction was used to determine test significance ($p = 0.05 \div 4 = .0125$).

The one-way repeated measures ANOVA resulted in a statistically significant main effect (F = (3, 144.233) = 5.804, $p < .0125$) with a medium effect (η_p^2 = .14; Cohen, 1988). Because of the significant main effect, pairwise comparisons were conducted to determine if there were statistically significant differences between pre-test fluency scores and post-test fluency scores following each of the intervention conditions. Pairwise comparisons were also conducted to examine differences among the intervention conditions' post-test fluency scores (i.e., IR post-test DCPM, TD post-test DCPM, SIR post-test DCPM). The pairwise comparison indicated there was a statistically significant difference between pre-test fluency and IR post-test fluency (p = .005), but there were not statistically significant differences between pre-test fluency and TD post-test fluency (p = .021) or between pre-test fluency and SIR post-test fluency (p = .173). Additionally, there were no statistically significant differences between IR and TD post-test fluency scores (p = 1.000), IR and SIR post-test fluency scores (p = .966), or TD and SIR post-test fluency scores (p = 1.000). Overall, these results suggested that students demonstrated greater fluency scores following the IR condition compared to their pre-test scores, but they demonstrated similar post-test fluency scores across the three intervention conditions. Results are presented in Table 4.

An ANCOVA was conducted to examine differences in fluency outcomes among the racial groups for each of the intervention conditions. The results indicated no statistically significant differences for IR (F = (2, 6.62) = .13, p = .88), TD (F = (2, 67.93) = 1.13, p = .34), and SIR (F = (2, 64.96) = 1.24, p = .30) across the racial the groups with small (Cohen, 1988) effect sizes of η_p^2 = .01 (IR) and η_p^2 = .07 (TD & SIR). Thus, students had similar post-test fluency scores across racial groups. Results are presented in Table 5

Multiplication fact maintenance. Similar to analyses of the retention and fluency assessments, preliminary analyses were conducted to determine if the maintenance variable residuals met the assumptions of normality and sphericity. First, visual analysis was used to examine the distribution of the maintenance data residuals, which suggested that the residuals were normally distributed for IR and TD, but that the SIR residuals were positively skewed. Next, the Shapiro-Wilks test was not statistically significant for IR (p = .012) and TD (p = .011), but was statistically significant for SIR (p = .005), suggesting that the assumption of normality was violated for the SIR maintenance residuals. Thus, the Friedman rank test was most appropriate for examining differences in multiplication fact maintenance across the three intervention conditions (van Gorp, Segers, & Verhoeven, 2014). The Friedman rank test was not statistically significant (x^2 = 2.935; df = 2; p = .231) with a small effect (W = .042; Tomczak & Tomczak, 2014). These results indicated there were no meaningful differences in students' multiplication fact maintenance across the intervention conditions. These results suggested that students demonstrated similar multiplication fact maintenance across the intervention conditions. Results are presented in Table 4.

A Kruskal-Wallis test was used to determine if there were any differences across racial groups in multiplication fact maintenance for each intervention condition. The results of the Kruskal-Wallis test indicated that there were no statistically significant differences in multiplication fact maintenance across the three racial groups for IR (p = .790), TD (p = .804), or SIR (p = .815) with effect sizes of η^2 = .05 (IR, TD, & SIR) indicating a small effect (Bakeman, 2005). These results suggested similar rates of multiplication fact maintenance across racial groups. Results are presented in Table 5.

Intervention Efficiency

Preliminary analyses were conducted to examine if the residuals of the intervention efficiency (i.e., facts retained per minute) data met the assumption of normality. Based on visual analysis of the efficiency residuals, the IR efficiency residuals appeared normally distributed, but TD and SIR efficiency residuals were positively skewed suggesting that the assumption of normality was violated. The Shapiro-Wilks test was not statistically significant for the efficiency data for IR (p = .100) and SIR (p = .063), but was statistically significant for TD (p = .004), suggesting that the assumption of normality was violated for TD intervention efficiency residuals. Because the assumption of normality was not met across all three conditions, the Friedman rank test (van Gorp, Segers, & Verhoeven, 2014) was conducted to examine differences in intervention efficiency across the intervention conditions. The Friedman rank test was statistically significant (x^2 = 27.722; df = 2; p < .05) with a medium effect (W = .385; Tomczak & Tomczak, 2014); thus, indicating that there were differences in intervention efficiency among the intervention conditions. A Wilcoxon signed-rank test (Hinkle, Wiersma, & Jurs, 2003) was then used to examine specific differences in efficiency among the intervention conditions. Because there were three different comparisons (i.e., TD vs. IR, SIR vs. IR, and SIR vs. TD), a Bonferroni correction was used to determine test significance (p = 0.05/3 = .017). The results of the Wilcoxon signed-rank test indicated that SIR (Z = -4.713, p < .017) and TD (Z = -4.587, p < .017) were more efficient than IR and that TD was more efficient (Z = -3.174, p = .002) than SIR. These results suggested that TD was the most efficient intervention condition, followed by SIR, and IR was least efficient. Results are presented in Table 4.

A Kruskal-Wallis test was used to determine if there were any differences in intervention efficiency across racial groups for each intervention. The results of the Kruskal-Wallis test indicated that there were no statistically significant differences in intervention efficiency across the three racial groups for IR (p = .688), TD (p = .689), or SIR (p = .089) with effect sizes of η^2 = .04 (IR & TD) and η^2 = .09 (SIR) indicating small (IR & TD) and small to moderate (SIR) effects (Bakeman, 2005). These findings suggested that similar rates of efficiency across racial groups. Results are presented in Table 5.

Treatment Acceptability

Residuals of the treatment acceptability data were examined to determine if they met the assumption of normality for a one-way repeated measures ANOVA. Based on visual analysis, treatment acceptability residuals across all three conditions were negatively skewed and leptokurtic, which suggested the assumption of normality was violated. The Shapiro-Wilks test was statistically significant for all three conditions (IR & SIR p = .000; TD p = .001), which also suggested the treatment acceptability residuals were not normally distributed and the assumption of normality was violated. Given the violation of the assumption of normality, the Friedman Rank test was conducted to examine differences in treatment acceptability across the intervention conditions. The Friedman rank test was not statistically significant (X^2 = 2.113; df = 2; p = .348) with a small effect (W = .032; Tomczak & Tomczak, 2014). These results suggested that students' treatment acceptability was similar across the intervention conditions. Results are presented in Table 4.

A Kruskal-Wallis test was used to examine differences in treatment acceptability across racial groups for each intervention. The results indicated that there were no statistically significant differences in treatment acceptability across the three racial groups for IR (p = .702), TD (p = .988), and SIR (p = .882) with small η^2 = .04 (IR) and small to moderate η^2= .06 (TD & SIR) effects (Bakeman, 2005). These results suggested that students that identified as White, Latinx, and Multiracial reported similar treatment acceptability for all intervention conditions. Results are presented in Table 5.

Table 4

Descriptive and Inferential Statistics for Whole Sample

Measure	*M (SD)*	95% CI
Pre-Test Fluency	20.79 (10.52)	[5.5, 49.5]

	IR		TD		SIR			
	M (SD)	95% CI	*M (SD)*	95% CI	*M (SD)*	95% CI	Statistic	Effect Size
Facts Retained	2.31 (1.26)	[0, 4.0]	2.14 (1.27)	[0, 4.0]	2.28 (1.00)	[0.0, 4.0]	x^2 = .695	*W* = .01
Fluency	25.19 (12.03)	[4.0, 59.0]	24.86 (14.20)	[5.0, 57.5]	23.58 (14.07)	[4.5, 55.5]	F = 5.804*	η_p^2 = .14
Maintenance	2.03 (1.18)	[0, 4.0]	1.81 (1.17)	[0, 4.0]	1.60 (1.19)	[0, 4.0]	x^2 = 2.935	*W* =.04
Efficiency	0.30 (0.19)	[0, .69]	1.20 (0.96)	[0, 4.0]	0.69 (0.39)	[0.0, 1.73]	x^2 = 27.722*	*W* =.39
T.A.	17.69 (1.69)	[11.0, 20.0]	17.26 (2.17)	[11.0, 20.0]	17.36 (2.23)	[11.0, 20.0]	x^2 = 2.113	*W* =.03

Note. Kendall's W: .1 = small, .3 = medium, .5 = large (Tomczak & Tomczak, 2014). Partial eta squared: .01 = small, .09 = medium, .25 = large (Cohen, 1988).

