# REMOTE PROFESSIONAL DEVELOPMENT FOR THE ROOTS INTERVENTION:

# A CONCEPTUAL REPLICATION STUDY

by

DAVID FAINSTEIN

A DISSERTATION

Presented to the Department of Special Education and Clinical Sciences
and the Division of Graduate Studies of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2022

DISSERTATION APPROVAL PAGE

Student: David Fainstein

Title: Remote Professional Development for the ROOTS Intervention: A Conceptual Replication Study

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Special Education and Clinical Sciences department by:

| | |
|---|---|
| Dr. Ben Clarke | Chairperson |
| Dr. Lauren Cycyk | Core Member |
| Dr. Hank Fien | Core Member |
| Dr. Derek Kosty | Core Member |
| Dr. David Liebowitz | Institutional Representative |

and

| | |
|---|---|
| Andrew Karduna | Interim Vice Provost for Graduate Studies |

Original approval signatures are on file with the University of Oregon Division of Graduate Studies.

Degree awarded June 2022

## DISSERTATION ABSTRACT

David Fainstein

Doctor of Philosophy

Special Education and Clinical Sciences

June 2022

Title: Remote Professional Development for the ROOTS Intervention: A Conceptual Replication Study

Early numeracy skills are vital for kindergarten students because they are predictive of later success in school. Unfortunately, many kindergarteners enter school with numeracy skills that fall below national standards, and our youngest learners begin school with an achievement gap in mathematics that widens over time. Intervening early on kindergarteners at-risk for poor mathematics learning outcomes is a promising approach to address this achievement gap. ROOTS is a small group evidence-based, early numeracy intervention that focuses on whole number development for at-risk learners. The current study explores the feasibility, acceptability, and impact of ROOTS with remote professional development in a small-scale randomized controlled trial. Findings indicate that interventionists receiving remote support for implementing ROOTS found training and coaching to be sufficient to deploy an early numeracy intervention, and strong implementation fidelity was consistent with previous research on ROOTS. Further, children who received the ROOTS intervention outperformed their control group peers on post-test measures of mathematics achievement. Outcomes and implications for using remote professional development as a format for supporting educators in their use of early numeracy intervention programs are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: David Fainstein

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
Mount Saint Vincent University, Halifax, NS
University of Victoria, Victoria, BC

DEGREES AWARDED:

Doctor of Philosophy, School Psychology, 2022, University of Oregon
Master of Science, Special Education, 2021, University of Oregon
Master of Arts, School Psychology, 2015, Mount Saint Vincent University
Bachelor of Science, Psychology, 2013, University of Victoria

AREAS OF SPECIAL INTEREST:

Supporting educators with the adoption and implementation of academic interventions
Instructional coaching as professional development for educators and families
Academic assessment with instructional utility for children, families, and educators

PROFESSIONAL EXPERIENCE:

Graduate Employee - Research, UO Center on Teaching and Learning, 2020-present

Screening Facilitator, Prevention Science Institute, 2020-2021

Graduate Employee - Administrative, Student Conduct and Community Standards in the Dean of Students Office, 2018-2020

Advanced Practice Student, UO Center on Teaching and Learning, 2019-2020

Academic Interventionist, CTL Academic Clinic, 2019

School Psychologist, Surrey School District, 2015-2018

Inclusion Facilitator, Mount Saint Vincent University Disability Services, 2014-2015

Special Education Assistant, Mosaic Learning Independent School, 2012-2013

Behavioural Interventionist, Child and Family Counselling Association, 2009-2012

Student Coordinator, University of Victoria Education Technology Integration and Evaluation Laboratory, 2011-2012

GRANTS, AWARDS, AND HONORS:

Doctoral Fellowship, Social Science and Humanities Research Council of Canada, 2020-present

Special Opportunities Training Grant, University of Oregon, 2020

Travel Grant, Surrey School District, 2018

General Research Grant, Mount Saint Vincent University, 2014

BMO Endowed Scholarship, Mount Saint Vincent University, 2013

President's Scholarship, University of Victoria, 2012

Undergraduate Research Award, University of Victoria, 2011

PUBLICATIONS:

Clarke, B., Sutherland, M., Doabler, C., Lesner, T., **Fainstein, D.**, Nolan, K, Landis, B., & Kosty, D. (2021). Developing and Investigating the Promise of Early Measurement Screeners. *School Psychology Review*, ahead-of-print, 1-13. 10.1080/2372966X.2021.1919493

**Fainstein, D.,** Whitney, M., Avelar, J., Swinehart, J. (in press). Exploring Prevalence Rates of Students with Exceptionalities in British Columbia. *International Journal of Technology and Inclusive Education*, 10(1). ISSN: 2047-0533

**Fainstein, D.,** Lussier, C., Cook, M., Men, V. (in press). Professional Learning Checklist for a Remote Delivery Format. *Literacy Information and Computer Education Journal*, 12(1). ISSN: 2040-2589

Sutherland, M., **Fainstein, D.**, Lesner, T., Kimmel, G. L., Clarke, B., & Doabler, C. T. (submitted for review). Teaching data analysis and statistical proficiency to students with mathematics difficulties. *Teaching Exceptional Children.*

# ACKNOWLEDGMENTS

I wish to express sincere gratitude to my dissertation committee, the ROOTS project staff, and all staff and students who participated in this study. To my committee, Drs. Ben Clarke, Derek Kosty, Lauren Cycyk, David Liebowitz, and Hank Fien, I thank you for your guidance, grace, and relentless pursuit for rigor in educational science. My advisor, Ben Clarke, deserves additional recognition for the insights and motivations that he provided to me over the past few years to enact a systems-level approach to finding solutions as a scientist and practitioner working in schools.

Running an early intervention pilot project during a global health crisis is ambitious. I am incredibly grateful for the support of so many skilled, passionate researchers and practitioners who supported this work. The following contributors were instrumental, and this pilot project could not have happened without them (in alphabetical order): Madison Cook, Chris Doabler, Georgia Kimmel, Taylor Lesner, Maria Longhi, Cayla Lussier, Marcia Moore, Janet Otterstedt, Teresa Schantin, Marah Sutherland, and Jessica Turtura. I owe perhaps the deepest debt of gratitude to the educators, students, and families in Montana who navigated a challenging school year and also found time to prioritize early numeracy. Thanks to everyone who contributed to this pilot project for your participation, learning, and laughing along the way.

The investigation was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324R200005 to the Center on Teaching and Learning, University of Oregon. Additionally, the Social Sciences and Humanities Research Council of Canada provided me with a fellowship to support my research line as a social scientist working in the education field, and I am deeply thankful for their support.

To my friends near and far - balancing work with play in the Pacific Northwest made this doctorate possible. Thank you for all of the adventures before, during, and after my time in Oregon.

To my friends and colleagues working in schools - thank you for your dedication to students and families. This doctorate has given me time to think deeply and grow to love our field more. I hope that our work on equity, promotion of resilience, and advancement of effective instructional practice continues forever.

To my family - I am excited to continue learning from you, with you, and for you. We are teachers, environmental stewards, administrators, counsellors, health professionals, and more, and I am convinced that I would have never found my passions without you. Thank you for everything. I dedicate this dissertation to you, my family.

TABLE OF CONTENTS


Chapter Page

I. INTRODUCTION .......... 1

Factors Influencing Mathematics Skills at School Entry .......... 2

The Importance of Kindergarten for Numeracy Development .......... 5

Early Numeracy Intervention in Kindergarten .......... 6

Research Applications of Early Numeracy Interventions .......... 8

Research Replication in an Educational Context .......... 9

Considerations in Conceptual Replication Research .......... 12

Interventionist Training and Coaching in Educational Research .......... 13

Piloting an Early Intervention Replication Study During a Pandemic .......... 15

Study Purpose and Research Questions .......... 17

II. METHOD .......... 21

Design .......... 21

Procedures .......... 22

Participants .......... 22

Schools .......... 22

Interventionists .......... 23

Students .......... 23

Demographic Characteristics .......... 24

ROOTS Intervention .......... 27

Control Condition .......... 28

PD and Coaching .......... 29

Chapter Page


PD Trainers and Coaches ... 31

Measures ... 31

Feasibility and Acceptability of remote PD ... 31

Implementation Fidelity ... 32

Student Mathematics Achievement Measures ... 34

Data Collection ... 35

Analysis ... 36

III. RESULTS ... 39

Attrition and Missing Data ... 39

Acceptability and Feasibility ... 40

Implementation Fidelity ... 44

Student Outcomes ... 45

Multilevel Model Analysis ... 47

IV. DISCUSSION ... 51

Further Considerations for Interpretation ... 55

Limitations and Future Directions ... 57

Conclusion ... 62

APPENDICES ... 64

A. CURRICULUM STANDARDS ... 64

B. ROOTS TRAINING WORKSHOP FEEDBACK ... 65

C. POST-INTERVENTION SURVEY RESULTS ... 66

D. RCMIS RESULTS ... 69

Chapter Page

E. SUPPLEMENTAL MULTILEVEL MODEL RESULTS ................................ 71

REFERENCES CITED........................................................................................ 73

LIST OF FIGURES

Figure Page

1. RAENS Time × Condition Model Results .......................................................... 49

2. ASPENS Time × Condition Model Results ......................................................... 50

LIST OF TABLES


Table Page

1. Student Demographics and Teacher Perceptions Data by Condition .................. 25

2. Teacher and Interventionist Demographics .......................................................... 26

3. Post-Training Survey Feedback for Interventionists ............................................ 43

4. Summary of Interventionist Perceptions on Acceptability of Online PD for ROOTS......................................................................................... 44

5. Descriptive Statistics for Student Mathematics Achievement Measures by Condition .......................................................................................... 46

6. Results of Partially Nested Analyses that Compared Winter-to-Spring Gains in Math Scores between ROOTS Students Nested within Groups to Unclustered Control Students ........................................................................... 48

E1. Supplemental Multilevel Modeling Results for Student Outcomes on RAENS and ASPENS........................................................................................ 72

# CHAPTER I

# INTRODUCTION

Difficulty learning mathematics at a young age is associated with persistent challenges in acquiring functional numeracy and later adverse outcomes in adult life (Ritchie & Bates, 2013; Watts et al., 2014). The National Assessment of Education Progress (NAEP; 2019) places approximately 59% of American fourth grade students below proficient in mathematics. This concerning level of students struggling in mathematics has held steady for over two decades. Disproportionalities among the "below proficient" group are even more concerning, as African American, Hispanic, Native American, and lower socioeconomic students overrepresent the "below basic" level of mathematics achievement in comparison to their higher socioeconomic or Caucasian peers (Duncan & Murnane, 2011; Flores, 2007; Quinn et al., 2016). Given these stark and consistent findings of the challenges students in the United States face in acquiring numeracy skills, examining factors impacting student mathematics achievement is warranted, including the importance of a successful early start in mathematics. Longitudinal studies indicate that early (e.g., preschool and kindergarten) mathematics difficulties strongly predict later mathematics difficulties through elementary, middle, and high-school settings (e.g., Geary et al., 2013; Lyons et al., 2014; Nguyen et al., 2016). The relationship between student mathematics proficiency at school entry and later mathematics performance is strong for discrete areas (e.g., number operations; Major et al., 2017) as well as broader areas and applications of mathematics (e.g., functional numeracy; Geary et al., 2013). Jordan and colleagues (2009) found that number competence at kindergarten entry is the strongest predictor of mathematics achievement and rate of growth of mathematics through third grade. Further, several authors have pointed to low numeracy skills in kindergarten as a strong predictor of several success indicators well beyond formal school and into adulthood (Mazzocco, 2016; Reigosa-Crespo & Castro Cañizares, 2015; Romano et al., 2010). Low student numeracy skills in the kindergarten year link to an arduous path of mathematics skill acquisition throughout K-12 education, and then links to downstream socioeconomic (e.g., income potential, navigating financial decisions), health (e.g., understanding medical information), and adaptive (e.g., understanding variability in real-world data) outcomes. Leveraging the predictive power of understanding a student's mathematics achievement in kindergarten provides a logical rationale for examining factors that contribute to the success of early intervention programs.

## Factors Influencing Mathematics Skill at School Entry

Children arrive at kindergarten with various educational, academic, socioemotional, and broader life experiences that impacts school readiness and potential need for an early numeracy intervention (Rimm-Kaufman & Pianta, 2000). Factors such as numeracy experiences in the home and community as well as formal and informal educational experiences relate to a child's facility with number when beginning kindergarten. Specifically, home numeracy environments play an important role in forming the foundation for children to become familiar with and then build on mathematical concepts (Kleemans et al., 2012). Missall (2015) discusses parental beliefs on the importance of mathematics, mathematics skill-level, and incorporation of math-based activities in the home as areas that may contribute to the trajectory of mathematics achievement for young children. Research on the home numeracy environment continues to expand the literature base with multiple lines of inquiry pointing to the importance of including numeracy experiences at an early age for children to start learning mathematical concepts (Lukie et al., 2014; Napoli & Purpura, 2018; Niklas & Schneider, 2014; Susperreguy et al., 2020; Thompson et al., 2017). From an early numeracy developmental theorist ideal perspective, parents incorporate early numeracy experiences (e.g., counting Cheerios during breakfast) as a mirror to early literacy experiences (e.g., reading a story to a child during their bedtime routine), such that children are receiving an informal learning environment in the home far ahead of a formal learning environment in school (i.e., kindergarten; Niklas et al., 2016). However, research shows that parents are much more likely to incorporate literacy activities with higher frequency in the home than numeracy experiences (Anders et al., 2012; Huntsinger et al., 2016; Skwarchuk, 2009). The absence of home numeracy experiences is a potentially compounding difficulty for children at-risk for math learning difficulty (MLD) because instruction for key ideas in mathematics (e.g., cardinality) can be quickly covered at the beginning of kindergarten. In contrast, other kindergarteners have encountered these concepts and skills for multiple months before school entry and may benefit from the increase in instructional pacing through known topics. Taken together, exposure to and experience with numeracy before kindergarten arrival is one major factor that speaks to the range of numeracy skills children acquire by their entry point to formal school.

To support students' acclimatization to a new learning environment, some school systems implement formalized kindergarten transition programs (e.g., family orientations, staggered starts, summer camps), but these are the exception rather than the norm (Purtell et al., 2020). Ideally, the transition to kindergarten would level the playing field for all kindergarteners to meaningfully participate and engage in rich learning experiences; though, kindergarten transition programs alone cannot prepare students with all required academic readiness skills (McIntyre et al., 2010). Instead, children can accumulate academic readiness skills before their kindergarten year through educational opportunities such as preschool. Children who attend formal learning environments before kindergarten have increased opportunities to form academic readiness skills, and recent research suggests that pre-kindergarten formal learning environments are strong predictors of school readiness in the kindergarten year (Ansari et al., 2021; Duncan et al., 2018). This is another factor related to early numeracy development which highlights potential inequities and space for intervention to occur. Preschool settings are essential in the context of providing formal education before kindergarten for low-income families (Duncan & Sojourner, 2013). Preschool can be highly beneficial for all students' academic readiness because it directly relates to kindergartener's mathematical competency (Throndsen et al., 2020). It is especially important for children from low-income families because socioeconomic status is positively related to mathematics skills at kindergarten. National data shows that learners from high-income families typically have initial mathematics skills that are more than one standard deviation higher than their low-income family peers (Reardon & Portilla, 2016). Pre-kindergarten formal learning environments and experiences vary widely (e.g., Head Start versus statewide pre-K; Jenkins et al., 2016), and, unfortunately, some students do not have any experience with formal learning environments before attending kindergarten. Further, factors such as alignment between preschool and kindergarten teacher pedagogical philosophies (Abry et al., 2015), peer relationships and social adjustment (McWayne et al., 2004), and selfregulatory abilities in the preschool setting (Gaias et al., 2016) are all linked to mathematical skill at school entry. Due to the high degree of variability in home numeracy environments, preschool experience, and other critical factors, school systems will welcome kindergarten students who have a range of background knowledge that primes the learning of mathematics. Schools must be ready to implement actions that support kindergarten numeracy development. Intensifying instruction through the use of kindergarten intervention programs may be best suited to support the growth and acquisition of foundational academic skills for students who lack access to transition programs, rich home numeracy environments, and preschool learning experiences (Shanley et al., 2017).

**The Importance of Kindergarten for Numeracy Development**

The kindergarten year is a widely accepted common-ground for formal education to occur and a critically important year for educators to set early learners on positive trajectories for life success (Chetty et al., 2011, 2014). As part of realizing an ideal learning trajectory that begins in kindergarten, Claessens (2009) suggests that student academic skills should be the primary focus for long-term success in school. Many practitioners and researchers have observed a shift towards increasing academic instruction and intervention efforts for younger students (e.g., kindergarten and preschool) to bolster student skills as early as possible (Bassok et al., 2016; Bassok & Latham, 2017). Nonetheless, an increasing nationwide focus on academic skill attainment in kindergarten has not closed the achievement gap between low- and high-income students, nor eradicated concerns for learners who arrive at kindergarten with opportunity gaps in early numeracy experiences (Huinker et al., 2020; NAEP, 2019).