Table 5

Descriptive and Inferential Statistics for Racial Groups

	IR			TD			SIR		
	M (SD)	Statistic	E.S.	*M (SD)*	Statistic	E.S.	*M (SD)*	Statistic	E.S.
Facts Retained	2.31 (1.26)	$H = 0.82$	$\eta^2 = .04$	2.14 (1.27)	$H = 0.31$	$\eta^2 = .05$	2.28 (1.00)	$H = 1.15$	$\eta^2 = .03$
White	2.41 (1.18)	-	-	2.10 (1.24)	-	-	2.34 (1.01)	-	-
Latinx	2.00 (1.41)	-	-	2.10 (1.41)	-	-	2.00 (.000)	-	-
Multiracial	1.80 (1.79)	-	-	2.40 (1.67)	-	-	2.00 (1.23)	-	-
Fluency	25.19 (12.03)	$F = 0.13$	$\eta_p^2 = .01$	24.86 (14.20)	$F = 1.13$	$\eta_p^2 = .07$	23.58 (14.07)	$F = 1.24$	$\eta_p^2 = .07$
White	25.14 (13.13)	-	-	24.88 (13.56)	-	-	23.90 (13.86)	-	-
Latinx	16.25 (10.25)	-	-	11.00 (5.66)	-	-	9.00 (5.66)	-	-
Multiracial	29.10 (20.78)	-	-	20.20 (18.33)	-	-	27.60 (15.87)	-	-
Maintenance	2.03 (1.18)	$H = 0.47$	$\eta^2 = .05$	1.81 (1.17)	$H = 0.44$	$\eta^2 = .05$	1.60 (1.19)	$H = 0.41$	$\eta^2 = .05$
White	2.07 (1.16)	-	-	1.76 (1.12)	-	-	1.64 (1.16)	-	-
Latinx	1.50 (0.71)	-	-	1.20 (2.12)	-	-	1.50 (0.71)	-	-
Multiracial	2.00 (1.58)	-	-	2.20 (1.30)	-	-	1.40 (1.67)	-	-

Efficiency	0.30 (0.19)	*H* = 0.75	η^2 = .04	1.20 (0.96)	*H* = 0.75	η^2 = .04	0.69 (0.39)	*H* = 4.85	η^2 = .09
White	0.32 (0.18)	-	-	1.17 (0.88)	-	-	0.72 (0.36)	-	-
Latinx	0.26 (0.23)	-	-	0.73 (0.70)	-	-	0.30 (0.08)	-	-
Multiracial	0.24 (0.23)	-	-	1.56 (1.49)	-	-	0.66 (0.61)	-	-
T.A.	17.69 (1.69)	*H* = 0.71	η^2 = .04	17.26 (2.17)	*H* = 0.02	η^2 = .06	17.36 (2.23)	*H* = 0.25	η^2 = .06
White	17.57 (1.79)	-	-	17.21 (2.22)	-	-	17.31 (2.36)	-	-
Latinx	18.50 (0.71)	-	-	17.50 (2.12)	-	-	15.50 (3.54)	-	-
Multiracial	18.00 (1.41)	-	-	17.40 (2.30)	-	-	18.00 (1.00)	-	-

Note. E.S. = effect size. Effect size for Fluency is Partial Eta Squared (η_p^2): .01 = small, .09 = medium, .25 = large (Cohen, 1988); Effect size for Retention, Maintenance, Efficiency, and Treatment Acceptability are Eta Squared (η^2): .02 = small, .13 = medium, .26 = large (Bakeman, 2005).

CHAPTER 5

DISCUSSION

The present study used a within-subject design to examine differences in students' retention and maintenance of and fluency with multiplication facts as well as intervention efficiency and treatment acceptability across three drill-based flashcard interventions: IR, TD, and SIR. Results showed that there were not statistically significant differences in next day retention across the three intervention conditions, suggesting students had similar rates of retention across interventions. Similarly, there were not statistically significant differences among the intervention conditions in multiplication fact fluency (i.e., DCPM) or maintenance of multiplication facts two weeks after intervention. Treatment acceptability results also indicated that there were no differences in students' preferences for the three interventions. In contrast, efficiency results indicated that TD was more efficient than IR and SIR, and SIR was more efficient than IR.

Intervention Effectiveness

Multiplication fact retention. Following all three intervention conditions, students demonstrated increased multiplication fact retention, which was consistent with previous research, IR, TD, and SIR were all shown to increase students' basic skill (e.g., sight words, math facts) acquisition and retention (Burns, 2005; Burns et al., 2016; January et al., 2017; Kupzyk et al., 2011). However, the present study found that differences were small but non-statistically significant for next day multiplication fact retention between IR and TD, which was consistent with Volpe, Mulé and colleagues' (2011) findings examining the comparative effectiveness of IR and TD; however, Volpe, Mulé et al. (2011) modified the IR procedure in order to compare specific intervention features (i.e., opportunities to respond, standardized intervention time). Therefore, Volpe, Mulé and colleagues' (2011) findings may not be directly comparable to the present study in which traditional IR procedures were implemented. In contrast, this study was inconsistent with most prior research that suggested IR was more effective than TD (Burns et al., 2019; Burns & Sterling-Turner, 2010; MacQuarrie et al., 2002; Nist & Joseph, 2008). Such findings may be reflective of differences in stimulus materials and individual student differences. Burns et al. (2019) did not include the use of commutative properties of multiplication facts (e.g., 8 x 6 and 6 x 8) in their comparison of retention between IR and TD, but the present study did. This procedural difference could explain why there were not retention differences between IR and TD because students may have had more opportunities to practice a fact which could have led to better retention outcomes across conditions. Furthermore, majority of the research has examined the retention of words rather than math facts (Burns & Sterling-Turner, 2010; MacQuarrie et al., 2002; Nist & Joseph, 2008). This difference in skill area being targeted may also be a possible explanation for the inconsistency in retention results between the present study and previous research. However, these explanations are hypotheses requiring further examination.

In this study, there were not differences in retention between IR and SIR, which was inconsistent with the preliminary research base for SIR suggesting that SIR was slightly more effective than IR when teaching sight words (January et al., 2017; Kupzyk et al., 2011). Differences in skill focus (i.e., sight words versus multiplication facts) could explain the differences in findings between January et al. and Kupzyk et al's findings and this study's findings; however, this explanation is a hypothesis in need of further examination, particularly because other studies have shown that IR was effective for teaching sight words and math facts (Burns, 2005; Joseph, 2006; McVancel et al., 2018; Nist & Joseph, 2008). This study was the first to directly compare SIR and TD on retention of multiplication facts, and thus, provided preliminary findings regarding their differential effectiveness to teach math facts. Future research could replicate or build upon the present study to examine differences among SIR, IR, and TD to teach math facts.

In addition to overall retention outcomes, there were small but non-statistically significant differences for multiplication fact retention among the racial groups. This study was the first to compare differences among racial groups in the context of fact retention for IR, TD, and SIR and, therefore, future research should seek to further examine racial group differences in the context of multiplication fact retention for IR, TD, and SIR.

Multiplication fact fluency (DCPM). There were small but non-statistically significant differences in students' post-test performance on multiplication fact fluency across the three intervention conditions as well as across the three racial groups. There is a dearth in the research base comparing IR and TD on math fact fluency outcomes (Cooke et al., 1993). Rather, almost all of the research of drill-based interventions has examined their effects on next day retention and generalization of skills (Burns & Sterling-Turner, 2010; Joseph et al., 2012; Lozy & Donaldson, 2019; MacQuarrie et al., 2002; Nist & Joseph, 2008), fluency specific to IR only (Burns, 2005; Volpe, Burns, et al., 2011), or targeted math fluency interventions (e.g., Taped problems, Cover Copy and Compare; Poncy, Skinner, & McCallum, 2012). Similarly, this study was the first to examine the impact of SIR on multiplication fact fluency as well as racial group differences for fluency outcomes. Therefore, additional research should further examine the impact of SIR, IR, and TD on math fact fluency and differences in fluency outcomes among racial groups.

Although there were not statistically significant differences in post-test multiplication fact fluency among the intervention conditions, students demonstrated significant growth in DCPM between pre-test and post-test in the IR condition. These findings were consistent with previous research that found that IR resulted in increased multiplication fact fluency (Burns, 2005; Codding et al., 2010; McVancel et al., 2018). Students may have demonstrated significant fluency growth in the IR condition given the high number of opportunities to respond to the unknown multiplication facts (i.e., between 8 and 32 times) whereas TD had only four OTRs per multiplication fact. Incorporation of high OTRs to unknown facts within IR was shown to be an underlying causal mechanism to increase students' skill fluency (e.g., DCPM; McVancel et al., 2018). In contrast, for SIR, the OTRs varied because the number of times facts were presented for practice was contingent upon each student's correct responding to the multiplication facts. Therefore, it is not clear how OTR may have impacted fluency in the SIR condition.

Multiplication fact maintenance. There were small but non-statistically significant differences in students' maintenance of multiplication facts across the three conditions. These findings were consistent with Joseph et al.'s (2012) findings that TD and IR resulted in similar maintenance of rehearsed words (Joseph et al., 2012). In contrast, the present study's findings were inconsistent with the majority of prior research that suggested that IR resulted in greater maintenance of math facts than TD (Burns & Sterling-Turner, 2010; MacQuarrie et al., 2002; Nist & Joseph, 2008). It is possible that differences in the time-points at which maintenance was assessed may explain these inconsistencies across studies. For example, this study assessed maintenance two weeks after intervention, which was longer than previous IR and TD research that assessed maintenance after 5 days (Nist & Joseph, 2008) or 1 week (Mulé et al., 2018; Burns & Sterling-Turner, 2010). Future research could examine maintenance at multiple timepoints to replicate previous research and further understanding of these interventions' impact on fact maintenance over time.

This study's findings were inconsistent with previous research that examined the impact of SIR on skill maintenance that found that SIR resulted in greater maintenance than IR (January et al., 2017; Kupzyk et al., 2011). In January and colleagues' (2017) study, in SIR, previously rehearsed stimuli were reintroduced when participants made errors during retention assessments, which resulted in high OTR to those reintroduced stimuli, which could explain higher maintenance in the SIR condition. In contrast, stimuli taught during the IR intervention sequence were not reintroduced, resulting in fewer OTR to stimuli in the IR condition compared to stimuli that were reintroduced in the SIR condition. However, because this reintroduction of stimuli during SIR was dependent upon whether or not the student responded correctly during the retention assessment, it is also possible that students had fewer OTR if they responded correctly to the stimulus and the stimulus was therefore not reintroduced. The present study did not reintroduce facts in any of the intervention conditions, which may explain this study's and January et al.'s inconsistent maintenance findings.

There were small but non-statistically significant differences in multiplication fact maintenance among the racial groups. However, this study was the first to compare differences among racial groups in the context of maintenance for IR, TD, and SIR. Therefore, future research should seek to further examine racial group differences in the context of multiplication fact maintenance for IR, TD, and SIR.

This study was also one of the first to compare the impact of SIR and TD on multiplication fact maintenance. Lozy and Donaldson (2019) compared maintenance outcomes when TD and SIR were implemented to teach letter sounds and found that TD resulted in greater maintenance than SIR, but they noted their findings were inconsistent with previous research that found that TD was less effective than interspersal techniques (e.g., IR) for maintenance. Because Lozy and Donaldson's (2019) study and this study were the first to compare maintenance outcomes of TD and SIR, future research could further examine differential maintenance effects for SIR and TD when targeting math facts.

Research procedures and effectiveness. The present study's findings may not have coincided with previous research due to differences in the research methodology. Previous research examined the differential effectiveness of SIR, IR, and TD following multiple intervention sessions (Codding et al., 2011; January et al., 2017; Kupzyk et al., 2011; Mulé et al., 2018; Nist & Joseph, 2008; Volpe, Mulé et al., 2011) whereas in the present study, students received each intervention only once. Previous research suggested that students required multiple opportunities (i.e., intervention sessions) for repeated practice of math facts in order to become fluent (Burns et al., 2015) and to demonstrate increased retention of words (Volpe, Mulé et al., 2011). This study examined the short-term efficacy of the interventions, which may not have allowed enough time to detect differences among conditions. Future research could compare these conditions within a multi-session, group design to determine if there are differences among the conditions after the interventions are implemented over an extended period of time.

Instructional Hierarchy (IH) and Intervention Effectiveness

Students who fall within the acquisition stage of the IH demonstrate slow and inaccurate responding. Previous research showed that fourth- and fifth-grade students should score between 24 and 49 DCPM on fluency measures (i.e., CBM-M or SSMM) to demonstrate instructional-level performance (Burns, VanDerHeyden, & Jiban, 2006) and that students who fell below the instructional range typically required additional support to reach grade-appropriate instructional level. In the present study, average student performance on pre-test fluency fell below the instructional range (i.e., mean pretest = 20.79 DCPM), which suggested below grade-level responding. Thus, on average, the present study's participants appeared to be representative of students in need of acquisition-based intervention support. Although there were no differences in intervention effectiveness across the intervention conditions, students demonstrated growth across all measures (i.e., mean retention of 2.24 multiplication facts, mean maintenance of 1.81 facts, and mean DCPM increase of 3.75) for all intervention conditions. These results provided support for use of the IH in informing implementation of acquisition-based interventions for students demonstrating slow and below-grade level responding.

The instructional features of the interventions may also have contributed to students' acquisition and retention of multiplication facts in the intervention conditions. During the acquisition stage of the IH, there are key instructional components (i.e., modeling, explicit instruction, and immediate corrective feedback; Haring & Eaton, 1978) that should be implemented to support students' acquisition of basic academic skills. These instructional features are important for the development and evaluation of instructional methods (Ellis, 2014). The present study used the IH to inform the implementation of and instructional methods used in the intervention conditions. For example, unknown facts were modeled and explicitly taught when first presented in all intervention conditions. Additionally, immediate corrective feedback, which included additional explicit instruction, was provided when errors occurred in all intervention sessions. Thus, the lack of differences in multiplication fact retention and maintenance across the intervention conditions may be attributed to the theoretically-informed instructional features that were incorporated in each of the interventions. However, this explanation is a hypothesis and previous research has identified differences in retention and maintenance when comparing IR, TD, and SIR. Future research could also compare these interventions within the context of IH and further examine the impact of the specific instructional components on students' skill outcomes.

Haring and Eaton (1978) also outlined specific instructional features necessary to promote skill fluency at the proficiency stage of the IH (e.g., explicit timing). In the present study, a time component was incorporated in the intervention sessions as students were given 3 s to respond to the multiplication facts presented. However, it is important to note that this timing procedure differed from typical explicit timing instruction in the proficiency stage in that most often, students are tasked with completing multiple multiplication problems within a set time frame (e.g., 2 min). Even in the absence of specific instructional features of the proficiency stage (e.g., explicit timing), descriptive statistics showed that students' fluency (i.e., DCPM) increased from pre-test to post-test following each intervention condition (i.e., range = 3 – 5 DCPM gained). These results coincided with prior research that also found students made gains in fluency in response to an acquisition-based intervention (e.g., IR; Burns, 2005), suggesting that drill-based interventions may also be effective for supporting math fluency growth. This study expanded the multiplication fact and drill-based intervention research base by incorporating theory within intervention implementation and results interpretation and future research examining math fact interventions should aim to do so as well.

Intervention Efficiency

In the present study, there were small to moderate statistically significant differences in intervention efficiency with TD having greater multiplication facts retained per instructional minute compared to IR and SIR. These results were consistent with previous research comparing the efficiency of IR and TD (Burns & Sterling-Turner, 2010; Nist & Joseph, 2008). TD uses only unknown stimuli (Mulé et al., 2018), which results in less required instructional time to practice the multiplication facts. Students demonstrated similar next day retention in the TD and IR conditions, which in combination with shorter implementation time for TD likely contributed to the greater efficiency in the TD condition compared to the IR condition. TD was also more efficient than SIR whereas previously, SIR demonstrated greater efficiency than TD (Lozy & Donaldson, 2019). However, Lozy and Donaldson (2019) defined efficiency differently than most prior research of intervention efficiency by examining the number of intervention sessions it took participants to meet mastery criteria, which could explain the differing results. Future research should use consistent definitions of efficiency in order to more directly compare outcomes across studies.

TD has consistently been shown to be more efficient and to require less instructional time than interventions using interspersal procedures (e.g., IR; Burns & Sterling-Turner, 2010; Nist & Joseph, 2008). In this study, TD took less time to implement (M = 2.13 minutes) compared to SIR (M = 4.14 minutes). In SIR, the use of interspersal techniques during the introduction and initial practice of unknown stimuli and the individualized approach to demonstration of fact mastery (i.e., the intervention was implemented until the student demonstrated mastery of all unknown facts) appeared to require more instructional time than TD. In contrast, TD presented all facts at once and had a predetermined number of times (i.e., four) students practiced each unknown fact. Thus, TD's shorter intervention duration likely contributed to its greater facts retained per instructional minute than SIR. This study was the first to compare efficiency outcomes between SIR and TD in the context of teaching multiplication facts, and therefore, additional research is needed to better understand the comparative efficiency of SIR and TD.

The results of the present study also suggested that SIR was more efficient than IR, which was consistent with previous research (January et al., 2017). Volpe, Mulé and colleagues (2011) asserted that IR's use of known stimuli results in a relatively longer intervention duration than other drill-based interventions. Because SIR uses only unknown stimuli, instructional time can be allotted to OTR to unknown stimuli rather than stimuli the student has already mastered and providing corrective feedback to unknown stimuli. High OTR led to improved acquisition and retention (Burns et al., 2010), which may have contributed to more multiplication facts retained per instructional minute (i.e., efficiency) in the SIR condition compared to the IR condition both in this study and previous research.

There were small but non-statistically significant differences in intervention efficiency among the racial groups. However, this study was the first to compare differences among racial groups in the context of intervention efficiency for IR, TD, and SIR. Therefore, future research should seek to further examine racial group differences in the context of intervention efficiency for IR, TD, and SIR.

Efficiency is an important, and growing, topic in drill-based intervention research given the limited time and resources schools often encounter when providing intervention services. However, previous research has operationalized efficiency with respect to both next day retention and maintenance one to two weeks after instruction (Burns & Sterling-Turner, 2010; McVancel et al., 2018). When efficiency was examined within the context of skill maintenance, IR was more efficient than TD (Burns & Sterling-Turner, 2010). In the present study, however, efficiency was operationalized as multiplication facts *retained* per instructional minute based on the next day retention assessment, which was in line with the most common operationalization of efficiency (Burns & Sterling-Turner, 2010; Nist & Joseph, 2008).

Given the inconsistencies in the definition of efficiency in prior research, it is difficult to compare this study's findings to other studies that have used maintenance data to calculate intervention efficiency. Future research should continue to examine intervention efficiency in an effort to identify the most appropriate operationalization when examining efficiency of these interventions. In doing so, researchers should consider the educational goals for student learning. For example, it is important that students are able to acquire and maintain early math skills in order to be successful in more complex math curriculum (e.g., algebra) later in schooling (Duncan et al., 2007; Gersten et al., 2005). Thus, although next day retention is currently the most common efficiency definition, using maintenance outcomes to evaluate efficiency may better align with educational goals of long-term skill maintenance.

Treatment Acceptability

In the present study, there were small but non-statistically significant differences in students' treatment acceptability among the three drill-based interventions used to target multiplication fact retention, maintenance, and fluency. Additionally, there were small but not-statistically significant differences in treatment acceptability across racial groups. There is a relative dearth of research examining treatment acceptability for math interventions (McVancel et al., 2018; Mong & Mong, 2010; McCallum, Skinner, Turner, & Saecker, 2006; Arra & Bahr, 2015). Although no prior research has examined student treatment acceptability of IR and TD when used to target math facts, treatment acceptability was previously examined when IR and TD were implemented to teach sight words, and results suggested that students did not prefer one intervention over the other (Volpe, Mulé et al., 2011). Additionally, the present study was the first to evaluate treatment acceptability of SIR (January et al., 2017; Kupzyk et al., 2011) as well as treatment acceptability across racial groups for IR, TD, and SIR. Additional research should therefore examine treatment acceptability of SIR to better understand students' preference for intervention procedures both for whole samples as well as across racial groups.

Student acceptability is important because when students are provided choices during instruction (e.g., getting to choose a preferred method of practice), they demonstrated greater task engagement and less off-task behavior (DiCarlo, Baumgartner, Stephens, & Pierce, 2013; Lane et al., 2015). Student engagement is important given the limited amount of time available for intervention implementation during the school day (Nist & Joseph, 2008). As such, researchers' continued examination of student acceptability for drill-based interventions is important to help identify students' preferred methods of practice so that educators are able to offer effective preferred interventions as choices to students in need of additional support. Furthermore, student treatment acceptability could support the effectiveness of an intervention (Mautone et al., 2009). It is possible that when students perceive the intervention as effective in improving their skill difficulties, they may also be more engaged in the intervention. However, these hypotheses should be further examined to better understand student acceptance of drill-based interventions and the impact of intervention preference on student engagement and skill growth.