In comparison to literacy, mathematics is often a second priority for kindergarten teacher instructional time and lesson delivery (Engel et al., 2013, 2016). Moreover, the delivery of mathematics instruction in kindergarten may not be well-suited to the instructional needs of students, as researchers continually locate a mismatch between what is taught (e.g., shapes) and student skill gaps (e.g., counting by multiples; Chiatovich & Stipek, 2016; Engel et al., 2013). Specifically, Engel et al. (2015) found that kindergarten teachers spend too much time on mathematics content that students already know and too little time on advanced mathematics content that is empirically understood to promote mathematics skill growth over the kindergarten year. This misalignment can be problematic because the scope of mathematics instruction may miss two marks: students do not encounter advanced concepts and skill prior to first grade and struggling students cannot close skill gaps. Given this, researchers suggest a renewed focus on advancing mathematics instructional delivery to address more challenging content and provide space to intervene with the students who enter kindergarten with significant skill gaps and are at-risk for MLD (Duncan et al., 2015; Engel et al., 2016; Jenkins et al., 2018). One potential solution to address this challenge is to design and deliver early numeracy interventions. In the following section, research to date on early numeracy interventions is reviewed.

**Early Numeracy Intervention in Kindergarten**

Early numeracy intervention in kindergarten can establish positive trajectories of mathematics skill acquisition (Dyson et al., 2013; Jordan et al., 2009; Jordan & Dyson, 2016; Morgan et al., 2009). As such, early intervention for students at-risk for mathematics learning difficulties (MLD) continues to be an empirical area of inquiry

(Bailey et al., 2020; Jordan et al., 2009; Libertus et al., 2011). There are several examples of empirically rigorous mathematics interventions designed for the kindergarten year (Clarke et al., 2016; Dyson et al., 2013; Klein et al., 2008; Toll & Van Luit, 2012); many of these are built to improve student proficiency in early mathematics to prevent MLD in future years. The design, content, approach, and incorporation of instructionally relevant variables (e.g., executive functioning training, and language differentiation) vary widely in kindergarten mathematics interventions (Nelson & McMaster, 2019b). Still, the growing body of literature on mathematics interventions for kindergarteners is becoming more robust and suggestive of the meaningful impact that early intervention can have on students at-risk for MLD.

The instructional design landscape of early mathematics interventions that are empirically validated is coming into focus as more programs are available in the literature base (Charitaki et al., 2020). Effect sizes are strongest for instructional techniques that are either known to be effective for academic interventions broadly (e.g., explicit instruction, corrective feedback) or instructional techniques for mathematics specifically (e.g., using concrete manipulatives, visual representations), as found in recent meta-analyses (Charitaki et al., 2020; Chodura et al., 2015; Gersten et al., 2009; Nelson & McMaster, 2019b). Explicit instruction is a powerful instructional approach within the universe of effective interventions for early mathematics (Bryant et al., 2011; Gersten et al., 2009, 2015; Shanley et al., 2017). It is critical for explicit instruction to occur within a coherent curriculum that intentionally builds student knowledge and skills systematically (Archer & Hughes, 2010). National councils of educators and researchers propose that early numeracy intervention should target whole-number concepts (e.g., operations and relations with numbers 1 through 10 inclusive) for intervention (Kilpatrick et al., 2001; National Council of Teachers of Mathematics, 2006). An increasing number of kindergarten mathematics interventions have a whole number focus to address skill gaps observed at school entry (Clarke et al., 2016, 2019; Codding et al., 2011; Jordan & Dyson, 2016; Sood & Jitendra, 2013; Sood & Mackey, 2015). Available evidence suggests that a promising kindergarten numeracy intervention includes explicit instruction with a focus on whole-number concepts.

***Research Applications of Early Numeracy Interventions***

The early numeracy intervention field includes a widening base of randomized-controlled, peer-reviewed studies demonstrating growth in mathematics skills for kindergarteners receiving treatment conditions (e.g., Clarke et al., 2016; DeFlorio et al., 2019; Jordan & Dyson, 2016; Toll & Van Luit, 2012). Nelson and McMaster (2019b) conducted a meta-analysis of preschool, kindergarten, and first-grade numeracy interventions, showing a moderate effect size ($g$ = 0.64) from 34 studies. While the effect size determined through meta-analysis may be influenced by publication bias and heterogeneity among kindergarten and first-grade numeracy interventions (Nelson & McMaster, 2019b), the variety of interventions still point to promise of a number of programs that directly support early numeracy development. One example of an early numeracy intervention from Nelson and McMaster's meta-analysis is the ROOTS program. ROOTS is an evidence-based intervention with an explicit instructional focus on developing whole number understanding (Clarke et al., 2016). The ROOTS intervention builds foundational whole number concepts that are important for kindergarteners at-risk for MLD by providing instruction to small groups (e.g., one interventionist for two to five students) who are likely to benefit from instructional intensification (Clarke et al., 2019). ROOTS is evidence-based in that three peer-reviewed, randomized controlled studies (Clarke et al., 2014, 2016; Doabler et al., 2016) resulted in strong effect sizes for treatment groups (average $g = 0.77$). As researchers continue to investigate early numeracy interventions, it is critical to systematically explore variables related to intervention implementation (Nelson & McMaster, 2019a).

While several factors make an early mathematics intervention such as ROOTS effective, the accumulation of evidence-based interventions is just one step to increase numeracy skills for students at risk. Research on the critical variables related to implementation in real-world settings is integral for the uptake and sustainability of intervening on student skill gaps (Farrell et al., 2019). Implementing researcher-developed mathematics interventions in authentic school settings is difficult without research-practice partnerships that help bring the tools, resources, and knowledge from research institutions into classrooms (Rosenquist et al., 2015). One primary driver to address the limitations of the academic intervention literature is a call for replication research to strengthen the evidence-base for feasibility and scalability (Onken et al., 2014). Specifically, replicating intervention studies is seen in psychological and educational research as a top priority for addressing policy and practice concerns related to intervention outcomes (Makel & Plucker, 2014). Specific to subpopulations within education (e.g., students with disabilities), demonstrating that research findings replicate across contexts is important because of the potential for biases to alter the scientist and practitioner understanding of effective interventions (Cook, 2014; Cook et al., 2018).

**Research Replication in an Educational Context**

Successfully intervening on student academic skills with the same program multiple times builds toward a convergence of evidence on effective instructional practice (Coyne et al., 2016; Doabler et al., 2016). A convergence of evidence is essential for practitioners to meaningfully select and implement interventions that they believe will be effective in their local context (i.e., generalizability; Schmidt, 2009). Despite the importance of leaning on a strong evidence-base for instruction and intervention practice in schools, scientifically important variables (e.g., random assignment to condition, implementation fidelity) in the research base are rarely considered when educators' make instructional adjustments in school settings (Powers et al., 2010; Sloboda et al., 2014). The disconnect between research-informed instructional practice and implementation in classrooms is part of the research-to-practice gap because there is a difference between "that which is published in the literature and that which schools actually carry out" (p. 162, McIntosh et al., 2009). Boosting practitioners' trust in the scientific grounding and replicability of research studies is one avenue to close the research-to-practice gap (Wingen et al., 2019). Replicating intervention research in school settings conceivably increases the likelihood that educators will view research-informed practices as feasible, generalizable, and valid for their local context. Knowing this, several prominent authors in the educational science field have called for an increase to research replication that applies to interventions for K-12 students (Cook et al., 2018; Lemons et al., 2016; McIntosh et al., 2009).

Replication research has only recently gained traction as a priority in human subjects research (Maxwell et al., 2015), and replication studies in the social sciences typically occur in one of two ways: direct replication or conceptual replication (Schmidt, 2009). Direct replication attempts to duplicate findings with the same sample, variables, design, and analysis as an original study. By design, direct replications test the reliability of a statistically significant effect (Simons, 2014). On the other hand, conceptual replication follows an original study closely, but systematically varies one or more components of the initial research study. In special education research, conceptual replication is often preferable for feasibility and generalizability (Coyne et al., 2016). Conceptual replication studies can point to the active ingredients, or essential elements, of an intervention. Knowledge of critical features housed within an intervention program is vital in education because we need to find the most “potent, well-understood, efficacious, effective, and implementable interventions” (p. 30, Onken et al., 2014) to address low academic proficiency rates in K-12 schools (Desilver, 2017).

Conducting a conceptual replication of an academic intervention program should occur as part of an intentional process advancing basic science to implementation and dissemination (Onken et al., 2014; Smolkowski et al., 2019). Onken (2014) provides a framework for the continuum of scientific intervention research that includes cornerstones of high-leverage practice, such as examining efficacy in real-world settings and change mechanisms. This framework (Onken, 2014) includes a specific description of advancing from efficacy trials to effectiveness trials using conceptual replication research. More specifically, researchers should “accept the responsibility of routinely and systematically creating and adapting interventions to the intervention delivery context” (p. 25). Academic interventions have many components worth consideration for systematic variation. For example, dosage, teacher professional development, geographical region, sample, and interventionist selection (e.g., teachers versus instructional assistants) are all examples of variables that can be systematically varied in a conceptual replication. Additionally, researchers are then positioned to study the altered component from the initial intervention for feasibility (e.g., if the professional development format changes, then examine the acceptability of that alteration), through examining factors such as fidelity of implementation to gain insight into how the component functioned and if the intervention was delivered comparably to earlier trials. A focus on the altered components of a conceptual replication study theoretically can provide clarity on the mechanisms driving treatment effects (Coyne et al., 2016).

***Considerations in Conceptual Replication Research***

Feasibility is a key consideration in the research and development of early numeracy interventions because intervention usability leads to authentic implementation (Doabler et al., 2015). Feasibility factors include practicality, ease of implementation, adaptation, user-friendliness, acceptability, and other perceptions from educators that point to realistically implementing an intervention (Bowen et al., 2009). In an early numeracy intervention research context, feasibility data is similar to fidelity data in that reporting ranges from anecdotal to quantitative across studies (e.g., Bryant et al., 2011; Chard et al., 2008). Survey or interview methods for collecting data on perceptions (e.g., social validity) and usability ratings (e.g., Likert scale questions probing intervention delivery difficulty) are typical for use in early numeracy intervention research (e.g., Chard et al., 2008; Doabler et al., 2015; Nozari, 2019; Shanley et al., 2013).

Implementing an academic skill intervention as intended by the program developers is challenging for several reasons (e.g., interventionist fidelity to implementation, delivering intervention uninterrupted in a naturalistic school setting, resolving researcher suggested dosage to intervention group scheduling constraints; Cochrane et al., 2019;

Sanetti & Luh, 2019; Sanetti & Collier-Meek, 2019). In Nelson and McMaster's (2019) meta-analysis, roughly two-thirds of early numeracy intervention studies measured treatment fidelity. However, operationalizing treatment fidelity varies widely across studies, and many authors reported fidelity qualitatively (e.g., vague in-text description of program delivery). Chard et al. (2008) offer a quantitative example of feasibility reporting, detailing the metrics of intervention completion (i.e., median 97% sessions completion, ranging from 86-100%) as well as randomly sampled observation data (i.e., interventionists complete 80% activities within a lesson). Combining observation data with an interventionist self-report of lesson completion rate is ideal for determining intervention response when working with students at-risk for academic skill deficits (Sanetti & Luh, 2019).

**Interventionist Training and Coaching in Educational Research**

Student academic achievement stands to benefit immensely from educators equipped with adequate skills and knowledge to provide effective instruction. Lemons and Toste (2019) drew attention to disparities in teacher professional development (PD) by calling for researchers to identify *how* to develop teacher skills and knowledge through PD to positively impact student achievement. In a recent meta-analysis, Kraft et al. (2018) suggest that an ongoing teacher coaching model of PD is more effective than one-day, large group workshop PD to positively impact teacher instruction and student performance in academic skill development. This holds true in an numeracy intervention context, where the current state of highly variable PD implementation has an unclear link to student mathematics achievement (Gersten et al., 2014). Without strong and sustained inservice PD for educators, the likelihood of strong intervention program fidelity of implementation is low (Kisa & Correnti, 2015).

Kraft and colleagues (2018) expanded on the challenges of scaling-up a coaching model of educator PD by leveraging a technological solution. Using a web-based coaching platform has the potential to provide high-quality PD and reduce financial and time constraints. Pairing a web-based coaching approach with more traditional forms of PD (e.g., initial training session) is especially promising in an educational intervention context because of the potential to individualize context-specific feedback and focus on improving instructional practice (Kraft et al., 2018; Visscher, 2020). Preliminary evidence in this area suggests that research-driven intervention practices can be delivered and improved upon by using mathematics coaches in school-settings (Campbell & Malkus, 2011; Kraft & Hill, 2020). Despite the practicality and initial evidence supporting virtual mathematics coaching, this research line is largely unexplored (Kraft & Hill, 2020).

Within a conceptual replication framework for education (Coyne, 2016), studying the professional development context is a variable worth exploring because it is one of several catalysts that drives intervention outcomes. PD theoretically progresses interventionist understanding and skill level in implementing a specified intervention, and, therefore, PD directly links to fidelity to implementation and potentially student outcomes. Equally crucial for consideration are several variables that confound our understanding of how teacher PD and coaching links to improved student outcomes. First, teacher practice is a complex system in which it is difficult to pinpoint the components of PD and coaching (e.g., knowledge and skills, quality of interactions, perceptions, and values on intervention packages or subject material) that instigate pedagogical enhancement (Lemons & Toste, 2019). Second, measuring teacher coaching efficacy through direct observation (e.g., coach interaction rating scale) and indirect measures (e.g., surveys on the perception of coach knowledge and skills) are in relatively early stages of development (Reddy et al., 2019). Third, school buildings deliver education in complex systems where administrative oversight, school climate, and community support interact with instruction adoption, implementation, and outcomes (Lemons & Toste, 2019). There are additional confounds in understanding the active ingredients in teacher PD and coaching that propel improved student outcomes; yet, examining novel modalities and approaches to PD and coaching is important with concurrent acknowledgment of contextual and measurement confounds. In the 2020-21 school year, the most important contextual factor for consideration was likely the educational and global health context.

**Piloting an Early Intervention Replication Study During a Pandemic**

The public health situation (e.g., the SARS-Cov-2 pandemic, also known as COVID-19) in which this study took place necessitated an online training and coaching platform. This creates both a unique opportunity for novel research and presents questions on the generalizability of outcomes from the current study given the public health and education context of 2020-21. For opportunities, there was a swift transition to online education platforms in March-April 2020 in American schools (Slavin & Storey, 2020), where educators implemented remote teaching with widely varying quality of instruction (Malkus, 2020). While the rigor of high-quality instruction was challenging to maintain, many educators developed a familiarity with digital tools and platforms for instruction

(Daniel, 2020) that would have otherwise been a lower priority for professional skill development in the absence of a global health crisis. For implications on generalizability, COVID-19 is thought to be a once per century disruption in daily life that has incredibly far-reaching impacts (e.g., health, education, industry, governmental functioning; Gates, 2020). We are at the beginning of capturing the effects COVID-19 has had on the systems, functioning, and people in education (Buonsenso et al., 2021).

Conducting intervention research during COVID-19 occurs within a unique context (e.g., personal protective equipment, fluctuating state regulations, societal response to exacerbated inequities, educator burnout) and findings should be interpreted with greater caution because of the dynamic and varied nature of how school systems respond to a pandemic (Wade et al., 2020). For example, in research studies conducted during this period, it is likely that the control condition does not constitute a typical "business-as-usual" counterfactual, due to the entire education system faced significant disruption beginning in March 2020. This disruption has several documented impacts thus far (e.g., educator burnout; Dabrowski, 2020; Pressley, 2021), with a deeper understanding of other impacts forthcoming (e.g., learning loss and recovery, mental health, systemic changes in response to highlighted inequities amid pandemic; Chaturvedi et al., 2021; Engzell et al., 2021). Thus, treatment effects from research studies conducted during this time period should be considered and interpreted with sufficient caution.

In addition to student outcomes, examining social validity outcomes (i.e., participant perceptions on achievement of the goals, procedures, outcomes in an applied research study; Wolf, 1978) that link with the feasibility of intervention program PD are important to consider within a pandemic context. Social validity remains a critical construct for evaluation because of the ties to social significance leading to authentic implementation and sustainability (Chung et al., 2020; Ledford et al., 2016; Spear et al., 2013). A myriad of pandemic-related factors (e.g., educator burnout, reprioritization of curricular foci, social distancing tolerance, and preference) could impact the relationship between perceived social validity and the successful adoption, training, and implementation of an early numeracy intervention. A pandemic context may be most suitable for gathering preliminary social validity data (e.g., acceptability of an online PD delivery format) as opposed to traditional social validity data (Snodgrass et al., 2018) to speak to the promise of novel techniques in early intervention adoption and training.

Fortunately, research on remote PD and coaching specific to early academic intervention settings began prior to the pandemic and showed initial promise in the adoption and use of online platforms for delivering training to educators for packaged intervention programs (e.g., Amendum et al., 2018; Kraft & Hill, 2020). Further, teacher education programs are increasing the prevalence of digital platforms for training, classes, and formal educational interactions (e.g., professional learning communities; Karchmer-Klein & Pytash, 2020). The years-long movement for preservice educators and rapid adoption for inservice educators to use a remote modality for receiving PD highlights the need to better understand educator perceptions on the acceptability and feasibility of this technologically enhanced approach. Recent educational research on remote PD and coaching for school contexts points to a need to better understand acceptability and feasibility as indicators of successful PD to increase the likelihood of strong implementation (Dede et al., 2009; Nese et al., 2020).