Implications for practice

Results from the present study provided implications for drill-based interventions targeting multiplication facts in school settings. Intervention efficiency was an important finding for school-based services given that there is limited time to implement interventions during the school day (Nist & Joseph, 2008). TD was the most efficient of the three interventions examined, a finding that was consistent with previous research comparing IR and TD (Burns & Sterling-Turner, 2010; Nist & Joseph, 2008). Educators may consider implementing TD if students demonstrate difficulties with multiplication facts and significantly limited intervention time is available to provide additional instruction.

Students were taught four multiplication facts per intervention session, which was slightly less than the average acquisition rate for fifth graders (i.e., 5; Burns, 2001). This lower acquisition rate was used in this study because the sample included students experiencing difficulties with multiplication facts as well as both fourth and fifth grade students. Individual student acquisition rate is an important consideration when teaching students with math difficulties (Burns et al., 2016), so educators may consider teaching more than four stimuli based on students' grade level and accuracy responding to unknown stimuli during the intervention (i.e., continue to teach unknown stimuli until the student makes 3 errors while practicing a new unknown fact; Burns, 2001; Haegele & Burns, 2015). Prior research examining SIR showed that some students acquired up to 10 unknown stimuli within a single intervention session, but students demonstrated varying levels of retention (i.e., approximately 5 to 9 stimuli retained) during a next day retention assessment (Kupzyk et al., 2011). Thus, there is likely variability in the number of unknown stimuli students can accurately acquire and retain, so educators may wish to adapt the number of unknown stimuli rehearsed during an intervention session based on each students' individual responding during intervention and retention outcomes.

Although there were no statistically significant differences in intervention effectiveness for retention, fluency, or maintenance across the conditions, there was an increase of mean DCPM in all intervention conditions, and in the IR condition, there was a statistically significant difference in pre- to post-test DCPM. If educators want to target student fluency in conjunction with skill acquisition, IR may be the most appropriate intervention to promote growth in math fact fluency (Burns, 2005; Codding et al., 2010). However, when fluency is not of focus, practitioners may also consider individual student factors (e.g., student acceptability, diversity factors such as disabilities, skills being targeted) when selecting which intervention to implement. For example, if a student has significant attention difficulties, educators may wish to consider implementing the intervention the student most prefers or in which they are most engaged. In this instance, students with attention difficulties may prefer TD if they are able to consistently concentrate during the intervention given that TD is a relatively quick intervention.

Limitations and Implications for Future Research

Although the results of this study provided novel and interesting findings regarding drill-based interventions targeting multiplication facts, such findings should be interpreted within the context of the study's limitations. First, not all students received study procedures with full fidelity due to absences, school closure, or interventionist error, which could have impacted the effectiveness and efficiency outcomes. The deviations in procedures included assessments (e.g., retention and maintenance) administered late or early on a few occasions across students and conditions. Additionally, one student was not administered the maintenance assessments because they moved to a different school during data collection. Although a limitation of the interpretation of the present study's results, these procedural deviations naturally occur in a school setting; thus, these findings may in fact reflect practical application to school settings. Future research could examine whether there are differences in outcomes when controlling for procedural differences related to the time points of retention, maintenance, and fluency assessments.

Second, the order of intervention implementation was not evenly counterbalanced across participants. Although the condition order was randomly assigned across participants, future research may wish to ensure counterbalancing of conditions so that condition order combinations are equal across participants. Third, the within-subject design provided short-term outcome data based solely on one time point of instruction for each intervention condition. Thus, it is possible that results may have differed if each intervention condition had been implemented over an extended period of time. Given that group designs allow for greater generalizability, future research may wish to evaluate SIR, IR, and TD's effectiveness and efficiency using a group design to identify outcomes based on multiple instruction time points. Implementing a group design over an extended period of time would enhance understanding of these three drill-based interventions' long-term outcomes.

Fourth, not all students in this study demonstrated significant difficulties with multiplication facts, even though they did meet the minimum inclusion criteria of 12 unknown facts. The participating students demonstrated a wide range of initial skill levels with some students having only 12 unknown facts at pre-test. The drill-based interventions examined in this study are typically implemented to support students with significant skill deficits. As such, it may have been easier for the students with initial higher skill level (e.g., only had 12 unknown facts at pre-test) to acquire, retain, and maintain the 12 unknown facts taught regardless of the intervention procedures implemented, which could have contributed to the lack of differences in intervention outcomes. Future research could modify the inclusion criteria in order to include a more homogenous sample of students demonstrating significant difficulties with multiplication facts. Last, as a result of time and personnel constraints, maintenance of facts was only evaluated at one time point, two weeks after intervention. Thus, it is unknown if students would maintain rehearsed multiplication facts more than two weeks after the interventions concluded. Future research could evaluate maintenance patterns across multiple time points following the implementation of each intervention. Last, the sample for the present study was predominantly

White (80.5%) which may have impacted the non-statistically significant findings. Given that other races and ethnicities had limited representation (i.e., Latinx, Multiracial), future researchers could examine differences in outcomes across racial/ethnic groups for IR, TD, and SIR with a more diverse sample.

In addition to the points discussed above, there are further considerations for future research regarding drill-based interventions for multiplication facts. First, given the differing operationalizations of intervention efficiency in the literature (i.e., next day retention vs. maintenance; McVancel et al., 2018), future research could examine efficiency outcomes across both maintenance and retention operationalizations. Such research could inform whether one intervention is more efficient than another intervention in the long-term, which is an important consideration given that ultimately students need to maintain skills over a long period of time. Second, future research could examine how receiving corrective feedback during next day retention assessments may impact maintenance of rehearsed multiplication facts. This feedback may be important given that corrective feedback is an essential instructional component outlined in the IH. Thus, the more frequently a student receives corrective feedback (i.e., during intervention and assessment), the more support the student receives for further skill development. Future studies could examine outcomes when corrective feedback is provided during intervention and assessment compared to feedback provided only during intervention sessions.

Conclusion

The present study examined differences in intervention effectiveness and efficiency across IR, TD, and SIR using a within-subject design. The majority of drill-based intervention research has focused on reading, including novel drill-based interventions (e.g., SIR; January et al., 2017; Burns, Zaslofsky et al., 2012). The present study also expanded upon drill-based math intervention research to include SIR, which had previously only been compared to IR and TD in the context of sight words. Examining the use of SIR to teach math facts was an important area of expansion given its promising outcomes when used to teach sight words (January et al., 2017). Although the present study did not result in differences in effectiveness among the conditions, IR had significant differences between pre-test fluency scores and post-test fluency scores, so it may be particularly useful for promoting students' multiplication fact fluency. Additionally, the findings showed that although equally effective as IR and SIR, TD was more efficient than IR which coincided with previous research (Burns & Sterling-Turner, 2010; Nist & Joseph, 2008). These findings are important for educators when identifying and implementing interventions that most efficiently promote optimal skill growth given the limited time for additional instruction within the school day. For example, TD may be most appropriate in instances when there is very little time for additional multiplication fact instruction. However, educators should balance intervention efficiency with intervention effectiveness, particularly in the long-term, when determining which intervention to implement.

The present study provides preliminary results for comparing IR, TD, and SIR to target multiplication facts, but continued examination of the differences among these interventions is needed. Examining instructional methods that promote the acquisition of early math skills (e.g., multiplication facts) is essential due to the impact these early skills have on later skill development and broader life outcomes (e.g., completing high school and attaining higher vocational preparedness). Given the broader implications of these early skills, it is also important that schools intervene to promote students' acquisition of these foundational skills early in schooling (e.g., third and fourth grade; Duncan et al., 2007; Gersten et al., 2005). Early intervention is important because it is easier to ameliorate skill difficulties when they first arise compared to when such difficulties have persisted over an extended period of time and have become more severe. In order to remediate early math difficulties, the continued examination of these interventions for targeting multiplication facts, as well as all basic math facts, is essential to assist school psychologists and other educators in identifying the most effective and efficient interventions for students struggling with math fact acquisition, retention, and maintenance.

References

Agrawal, J., & Morin, L. L. (2016). Evidence-based practices: Applications of concrete representational abstract framework across math concepts for students with mathematics disabilities. *Learning Disabilities Research & Practice*, *31*(1), 34-44.

Algozzine, B., Wang, C., & Violette, A. S. (2011). Reexamining the relationship between academic achievement and social behavior. *Journal of positive behavior interventions*, *13*(1), 3-16.

Apter, S. J., & Conoley, J. C. (1984). *Childhood behavior disorders and emotional disturbance: An introduction to teaching troubled children*. Prentice-Hall.

Archambault, I., Janosz, M., Fallu, J., & Pag, L. S. (2009). Student engagement and its relationship with early high school dropout. Journal of Adolescence, 32, 651–670. doi:10.1016/j.adolescence.2008.06.007

Ardoin, S. P., Binder, K. S., Zawoyski, A. M., & Foster, T. E. (2018). Examining the maintenance and generalization effects of repeated practice: A comparison of three interventions. *Journal of school psychology*, *68*, 1-18.

Ardoin, S. P., & Daly, E. J. (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), 1-6.

Ardoin, S. P., McCall, M., & Klubnik, C. (2007). Promoting generalization of oral reading fluency: Providing drill versus practice opportunities. *Journal of Behavioral Education*, *16*(1), 54-69.

Ardoin, S. P., Witt, J. C., Connell, J. E., & Koenig, J. L. (2005). Application of a three-tiered response to intervention model for instructional planning, decision making, and the identification of children in need of services. *Journal of Psychoeducational Assessment*, *23*(4), 362-380.

Arra, C. T., & Bahr, M. W. (2005). Teachers' and students' preferences for mathematics interventions: Implications for teacher acceptability in consultation. *Journal of Educational and Psychological Consultation*, *16*(3), 157-174.