**Study Purpose and Research Questions**

The purpose of the current study was to expand on existing research in mathematics intervention and PD by providing virtual PD activities (i.e., initial training and ongoing coaching) to educators within a conceptual replication framework of the ROOTS early numeracy intervention. ROOTS is an efficacious early numeracy intervention that meets rigorous criteria for research design, program implementation, and initial evidence of replicability (Clarke et al., 2016; Doabler et al., 2019; Nelson & McMaster, 2019b). The current study also addresses an earlier call for distal replications of the ROOTS program to demonstrate generalizability across contexts (Doabler et al., 2016). This conceptual replication was distal by design, given the intentional manipulation of a theoretically important contextual variable (i.e., PD was delivered entirely online, versus previous ROOTS program research where all PD was in-person; Coyne, 2016; Schmidt, 2009). Other aspects of ROOTS adoption, delivery, and program assessment were held constant (e.g., intervention was provided in-person, dosage of intervention, group size). Notably, the public health situation in which the current study took place was unique and will be important for consideration when addressing outcomes for interventionists and students. The current study investigates three research questions:

1. *To what degree do interventionists find remote training and coaching acceptable to support their implementation of the ROOTS with remote PD?* Preliminary research on remote PD in an educational intervention context supports the idea that interventionists perceive an online environment as acceptable for training on academic interventions (Amendum et al., 2018; Kraft & Hill, 2020; Kurz et al., 2017; Vernon-Feagans et al., 2013). Given the success in prior academic intervention studies and increased educator familiarity with the tools required for

online training due to COVID-19, it is hypothesized that interventionists will find remote PD (i.e., training and coaching) is acceptable for supporting the implementation of ROOTS. It is noteworthy that remote PD is hypothesized as acceptable, though not preferred over in-person PD, as educators and interventionists may experience burnout from online activities under pandemic conditions (Dabrowski, 2020).

2. *To what degree do interventionists implement the ROOTS intervention program with fidelity after receiving remote PD?* It is hypothesized that implementation fidelity will be similar to previous ROOTS intervention studies (Clarke et al., 2016; Doabler et al., 2016) given the PD and coaching session structures are held constant with in-person delivery of ROOTS (e.g., two synchronous workshops to orient to the ROOTS intervention program, and three coaching sessions throughout ROOTS delivery). Fidelity data was collected on self-reported program completion by instructional logs as well as remote researcher observations of intervention delivery. Taken together, both implementation fidelity data sources are hypothesized to indicate a high degree of implementation fidelity (e.g., all lessons delivered, direct remote observations indicating satisfactory delivery of ROOTS lesson components).
3. *Do students at-risk for MLD who receive the ROOTS intervention program improve mathematics outcomes more than students at-risk for MLD in a control condition following remote PD?* It is hypothesized that students who receive the ROOTS intervention will have higher post-test scores on early numeracy measures than their peers who receive core classroom instruction without ROOTS. Similar to previous findings from research on the ROOTS intervention (Clarke et al., 2016, 2017, 2020;

Doabler et al., 2016), kindergarteners will demonstrate early numeracy skill development when the intervention is delivered as intended (i.e., full 50 lessons of high-quality instruction promoted by ongoing interventionist coaching).

# CHAPTER II

# METHOD

## Design

The COVID-19 pandemic necessitated conditions that fit best with a smaller-scale pilot study. The pilot nature of the current study is intentionally positioned to speak to promising practices of remote PD commensurate with student outcomes in an early numeracy intervention program. A randomized controlled trial (RCT) design was used to conduct the current pilot study. Students were randomly assigned to a treatment (ROOTS with remote PD) or control (business-as-usual math instruction) condition, blocking on classroom, if they were identified as at-risk for MLD by scoring in roughly the bottom half (i.e., lowest 8 students were screened in) of their in-person classroom on an initial screener. Categorizing students in the current study as "at-risk" was intentionally broadened given the malleability of risk status for students at school entry (Morgan et al., 2016). We used a partially nested structure, where students who were deemed at-risk were assigned to an intervention condition or an unnested control condition. Partially nested RCT designs weaken internal validity because grouping effects threaten a robust test of treatment effects (e.g., grouping participants together can conflate treatment effects). However, partially nested designs counter the loss in internal validity with an improvement in external validity because they mimic real-world settings and address ethical concerns of grouping students in the control condition without an academic intervention (Bauer et al., 2008). It is also important to highlight the potential impact of "teacher effects" because teachers were not randomly assigned to a condition in the current study (Weiss, 2010). These threats to validity are important limitations that will be addressed further in the discussion section in concert with implications for research and practice.

## Procedures

The pilot study began in November 2020 with teacher PD and coordination, then pre-test data collection occurred in January 2021, the intervention took place from February through May, 2021, and post-test data collection occurred at the end of the school year, May-June 2021. Given the conditions of COVID-19, some flexibility was anticipated with respect to timelines for data collection, professional development, intervention, and group assignment due to anticipated fluctuations in school, district, and state policies. However, pre-testing and post-testing occurred within scheduled two-week windows, and ROOTS program delivery (i.e., all 50 lessons) occurred in under 12 instructional weeks for all intervention groups.

### *Participants*

**Schools.** Eight interventionists from two schools in different districts located in the Montana participated in the current study. Participating schools were directly recruited through an existing PD arrangement with a University of Oregon Center on Teaching and Learning employee. Educators and administrators agreed to participate for the duration of the study in 2021. Specific to the state where the current study took place, Montana Office of Public Instruction (2020), elementary student ethnicity is predominantly Caucasian (89-92%) with nationally representative identification rates for students receiving special education services (8-9%), and cohorts of approximately 60-70 kindergarteners per school. Participating interventionists, classroom teachers, and schools were are given remuneration in addition to free high-quality mathematics education resources for participating.

**Interventionists.** Participating interventionists are either current kindergarten teachers or paraprofessionals working in primary school classrooms at the participating schools. All participating interventionists were female, either certified teachers or paraprofessional educators, and indicated prior experience teaching math to elementary school students on a post-test survey. Interventionist highest educational attainment ranged broadly: high school diploma (3), associate's degree (1), bachelor's degree (3), and master's degree (1). Similarly, there was a wide range of years of experience working in an educational setting ($M$ = 5.4 years, $SD$ = 6.3 years, range = 1 - 16 years of experience).

**Students.** A two-step process was used to identify students who qualified as a candidate to receive the ROOTS intervention program in the current study. First, students were initially screened for eligibility by teacher report of students participating in in-person instruction and core classroom activities (i.e., kindergarteners who attended school in-person with a classroom of peers). Then, an early numeracy measure, Assessing Student Proficiency in Early Number Sense (ASPENS; Clarke et al., 2011) was used to identify students at-risk for MLD. All students performing in the *strategic* or *intensive* range of the ASPENS middle-of-year composite score, with a maximum of eight per classroom, were screened in for each participating classroom. Previous research points to the *strategic* and *intensive* ranges on ASPENS as being an indicator of risk-status for MLD (Clarke et al., 2011).

At the beginning of the current study, two teachers in one participating school reorganized their classroom composition of students, and seven students in the intervention condition traded core classroom teachers but retained the same interventionists and ROOTS groups. Group attendance data indicated that on average students attended 45 out of 50 lessons (ROOTS lesson attendance: $M = 45.4$, $SD = 4.0$, range = 32.0 - 50.0). When a student receiving ROOTS missed a lesson, no supplementary instruction was provided unless the interventionist consulted directly with their coach to address concerns on missing significant intervention time (e.g., a week, or five lessons, of ROOTS intervention). Two of 28 students in the intervention condition missed at least five consecutive lessons, and both coach-interventionist dyads for these two students decided on individual make-up lessons for students using activities from missed ROOTS lessons.

### *Demographic Characteristics*

Demographics and data on teacher perception of skill for students in the intervention and control conditions are presented in Table 1. Descriptive data on age, sex, race, ethnicity, eligibility for school-based services (e.g., special education, free and reduced price lunch), and attendance was provided by school district staff. Classroom teachers provided survey data on their perceptions of student skill in reading, writing, math, and oral language development on the post-intervention survey. Due to the relatively low number of students in each condition, tests of group mean differences were not conducted to establish baseline equivalence for individual demographic categories (e.g., race/ethnicity, free and reduced priced lunch).

**Table 1**

Student Demographics and Teacher Perceptions Data by Condition

| | Student Group by Condition | | | |
|---|---|---|---|---|
| | Control | | Intervention | |
| Variable | *n* | % | *n* | % |
| Males | 7 | 27 | 11 | 39 |
| Females | 18 | 69 | 17 | 61 |
| Race/Ethnicity | | | | |
| Hispanic | 1 | 4 | 1 | 3 |
| White | 22 | 84 | 26 | 94 |
| Indigenous / Native American | 1 | 4 | 1 | 3 |
| Multi-racial | 1 | 4 | 0 | 0 |
| Not Specified | 1 | 4 | 0 | 0 |
| Special Education Eligibility | 13 | 50 | 16 | 57 |
| Language Service Eligibility | 13 | 50 | 16 | 57 |
| Absent more than 10 Instructional Days during Academic Year | 9 | 35 | 12 | 43 |
| Free-Reduced Price Lunch Eligible | 4 | 15 | 3 | 11 |
| | *M* | *(SD)* | *M* | *(SD)* |
| Age | 5.9 | (0.3) | 5.9 | (0.4) |
| Teacher Perceptions | | | | |
| Student Reading skills | 2.7 | (0.9) | 2.5 | (0.8) |
| Student Writing skills | 2.6 | (0.8) | 2.4 | (0.8) |
| Student Oral Language skills | 2.9 | (0.8) | 2.8 | (0.6) |
| Student Math skills | 2.7 | (0.7) | 2.5 | (0.8) |

*Note*. *M* = mean, *SD* = standard deviation. Sample sizes did not always sum to *N* for each item due to incomplete demographic information available or teacher reporting on student characteristics. Age is reported in years as of the first day that the ROOTS intervention began during Winter 2021. Teacher perception questions used a scale: 1 = "well below grade level"; 2 = "below grade level"; 3 = "solidly average"; 4 = "above grade level"; 5 = "well above grade level". Percentages are rounded to sum to 100 for demographic data.

Teacher and interventionist demographics were collected on the post-intervention survey. Demographic information on sex, race/ethnicity, years of experience, and highest academic degree or diploma obtained are presented in Table 2.

**Table 2**

Participating Teacher and Interventionist Demographics

| Demographic | *n* | % |
|---|---|---|
| Classroom Teachers | | |
| Females | 6 | 75 |
| White, non-Hispanic | 8 | 100 |
| Years of Experience teaching Kindergarten | | |
| 1-5 years | 2 | 25 |
| 6-10 years | 4 | 50 |
| More than 10 years | 2 | 25 |
| Highest Degree Obtained | | |
| Bachelor’s Degree | 6 | 75 |
| Graduate Degree | 2 | 25 |
| Interventionists | | |
| Females | 8 | 100 |
| White, non-Hispanic | 5 | 63 |
| Indigenous or Native American, non-Hispanic | 1 | 13 |
| Years of Experience in Education | | |
| 1-5 years | 6 | 75 |
| 6-10 years | 0 | 0 |
| More than 10 years | 2 | 25 |
| Highest Level of Educational Obtained | | |
| High School Diploma | 3 | 37 |
| Associates Degree | 1 | 13 |
| Bachelor’s Degree | 3 | 37 |
| Graduate Degree | 1 | 13 |

*Note*. Interventionist demographics are incomplete for race/ethnicity, where two participants did not report their race and ethnicity.

***ROOTS Intervention***

ROOTS is a 50-lesson, approximately 20 minutes per lesson, early numeracy intervention that promotes whole number understanding (Clarke et al., 2016). Whole number understanding for kindergarten-aged children links closely with developmentally appropriate number sense skills, which are foundational in developing later skills in mathematics (Clarke et al., 2016). ROOTS enhances number sense (e.g., understanding relationships between quantities; Jordan et al., 2007) by building whole number skills for students at-risk for MLD. In ROOTS, explicit instruction and offering repeated practice opportunities with number and operations (e.g., counting and cardinality, operations and algebraic thinking, operations in base 10) promotes whole number skill acquisition for young learners. Explicit instructional techniques are derived from longstanding research on methods of direct instruction, which seeks to maximize effective and efficient means of teaching students academic content (Carnine et al., 2004). ROOTS is ideal for small groups of kindergarteners (e.g., groups of two to five) multiple times per week (Doabler et al., 2019). Lessons typically include a warm-up routine (e.g., counting to the numbered lesson of the day), review of past content (e.g., vocabulary, numeral identification), math activities (e.g., identifying missing numbers on a number line), and opportunities for practice (e.g., group responses, individual seat work). In the current study, all lessons were delivered in-person, where students worked in small groups with their interventionist in a separate space from the kindergarten classroom.

ROOTS aligns with the Common Core State Standards (CCSS; National Governors Association, 2020) for whole number standards (Clarke et al., 2016). Specifically, the Counting and Cardinality area is covered extensively throughout

ROOTS lessons (1- 50), Operations and Algebraic Thinking curriculum coverage occurs for lessons 11-50, and Number and Operations in Base Ten content coverage is from lesson 21 onwards. Note that the state of Montana does not adopt CCSS as part of their state-level curriculum. However, Montana's standards for kindergarten mathematics map on directly to the CCSS for the content coverage areas of ROOTS. Book one (lessons 1 - 25) and two (lessons 26 - 50) comprise content from both Montana and CCSS standards. See Appendix A for a complete listing of Kindergarten math content areas aligned to CCSS standards, and see Clarke et al. (2016) for a complete description of the ROOTS intervention scope and sequence.

### *Control Condition*

Classroom teachers primarily reported providing 60 minutes of core math instruction per day, four to five days per week. The only exception to this trend was one teacher who reported providing 45 minutes of core math instruction per day. There was a range of responses for the math activities that took place during numeracy instructional blocks; though, all teachers reported providing teacher-led instruction (e.g., whole group lessons), incorporating counting, cardinality, and operations to daily calendar time, and online tools for math practice (e.g., Starfall). There was no alternate intervention program allocated to students in the control condition. As a basis of comparison, students in the control condition received core math instruction, and students in the intervention condition received core math instruction in addition to ROOTS intervention. Classroom teachers verified the delivery of core instruction to all students by sharing their instructional schedule with the interventionists or coaches and scheduling ROOTS groups to occur outside of classroom numeracy instruction or activities (e.g., calendar time).

### *PD and Coaching*

Participating interventionists attended two remote PD training sessions focused on the ROOTS curriculum. These remote trainings were modelled after in-person trainings for ROOTS, which include two four-hour sessions. Previous ROOTS trainings occurred in-person, with one or two content experts presenting to several interventionists (e.g., teachers, instructional assistants). These sessions typically included workshops on the design, delivery, and logistics of implementing ROOTS. Consistent with best-practices in PD for educators (Darling-Hammond et al., 2017; Desimone & Garet, 2015; Hochberg & Desimone, 2010), the in-person workshops included modeling, practice, and review.

The first remote training targeted instructional objectives and routines for lessons 1-25 in conjunction with ancillary training topics (e.g., small group management techniques, logistics, pilot project overview). This training occurred synchronously and interventionists were co-located (i.e., in the same school, at the same time) with trainers located elsewhere. The follow-up training covered lessons 26-50 and followed a similar format to the initial training. Ongoing coaching was provided from February through May 2021. There was a minimum of three and maximum of five coaching visits for all interventionists. Coaching visits focused on post-observation feedback to improve implementation fidelity, and coaches documented their primary activities during their one-on-one visits with interventionists. All PD in the current study upheld research-informed practices for effective professionalized educator training (e.g., active participation within coherent and content-focused training; Desimone, 2009; Desimone & Pak, 2017).

The remote nature of the PD for the current study was designed to mimic in-person PD. For example, PD trainers provided models and examples of ROOTS lesson routines through a live demonstration using a document camera. Many aspects of the in-person ROOTS PD remained consistent with prior ROOTS efficacy studies (e.g., materials, training topics, active engagement). The notable differences between remote PD and in-person PD were physical co-location of trainers with interventionists, technology use, and navigation of technical issues (e.g., internet connectivity). Regarding technology, the ROOTS PD trainers and coaches used the web-based platforms for video conferencing that were already familiar to interventionists, such as *Google Classroom*. A significant benefit of video recording ROOTS implementation is the opportunity for multiple coaches to observe lessons to ensure inter-observer agreement on implementation fidelity using a structured rating scale. Interventionists provided more video recordings during the first half of ROOTS implementation (i.e., two to three recordings for lessons 1-25) to adjust ROOTS implementation habits early, then later video recordings were not required as frequently (i.e., one recording for lessons 26-50).

Interventionists rated the acceptability of the remote ROOTS training after both major training sessions and the acceptability of online coaching after the study. Survey items on acceptability were previously used in studies of the ROOTS intervention (Clarke et al., 2016) for consistency of using the same proxies for social validity on the procedures of an intervention (Foster & Marsh, 1999). In addition, coaching perception data were used in concert with interventionist survey data to triangulate a broader picture of the degree to which interventionists viewed remote support as acceptable for training on and delivering the ROOTS program. For example, coaches provided data after each coaching session on the perceived likelihood that coaching recommendations would translate to instructional adjustment, providing one data point indicating that interventionists find the remote support acceptable enough to take actionable steps in ROOTS delivery.