Axtell, P. K., McCallum, R. S., Bell, S. M., & Poncy, B. (2009). Developing math automaticity using a classwide fluency building procedure for middle school students: A preliminary study. Psychology in the Schools, 46(6), 526-538. doi:10.1002/pits.20395

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior research methods*, *37*(3), 379-384.

Benner, G. J., Stage, S. A., Nelson, J. R., Laederich, M., & Ralston, N. C. (2010). Predicting the cumulative recidivism of juvenile detainees. *The Journal of Behavior Analysis of Offender and Victim Treatment and Prevention*, *2*(1), 51.

Benner, G. J., Nelson, J. R., Stage, S. A., & Ralston, N. C. (2011). The influence of fidelity of implementation on the reading outcomes of middle school students experiencing reading difficulties. *Remedial and Special Education*, *32*(1), 79-88.

Berrueta-Clement, J. R., Schweinhart, L. J., Barnett, W. S., Epstein, A. S., & Weikart, D. P. (1984). Changed lives. *Ypsilanti, MI: High/Scope*, *310*, 312-321.

Bineham, S. C., Shelby, L., Pazey, B. L., & Yates, J. R. (2014). Response to intervention: Perspectives of general and special education professionals. *Journal of School Leadership*, *24*(2), 230-252.

Bouck, E. C., & Cosby, M. D. (2017). Tier 2 response to intervention in secondary mathematics education. *Preventing School Failure: Alternative Education for Children and Youth*, *61*(3), 239-247.

Bouck, E. C., Satsangi, R., & Park, J. (2018). The concrete–representational–abstract approach for students with learning disabilities: an evidence-based practice synthesis. *Remedial and Special Education*, *39*(4), 211-228.

Bradshaw, C. P., Pas, E. T., Debnam, K. J., & Lindstrom Johnson, S. (2015). A focus on implementation of Positive Behavioral Interventions and Supports (PBIS) in high schools: Associations with bullying and other indicators of school disorder. *School Psychology Review*, *44*(4), 480-498.

Bramlett, R., Cates, G. L., Savina, E.,& Lauinger, B. (2010). Assessing effectiveness and efficiency of academic interventions in school psychology journals: 1995.2005. Psychology in the Schools, 47(2), 114 . 125.

Browder, D. M., & Shear, S. M. (1996). Interspersal of known items in a treatment package to teach sight words to students with behavior disorders. *The Journal of Special Education*, *29*(4), 400-413.

Brown-Chidsey, R., & Steege, M. W. (2011). *Response to intervention: Principles and strategies for effective practice*. Guilford Press.

Bryant, D. P., Bryant, B. R., Gersten, R., Scammacca, N., & Chavez, M. M. (2008). Mathematics intervention for first-and second-grade students with mathematics difficulties: The effects of tier 2 intervention delivered as booster lessons. *Remedial and special education*, *29*(1), 20-32.

Burns, M. K. (2007). Comparison of opportunities to response within a drill model when rehearsing sight words with a child with mental retardation. School Psychology Quarterly, 22, 250–263.

Burns, M. K. (2004). Empirical analysis of drill ratio research: Refining the instructional level for drill tasks. *Remedial and Special Education*, *25*(3), 167-173.

Burns, M. K. (2001). Measuring acquisition and retention rates with curriculum-based assessment. *Journal of Psychoeducational Assessment, 19*, 148–157.

Burns, M. K. (2011). School psychology research: Combining ecological theory and prevention science. *School Psychology Review*, *40*(1), 132-140.

Burns, M. K. (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*, 237-249.

Burns, M. K., Aguilar, L. N., Young, H., Preast, J. L., Taylor, C. N., & Walsh, A. D. (2019). Comparing the effects of incremental rehearsal and traditional drill on retention of mathematics facts and predicting the effects with memory. *School Psychology*, *34*(5), 521.

Burns, M. K., Codding, R. S., Boice, C. H., & Lukito, G. (2010). Meta-analysis of acquisition and fluency math interventions with instructional and frustration level skills: Evidence for a skill-by-treatment interaction. *School Psychology Review*, *39*(1), 69.

Burns, M. K., & Gibbons, K. (2012). School-based practice in action series. Implementing response-to-intervention in elementary and secondary schools: Procedures to assure scientific-based practices. New York, NY, US.

Burns, M. K., Maki, K. E., Warmbold-Brann, K., & Preast, J. L. (2018). H one of the most frequently and passionately debated issues in the history of education or psychology. The notable shortcomings of the discrepancy model left the field without a widely accepted method to identify students with SLD, when in 2004 the federal special education law changed. The 2004 amendments to the Individuals with Disabilities Education Act (IDEA) responded to long-standing criticisms of the ability-achievement discrepancy. *Essentials of Specific Learning Disability Identification*, 257.

Burns, M. K., & Parker, D. C. (2014). Curriculum-based assessment for instructional design: Using data to design instruction and intervention. New York, NY: Guilford Press.

Burns, M. K., Riley-Tillman, T. C., & VanDerHeyden, A. M. (2012). RT1 applications: Academic and behavioral interventions (Vol. 1). New York, NY: The Guilford Press.

Burns, M. K., & Sterling-Turner, H. E. (2010). Comparison of efficiency measures for academic interventions based on acquisition and maintenance. *Psychology in the Schools*, *47*(2), 126-134.

Burns, M. K., Ysseldyke, J., Nelson, P. M., & Kanive, R. (2015). Number of repetitions required to retain single-digit multiplication math facts for elementary students. *School Psychology Quarterly*, *30*(3), 398.

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

Burns, M. K., Zaslofsky, A. F., Kanive, R., & Parker, D. C. (2012). Meta-analysis of incremental rehearsal using phi coefficients to compare single-case and group designs. *Journal of Behavioral Education*, *21*(3), 185-202.

Burns, M. K., Zaslofsky, A. F., Maki, K. E., & Kwong, E. (2016). Effect of modifying intervention set size with acquisition rate data while practicing single-digit multiplication facts. *Assessment for Effective Intervention*, *41*(3), 131-140.

Campbell, F. A., & Ramey, C. T. (1995). Cognitive and school outcomes for high-risk African-American students at middle adolescence: Positive effects of early intervention. *American educational research journal*, *32*(4), 743-772.

Castro-Villarreal, F., Rodriguez, B. J., & Moore, S. (2014). Teachers' perceptions and attitudes about Response to Intervention (RTI) in their schools: A qualitative analysis. *Teaching and teacher education*, *40*, 104-112.

Cates, G. L., Dunne, M., Erkfritz, K. N., Kivisto, A., Lee, N., & Wierzbicki, J. (2007). Differential effects of two spelling procedures on acquisition, maintenance and adaption to reading. *Journal of Behavioral Education*, *16*(1), 70-81.

Cates, G. L., & Rhymer, K. N. (2003). Examining the relationship between mathematics anxiety and mathematics performance: An instructional hierarchy perspective. *Journal of Behavioral Education*, *12*(1), 23-34.

Cates, G. L,, & Skinner, C. H, (2002), Getting remedial mathematics students to prefer homework with 20% and 40% more problems: An investigation of the strengths of the interspersing procedure. Psychology in the Schools, 37, 339-347,

Cates, G. L., Skinner, C. H., Watson, T. S., Meadows, T. J., Weaver, A., & Jackson, B. (2003). Instructional effectiveness and instructional efficiency as considerations for data-based decision making: An evaluation of interspersing procedures. *School Psychology Review*, *32*(4), 601-616.

Codding, R. S., Archer, J., & Connell, J. (2010). A systematic replication and extension of using incremental rehearsal to improve multiplication skills: An investigation of generalization. *Journal of Behavioral Education*, *19*(1), 93-105.

Codding, R. S., Burns, M. K., & Lukito, G. (2011). Meta-analysis of mathematic basic-fact fluency interventions: A component analysis. *Learning Disabilities Research & Practice*, *26*(1), 36-47.

Codding, R. S., Hilt-Panahon, A., Panahon, C. J., & Benson, J. L. (2009). Addressing mathematics computation problems: A review of simple and moderate intensity interventions. *Education and treatment of Children*, 279-312.

Codding, R. S., & Martin, R. (2016). Tier 3: Intensive mathematics intervention strategies. In *Handbook of Response to Intervention* (pp. 375-388). Springer, Boston, MA.

Codding, R. S., & Poncy, B. C. (2010). Introduction to the special issue: Toward an explicit technology for generalizing academic behavior. *Journal of Behavioral Education*, *19*(1), 1-6.

Codding, R. S., Shiyko, M., Russo, M., Birch, S., Fanning, E., & Jaspen, D. (2007). Comparing mathematics interventions: Does initial level of fluency predict intervention effectiveness?. *Journal of School Psychology*, *45*(6), 603-617.

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

Cooke, N. L., Guzaukas, R., Pressley, J. S., & Kerr, K. (1993). Effects of using a ratio of new items to review items during drill and practice: Three experiments. Education and Treatment of Children, 16, 212–234.

Coulter, W. A., & Coulter, E. M. (1989). *Curriculum-based assessment for instructional design: Trainer's manual.* (Unpublished training manual available from Directions and Resources, P.O. Box 57113, New Orleans, LA 70157).

Cowan, R. J., & Sheridan, S. M. (2003). Investigating the acceptability of behavioral interventions in applied conjoint behavioral consultation: Moving from analog conditions to naturalistic settings. *School Psychology Quarterly*, *18*(1), 1.

Cusumano, D. L., Algozzine, K., & Algozzine, B. (2014). Multitiered system of supports for effective inclusion in elementary schools. In J. McLeskey, N. L. Waldron, F. Spooner, & B. Algozzine (Eds.), Handbook of effective inclusive schools: Research and practice. New York, NY: Routledge, pp. 197-209.

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

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

Daly, E. J. & L McCurdy, M. (2002). Getting it right so they can get it right: An overview of the special series. School Psychology Review, 31, 453-458.

DiCarlo, C. F., Baumgartner, J., Stephens, A., & Pierce, S. H. (2013). Using structured choice to increase child engagement in low-preference centres. *Early Child Development and Care*, *183*(1), 109-124.