**PD Trainers and Coaches.** The trainers and coaches for this study were four graduate-level trained professionals on the implementation of ROOTS. With the exception of one additional trainer, all coaches also provided training on lessons 1-25 and 26-50 in workshop formats. All coaches had experience conducting formal PD training for educators and establishing relationships with primary school educators' to encourage strong implementation fidelity of intervention programs. Two coaches were assigned to each school to provide continuity for support in intervention implementation. Coaches used intra-team weekly meetings to problem-solve common issues and ensure consistency of coaching activities and language.

***Measures***

**Feasibility and Acceptability of remote PD.** Interventionists completed researcher-developed surveys to rate each of acceptability and feasibility of ROOTS after receiving remote PD. Group training (e.g., initial training on ROOTS lessons 1 - 25) was followed up with an acceptability survey for participants proximal to the training delivery (e.g., Likert style rating questions on training organization, clarity, and utility alongside qualitative questions on acceptability). Similarly, a combined survey on acceptability and feasibility was delivered to interventionists and classroom teachers at the completion of the ROOTS intervention ("post-intervention survey"). Items on the feasibility survey are designed to assess the practicality of implementing ROOTS after the RCT (e.g., in the 2021-2022 school year) on multiple dimensions of feasibility (e.g., intervention format, fit with at-risk student learning needs, language complexity) using Likert style questions (see Appendix C). Items from this survey were drawn and adapted from existing measures on acceptability and feasibility of early childhood interventions (e.g., Chen et al., 2014; Gwazdauskas, 2009; McCrary, 2011). Due to the unique context in which this pilot study took place, several items were uniquely constructed as they related to COVID-19 (e.g., Likert rating question: "During COVID-19, how important was early numeracy was as an instructional focus for kindergarten students?").

**Implementation Fidelity.** After observing lesson delivery on recorded video Trained coaches completed rating forms on ROOTS program implementation fidelity and student-teacher interactions. The rating forms are adapted from previous ROOTS work (Doabler et al., 2020; Strand Cary et al., 2017) to the current version by modifying the scale from a 4-point Likert scale to a 5-point Likert scale and intended for research use only. More specifically, the Ratings of Classroom Management and Instructional Support (RCMIS; Doabler & Nelson-Walker, 2009) was used as the direct observation measure. The RCMIS is designed for rating early numeracy instruction quality (Doabler et al., 2015). The RCMIS includes 14 items proximal to effective math instruction, and has demonstrated a strong psychometric validity (Chronbach's $\alpha$ = .92; Doabler & Nelson-Walker, 2009). In practice, raters using the RCMIS completed a Likert scale rating for which interventionists met all of the lesson instructional objectives, followed lesson scripting, and incorporated elements of explicit instruction (e.g., offering positive feedback, immediate corrections, practice opportunities). The RCMIS includes two outcomes of interest: a composite score indicating overall implementation quality for an early numeracy intervention, and a subscale specific to explicit instructional techniques (quality of explicit instruction, QEI). Coaches completed co-fidelity checks (i.e., initial coach rating was compared with a different coach rating at a later point in time) for interobserver agreement using video recorded lessons. As an initial calibration, all coaches rated the same video recorded lesson, and the intraclass correlation for the composite RCMIS and QEI indicated interobserver agreement in the excellent and good reliability ranges (ICCs = .92 and .87, respectively; Koo & Li, 2016).

As an indirect measure of implementation fidelity, interventionists completed instructional logs to document lessons completed for ROOTS, which provide data for ROOTS completion rate and pacing. All ROOTS lessons were delivered by each interventionist in under 12 weeks, with the average pace of lesson delivery just under one lesson per instructional day. Common reasons for missing ROOTS lessons on instructional days included school-wide events (e.g., assemblies) and student absences (e.g., two out of three group members absent). The one exception to delivering one ROOTS lesson per instructional day was an interventionist who had a medical emergency, where a substitute interventionist acted quickly to mitigate any learning loss time by providing two ROOTS lessons per day for approximately two weeks.

Coaches observed three lessons per interventionist ($n$ = 8) for a combined 24 total direct observations of adherence to manualized procedures for intervening with the ROOTS early numeracy program. One observation recording was disrupted and not appropriately saved to the online repository, and there was no video recording available to retrieve sufficient data to conduct an observation post-hoc. Therefore, observation data were collected for 23 of 24 coaching occasions. In addition, data were collected on the timing and dosage of ROOTS implementation by both interventionists and coaches.

Interventionists provided project staff with lesson completion log data, which links to the sequence of delivering ROOTS as intended (i.e., one 20-minute lesson per instructional day). All interventionists delivered all 50 ROOTS lessons within the anticipated 12-week treatment window. Coaches observed an average lesson duration of 22.1 minutes (*SD* = 6.0).

Along with using the RCMIS to rate instructional interactions, coaches completed brief survey items on their perceptions of interventionist fidelity to implementation after each coaching visit. There were four items in the Fidelity to Implementation survey section indicating interventionist adherence to ROOTS delivery as intended (see Appendix D for a complete listing of RCMIS, QEI, and fidelity items). Interventionists were rated as using none (1), some (2), most (3), or all (4) components of ROOTS lesson delivery for each fidelity survey item.

**Student Mathematics Achievement Measures.** For a brief, standardized assessment of student mathematics skills, the Assessing Student Proficiency in Early Number Sense (ASPENS; Clarke et al., 2011) was used in the current study. ASPENS includes multiple one-minute fluency- and curriculum-based measures of important mathematics skills for primary students. Specifically, ASPENS includes subtests of number identification, magnitude comparison, and missing number identification. Test-retest reliabilities are strong for kindergarten ASPENS (.74 - .85). Predictive validities use the TerraNova 3 as a standard, and ranges from .45 to .52 from fall to spring time points (Clarke et al., 2017). ASPENS was used as the screening instrument as part of the determination if students are eligible for participation in the current study as well as a pre- to post-test measure to assess the impact of ROOTS on student outcomes.

In addition to ASPENS, the current study used the ROOTS Assessment of Early Numeracy Skills (RAENS; Doabler, Clarke, & Fien, 2012) as a researcher-developed measure of student mathematics achievement and is proximal to the intended ROOTS program outcomes. The RAENS is individually administered with content coverage that aligns with Common Core State Standards Number and Operations strand in Kindergarten (e.g., composing and decomposing numbers 11-19; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). There are 32 items on RAENS, and students have no time limit to complete this measure. The Test of Early Mathematics Ability - Third Edition (TEMA-3) was used for calculating predictive validity with the RAENS, and the resulting correlations ranged from .68 to .83. In previous studies, the RAENS had 100% inter-rater reliability (Clarke et al., 2016). RAENS was used as an early numeracy measure at both pre-test and post-test.

### *Data Collection*

Data collection occurred from January through June, 2021. The University of Oregon Institutional Review Board approved all components and amendments of the current study in November 2020, prior to conducting any human subjects research. Data acquisition procedures are addressed below in order of the abovementioned research questions (i.e., PD acceptability, implementation fidelity, and student achievement):

1. Acceptability surveys of remote PD were delivered via *Qualtrics* and correspond to coaching and training activities. After delivering ROOTS, interventionists provided feedback on both acceptability and feasibility. Together, there were three feedback

points from interventionists (minimum two from group training and one on feasibility post-ROOTS).

2. A minimum of three coaching sessions with incorporated fidelity observations of each interventionist took place in the current study. All fidelity ratings were stored in a collaborative, de-identified spreadsheet for ROOTS coaches to enter their data (i.e., ratings) and cross-check fidelity ratings where appropriate. Additionally, for instructional logs, interventionists had a tracking sheet to share their weekly progress of ROOTS lessons, attendance, and any other relevant data that they wanted to share with their coach and sent that data in by mail at the completion of the intervention.
3. Interventionists administered both ASPENS and RAENS in the current study. Training occurred in December 2020 to ensure that school staff administering student assessments had practice with ASPENS and RAENS before data collection with participating students. All staff achieved a minimum of 90% accuracy in standardized test administration via *Qualtrics* assessment before student assessments occurred. Pre-test and post-test data collection occurred within a two-week window (i.e., ten instructional days) before ROOTS intervention groups were formed (January 2021), and after completion of the last in-program assessment following lesson 50 (May 2021). Interventionists scored all ASPENS and RAENS measures in participating schools, and complete scoring pages were uploaded to an online platform (e.g., *OneDrive*) to ensure the accuracy of scoring and then data entry.

***Analysis***

Data for the first two research questions are examined descriptively and data for the third research question are analyzed using descriptive and inferential statistics. Descriptive analyses for interventionist outcomes (i.e., acceptability of online PD, implementation fidelity) are primarily reported out using central tendency data alongside supplemental qualitative data (e.g., text-response feedback to survey questions). Effects of ROOTS on student outcomes was assessed using a mixed model (multilevel) Time × Condition analysis (Murray, 1998) designed to account for students partially nested within small groups (Baldwin, Bauer, Stice, & Rohde, 2011; Bauer, Sterba, & Hallfors, 2008). The study design called for the randomization of individual students to receive ROOTS, nested within ROOTS groups, or a nonnested comparison condition, and the analytic model must account for the potential heterogeneity among variances across conditions (Roberts & Roberts, 2005). In particular, the ROOTS groups required a group-level variance while the unclustered controls did not. Furthermore, because the residual variances may have differed between conditions, we tested the assumption of homoscedasticity of residuals. The analysis tested for differences between conditions on gains in outcomes from pretest to posttest and is described in detail by Clarke et al. (2016) and Doabler et al. (2016). The statistical model included time, coded 0 at pretest and 1 at posttest, condition, coded 0 for control and 1 for ROOTS, and the interaction between the two. These models test for net differences between conditions (Murray, 1998), which provide an unbiased and straightforward interpretation of the results (Allison, 1990; Jamieson, 1999). Models were fit using SAS statistical software and robust standard errors for fixed effects were estimated using the empirical method (SAS

Institute Inc, 2010). Effect sizes for student outcomes are reported using Hedges' *g* (Cohen, 1988; Hedges, 1981).

# CHAPTER III

# RESULTS

The current early numeracy intervention study examined two primary units for analysis: educator outcomes (i.e., acceptability, feasibility, and fidelity to implementation) and student outcomes (i.e., mathematics achievement). Due to the relatively small sample size and exploratory nature of the current study, results are primarily offered descriptively. Sample characteristics, attrition, and missing data are reported for the whole sample as well as by condition. Student outcomes include inferential statistics contrasting intervention versus control groups to examine the impact of ROOTS as an intervention program for students experiencing difficult with early numeracy acquisition. Results are summarized below with item-level data made available in the appendices.

## Attrition and Missing Data

Three groups actively participated in the current study: (1) students within control or intervention conditions, (2) educators providing ROOTS intervention or core instruction, and (3) coaches providing direct support to ROOTS interventionists. First, regarding student-level data, all of the 54 students randomized to either control ($n$ = 26) or intervention ($n$ = 28) condition completed pre- and post-test mathematics achievement measures. Student attendance data was provided for nearly all of the lessons for all students in the intervention condition (i.e., 216/224 instances of attendance were reported to research staff).

Second, for interventionist-level data, a total of ten interventionists were trained to deliver the ROOTS intervention where eight interventionists were assigned groups to run the ROOTS program, and two interventionists were trained as substitute interventionists to provide coverage for ROOTS groups when the primary interventionist was absent. One interventionist had an accident during the first third of the ROOTS program, and a trained substitute was assigned so that students did not face a gap in intervention service delivery. All other interventionists completed at least 42 of 50 ROOTS lessons (interventionist attendance for lesson delivery: $M$ = 48.0, $SD$ = 2.7, range = 42.0 - 50.0) without significant disruption due to absence (i.e., no more than two consecutive days missed by interventionists).

Third, the four coaches for the current study mostly worked with the same interventionists for the duration of the study. The only exception was a coaching coverage adjustment after the substitute interventionist filled in for the primary interventionist who discontinued.

**Acceptability and Feasibility**

Survey feedback from interventionists was gathered with participants remaining anonymous after each major training (i.e., Book 1 and 2 workshops, 4-hour remote sessions in January and March, respectively) and the results indicated an overall positive impression of synchronous, online training. The first question that interventionists were asked was a perception question on the overall benefit of the trainings with the following scale: 1= "not worthwhile"; 2 = "worthwhile"; 3 = "very worthwhile." All interventionists who completed the post-training surveys ($n$ = 11 total responses from seven interventionists) rated training as "very worthwhile," ($M$ = 3.0, $SD$ = 0). Note that the higher number of responses than interventionists on post-training surveys is a result of two different occasions where interventionists were able to provide feedback (i.e., Book 1 and Book 2 workshops), and because survey responses were submitted anonymously the results are presented as a total mean score per item. Next on the post-training survey, acceptability was probed further through interventionist responses to survey items on perceptions of training delivery commensurate with components of effective PD (e.g., clear organization and structure, presenters developed a rapport, confidence in newly gained instructional skills for delivering ROOTS). Interventionists indicated broad agreement with the inclusion of effective PD components with a mean rating of 4.7 ($SD$ = 0.5) on items related to PD acceptability on the following scale: 1 = "strongly disagree"; 2 = "disagree"; 3 = "neither agree nor disagree"; 4 = "agree"; 5 = "strongly agree." See Table 3 and Appendix B for a summary and complete listing of interventionist responses.

Qualitative data captured by text entry on the post-training surveys pointed to an overall positive impression of the training delivery, with participants remarking "it was fast and to the point!," "the support was offered in a positive way," and "I really found the breakrooms helpful! It was so beneficial to have a smaller group with a coach." Few participants offered suggestions via text-entry; though, those that did most commonly suggested logistical alterations (e.g., different break formats, using a document camera to look over all materials at once instead of a portion of the materials).

After interventionists delivered all 50 ROOTS lessons and student post-testing was complete, a comprehensive post-intervention survey was delivered and completed by all participating interventionists plus one trained, substitute interventionist ($n$ = 9) as well as classroom teachers who had students participating in ROOTS groups ($n$ = 8). The post-intervention survey feedback from interventionists was consistent with the post-training survey feedback in that the responses indicated that ROOTS coaching and training was highly acceptable (i.e., answers to survey items on overall acceptability of remote PD for all respondents was “agree” or “strongly agree”). Further, ROOTS was viewed positively as an appropriate intervention for kindergarteners struggling in early mathematics. Survey items were derived from previously used measures constructed for interventionists to offer their perspective on coaching and PD acceptability during training on a new early intervention program (Maxwell, 2011; McCrary, 2011). In addition to the amended items from prior coaching on early intervention research, several items on the post-intervention survey were uniquely constructed to capture information on interventionist perceptions of the remote or pandemic context specifically. The outcomes on remote PD acceptability again indicated that the virtual format of delivery was not a hinderance to the successful acquisition and improvement of early numeracy instruction using ROOTS. In line with interventionist survey responses on the acceptability of ROOTS overall, classroom teachers viewed ROOTS as feasible to incorporate within their school day (*M* = 4.4, *SD* = 1.1; see response scale in Table 4) and a suitable early numeracy intervention for kindergarteners (*M* = 4.4, *SD* = 1.2). Note that classroom teachers did not receive remote PD, and therefore no data was collected on acceptability of the remote PD format for teachers. See Table 4 for a summary of perceived acceptability from interventionists, and Appendix C for a complete list of acceptability items used for descriptive analysis.

**Table 3**

*Post-Training Survey Feedback for Interventionists*

| Item | Rating ($n$ = 11) | | |
|---|---|---|---|
| | *M* | *(SD)* | Range |
| The training was logical and well organized | 4.7 | (0.5) | 4-5 |
| The information presented was clear | 4.8 | (0.4) | 4-5 |
| The presenters had good rapport with the participants | 4.7 | (0.5) | 4-5 |
| The information I received was relevant and will be useful in my instructional practice | 4.8 | (0.4) | 4-5 |
| I feel that I can teach ROOTS with fidelity | 4.5 | (0.7) | 3-5 |

*Note*. *M* = mean, *SD* = standard deviation. Survey questions used the following scale: 1 = “strongly disagree”; 2 = “disagree”; 3 = “neither agree nor disagree”; 4 = “agree”; 5 = “strongly agree”

Qualitative post-intervention survey data continued the patterns identified from the quantitative data suggesting that interventionists found remote PD acceptable. When asked about their preference for future PD formats absent of a pandemic context, one-third ($n$ = 3) of interventionists indicated a preference for in-person PD, two interventionists indicated that remote training was preferred, and the remaining responses were ambiguous or “not applicable.” It is noteworthy that the informal trend for interventionists who responded in favor of shifting back to in-person PD spoke to the importance of in-person interactions (e.g., “practicing face-to-face is better…”), and the informal trend for interventionists who sided with maintaining a remote PD format spoke to the equivalency and efficiency of digital tools for training (e.g., “I’m not sure what could have been achieved differently in person that wasn’t covered in zoom meetings”).

**Table 4**

*Summary of Interventionist Perception on Acceptability of online PD for ROOTS*

| Item | Rating ($n$ = 9) | | |
|---|---|---|---|
| | *M* | *(SD)* | Range |
| The ROOTS intervention is appropriate for kindergarten students who are struggling in early mathematics. | 4.8 | (0.4) | 4-5 |
| My coach helped me improve my delivery of ROOTS lessons | 4.8 | (0.4) | 4-5 |
| Overall, I benefitted from the support from a coach | 4.8 | (0.4) | 4-5 |
| Having a coach improved my fidelity to the ROOTS program (e.g., following the lesson, using the correct materials, lesson pacing) | 4.7 | (0.5) | 4-5 |
| Virtual/remote coaching activities (i.e., discussion, feedback, or modeling) provided constructive feedback in areas that I wanted to improve | 4.6 | (0.7) | 3-5 |
| Virtual/remote coaching was sufficient for communicating with my coach | 4.6 | (0.7) | 3-5 |
| My skills with technology affected my ability to effectively communicate with my coach | 3.4 | (1.0) | 2-5 |
| Overall, the online format for ROOTS coaching was effective for me | 4.6 | (0.5) | 4-5 |

*Note*. *M* = mean, *SD* = standard deviation. Survey questions used the following scale: 1 = "strongly disagree"; 2 = "disagree"; 3 = "neither agree nor disagree"; 4 = "agree"; 5 = "strongly agree"

**Implementation Fidelity**

The RCMIS serves as the primary observation tool for overall ROOTS delivery, with additional analysis performed on the QEI subscale to document adherence to explicit instruction principles. Intraclass correlations for the composite RCMIS and QEI scores on 26% of available recordings indicate interobserver agreement in the good to moderate ranges (ICCs = .79 and .57, respectively; Koo & Li, 2016). For overall ratings of implementation fidelity, coaches viewed interventionists as having strong adherence to ROOTS intervention protocols by largely meeting math objectives ($M$ = 3.5, $SD$ = 0.6), including lesson components ($M$ = 3.6, $SD$ = 0.6), and using teacher scripting ($M$ = 3.5, $SD$ = 0.7). Coaches also reported that interventionists used the prescribed math models most of the time (M = 3.2, SD = 0.8).

RCMIS items are on a Likert scale ranging from 1 = *not present* to 5 = *highly present*, and the five items within the RCMIS that produce a QEI score use the same scale. Mean scores for individual items on the RCMIS ranged from 3.7 ($SD$ = 1.0) to 4.5 ($SD$ = 0.5), indicating that interventionists predominantly incorporated classroom management and instructional support during ROOTS groups (Appendix D). Based on 23 observations, mean RCMIS scores demonstrate an overall high-quality of math instruction during ROOTS intervention ($M$ = 4.0, $SD$ = 0.7), and the QEI scores were comparably strong in demonstrating high-quality explicit instruction by interventionists ($M$ = 3.9, $SD$ =0.7), which is consistent with ratings of ROOTS interventionists from prior research (Doabler et al., 2016). Each coaching interaction with a RCMIS rating was based on an observation that was arranged by the interventionist-coach dyad (e.g., arranged for two coaching sessions in the first 25 lessons, one coaching session in the last 25 lessons). All sessions included an observation with RCMIS rating.

**Student Outcomes**

Almost all student outcome data (i.e., ASPENS and RAENS) in the current study did not violate significant skew or kurtosis (e.g., $\mp$ 2.0) at pre- or post-test, indicating that most measures across timepoints represented univariate normality. The one exception to likely normally distributed outcome measures is RAENS post-test (skew = -2.37; kurtosis = 6.47). RAENS contains a maximum score of 32, and, at post-test, both intervention and control groups were within one standard deviation of the maximum score, which likely contributes to the exceptional skew and kurtosis. Independent samples *t*-tests on pre-test measures revealed non-significant differences across groups on pre-test ASPENS ($t(52) = 0.77$, $p = 0.442$, $g = -0.21$) and pre-test RAENS ($t(52) = -0.903$, $p = 0.729$, $g = -0.25$). Descriptive statistics for student outcomes by condition are reported in Table 5.

**Table 5**

*Descriptive Statistics for Student Mathematics Achievement Measures by Condition*

| | Control ($n$ = 26) | | Intervention ($n$ = 28) | |
|---|---|---|---|---|
| Measure | *M* | *(SD)* | *M* | *(SD)* |
| ASPENS pre-test | 61.0 | (17.1) | 57.2 | (18.4) |
| ASPENS post-test | 96.0 | (32.5) | 115.4 | (27.6) |
| RAENS pre-test | 21.2 | (6.0) | 22.6 | (5.2) |
| RAENS post-test | 27.5 | (4.9) | 31.0 | (1.7) |

*Note. M* = mean, *SD* = standard deviation. ASPENS = Assessing Student Proficiency in Early Number Sense; RAENS = ROOTS Early Numeracy Assessment.

***Multilevel Model Analysis***

The Time × Condition model results (see Table 6) on ASPENS and RAENS data test for effects of condition, time, and their interaction. Outcomes for ASPENS and RAENS are described in text below; also see Figures 1 and 2 for an illustration of Time × Condition model results. Intervention effects were also examined using multilevel ANCOVA model, which is reported in full in Appendix E.

For RAENS, the intercept indicates that the average pretest score for control students was 21.2. The effect of condition suggests that students in the intervention condition scored 1.4 points higher at pretest than students in the control condition, but the difference was not statistically significant ($t$ = 1.08, $df$ = 72, $p$ = .2838). Results showed a statistically significant effect of time, suggesting that the average gain for control students was 6.2 points from pretest to post-test ($t$ = 6.52, $df$ = 26, $p$ < .0001). Although the Time × Condition interaction was not statistically significant ($t$ = 1.69, $df$ = 53, $p$ = .0969), the estimated difference in gains between ROOTS and control students was 2.2 points and represents a medium effect size (Hedges' $g$ = 0.61, 95% CI [-0.14, 1.36]).

For ASPENS, the intercept indicates that the average pretest score for control students was 61.0. The effect of condition suggests that students in the intervention condition scored 3.7 points lower at pretest than students in the control condition, but the difference was not statistically significant ($t$ = -0.57, $df$ = 91, $p$ = .5717). Results showed a statistically significant effect of time, suggesting that the average gain in ASPENS scores for control students was 35.0 points from pretest to post-test ($t$ = 6.91, $df$ = 54, $p$ < .0001). The Time × Condition interaction fixed effect also achieved for significance for ASPENS ($t$ = 3.29, $df$ = 54, $p$ = . 0018). The Time × Condition model estimated a difference in gains by condition of 23.1 for ASPENS (Hedges' $g = 0.77$, 95% CI [0.34, 1.20]).

**Table 6**

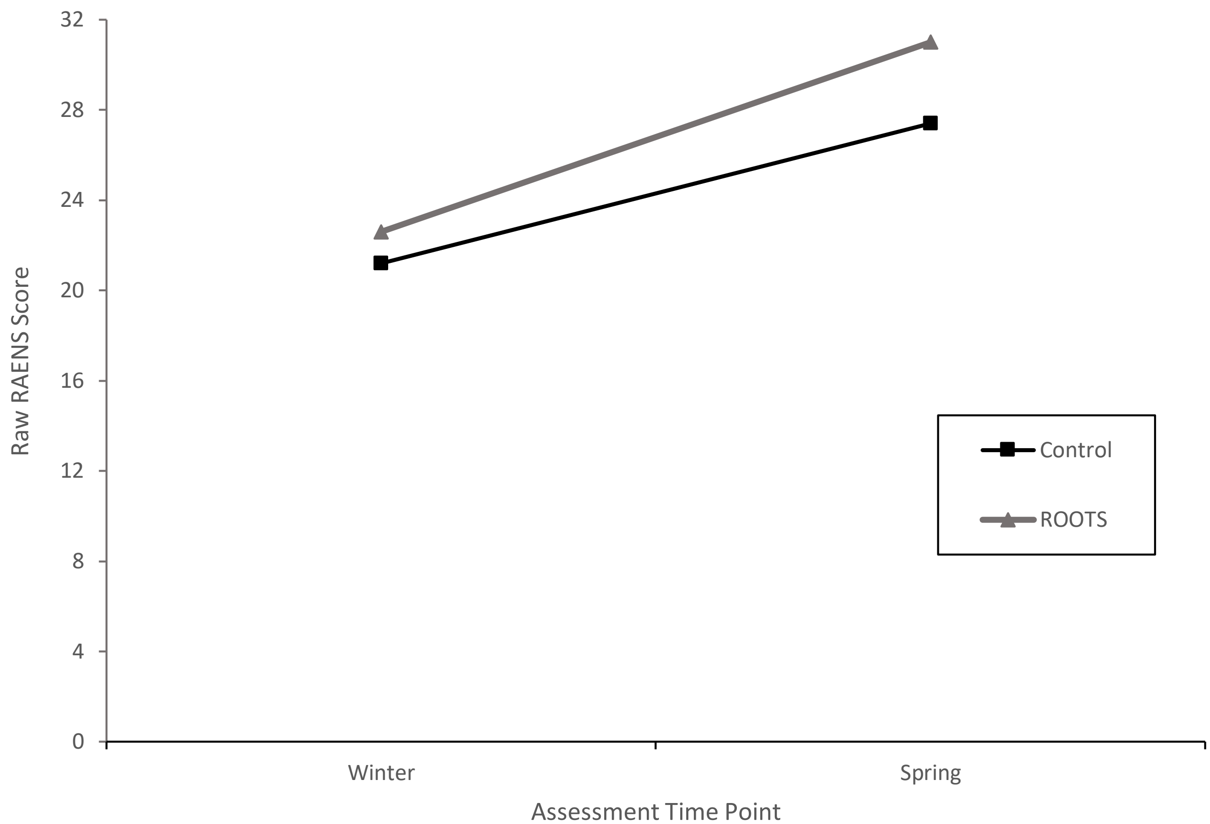
*Results of Partially Nested Time × Condition Analyses that Compared Winter-to-Spring Gains in Math Scores between ROOTS Students Nested within Groups to Unclustered Control Students*

| | | RAENS | ASPENS |
|---|---|---|---|
| Fixed Effects | Intercept | 21.2**** (1.2) | 61.0**** (3.3) |
| | Time | 6.2**** (1.0) | 35.0**** (4.8) |
| | Condition (ROOTS) | 1.4 (1.5) | -3.8 (4.3) |
| | Time × Condition | 2.2 (1.3) | 23.1** (6.4) |
| Variances | T × C between ROOTS Groups | 0.0 (Bounded) | 0.0 (Bounded) |
| | Pre-Post Covariance | | 254.1** (87.1) |
| | Residual | | 333.4**** (64.2) |
| | ROOTS Residual | 10.9**** (2.9) | |
| | ROOTS Pre-Post Covariance | 3.6 (2.8) | |
| | Control Residual | 11.9*** (3.3) | |
| | Control Pre-Post Covariance | 16.9** (6.5) | |
| Hedges' *g* & 95% CI | Time × Condition | 0.61 [-0.14, 1.36] | 0.77 [0.34, 1.20] |
| *p*-values | Time × Condition | .1083 | .0010 |
| *df* | Time × Condition | 32 | 32 |
| Likelihood ratio $\chi^2$ | | 5.64 | 1.05 |
| *p*-values | | .0597 | .5924 |

*Note*. Fixed effects and variances shown with standard errors in parentheses. The models nested only ROOTS students within groups. Bounded = variances were constrained to zero to achieve model fit. Degrees of freedom (*df*) for tests of fixed effects based on the Satterthwaite approximation. Likelihood ratio tests compared homoscedastic to heteroscedastic residuals (α = .20, 2 degrees of freedom).
~$p < .10$. *$p < .05$. **$p < .01$. ***$p < .001$. ****$p < .0001$

**Figure 1**

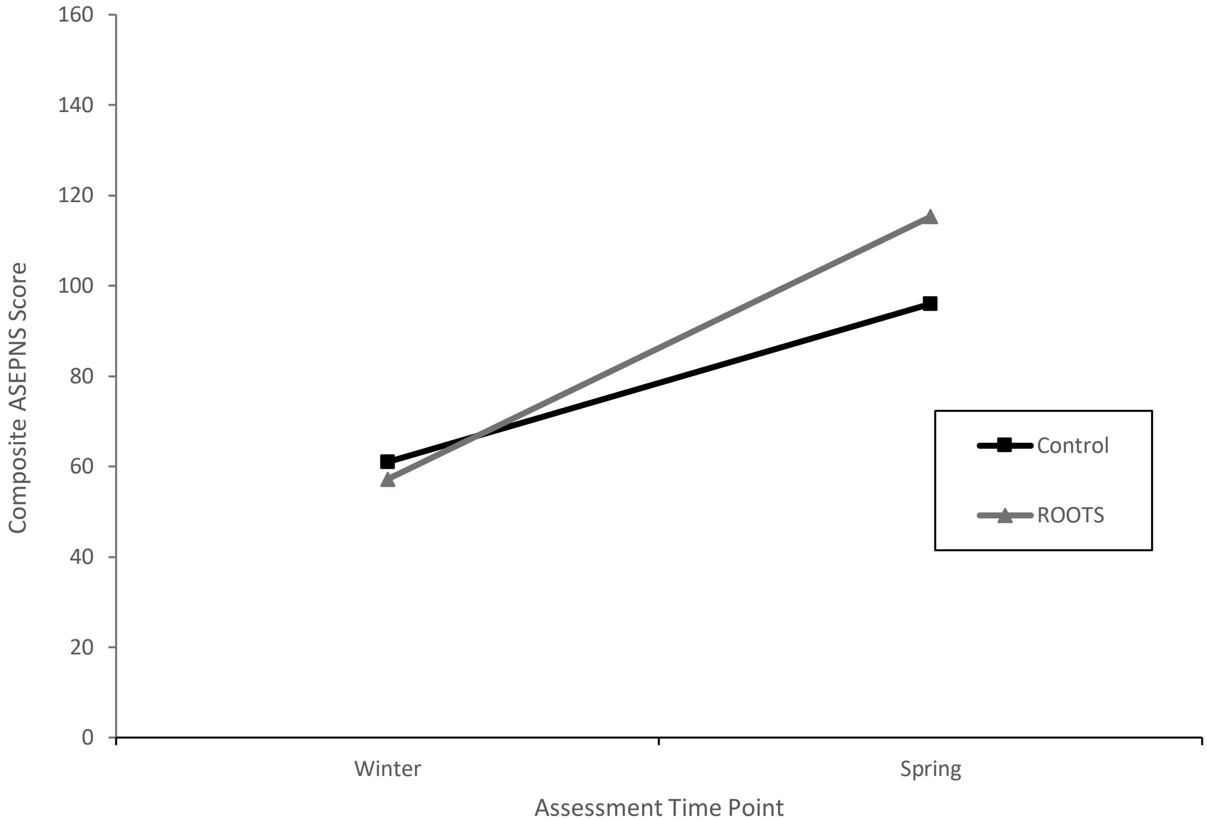
*RAENS Time × Condition Model Results*




*Note*. This figure represents student gains by condition on RAENS per the Time × Condition model results. The y-axis is capped at the highest possible raw score on the RAENS assessment.

**Figure 2**

*ASPENS Time × Condition Model Results*




*Note*. This figure represents student gains by condition on ASPENS per the Time × Condition model results. ASPENS computes a composite score by multiplying raw scores by preset weights. The y-axis is capped at the highest possible composite score on ASPENS.

## CHAPTER IV

## DISCUSSION

The current study is a conceptual replication of earlier research on ROOTS, an early numeracy intervention, which focused on developing early numeracy concepts and skills. Amid a pandemic, interventionists were supported remotely by coaches to acquire instructional techniques and deliver ROOTS to kindergarteners at-risk for experiencing mathematics difficulty in school. This pilot study of supporting educators remotely to implement the ROOTS early numeracy intervention showed positive outcomes in acceptability and implementation fidelity for interventionists in conjunction with promising results for students. The first two research questions examine the suitability of remote PD for an existing evidence-based mathematics intervention for kindergarteners. The third research question examines whether or not students who received ROOTS intervention demonstrated improved outcomes over students in a control condition. In short, the outcomes (1) pointed to an online approach for ROOTS PD as acceptable (i.e., interventionist ratings at multiple timepoints indicated their belief that they received high-quality training); (2) the intervention was primarily implemented as intended (i.e., both direct observation and indirect measures through surveys and logs indicate that ROOTS was delivered per the manualized instructions); and (3) On both post-test measures, ROOTS students outperformed their control group peers ($g = 0.77$ and 0.61 for ASPENS and RAENS, respectively), with the result on ASPENS reaching statistical significance ($p = 0.0010$) and RAENS not reaching statistical significance ($p = 0.1083$). Taken together, the outcomes from the current pilot study suggest that remote PD for ROOTS is a viable alternative to traditional face-to-face PD and warrants continued and more in-depth investigation. Key findings for each research question, limitations, considerations for scalability, and additional future directions are discussed below.

Previous research on remote PD before the COVID-19 pandemic demonstrates widely varying outcomes on the translation from professional learning (e.g., educator workshops, coaching) to improvements in instruction and student outcomes for mathematics (e.g., Kraft & Hill, 2020) and literacy (Vernon-Feagans et al., 2013; Walsh et al., 2020). The current study extends prior work on remote PD to a pandemic context, where an online format for teacher support was a necessity. Due to the pilot nature of the current study and occurrence during a global health crisis, the results are best interpreted cautiously.