Doabler, C. T., Cary, M. S., Jungjohann, K., Clarke, B., Fien, H., Baker, S., ... & Chard, D. (2012). Enhancing core mathematics instruction for students at risk for mathematics disabilities. *Teaching Exceptional Children*, *44*(4), 48-57.

Duhon, G. J., House, S. H., & Stinnett, T. A. (2012). Evaluating the generalization of math fact fluency gains across paper and computer performance modalities. *Journal of school psychology*, *50*(3), 335-345.

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., ... & Sexton, H. (2007). School readiness and later achievement. *Developmental psychology*, *43*(6), 1428.

Drevon, D. D., & Reynolds, J. L. (2018). Comparing the Effectiveness and Efficiency of Response Repetition to Simultaneous Prompting on Acquisition and Maintenance of Multiplication Facts. *Journal of Behavioral Education*, *27*(3), 358-374.

Eckert, T. L., & Shapiro, E. S. (1999). Methodological issues in analog acceptability research: Are teachers' acceptability ratings of assessment methods influenced by experimental design?. *School Psychology Review*, *28*(1), 5.

Elbaum, B., Vaughn, S., Tejero Hughes, M., & Watson Moody, S. (2000). How effective are one-to-one tutoring programs in reading for elementary students at risk for reading failure? A meta-analysis of the intervention research. *Journal of educational psychology*, *92*(4), 605.

Ellis, A. K. (2014). *Research on educational innovations*. Routledge.

Elliott, S. N. (1986). Children's ratings of the acceptability of classroom interventions for misbehavior: Findings and methodological considerations. *Journal of School Psychology*, *24*(1), 23-35.

Evans-Hampton, T. N., Skinner, C. H., Henington, C., Sims, S., & McDaniel, C. E. (2002). An investigation of situational bias: Conspicuous and covert timing during curriculum-based measurement of mathematics across African American and Caucasian students. *School Psychology Review*, *31*(4), 529-540.

Every Student Succeeds Act (ESSA) of 2015. :114-85. Pub. L.

Fan, W., & Wolters, C. A. (2014). School motivation and high school dropout: The mediating role of educational expectation. *British Journal of Educational Psychology*, *84*(1), 22-39.

Fentress, G. M., & Lerman, D. C. (2012). A comparison of two prompting procedures for teaching basic skills to children with autism. *Research in Autism Spectrum Disorders*, *6*(3), 1083-1090.

Flores, M. M., Hinton, V. M., & Burton, M. E. (2016). Teaching problem solving to students receiving tiered interventions using the concrete-representational-abstract sequence and schema-based instruction. *Preventing School Failure: Alternative Education for Children and Youth*, *60*(4), 345-355.

Fuchs, L.S. (2004). The past, present, and future of curriculum-based measurement research. School Psychology Review, 33, 188-192.

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*(3), 493.

Fuchs, L. S., Fuchs, D., & Gilbert, J. K. (2019). Does the Severity of Students' Pre-Intervention Math Deficits Affect Responsiveness to Generally Effective First-Grade Intervention?. *Exceptional children*, *85*(2), 147-162.

Fuchs, L. S., Fuchs, D., & Hollenbeck, K. N. (2007). Extending responsiveness to intervention to mathematics at first and third grades. *Learning Disabilities Research & Practice*, *22*(1), 13-24.

Fuchs, L. S., Fuchs, D., & Prentice, K. (2004). Responsiveness to mathematical problem-solving instruction: Comparing students at risk of mathematics disability with and without risk of reading disability. *Journal of Learning Disabilities*, *37*(4), 293-306.

Fuchs, L. S., Fuchs, D., Powell, S. R., Seethaler, P. M., Cirino, P. T., & Fletcher, J. M. (2008). Intensive intervention for students with mathematics disabilities: Seven principles of effective practice. *Learning Disability Quarterly*, *31*(2), 79-92.

Fuchs, L. S., Fuchs, D., Yazdian, L., & Powell, S. R. (2002). Enhancing first-grade children's mathematical development with peer-assisted learning strategies. *School Psychology Review*, *31*(4), 569.

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of learning disabilities*, *45*(3), 195-203.

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting Students Struggling with Mathematics: Response to Intervention (RtI) for Elementary and Middle Schools. NCEE 2009-4060. *What Works Clearinghouse*.

Gersten, R., Jordan, N. C., & Flojo, J. R. (2005). Early identification and interventions for students with mathematics difficulties. *Journal of learning disabilities*, *38*(4), 293-304.

Gickling, E. E., & Thompson, V. P. (1985). A personal view of curriculum-based assessment. *Exceptional Children*, *52*(3), 205-218.

Glover, T. A., & DiPerna, J. C. (2007). Service delivery for response to intervention: Core components and directions for future research. *School Psychology Review*, *36*(4), 526-541.

Good, R. H., Kaminski, R. A., & Moats, L. C. (2003). *DIBELS: Dynamic indicators of basic early literacy skills* (Vol. 180). Sopris West Incorporated.

Good, R. H., Kaminski, R. A., Smith, S., Simmons, D., Kame'enui, E., & Wallin, J. (2003). Reviewing outcomes: Using DIBELS to evaluate a school's core curriculum and system of additional intervention in kindergarten. *Reading in the classroom: Systems for observing teaching and learning. Baltimore: Paul H. Brookes*.

Gravois, T. A., & Gickling, E. (2002). Best practices in curriculum-based assessment. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology (Vol. IV, pp. 885–898). Bethesda, MD: National Association of School Psychologists.

Gresham, G., & Little, M. (2012). RtI in math class. *Teaching Children's Mathematics*, *19*(1), 20-29.

Griffin, S. (2004). Building number sense with Number Worlds: A mathematics program for young children. *Early childhood research quarterly*, *19*(1), 173-180.

Griffiths, A., Parson, L., Burns, M., VanDerheyden, A., & Tilly, W. (2007). Response to Intervention: Research for practice. Alexandria, VA: National Association of State Director of Special Education.

Haegele, K., & Burns, M. K. (2015). Effect of modifying intervention set size with acquisition rate data among students identified with a learning disability. *Journal of Behavioral Education*, *24*(1), 33-50.

Hammerschmidt-Snidarich, S. M., Maki, K. E., & Adams, S. R. (2019). Evaluating the effects of repeated reading and continuous reading using a standardized dosage of words read. *Psychology in the Schools*.

Harber, M., Henington, C., Dufrene, B. A., Baylot, L., Townsend, A. E., Fancher, M., et al. (2004). Academic skill-builder summer clinic for elementary school-aged children. Symposium presented at the annual meeting of the National Association of School Psychologists. Dallas, TX.

Haring, N. G., & Eaton, M. D. (1978). Systematic instructional procedures: An instructional hierarchy. *The fourth R: Research in the classroom*, 23-40.

Higgins-Averill, O., & Rinaldi, C. (2013a) Research Brief: Multi-tier System of Supports (MTSS) Urban Special Education Leadership Collaborative: From RTI and PBIS to MTSS.

Higgins-Averill, O., & Rinaldi, C. (2013b). Response to intervention and educator relations: a contrivance of collegiality or path to professional collaboration. *Communiqué*, *42*(4), 4-5.

Hinkle, D.E., Wiersma, W., & Jurs, S.G. (2003). Other nonparamentric tests. In K.T. Baruth, S. Wise, & K. Dinovo (Eds.), *Applied statistics for the behavioral sciences* (pp. 572-586). Houghton Mifflin Company.

Hosp, J. L., Huddle, S., Ford, J. W., & Hensley, K. (2016). Learning disabilities/special education. In *Handbook of response to intervention* (pp. 43-58). Springer, Boston, MA.

Howell, D. C. (2012). *Statistical methods for psychology (8 Ed.)*. Cengage Learning, Belmont, CA.

Hughes, C. J., Beverley, M., & Whitehead, J. (2007). Using precision teaching to increase the fluency of word reading with problem readers. *European Journal of Behavior Analysis*, *8*(2), 221-238.

Hughes, C. A., & Dexter, D. D. (2011). Response to intervention: A research-based summary. *Theory into practice*, *50*(1), 4-11.

Janosz, M., Archambault, I., Morizot, J., & Pagani, L. S. (2008). School engagement trajectories and their differential predictive relations to dropout. Journal of Social Issues, 64, 21–40. doi:10.1111/j.1540-4560.2008.00546.x

January, S. A. A., Lovelace, M. E., Foster, T. E., & Ardoin, S. P. (2017). A comparison of two flashcard interventions for teaching sight words to early readers. *Journal of Behavioral Education*, *26*(2), 151-168.

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., & Karl, S. R. (2017). A randomized trial of the effects of schema-based instruction on proportional problem-solving for students with mathematics problem-solving difficulties. *Journal of learning disabilities*, *50*(3), 322-336.

Jitendra, A. K., Harwell, M. R., Dupuis, D. N., Karl, S. R., Lein, A. E., Simonson, G., & Slater, S. C. (2015). Effects of a research-based intervention to improve seventh-grade students' proportional problem solving: A cluster randomized trial. *Journal of Educational Psychology*, *107*(4), 1019.

Joseph, L. M. (2006). Incremental rehearsal: A flashcard drill technique for increasing retention of reading words. *The Reading Teacher*, *59*(8), 803-807.

Joseph, L., Eveleigh, E., Konrad, M., Neef, N., & Volpe, R. (2012). Comparison of the efficiency of two flashcard drill methods on children's reading performance. *Journal of Applied School Psychology*, *28*(4), 317-337.

Joseph, L. M., Konrad, M., Cates, G., Vajcner, T., Eveleigh, E., & Fishley, K. M. (2012). A meta-analytic review of the cover-copy-compare and variations of this self-management procedure. *Psychology in the Schools*, *49*(2), 122-136.

Joseph, L. M., & Nist, L. M. (2006). Comparing the effects of unknown-known ratios on word reading learning versus learning rates. *Journal of Behavioral Education*, *15*(2), 69.