The first research question, on acceptability, relates to social validity for the procedures and goals (see Wolf, 1978) of remote PD in the current study. Both forms of remote PD (i.e., training workshops and coaching sessions) were highly rated by interventionists for acceptability (i.e., interventionists almost entirely "strongly agreeing" that remote training was effective). Concurrently, data on implementation fidelity points to interventionists successfully acquiring instructional skills and running the ROOTS program successfully (i.e., strong direct observation ratings for high-quality instruction on RCMIS and indirect measures of treatment adherence through surveys and logs). It should be noted that while the results on the acceptability of remote PD and implementation fidelity are encouraging, an RCT-design blocking on classroom prohibits causal inferences about the impact of remote PD because the independent variable is the totality of the ROOTS intervention including the intervention, training, and coaching. Also note that three interventionists indicated through survey responses that facility with technology may have impacted their ability to effectively communicate with their coach. Despite the relatively small share of interventionists who indicated that technology was a challenge, this pattern is likely to be found in future PD, and researchers should consider how to support intervention delivery with interventionists who may not be as facile with technology. Still, the broad positive findings on acceptability and implementation fidelity in the current study contributes to an important, initial understanding of contemporary methods to support educators' use and implementation of high-quality mathematics intervention programs.

In addition to the abovementioned contextual factors for the current study, it is noteworthy that the ROOTS intervention achieved favorable outcomes for students in the intervention condition within a conceptual replication context. Student outcome data in the current study is consistent with data from a previous federally funded efficacy trial on the ROOTS program occurring across two sites and four cohorts (Clarke et al., 2020); effect sizes for student outcomes on the RAENS and ASPENS were comparable to the current study at $g = .81$ and $g = .49$, respectively. Additionally, the Clarke et al. (2020) study included the use of broader measures for student mathematics achievement gains, demonstrating the comprehensiveness of student outcomes when using ROOTS (e.g., $g = .23$ on the Tests of Early Mathematics Achievement). Our outcomes add to the existing research on ROOTS (Clarke et al., 2016, 2017; Doabler et al., 2016) with a statistically significant difference on the ASPENS and positive effect sizes on both measures (i.e., effect sizes of 0.61 and 0.77 on RAENS and ASPENS measures, respectively). Prior replication research on ROOTS (Doabler et al., 2016) discusses the "changing landscape of the counterfactual," which coincides with our position that the current replication trial had positive results for student outcomes in a dramatically shifted counterfactual environment due to the COVID-19 pandemic. While instruction remained relatively constant in comparison to previous iterations of ROOTS research (e.g., instruction was provided in-person, roughly 60 minutes of core math instruction per day, incorporation of numeracy activities in daily classroom routines such as calendar time), there were likely additional challenges faced by educators providing instruction for both intervention and control conditions (e.g., teaching while wearing personal protective equipment, personal health anxieties). The ROOTS interventionists did not receive any special affordances to navigate the potential pandemic-related challenges in delivering an early numeracy program to kindergarteners (e.g., physical distancing), so there is an assumed equivalency across conditions for managing COVID-19 concerns. The pandemic likely disrupted student acquisition of numeracy skills in the fall of their kindergarten year (Bailey et al., 2021), which was prior to the initiation of the ROOTS intervention in winter 2021. However, the initial skill impact was realized comparably across intervention and control conditions as both early numeracy measures demonstrated equivalent initial pre-test skill by condition. Despite the potential impacts of the global health crisis on delivering instruction and intervention during the 2020-21 school year, the ROOTS program resulted in positive outcomes for students in the intervention condition.

Analyses revealed that student outcomes were slightly different by measure. RAENS was the most proximal measure to ROOTS, and yet when intervention and control conditions were contrasted, students demonstrated greater gains on ASPENS. One hypothesis to explain this result is that post-test scores on the RAENS either approached or hit a ceiling effect depending on the condition (i.e., descriptive statistics revealed that mean post-test raw scores for intervention students was roughly 31 points on the RAENS, where the highest possible raw score is 32). A ceiling effect derives from a significant negative skew in the RAENS post-test data, and the capacity to measure student growth is inhibited from the substantial skew (Ho & Yu, 2015). The ceiling effect is potentially a result of beginning ROOTS mid-school year, where students had stronger initial and resulting skills compared to running ROOTS from the beginning-to-middle of the school year.

***Further Considerations for Interpretation***

One final consideration regarding the interpretation of results is the threats to internal validity posed by the partially-nested condition structure of the current study. Without adequate protection against threats to the validity of treatment effects, making causal inferences are inappropriate (Kim & Steiner, 2016). More specifically, the Rubin causal model (Rubin, 1986) identifies a common expectation in RCT-structured research, where a stable unit of treatment (i.e., the potential outcome of a student is independent of the condition of others, described as a the Stable Unit Treatment Value Assumption or SUTVA) is assumed. Adjusting the structure of intervention deployment can mitigate potential spillover effects which violate the SUTVA (Magill et al., 2019). Specific to the context in which the ROOTS intervention took place, there is little evidence to suggest that the SUTVA was violated. The majority of ROOTS groups were run by paraprofessional interventionists who exclusively provided small group intervention in reading or math. For the classroom teachers who provided ROOTS intervention in addition to core math instruction (three of eight), everyone reported running ROOTS during a separate time of day from core instruction in a space outside the kindergarten classroom. Further, the ROOTS intervention program uses scripted lessons with specific manipulatives, routines, and activities that are not necessarily appropriate for core classroom instruction because the group size is too large, and therefore the transferability of ROOTS lessons to a class-wide context seems unlikely.

Threats to internal validity through spillover between units in different conditions can also occur among participants (Magill et al., 2019). In the current study, students learned math content aligned with the Kindergarten curriculum, and therefore it is theoretically feasible for a student to use techniques they learned in a ROOTS group and generalize those skills to the class-wide context. In the scenario where students apply ROOTS skills in the broader classroom with their control peers observing, this would theoretically dilute the treatment impact for students in the intervention condition because students in the control condition would gain access to conceptual and procedural knowledge via treatment contamination. In fact, Rubin's later work promotes the importance of not overwhelming an environment with individuals receiving a treatment condition because of the increased likelihood of violating the SUTVA (Imbens & Rubin, 2010). In an early numeracy intervention context, this adds some defensibility to a partially-nested research design, wherein ROOTS eligible students (i.e., intervention and control conditions) share the same classroom with ROOTS ineligible students (i.e., students without early numeracy development concerns) to ensure that students in the intervention condition do not compose a majority of the class. While we cannot rule out potential spillover of ROOTS program elements into the core kindergarten classroom, it is hypothesized that SUTVA violations inflating treatment effects is unlikely given logistical and structural factors.

### Limitations and Future Directions

The pandemic context influenced the design of the research study and implementation of the ROOTS intervention. First, related to design, the current study is a pilot study with a relatively small sample size (54 students from eight classrooms in two schools) of schools that were willing to run this research study amid a pandemic. The likelihood of a participation bias during a pandemic is theoretically exacerbated by the fact that many larger schools in metropolitan areas experienced school shutdowns due to the high incidence of COVID-19 transmission (Grooms & Childs, 2021) and thus were not available or considered for participation in the current study. Another design limitation is that the control condition did not have an equal-dosage academic treatment, sometimes called an "active control" condition (e.g., a daily 20-minute reading intervention), to account for potential gains resulting from increased time on academic skill development. A no-treatment control group is sub-optimal for RCT designs with two arms (i.e., intervention and control) when evaluating interventions because of the assumption that an absence of intervention equates to absence of effect, in addition to the bias related to interventionist alliance (Mohr et al., 2009). Further, using a partially-nested design (e.g., some students within classroom assigned to condition) may be a better match for practical student support systems in schools (e.g., multi-tiered system of supports; Kearney & Childs, 2021), but a weaker study design in comparison to fully-nested conditions (e.g., all students within classroom assigned to condition). There is no evidence available to suggest that screening students into intervention or control conditions resulted in post-assignment grouping effects, but the possibility of this confound cannot be ruled out. While the benefits of fully-nested designs for future research should be considered, it is also vital to consider them within the need for social scientists to acknowledge that the increased external validity achieved with partially-nested designs has merit for practitioners (Bauer et al., 2008). Another limitation of the current study was the limited student outcome measurement net. Previous studies of ROOTS have included proximal and distal measures including broad measures of overall mathematics such as the TEMA-3. Future conceptual replication research for academic interventions would ideally include a comparable measurement net.

With respect to data acquisition limitations, data on the counterfactual was minimal. Remote data collection on interventionists was logistically taxing to manage with five coaches and eight interventionists, and for this reason, a more robust data collection system for core instruction was impractical in the current study. With broader human and financial resources, future research should examine the counterfactual to enable a greater understanding of treatment effects and the contexts in which they occur (see Lemons et al., 2014). Given the pandemic context, the student achievement measures and interventionist rating or survey items were constructed with practicality as a priority (e.g., the student outcome measures included were relatively easy to train novice assessors on). This resulted in student achievement measures that did not include a distal, broad measure of mathematics achievement (e.g., the Tests of Early Math Abilities were used in prior research on ROOTS). It is important to note that the observation instrument for the current study (i.e., RCMIS) was validated on in-person observations. Therefore, it is unknown if using the RCMIS tool remotely impacts the accuracy of the observation data. Another data acquisition limitation is that the interventionists collected all pre- and post-test measures themselves as opposed to a third-party assessor. Interventionists were trained to a high standard before administering the ASPENS and RAENS; however, this is a threat to the validity of the results as the assessors (i.e., interventionists) were not blinded to the condition and could have been biased. Finally, interventionists had two opportunities to complete surveys after workshops at two points in time (i.e., Book 1 and Book 2 trainings), and because survey responses were submitted anonymously and respondents submitted responses once or twice, survey outcomes may have favored respondents who completed the training survey at both timepoints. In contrast, the post-intervention survey had one opportunity for response which all interventionists ($n$ = 8) completed, thereby offering equal weighting for responses to mean scores.

There are two overarching considerations for future research based on the outcomes from the current study. First, conceptual replications are ideally designed to systematically vary one element at a time (Coyne et al., 2016). In the current study, at least three important elements varied from the original ROOTS study (i.e., geographical location, remote PD, public education and health context related to pandemic). It is challenging to attribute any differences in outcomes with previous research to one specific altered variable. Prospective conceptual replication research study designers can strengthen their approach by narrowing the elements which are altered from the initial study. Specifically, future research on remote PD for academic intervention programs would ideally isolate PD as the only factor that systematically varies from an initial research study. By varying PD in isolation, researchers can then design a conceptual replication study that centers around the independent variable of interest (i.e., teacher PD is the primary research question, and researchers conduct random assignment at the school level to address spillover effects among educators). Positive outcomes from a conceptual replication study where remote PD is the independent variable bolsters the argument that remote PD is a viable mechanism to scale up efficient and effective support for educators in the adoption and use of early numeracy interventions. Further, future research in early numeracy intervention that isolates PD as an independent variable adds to the initial promise of the current study, Kraft and Hill's (2020) work, and increase the potential for causal inferences using a scientifically rigorous research design (Kraft et al., 2018). Author overlap (i.e., the same author or author team conducts the initial study and subsequent replication trials) presents a second future direction for replication research, as the current study occurred with author overlap on two other RCTs on ROOTS (see Clarke et al., 2016, 2017; Doabler et al., 2016). Author overlap is problematic because it presents an opportunity for unintentional bias (Makel & Plucker, 2014). In order to mitigate any biases stemming from author overlap in the current study, we took an approach similar to previous ROOTS work where school-based staff, initially unfamiliar with ROOTS, conducted the assessments and interventions, and research staff provided supplemental support and then post-intervention analyses. We believe that our procedures mimicked an authentic academic intervention adoption where school staff are given the latitude to learn, implement, and then provide feedback on the utility of the intervention components and receive support when appropriate. Despite our best efforts to mitigate author overlap, future conceptual replication work is recommended to consider infusing third-party researchers and staff to the greatest extent possible. Integrating new social scientists and practitioners to conduct replication trials in their local context is an excellent match for scalability needs, where endogenous providers have a higher likelihood of understanding local contextual factors that can drive the successful implementation of a new intervention (Brunwasser & Garber, 2016).

Scaling up PD practices that save money and time (i.e., remote support instead of trainers flying to local school districts) has the potential to more effectively leverage the estimated tens of billions of dollars annually in the United States spent on teacher PD (cost estimated in Kraft et al., 2018). One significant benefit to an online format for educator coaching, in particular, is the potential for coaches being flexible and responsive to teacher needs beyond what would be possible in an in-person format. For example, teacher coaches could theoretically reschedule and reformat their remote PD activities more easily because of decreased travel (e.g., broader availability instead of only being able to travel to school sites on certain days) and increased access to coaching materials (e.g., manipulatives for lessons immediately available instead of traveling to a site with a portion of coaching materials). It is also important to give consideration to future research on mitigating the potential barriers with navigating technology for PD, which was an issue noted by some interventionists in the current study. Addressing technological barriers could be done directly (e.g., training on technology used for PD with initial in-person support) or indirectly (e.g., email offers of support). Addressing the technology barrier may be crucial to the successful uptake and sustainability of remote PD because practitioners would ideally be able to access expert guidance on intervention implementation without undue burden. Our brief qualitative survey responses indicated that several of the interventionists would prefer to return to in-person PD in a non-pandemic future; though, those who experienced technological issues might have been more likely to prefer the in-person PD experience. Future research is well-positioned to examine this question through PD design with a larger sample size and more specific survey items on interventionist preference for PD format alongside comfort with technology use. Recent research also suggests that remote coaching support for teachers working in rural school districts addresses the barrier of access to academic intervention experts to enhance instructional techniques (e.g., Glover, 2017; Lee et al., 2018) which can serve to address equity issues in PD quality for rural communities. Additionally, Kraft and Hill (2020) note that their 2014-16 study was the first available peer-reviewed evidence of a web-based coaching platform specific to math instruction; therefore, it is likely that the current study is one of very few to show the promise of remote coaching for an early numeracy intervention program. Researchers can make progress on closing the research-to-practice gap by evaluating an online coaching format that better serves teacher needs, as much of the research to date on teacher coaching for academic intervention demonstrates instructional change without direct improvement in student outcomes (Hamre et al., 2017; Kraft et al., 2018; Reddy et al., 2021). Relatedly, this research line has potential policy implications as well based on recent work synthesizing the importance of using effective PD practices that result in increased student achievement (Darling-Hammond et al., 2017). Policy positions in support of remote PD may be accompanied by systemic challenges (e.g., policies on staff technology use or parental consent for live videoconference recordings of instructional interactions); though, absorbing the upfront cost for long-term savings in time and money is a potentially worthwhile endeavor (Rodgers et al., 2019).

**Conclusion**

Results from the current study support the promise of remote PD for educators' delivery of an early intervention program focused on student's mathematics skill development. The combination of positive ratings by educators on the acceptability of remote PD and strong implementation fidelity outcomes in the current study dovetails with favorable achievement outcomes for students receiving the ROOTS intervention, which suggests that this pilot study adds to the literature in a couple of ways. First, we add to the initial promise for a virtual format of educator training on manualized early intervention programs in numeracy. Second, our work highlights another data point to support the robustness of ROOTS as an effective early numeracy intervention. Note that the current study used an underpowered design with limitations requiring consideration before making claims about how positive effects of remote PD may have directly translated to high implementation fidelity, and then translated again through high-quality, explicit instruction to positive student outcomes for students receiving the ROOTS intervention. Based on the limitations noted above, scaling this work on remote PD from the "initial promise" stage to efficacy and effectiveness stages of research will require significant technical assistance and the use of research designs that examine the format of PD's impact on student outcomes in a casual manner.

# APPENDIX A

## CURRICULUM STANDARDS

*ROOTS alignment with the Common Core State Standards for mathematics (2010)*

| | **Counting and Cardinality** | L1-10 | L11-20 | L21-30 | L31-40 | L41-50 |
|---|---|---|---|---|---|---|
| | *Know number names and the count sequence.* | | | | | |
| 1 | Count to 100 by ones and by tens. | to 5 | to 8 | to 20 | to 20 | to 20 |
| 2 | Count forward beginning from a given number within the known sequence (instead of having to begin at 1). | | ✓ | ✓ | ✓ | ✓ |
| 3 | Write numbers from 0 to 20. Represent a number of objects with a written numeral 0-20 (with 0 representing a count of no objects). | ✓ | ✓ | ✓ | ✓ | ✓ |
| | *Count to tell the number of objects* | | | | | |
| 4 | Understand the relationship between numbers and quantities; connect counting to cardinality. | ✓ | ✓ | ✓ | ✓ | ✓ |
| 4.a | When counting objects, say the number names in the standard order, pairing each object with one and only one number name and each number name with one and only one object. | | ✓ | ✓ | ✓ | ✓ |
| 4.b | Understand that the last number name said tells the number of objects counted. The number of objects is the same regardless of their arrangement or the order in which they were counted. | ✓ | ✓ | ✓ | ✓ | ✓ |
| 4.c | Understand that each successive number name refers to a quantity that is one larger. | | ✓ | ✓ | ✓ | ✓ |
| 5 | Count to answer "how many?" questions about as many as 20 things arranged in a line, a rectangular array, or a circle, or as many as 10 things in a scattered configuration; given a number from 1–20, count out that many objects. | ✓ | ✓ | ✓ | ✓ | ✓ |
| | *Compare numbers.* | | | | | |
| 6 | Identify whether the number of objects in one group is greater than, less than, or equal to the number of objects in another group, e.g., by using matching and counting strategies (groups up to 10 numbers). | ✓ | ✓ | ✓ | ✓ | ✓ |
| 7 | Compare two numbers between 1 and 10 presented as written numerals. | | ✓ | ✓ | ✓ | ✓ |
| | **Operations and Algebraic Thinking** | | | | | |
| | *Understand addition as putting together and adding to, and understand subtraction as taking apart and taking from.* | | | | | |
| 1 | Represent addition and subtraction with objects, fingers, mental images, drawings, sounds (e.g., claps), acting out situations, verbal explanations, expressions, or equations. | | ✓ | ✓ | ✓ | ✓ |
| 2 | Solve addition and subtraction word problems, and add and subtract within 10, e.g., by using objects or drawings to represent the problem. | | ✓ | ✓ | ✓ | ✓ |
| 3 | Decompose numbers less than or equal to 10 into pairs in more than one way, e.g., by using objects or drawings, and record each decomposition by a drawing or equation (e.g., 5 = 2 + 3 and 5 = 4 + 1). | | ✓ | ✓ | ✓ | ✓ |
| 4 | For any number from 1 to 9, find the number that makes 10 when added to the given number, e.g., by using objects or drawings, and record the answer with a drawing or equation. | | | | | ✓ |
| 5 | Fluently add and subtract within 5. | | add | add | add | add |
| | **Number and Operations in Base Ten** | | | | | |
| | *Work with numbers 11-19 to gain foundations for place value* | | | | | |
| 1 | Compose and decompose numbers from 11 to 19 into ten ones and some further ones, e.g., by using objects or drawings, and record each composition or decomposition by a drawing or equation (e.g., 18 = 10 + 8); understand that these numbers are composed of ten ones and one, two, three, four, five, six, seven, eight, or nine ones. | | | ✓ | ✓ | ✓ |

**APPENDIX B**

ROOTS TRAINING WORKSHOP FEEDBACK

*Items on Interventionist Perception of Training for ROOTS Book 1 (January, 2021) and Book 2 (March, 2021) Content with Online PD*

| Item | Rating (*n* = 9) | |
|---|---|---|
| | *M* | *(SD)* |
| I have the necessary resources to effectively implement the ROOTS intervention in my school. | 4.7 | (0.7) |
| The ROOTS intervention is appropriate for kindergarten students who are struggling in early mathematics. | 4.8 | (0.4) |
| The ROOTS intervention is appropriate for kindergarten students who are achieving at grade level in mathematics. | 3.4 | (1.2) |
| The ROOTS intervention is appropriate for kindergarten students who are English learners (Els). | 4.4 | (0.7) |
| The ROOTS intervention is appropriate for kindergarten students who receive special education in mathematics. | 4.7 | (0.5) |

*Note*. These items were included on the post-training surveys for interventionists immediately following the remote ROOTS workshops. Surveys were delivered via Qualtrics.

**APPENDIX C**

POST-INTERVENTION SURVEY RESULTS

*Items on Interventionist Perception of ROOTS Implementation, Support, and Training with Online PD*

| | Rating ($n$ = 9) | | |
|---|---|---|---|
| Item | *M* | *(SD)* | Range |
| I have the necessary resources to effectively implement the ROOTS intervention in my school. | 4.7 | (0.7) | 3-5 |
| The ROOTS intervention is appropriate for kindergarten students who are struggling in early mathematics. | 4.8 | (0.4) | 4-5 |
| The ROOTS intervention is appropriate for kindergarten students who are achieving at grade level in mathematics. | 3.4 | (1.2) | 1-5 |
| The ROOTS intervention is appropriate for kindergarten students who are English learners (Els). | 4.4 | (0.7) | 3-5 |
| The ROOTS intervention is appropriate for kindergarten students who receive special education in mathematics. | 4.7 | (0.5) | 4-5 |
| My coach helped me improve my delivery of ROOTS lessons | 4.8 | (0.4) | 4-5 |
| My coach was easy to collaborate with | 4.8 | (0.4) | 4-5 |
| My coach was constructive when giving feedback on my ROOTS instruction | 4.8 | (0.4) | 4-5 |
| My coach was willing to work with me on group instruction factors (e.g., behavior management, attendance concerns) | 4.8 | (0.4) | 4-5 |

| | | | |
|---|---|---|---|
| Overall, I benefitted from the support from a coach | 4.8 | (0.4) | 4-5 |
| Having a coach improved my fidelity to the ROOTS program (e.g., following the lesson, using the correct materials, lesson pacing) | 4.7 | (0.5) | 4-5 |
| Having a coach improved my understanding of math content that will help with teaching other math groups | 4.6 | (0.7) | 3-5 |
| Having a coach improved my instructional skills that will help with teaching other content areas (e.g., reading) or other small groups | 4.6 | (0.7) | 3-5 |
| ROOTS Trainings (instructional books 1 & 2) improved my ROOTS instruction | 4.6 | (0.7) | 3-5 |
| Discussing ROOTS with other interventionists and teachers within your school improved my ROOTS instruction | 4.4 | (0.7) | 3-5 |
| Gathering information from websites or social media (e.g., Facebook, Pintrest, Twitter) improved my ROOTS instruction | 3.0 | (1.0) | 2-5 |
| Coaching sessions were structured effectively so we were able to cover critical content | 4.6 | (0.7) | 3-5 |
| Virtual/remote coaching activities (i.e., discussion, feedback, or modeling) provided constructive feedback in areas that I wanted to improve | 4.6 | (0.7) | 3-5 |
| Virtual/remote coaching was sufficient for communicating with my coach | 4.6 | (0.7) | 3-5 |
| My skills with technology affected my ability to effectively communicate with my coach | 3.4 | (1.0) | 2-5 |
| I wish the coaching session(s) would have lasted longer | 2.9 | (1.1) | 1-5 |

| | | | |
|---|---|---|---|
| Overall, the online format for ROOTS coaching was effective for me | 4.6 | (0.5) | 4-5 |

*Note*. These items were included on the post-ROOTS survey of interventionists delivered via Qualtrics.

## APPENDIX D

RCMIS RESULTS

*Complete RCMIS Items with Scores Across All Items for All Interventionists*

| | Rating (*n* = 23) | |
|---|---|---|
| Item | *M* | *(SD)* |
| 1. Community of positive learning (Rapport, Respect, Positive Attitude) | 4.3 | (0.7) |
| 2. Organization of instructional materials and learning tasks (Preparation, teacher-initiated transitions, accessibility) | 3.7 | (0.8) |
| 3. Effective small-group management techniques (Sets clear expectations; maximizes instr. time; addresses appropriate behavior) | 3.7 | (0.8) |
| 4. Support of students' emotional needs (Sensitivity, Respect, Support) | 4.3 | (0.6) |
| *5. Efficient delivery of instruction (Uses appropriate pacing; consistent language; minimizes student confusion) | 3.9 | (0.8) |
| *6. Student participation and engagement (Active involvement, compliance, completion of work) | 4.0 | (0.7) |
| *7. Effective teacher modeling and demonstrations (Models skills and concepts clearly; uses math representations effectively) | 3.7 | (1.0) |
| *8. High-quality opportunities for group practice (Offers frequent and rich opportunities for guided & independent practice) | 3.9 | (0.7) |
| *9. Checks of student understanding (Provides timely academic feedback; actively monitors practice opportunities) | 4.3 | (0.7) |
| *10. High-quality practice opportunities for individuals (Distributes individual practice opportunities, both guided and independent) | 4.0 | (0.6) |
| *11. Instructional Scaffolding and Support (Provides adequate think/response time and independent learning opportunities) | 3.8 | (0.7) |
| 12. Productive disposition of mathematical learning (Positive outlook on math; views math as important; confidence) | 4.5 | (0.5) |
| 13. Accomplishment of instructional tasks and activities (Completes tasks; uses time efficiently, student-initiated routines) | 3.9 | (0.9) |

| | | |
|---|---|---|
| 14. Teaching for mathematical proficiency<br>(States purpose of lesson; addresses big ideas; effective teaching examples; anticipates student misconceptions; frequent instructional interactions) | 4 | (0.9) |

*Note*. These items were included on the post-ROOTS survey of interventionists delivered via Qualtrics. * = items used for the quality of explicit instruction (QEI) subscale.

**APPENDIX E**

SUPPLEMENTAL MULTILEVEL ANCOVA MODEL RESULTS

In addition to the Time × Condition model analyses in the Results section, student outcomes were also examined using a multilevel ANCOVA model to test for group differences at post-test adjusting for pre-test achievement scores on both ASPENS and RAENS. In order to account for the partially nested structure of the data, analyses included student ROOTS group ($n$ = 8) as a random effect, and then pretest score and condition (intervention, $n$ = 28, versus control, $n$ = 26) were included as fixed effects. Models were fit using R statistical software (R Core Team, 2021) with the Linear Mixed-Effects Models package (Bates et al., 2021). See Table E1 for complete modeling results.

First, using an multilevel ANCOVA to examine group differences on post-test RAENS scores when adjusting for RAENS pre-test scores, students in the control condition with an average pretest score had a post-test score of 27.6 on RAENS post-test ($t$ = 38.35, $p$ < 0.0001). A single point increase on the pretest RAENS score was associated with a 0.25 increase in post-test RAENS score ($t$ = 4.18, $p$ < 0.0002). The intervention condition also had a significant effect, where students in the intervention condition scored 3.3 points higher on post-test ($t$ = 2.33, $p$ = 0.0288) in comparison to their control peers. There was an effect size of group differences based on student's receiving the ROOTS intervention ($g$ = 0.58, 95% CI = [0.33, 0.83]).

Second, using ASPENS to compare group differences at post-test when adjusting for pre-test scores, students in the control condition with an average pretest score had an ASPENS composite score of 94.3 points on post-test ($t$ = 18.16, $p$ < 0.0001). A single point increase on the pretest ASPENS score was associated with a 0.8 increase in posttest

ASPENS score ($t = 4.05$, $p = 0.0002$). Students in the intervention condition outperformed their control group peers by 22.5 points on post-test ($t = 3.11$, $p = 0.0030$). Similar to RAENS, the effect size of intervention using ASPENS resulted in group differences ($g = 0.75$, 95% CI = [0.51, 0.99]).

**Table E1**

*Supplemental Multilevel Modeling Results for Student Outcomes on RAENS and ASPENS*

| | Measure | |
|---|---|---|
| | RAENS | ASPENS |
| | Fixed Effects | |
| Intercept | 22.0***<br>(1.5) | 45.1**<br>(13.6) |
| Pretest | 0.3***<br>(0.1) | 0.8***<br>(0.2) |
| Condition | 3.3*<br>(1.4) | 22.5**<br>(7.2) |
| | Random Effects | |
| Group-level | 10.7<br>(3.3) | 64.2<br>(8.0) |
| Residual | 2.7<br>(1.7) | 697.3<br>(26.4) |
| | Condition Effect Sizes | |
| $R^2$ | 0.28 | 0.31 |
| *Hedge's g* | 0.60 | 0.75 |

*Note*. $^*p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. Analyses are presented for ASPENS and RAENS as the models included the same fixed and random effects for each measure. Standard Errors are reported in parentheses.

# REFERENCES CITED

Abry, T., Latham, S., Bassok, D., & LoCasale-Crouch, J. (2015). Preschool and kindergarten teachers' beliefs about early school competencies: Misalignment matters for kindergarten adjustment. *Early Childhood Research Quarterly*, *31*, 78–88. https://doi.org/10.1016/j.ecresq.2015.01.001

Allison, P. D. (1990). Change Scores as Dependent Variables in Regression Analysis. *Sociological Methodology*, *20*, 93–114. https://doi.org/10.2307/271083

Amendum, S. J., Bratsch-Hines, M., & Vernon-Feagans, L. (2018). Investigating the Efficacy of a Web-Based Early Reading and Professional Development Intervention for Young English Learners. *Reading Research Quarterly*, *53*(2), 155–174. https://doi.org/10.1002/rrq.188

Anders, Y., Rossbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., & von Maurice, J. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, *27*(2), 231–244. https://doi.org/10.1016/j.ecresq.2011.08.003

Ansari, A., Pianta, R. C., Whittaker, J. E., Vitiello, V., & Ruzek, E. (2021). Enrollment in public-prekindergarten and school readiness skills at kindergarten entry: Differential associations by home language, income, and program characteristics. *Early Childhood Research Quarterly*, *54*, 60–71. https://doi.org/10.1016/j.ecresq.2020.07.011

Archer, A. L., & Hughes, C. A. (2010). *Explicit Instruction: Effective and Efficient Teaching*. Guilford Publications.

Bailey, D. H., Duncan, G. J., Murnane, R. J., & Au Yeung, N. (2021). Achievement Gaps in the Wake of COVID-19. *Educational Researcher*, *50*(5), 266–275. https://doi.org/10.3102/0013189X211011237

Bailey, D. H., Fuchs, L. S., Gilbert, J. K., Geary, D. C., & Fuchs, D. (2020). Prevention: Necessary But Insufficient? A 2-Year Follow-Up of an Effective First-Grade Mathematics Intervention. *Child Development*, *91*(2), 382–400. https://doi.org/10.1111/cdev.13175

Baldwin, S. A., Bauer, D. J., Stice, E., & Rohde, P. (2011). Evaluating models for partially clustered designs. Psychological Methods, 16, 149–165. https://doi.org/10.1037/a0023464

Bassok, D., & Latham, S. (2017). Kids Today: The Rise in Children's Academic Skills at Kindergarten Entry. *Educational Researcher*, *46*(1), 7–20. https://doi.org/10.3102/0013189X17694161

Bauer, D. J., Sterba, S. K., & Hallfors, D. D. (2008). Evaluating Group-Based Interventions When Control Participants Are Ungrouped. *Multivariate Behavioral Research*, *43*(2), 210–236. https://doi.org/10.1080/00273170802034810

Bassok, D., Latham, S., & Rorem, A. (2016). Is Kindergarten the New First Grade? *AERA Open*, *2*(1). https://doi.org/10.1177/2332858415616358

Bowen, D. J., Kreuter, M., Spring, B., Cofta-Woerpel, L., Linnan, L., Weiner, D., Bakken, S., Kaplan, C. P., Squiers, L., Fabrizio, C., & Fernandez, M. (2009). How We Design Feasibility Studies. *American Journal of Preventive Medicine*, *36*(5), 452–457. https://doi.org/10.1016/j.amepre.2009.02.002

Brunwasser, S. M., & Garber, J. (2016). Programs for the Prevention of Youth Depression: Evaluation of Efficacy, Effectiveness, and Readiness for Dissemination. *Journal of Clinical Child & Adolescent Psychology*, *45*(6), 763–783. https://doi.org/10.1080/15374416.2015.1020541

Bryant, D. P., Bryant, B. R., Roberts, G., Vaughn, S., Pfannenstiel, K. H., Porterfield, J., & Gersten, R. (2011). Early Numeracy Intervention Program for First-Grade Students with Mathematics Difficulties. *Exceptional Children*, *78*(1), 7–23. https://doi.org/10.1177/001440291107800101

Bull, R., & Lee, K. (2014). Executive Functioning and Mathematics Achievement. *Child Development Perspectives*, *8*(1), 36–41. https://doi.org/10.1111/cdep.12059

Buonsenso, D., Roland, D., De Rose, C., Vásquez-Hoyos, P., Ramly, B., Chakakala-Chaziya, J. N., Munro, A., & González-Dambrauskas, S. (2021). Schools Closures During the COVID-19 Pandemic: A Catastrophic Global Situation. *The Pediatric Infectious Disease Journal*, *40*(4), e146. https://doi.org/10.1097/INF.0000000000003052

Campbell, P. F., & Malkus, N. N. (2011). The Impact of Elementary Mathematics Coaches on Student Achievement. *The Elementary School Journal*, *111*(3), 430–454. https://doi.org/10.1086/657654

Carnine, D., Silbert, J., Kame'enui, E., & Tarver, S. (2004). *Direct instruction reading* (4th ed.). Upper Saddle River, NJ: Pearson.