Joseph, L. M., & Schisler, R. A. (2007). Getting the "Most Bang for Your Buck" Comparison of the Effectiveness and Efficiency of Phonic and Whole Word Reading Techniques During Repeated Reading Lessons. *Journal of Applied School Psychology*, *24*(1), 69-90.

Kauffman, J. M., Bruce, A., & Lloyd, J. W. (2012). Response to intervention (RTI) and students with emotional and behavioral disorders. *Behavioral disorders: Practice concerns and students with EBD*, *23*, 107-128.

Kazdin, A. E. (1981). Acceptability of child treatment techniques: The influence of treatment efficacy and adverse side effects. Behavior Therapy, 11, 329–344.

Ketterlin-Geller, L. R., & Yovanoff, P. (2009). Diagnostic assessments in mathematics to support instructional decision making. *Practical Assessment, Research & Evaluation*, *14*(16), 1-11.

Kleinert, W. L., Codding, R. S., Minami, T., & Gould, K. (2018). A Meta-analysis of the Taped Problems Intervention. *Journal of Behavioral Education*, *27*(1), 53-80.

Klingbeil, D. A., Bradley, T. L., & McComas, J. J. (2016). Progress monitoring for students receiving intensive academic intervention. In *Handbook of response to intervention* (pp. 343-360). Springer, Boston, MA.

Klingbeil, D. A., Moeyaert, M., Archer, C. T., Chimboza, T. M., & Zwolski Jr, S. A. (2017). Efficacy of peer-mediated incremental rehearsal for English language learners. *School Psychology Review*, *46*(1), 122-140.

Kilpatrick, J., Swafford, J., & Findell, B. (2001). Adding it up: Helping children learn mathematics. Washington, DC: National Academies Press.

Kroesbergen, E. H., & Van Luit, J. E. (2003). Mathematics interventions for children with special educational needs: A meta-analysis. *Remedial and special education*, *24*(2), 97-114.

Kupzyk, S., Daly III, E. J., & Andersen, M. N. (2011). A comparison of two flash-card methods for improving sight-word reading. *Journal of applied behavior analysis*, *44*(4), 781-792.

Lane, K. L., Royer, D. J., Messenger, M. L., Common, E. A., Ennis, R. P., & Swogger, E. D. (2015). Empowering teachers with low-intensity strategies to support academic engagement: Implementation and effects of instructional choice for elementary students in inclusive settings. *Education and Treatment of Children*, *38*(4), 473-504.

Leaf, J. B., Sheldon, J. B., & Sherman, J. A. (2010). Comparison of simultaneous prompting and no-no prompting in two-choice discrimination learning with children with autism. *Journal of Applied Behavior Analysis*, *43*(2), 215–228. https://doi.org/10.1901/jaba.2010.43-215.

Lembke, E. S., Hampton, D., & Beyers, S. J. (2012). Response to intervention in mathematics: Critical elements. *Psychology in the Schools*, *49*(3), 257-272.

Libby, M. E., Weiss, J. S., Bancroft, S., & Ahearn, W. H. (2008). A comparison of most-to-least and least-to-most prompting on the acquisition of solitary play skills. *Behavior Analysis in Practice*, *1*(1), 37-43.

Lozy, E. D., & Donaldson, J. M. (2019). A comparison of traditional drill and strategic incremental rehearsal flashcard methods to teach letter–sound correspondence. *Behavioral Development*, *24*(2), 58.

MacQuarrie, L. L., Tucker, J. A., Burns, M. K., & Hartman, B. (2002). Comparison of retention rates using traditional, drill sandwich, and incremental rehearsal flash card methods. *School Psychology Review*, *31*(4), 584-595.

Malouf, R., Reisener, C., Gadke, D., Wimbish, S., & Frankel, A. (2014). The effect of helping early literacy with practice strategies on reading fluency for children with severe reading impairments. *Reading Improvement*, *51*(2), 269-279.

Martens, B. K., Werder, C. S., Hier, B. O., & Koenig, E. A. (2013). Fluency training in phoneme blending: A preliminary study of generalized effects. *Journal of Behavioral Education*, *22*(1), 16-36.

Marvin, K. L., Rapp, J. T., Stenske, M. T., Rojas, N. R., Swanson, G. J., & Bartlett, S. M. (2010). Response repetition as an error-correction procedure for sight-word reading: A replication and extension. *Behavioral Interventions: Theory & Practice in Residential & Community-Based Clinical Programs*, *25*(2), 109-127.

Mautone, J. A., DuPaul, G. J., Jitendra, A. K., Tresco, K. E., Junod, R. V., & Volpe, R. J. (2009). The relationship between treatment integrity and acceptability of reading interventions for children with Attention-Deficit/Hyperactivity Disorder. *Psychology in the Schools*, *46*(10), 919-931.

McCallum, E., Skinner, C. H., Turner, H., & Saecker, L. (2006). The taped-problems intervention: Increasing multiplication fact fluency using a low-tech, classwide, time-delay intervention. *School Psychology Review*, *35*(3), 419.

McCurdy, E. E., & Cole, C. L. (2014). Use of a peer support intervention for promoting academic engagement of students with autism in general education settings. *Journal of autism and developmental disorders*, *44*(4), 883-893.

McCurdy, M., Skinner, C. H., Grantham, K., Watson, T. S., & Hindman, P. M. (2001). Increasing on-task behavior in an elementary student during mathematics seat-work by interspersing additional brief problems. School Psychology Review, 30, 23–32.

McMaster, K. L., & Wagner, D. (2007). Monitoring response to general education instruction. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.). Handbook of response to intervention: The science and practice of assessment and intervention (pp. 223-233). New York: Springer.

McVancel, S. M., Missall, K. N., & Bruhn, A. L. (2018). Examining Incremental Rehearsal: Multiplication Fluency with Fifth-Grade Students with Math IEP Goals. *Contemporary School Psychology*, *22*(3), 220-232.

Mellard, D., McKnight, M., & Jordan, J. (2010). RTI tier structures and instructional intensity. *Learning Disabilities Research & Practice*, *25*(4), 217-225.

Mercer, S. H., Idler, A. M., & Bartfai, J. M. (2014). Theory-driven evaluation in school psychology intervention research: 2007-2012. *School Psychology Review*, *43*(2), 119-132.

Methe, S. A., Hojnoski, R., Clarke, B., Owens, B. B., Lilley, P. K., Politylo, B. C., ... & Marcotte, A. M. (2011). Innovations and future directions for early numeracy curriculum-based measurement: Commentary on the special series. *Assessment for Effective Intervention*, *36*(4), 200-209.

Miller, K. C., Skinner, C. H., Gibby, L., Galyon, C. E., & Meadows-Allen, S. (2011). Evaluating generalization of addition-fact fluency using the taped-problems procedure in a second-grade classroom. *Journal of Behavioral Education*, *20*(3), 203-220.

Mong, M. D., & Mong, K. W. (2010). Efficacy of two mathematics interventions for enhancing fluency with elementary students. *Journal of Behavioral Education*, *19*(4), 273-288.

Mulé, C. M., Daniels, B., Volpe, R. J., Briesch, A. M., Joseph, L. M., Harris, K., ... & Leslie, L. K. (2018). A comparative effectiveness study of two high-frequency word interventions: Traditional drill and WordSheets. *Journal of Behavioral Education*, *27*(2), 240-261.

Musti-Rao, S., & Plati, E. (2015). Comparing two classwide interventions: Implications of using technology for increasing multiplication fact fluency. *Journal of Behavioral Education*, *24*(4), 418-437.

National Center for Educational Statistics (2018). *Digest of educational statistics.* Washington, DC: U.S. Department of Education.

National Center for Education Statistics (2017). The nation’s report card: 2017 mathematics results. Washington, DC: Institute for Education Sciences (IES), US Department of Education.

National Center on Response to Intervention. (2010, March). Essential components of RtI – A closer look at response to intervention. Washington, DC: U.S. Department of Education, Office of Special Education Programs, National Center on Response to Intervention.

National Mathematics Advisory Panel (2008). Foundations for success: The final report of the National Mathematics Advisory Panel. Washington, DC: U.S. Department of

Education.

Neef, N. A., Iwata, B. A., & Page, T. J. (1980). The effects of interspersal training versus high-density reinforcement on spelling acquisition and retention. *Journal of Applied Behavior Analysis*, *13*(1), 153-158.

Nist, L., & Joseph, L. M. (2008). Effectiveness and efficiency of flashcard drill instructional methods on urban first-graders' word recognition, acquisition, maintenance, and generalization. *School Psychology Review*, *37*(3), 294.

Parker, D. C., & Burns, M. K. (2014). Using the instructional level as a criterion to target reading interventions. *Reading & Writing Quarterly*, *30*(1), 79-94.

Poncy, B. C., Fontenelle, S. F., & Skinner, C. H. (2013). Using detect, practice, and repair (DPR) to differentiate and individualize math fact instruction in a class-wide setting. Journal of Behavioral Education, 22(3), 211–228.

Poncy, B. C., & Skinner, C. H. (2011). Enhancing first-grade students' addition-fact fluency using classwide cover, copy, and compare, a sprint, and group rewards. *Journal of Applied School Psychology*, *27*(1), 1-20.

Poncy, B. C., Skinner, C. H., & Axtell, P. K. (2010). An investigation of detect, practice, and repair to remedy math-fact deficits in a group of third-grade students. Psychology in the Schools, 47(4), 342–353.

Poncy, B. C., Skinner, C. H., & Jaspers, K. E. (2007). Evaluating and comparing interventions designed to enhance math fact accuracy and fluency: Cover, copy, and compare versus taped problems. *Journal of Behavioral Education*, *16*(1), 27-37.

Poncy, B. C., Skinner, C. H., & McCallum, E. (2012). A comparison of class-wide taped problems and cover, copy, and compare for enhancing mathematics fluency. *Psychology in the Schools*, *49*(8), 744-755.

Poncy, B. C., Skinner, C. H., & O'Mara, T. (2006). Detect, practice, and repair: The effects of a classwide intervention on elementary students' math-fact fluency. Journal of Evidence-Based Practices for Schools, 7(1), 47–68.

Poncy, B. C., Solomon, B., Duhon, G., Skinner, C., Moore, K., & Simons, S. (2015). An analysis of learning rate and curricular scope: Caution when choosing academic interventions based on aggregated outcomes. *School Psychology Review*, *44*(3), 289-305.

Rahn, N. L., Wilson, J., Egan, A., Brandes, D., Kunkel, A., Peterson, M., & McComas, J. (2015). Using incremental rehearsal to teach letter sounds to English language learners. *Education and Treatment of Children*, *38*(1), 71-91.

Rao, S., & Mallow, L. (2009). Using simultaneous prompting procedure to promote recall of multiplication facts by middle school students with cognitive impairment. *Education and Training in Developmental Disabilities*, 80-90.

Reynolds, J. L., Drevon, D. D., Schafer, B., & Schwartz, K. (2016). Response Repetition as an Error-Correction Strategy for Teaching Subtraction Facts. In *School Psychology Forum* (Vol. 10, No. 4).

Rich, S. E. H., Duhon, G. J., & Reynolds, J. (2017). Improving the generalization of computer-based math fluency building through the use of sufficient stimulus exemplars. *Journal of Behavioral Education*, *26*(2), 123-136.

Robinson, S. L., & Skinner, C. H. (2002). Interspersing additional easier items to enhance mathematics performance on subtests requiring different task demands. *School Psychology Quarterly*, *17*(2), 191.

Rogers, L. S., & Ardoin, S. P. (2018). Investigating the benefit of adding listening passage preview to repeated readings. *School Psychology Quarterly*, *33*(3), 439.

Ross, J. A., & Maynes, F. J. (1983). Experimental problem solving: An instructional improvement field experiment. *Journal of Research in Science Teaching*, *20*(6), 543-556.

Ryder, J. F., Tunmer, W. E., & Greaney, K. T. (2008). Explicit instruction in phonemic awareness and phonemically based decoding skills as an intervention strategy for struggling readers in whole language classrooms. *Reading and Writing*, *21*(4), 349-369.

Sarason, S. B. (1990). *The Predictable Failure of Educational Reform: Can We Change Course before It's Too Late? The Jossey-Bass Education Series and the Jossey-Bass Social and Behavioral Science Series*. Jossey-Bass, Inc., Publishers, PO Box 44305, San Francisco, CA 94144-4305.

Sander, J. B. (in press). Only Disproportionality in School Discipline and Juvenile Justice Systems: The Past and Future Role of Psychological Consultation to Promote Equity. *Journal of Educational & Psychological Consultation.*

Shapiro, E. S. (2015). Improving student achievement by infusing highly effective instructional strategies into multi-tiered support systems (MTSS)—Response to intervention (RtI) tier 2 instruction. *The National Council of Supervisors of Mathematics Improving Students Achievement Series*, *15*.

Shinn, M. R. (1989). Curriculum-based measurement: Assessing special children. New York: The Guilford Press.

Silber, J. M., & Martens, B. K. (2010). Programming for the generalization of oral reading fluency: Repeated readings of entire text versus multiple exemplars. *Journal of Behavioral Education*, *19*(1), 30-46.

Silberglitt, B., Parker, D., & Muyskens, P. (2016). Assessment: Periodic assessment to monitor progress. In *Handbook of response to intervention* (pp. 271-291). Springer, Boston, MA.

Singer-Dudek, J., & Greer, R. D. (2005). Long-term analysis of the relationship between fluency and the training and maintenance of complex math skills. The Psychological Record, 55, 361–376.

Skinner, C. H., & Daly, E. J. (2010). Improving generalization of academic skills: Commentary on the special issue. *Journal of Behavioral Education*, *19*(1), 106-115.

Skinner, C. H., Johnson, C. W., Larkin, M. J., Lessley, D. J., & Glowacki, M. L. (1995). The influence of rate of presentation during taped-words interventions on reading performance. *Journal of Emotional and Behavioral Disorders*, *3*(4), 214-223.

Skinner, C. H., Logan, P., Robinson, S. L., & Robinson, D. H. (1997). Demonstration as a reading intervention for exceptional learners. *School Psychology Review*, *26*(3), 437-447.

Stoiber, K. C., & DeSmet, J. L. (2010). Guidelines for evidence-based practice in selecting interventions. *Practical handbook of school psychology: effective practices for the 21st century*, 213-234.

Stoiber, K. C., & Gettinger, M. (2016). Multi-tiered systems of support and evidence-based practices. In *Handbook of response to intervention* (pp. 121-141). Springer, Boston, MA.

Strait, G. G., Smith, B. H., Pender, C., Malone, P. S., Roberts, J., & Hall, J. D. (2015). The reliability of randomly generated math curriculum-based measurements. *Assessment for Effective Intervention*, *40*(4), 247-253.

Swehla, S. E., Burns, M. K., Zaslofsky, A. F., Hall, M. S., Varma, S., & Volpe, R. J. (2016). Examining the use of spacing effect to increase the efficiency of incremental rehearsal. *Psychology in the Schools*, *53*(4), 404-415.

Szadokierski, I., & Burns, M. K. (2008). Analogue evaluation of the effects of opportunities to respond and ratios of known items within drill rehearsal of Esperanto words. *Journal of School Psychology*, *46*(5), 593-609.

Tabachnick, B. G., & Fidell, L. S. (2014). *Using Multivariate* Statistics (6th ed). Pearson.

Therrien, W. J. (2004). Fluency and comprehension gains as a result of repeated reading: A meta-analysis. Remedial and Special Education, 25, 252–261. http://dx.doi.org/10.1177/07419325040250040801

Tilly III, W. D. (2008). The evolution of school psychology to science-based practice. In A. Thomas & J. Grimes (Eds.), Best practices in school psychology V (pp. 18–32). Washington, DC: National Association of School Psychologists.

Tolar, T. D., Lederberg, A. R., & Fletcher, J. M. (2009). A structural model of algebra achievement: computational fluency and spatial visualisation as mediators of the effect of working memory on algebra achievement. *Educational Psychology*, *29*(2), 239-266.

Tomczak, M., & Tomczak, E. (2014). The need to report effect size estimates revisited. An overview of some recommended measures of effect size. *Trends in Sport Sciences, 1*(21), 19 – 25.

Truckenmiller, A. J., Eckert, T. L., Codding, R. S., & Petscher, Y. (2014). Evaluating the impact of feedback on elementary aged students' fluency growth in written expression: A randomized controlled trial. *Journal of school psychology*, *52*(6), 531-548.

Tucker, J. A. (1989). Basic flashcard technique when vocabulary is the goal. Unpublished teaching materials. University of Tennessee at Chattanooga. Chattanooga, TN: Author.

VanDerHeyden, A. M., Witt, J. C., & Gilbertson, D. (2007). A multi-year evaluation of the effects of a response to intervention (RTI) model on identification of children for special education. *Journal of School Psychology*, *45*(2), 225-256.

van Gorp, K., Segers, E., & Verhoeven, L. (2014). Repeated reading intervention effects in kindergartners with partial letter knowledge. *International Journal of Disability, Development and Education*, *61*(3), 225-239.

Van Houten, R., & Little, G. (1982). Increased response rate in special education children following an abrupt reduction in time limit in the absence of a token economy. *Education and Treatment of Children*, 23-32.

Varma, S., & Schleisman, K. B. (2014). The cognitive underpinnings of incremental rehearsal. *School Psychology Review*, *43*(2), 222-229.

Vaughn, S., Linan-Thompson, S., Woodruff, A. L., Murray, C. S., Wanzek, J., Scammacca, N., . . . Elbaum, B. (2009). Effects of professional development on improving at-risk students' performance in reading. In C. R. Greenwood, T. R. Kratochwill, & M. Clements (Eds.), Schoolwide prevention models: Lessons learned in elementary schools (pp. 115–142). New York, NY: Guilford.

Volpe, R. J., Burns, M. K., DuBois, M., & Zaslofsky, A. F. (2011). Computer-assisted tutoring: Teaching letter sounds to kindergarten students using incremental rehearsal. *Psychology in the Schools*, *48*(4), 332-342.

Volpe, R. J., Mulé, C. M., Briesch, A. M., Joseph, L. M., & Burns, M. K. (2011). A comparison of two flashcard drill methods targeting word recognition. *Journal of Behavioral Education*, *20*(2), 117-137.

Volpe, R. J., & Suldo, S. M. (2014). Introduction to the special issue on theoretical frameworks in school psychological intervention research: Interdisciplinary perspectives and future directions. *School Psychology Review*, *43*(2), 115-119.

Wagner, D. L., & Espin, C. A. (2015). The reading fluency and comprehension of fifth-and sixth-grade struggling readers across brief tests of various intervention approaches. *Reading Psychology*, *36*(7), 545-578.

Wang, M. T., & Fredricks, J. A. (2014). The reciprocal links between school engagement, youth problem behaviors, and school dropout during adolescence. *Child development*, *85*(2), 722-737.

Wendling, B. J., & Mather, N. (2009). *Essentials of evidence-based academic interventions* (Vol. 57). John Wiley & Sons.

Wixson, K. (2011). A systemic view of RTI research: Introduction to the special issue. *The Elementary School Journal*, *111*(4), 503-510.

Ysseldyke, J., Burns, M. K., Scholin, S. E., & Parker, D. C. (2010). Instructionally valid assessment within response to intervention. *Teaching Exceptional Children*, *42*(4), 54-61.

Zaslofsky, A. F., Scholin, S. E., Burns, M. K., & Varma, S. (2016). Comparison of opportunities to respond and generation effect as potential causal mechanisms for incremental rehearsal with multiplication combinations. *Journal of school psychology*, *55*, 71-78.