Chard, D. J., Baker, S. K., Clarke, B., Jungjohann, K., Davis, K., & Smolkowski, K. (2008). Preventing Early Mathematics Difficulties: The Feasibility of a Rigorous Kindergarten Mathematics Curriculum. *Learning Disability Quarterly*, *31*(1), 11–20. https://doi.org/10.2307/30035522

Charitaki, G., Tzivinikou, S., Stefanou, G., & Soulis, S.-G. (2020). A Meta-Analytic Synthesis of Early Numeracy Interventions for Low-Perfoming Young Children. *SN Social Sciences*, *1(5)*, *105*. https://doi.org/10.31234/osf.io/embd7

Chaturvedi, K., Vishwakarma, D. K., & Singh, N. (2021). COVID-19 and its impact on education, social life and mental health of students: A survey. *Children and Youth Services Review*, *121*, 105866. https://doi.org/10.1016/j.childyouth.2020.105866

Chen, J.-Q., McCray, J., Adams, M., & Leow, C. (2014). A Survey Study of Early Childhood Teachers' Beliefs and Confidence about Teaching Early Math. *Early Childhood Education Journal*, *42*(6), 367–377. https://doi.org/10.1007/s10643-013-0619-0

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2011). How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project Star. *The Quarterly Journal of Economics*, *126*(4), 1593–1660. https://doi.org/10.1093/qje/qjr041

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood. *American Economic Review*, *104*(9), 2633–2679. https://doi.org/10.1257/aer.104.9.2633

Chiatovich, T., & Stipek, D. (2016). Instructional Approaches in Kindergarten What Works for Whom? *The Elementary School Journal*, *117*(1), 1–29. https://doi.org/10.1086/687751

Chodura, S., Kuhn, J.-T., & Holling, H. (2015). Interventions for children with mathematical difficulties: A meta-analysis. *Zeitschrift Für Psychologie*, *223*(2), 129–144. https://doi.org/10.1027/2151-2604/a000211

Chung, M. Y., Meadan, H., Snodgrass, M. R., Hacker, R. E., Sands, M. M., Adams, N. B., & Johnston, S. S. (2020). Assessing the Social Validity of a Telepractice Training and Coaching Intervention. *Journal of Behavioral Education*, *29*(2), 382–408. https://doi.org/10.1007/s10864-020-09372-8

Claessens, A., Duncan, G., & Engel, M. (2009). Kindergarten skills and fifth-grade achievement: Evidence from the ECLS-K. *Economics of Education Review*, *28*(4), 415–427. https://doi.org/10.1016/j.econedurev.2008.09.003

Clarke, B., Doabler, C. T., Kosty, D., Kurtz Nelson, E., Smolkowski, K., Fien, H., & Turtura, J. (2017). Testing the Efficacy of a Kindergarten Mathematics Intervention by Small Group Size. *AERA Open*, *3*(2). https://doi.org/10.1177/2332858417706899

Clarke, B., Doabler, C. T., Smolkowski, K., Baker, S. K., Fien, H., & Strand Cary, M. (2016). Examining the Efficacy of a Tier 2 Kindergarten Mathematics Intervention. *Journal of Learning Disabilities*, *49*(2), 152–165. https://doi.org/10.1177/0022219414538514

Clarke, B., Doabler, C. T., Smolkowski, K., Turtura, J., Kosty, D., Kurtz-Nelson, E., Fien, H., & Baker, S. K. (2019). Exploring the Relationship Between Initial Mathematics Skill and a Kindergarten Mathematics Intervention. *Exceptional Children*, *85*(2), 129–146. https://doi.org/10.1177/0014402918799503

Clarke, B., Doabler, C. T., Strand Cary, M., Kosty, D., Baker, S., Fien, H., & Smolkowski, K. (2014). *Preliminary Evaluation of a Tier 2 Mathematics Intervention for First-Grade Students: Using a Theory of Change to Guide Formative Evaluation Activities* (Vol. 43). https://eric.ed.gov/?id=ED567775

Clarke, B., Doabler, C. T., Turtura, J., Smolkowski, K., Kosty, D. B., Sutherland, M., Kurtz-Nelson, E., Fien, H., & Baker, S. K. (2020). Examining the Efficacy of a Kindergarten Mathematics Intervention by Group Size and Initial Skill: Implications for Practice and Policy. *The Elementary School Journal*, *121*(1), 125–153. https://doi.org/10.1086/710041

Clarke, B., Gersten, R. M., Dimino, J., & Rolfhus, E. (2011). *Assessing Student Proficiency in Early Number Sense (ASPENS)*. Longmont, CO: Cambium Learning Group.

Clarke, B., Smolkowski, K., Baker, S. K., Fien, H., Doabler, C. T., & Chard, D. J. (2011). The Impact of a Comprehensive Tier I Core Kindergarten Program on the Achievement of Students at Risk in Mathematics. *The Elementary School Journal*, *111*(4), 561–584. https://doi.org/10.1086/659033

Cochrane, W. S., Sanetti, L. M. H., & Minster, M. C. (2019). School psychologists' beliefs and practices about treatment integrity in 2008 and 2017. *Psychology in the Schools*, *56*(3), 295–305. https://doi.org/10.1002/pits.22177

Codding, R. S., Chan-Iannetta, L., George, S., Ferreira, K., & Volpe, R. (2011). Early Number Skills: Examining the Effects of Class-Wide Interventions on Kindergarten Performance. *School Psychology Quarterly*, *26*(1), 85–96. https://doi.org/10.1037/a0022661

Cook, B. G. (2014). A Call for Examining Replication and Bias in Special Education Research. *Remedial and Special Education*, *35*(4), 233–246. https://doi.org/10.1177/0741932514528995

Cook, B. G., Lloyd, J. W., Mellor, D., Nosek, B. A., & Therrien, W. J. (2018). Promoting Open Science to Increase the Trustworthiness of Evidence in Special Education. *Exceptional Children*, *85*(1), 104–118. https://doi.org/10.1177/0014402918793138

Coyne, M. D., Cook, B. G., & Therrien, W. J. (2016). Recommendations for Replication Research in Special Education: A Framework of Systematic, Conceptual Replications. *Remedial and Special Education*, *37*(4), 244–253. https://doi.org/10.1177/0741932516648463

Dabrowski, A. (2020). Teacher wellbeing during a pandemic: Surviving or thriving? *Teacher Workforce and Careers*. https://doi.org/10.37256/ser.212021588

Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). Effective Teacher Professional Development. In *Learning Policy Institute*. Learning Policy Institute. https://eric.ed.gov/?id=ED606743

Dede, C., Jass Ketelhut, D., Whitehouse, P., Breit, L., & McCloskey, E. M. (2009). A Research Agenda for Online Teacher Professional Development. *Journal of Teacher Education*, *60*(1), 8–19. https://doi.org/10.1177/0022487108327554

Daniel, S. J. (2020). Education and the COVID-19 pandemic. *Prospects*, 1. https://doi.org/10.1007/s11125-020-09464-3

DeFlorio, L., Klein, A., Starkey, P., Swank, P. R., Taylor, H. B., Halliday, S. E., Beliakoff, A., & Mulcahy, C. (2019). A study of the developing relations between self-regulation and mathematical knowledge in the context of an early math intervention. *Early Childhood Research Quarterly*, *46*, 33–48. https://doi.org/10.1016/j.ecresq.2018.06.008

Desilver, D. (2017, February 15). *U.S. students' academic achievement still lags that of their peers in many other countries*. Pew Research Center. https://cpb-us-w2.wpmucdn.com/blogs.cofc.edu/dist/0/348/files/2018/11/U.S.-academic-achievement-lags-that-of-many-other-countries-2iqc8mk.pdf

Desimone, L. M. (2009). Improving Impact Studies of Teachers' Professional Development: Toward Better Conceptualizations and Measures. *Educational Researcher*, *38*(3), 181–199. https://doi.org/10.3102/0013189X08331140

Desimone, L., & Garet, M. S. (2015). Best Practices in Teachers' Professional Development in the United States. *Psychology, Society & Education*, *7*(3), 252–263. https://dialnet.unirioja.es/servlet/articulo?codigo=6360142

Desimone, L. M., & Pak, K. (2017). Instructional Coaching as High-Quality Professional Development. *Theory Into Practice*, *56*(1), 3–12. https://doi.org/10.1080/00405841.2016.1241947

Doabler, C. T., Baker, S. K., Kosty, D. B., Smolkowski, K., Clarke, B., Miller, S. J., & Fien, H. (2015). Examining the Association between Explicit Mathematics Instruction and Student Mathematics Achievement. *The Elementary School Journal*, *115*(3), 303–333. https://doi.org/10.1086/679969

Doabler, C. T., Clarke, B., & Fien, H. (2012). *ROOTS Assessment of Early Numeracy Skills (RAENS)* [Unpublished measurement instrument]. Eugene, OR: Center on Teaching and Learning.

Doabler, C. T., Clarke, B., Fien, H., Baker, S. K., Kosty, D. B., & Cary, M. S. (2015). The Science Behind Curriculum Development and Evaluation: Taking a Design Science Approach in the Production of a Tier 2 Mathematics Curriculum. *Learning Disability Quarterly*, *38*(2), 97–111. https://doi.org/10.1177/0731948713520555

Doabler, C. T., Clarke, B., Kosty, D. B., Kurtz-Nelson, E., Fien, H., Smolkowski, K., & Baker, S. K. (2016). Testing the Efficacy of a Tier 2 Mathematics Intervention: A Conceptual Replication Study. *Exceptional Children*, *83*(1), 92–110. https://doi.org/10.1177/0014402916660084

Doabler, C. T., Clarke, B., Kosty, D., Kurtz-Nelson, E., Fien, H., Smolkowski, K., & Baker, S. K. (2019). Examining the Impact of Group Size on the Treatment Intensity of a Tier 2 Mathematics Intervention Within a Systematic Framework of Replication. *Journal of Learning Disabilities*, *52*(2), 168–180. https://doi.org/10.1177/0022219418789376

Doabler, C. T., Clarke, B., Kosty, D., Turtura, J. E., Sutherland, M., Maddox, S. A., & Smolkowski, K. (2020). Using Direct Observation to Document “Practice-Based Evidence” of Evidence-Based Mathematics Instruction. *Journal of Learning Disabilities*, 0022219420911375. https://doi.org/10.1177/0022219420911375

Doabler, C., & Nelson-Walker, N. J. (2009). Ratings of classroom management and instructional support. [Unpublished observation instrument]. *Center on Teaching and Learning, College of Education, University of Oregon, Eugene*.

Duncan, G. J., Jenkins, J. M., Watts, T. W., Magnuson, K., Clements, D., Sarama, J., Wolfe, C. B., & Spitler, M. E. (2015). Preventing Preschool Fadeout through Instructional Intervention in Kindergarten and First Grade. In *Society for Research on Educational Effectiveness*. Society for Research on Educational Effectiveness. https://eric.ed.gov/?id=ED562417

Duncan, G. J., & Murnane, R. J. (2011). *Whither Opportunity?: Rising Inequality, Schools, and Children’s Life Chances*. Russell Sage Foundation.

Duncan, G. J., & Sojourner, A. J. (2013). Can Intensive Early Childhood Intervention Programs Eliminate Income-Based Cognitive and Achievement Gaps? *Journal of Human Resources*, *48*(4), 945–968. https://doi.org/10.3368/jhr.48.4.945

Duncan, R. J., Schmitt, S. A., Burke, M., & McClelland, M. M. (2018). Combining a kindergarten readiness summer program with a self-regulation intervention improves school readiness. *Early Childhood Research Quarterly*, *42*, 291–300. https://doi.org/10.1016/j.ecresq.2017.10.012

Dyson, N. I., Jordan, N. C., & Glutting, J. (2013). A Number Sense Intervention for Low-Income Kindergartners at Risk for Mathematics Difficulties. *Journal of Learning Disabilities*, *46*(2), 166–181. https://doi.org/10.1177/0022219411410233

Engel, M., Claessens, A., Watts, T. W., & Farkas, G. (2015). The misalignment of kindergarten mathematics content. Paper presented at the annual meeting of the American Education Research Association, Chicago, IL.

Engel, M., Claessens, A., & Finch, M. A. (2013). Teaching Students What They Already Know? The (Mis)Alignment Between Mathematics Instructional Content and Student Knowledge in Kindergarten. *Educational Evaluation and Policy Analysis*, *35*(2), 157–178. https://doi.org/10.3102/0162373712461850

Engel, M., Claessens, A., Watts, T., & Farkas, G. (2016). Mathematics Content Coverage and Student Learning in Kindergarten. *Educational Researcher*, *45*(5), 293–300. https://doi.org/10.3102/0013189X16656841

Engzell, P., Frey, A., & Verhagen, M. D. (2021). Learning loss due to school closures during the COVID-19 pandemic. *Proceedings of the National Academy of Sciences*, *118*(17). https://doi.org/10.1073/pnas.2022376118

Farrell, C. C., Coburn, C. E., & Chong, S. (2019). Under What Conditions Do School Districts Learn From External Partners? The Role of Absorptive Capacity. *American Educational Research Journal*, *56*(3), 955–994. https://doi.org/10.3102/0002831218808219

Flores, A. (2007). Examining Disparities in Mathematics Education: Achievement Gap or Opportunity Gap? *High School Journal*, *91*(1), 29–42. https://doi.org/10.1353/hsj.2007.0022

Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*, 67(3), 308.

Gaias, L. M., Abry, T., Swanson, J., & Fabes, R. A. (2016). Considering child effortful control in the context of teacher effortful control: Implications for kindergarten success. *Learning and Individual Differences*, *45*, 199–207. https://doi.org/10.1016/j.lindif.2015.11.016

Gates, B. (2020). Responding to Covid-19—A Once-in-a-Century Pandemic? *New England Journal of Medicine, 382*(18), 1677–1679. https://doi.org/10.1056/NEJMp2003762

Geary, D. C., Hoard, M. K., Nugent, L., & Bailey, D. H. (2013). Adolescents' Functional Numeracy Is Predicted by Their School Entry Number System Knowledge. *PLOS ONE, 8*(1), e54651. https://doi.org/10.1371/journal.pone.0054651

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics Instruction for Students With Learning Disabilities: A Meta-Analysis of Instructional Components. *Review of Educational Research*, *79*(3), 1202–1242. https://doi.org/10.3102/0034654309334431

Gersten, R., Rolfhus, E., Clarke, B., Decker, L. E., Wilkins, C., & Dimino, J. (2015). Intervention for First Graders With Limited Number Knowledge: Large-Scale Replication of a Randomized Controlled Trial. *American Educational Research Journal*, *52*(3), 516–546. https://doi.org/10.3102/0002831214565787

Gersten, R., Taylor, M. J., Keys, T. D., Rolfhus, E., & Newman-Gonchar, R. (2014). Summary of research on the effectiveness of math professional development approaches. *Instructional Research Group*, 16.

Glover, T. A. (2017). Investigating Teacher Professional Development with Distance Coaching to Promote Students' Response to Reading Interventions in Rural Schools. In G. C. Nugent, G. M. Kunz, S. M. Sheridan, T. A. Glover, & L. L. Knoche (Eds.), *Rural Education Research in the United States: State of the Science and Emerging Directions* (pp. 167–179). Springer International Publishing. https://doi.org/10.1007/978-3-319-42940-3_9

Grooms, A. A., & Childs, J. (2021). "We Need to Do Better by Kids": Changing Routines in U.S. Schools in Response to COVID-19 School Closures. *Journal of Education for Students Placed at Risk (JESPAR)*, *26*(2), 135–156. https://doi.org/10.1080/10824669.2021.1906251

Gwazdauskas, A. J. (2009). *Examination of sustained instructional improvement for teachers supported by an instructional coach* [Ed.D., Northcentral University]. https://www.proquest.com/docview/305177540/abstract/74E1E201AA24A09PQ/1

Hamre, B. K., Partee, A., & Mulcahy, C. (2017). Enhancing the Impact of Professional Development in the Context of Preschool Expansion. *AERA Open*, *3*(4), 2332858417733686. https://doi.org/10.1177/2332858417733686

Hedges, L. V. (1981). Distribution theory for Glass's estimator of effect size and related estimators. *Journal of Educational Statistics*, *6*(2), 107–128. https://doi.org/10.2307/1164588

Ho, A. D., & Yu, C. C. (2015). Descriptive Statistics for Modern Test Score Distributions: Skewness, Kurtosis, Discreteness, and Ceiling Effects. *Educational and Psychological Measurement*, *75*(3), 365–388. https://doi.org/10.1177/0013164414548576

Hochberg, E. D., & Desimone, L. M. (2010). Professional Development in the Accountability Context: Building Capacity to Achieve Standards. *Educational Psychologist*, *45*(2), 89–106. https://doi.org/10.1080/00461521003703052

Huinker, D., Sarah B. Bush, & Graham, K. J. (2020). Catalyzing Change in School Mathematics: Creating the Opportunities Our Students Deserve. *Mathematics Teacher: Learning and Teaching PK-12*, *113*(10), 780–790. https://doi.org/10.5951/MTLT.2020.0053

Huntsinger, C. S., Jose, P. E., & Luo, Z. (2016). Parental facilitation of early mathematics and reading skills and knowledge through encouragement of home-based activities. *Early Childhood Research Quarterly*, *37*, 1–15. https://doi.org/10.1016/j.ecresq.2016.02.005

Imbens, G. W., & Rubin, D. B. (2010). Rubin Causal Model. In S. N. Durlauf & L. E. Blume (Eds.), *Microeconometrics* (pp. 229–241). Palgrave Macmillan UK. https://doi.org/10.1057/9780230280816_28

Jamieson, J. (1999). Dealing with baseline differences: Two principles and two dilemmas. *International Journal of Psychophysiology*, *31*(2), 155–161. https://doi.org/10.1016/S0167-8760(98)00048-8

Jenkins, J. M., Farkas, G., Duncan, G. J., Burchinal, M., & Vandell, D. L. (2016). Head Start at Ages 3 and 4 Versus Head Start Followed by State Pre-K: Which Is More Effective? *Educational Evaluation and Policy Analysis*, *38*(1), 88–112. https://doi.org/10.3102/0162373715587965

Jenkins, J. M., Watts, T. W., Magnuson, K., Gershoff, E. T., Clements, D. H., Sarama, J., & Duncan, G. J. (2018). Do High-Quality Kindergarten and First-Grade Classrooms Mitigate Preschool Fadeout? *Journal of Research on Educational Effectiveness*, *11*(3), 339–374. https://doi.org/10.1080/19345747.2018.1441347

Jordan, N. C., & Dyson, N. (2016). Chapter 3 - Catching Math Problems Early: Findings From the Number Sense Intervention Project. In A. Henik (Ed.), *Continuous Issues in Numerical Cognition* (pp. 59–79). Academic Press. https://doi.org/10.1016/B978-0-12-801637-4.00003-2

Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting First-Grade Math Achievement from Developmental Number Sense Trajectories. *Learning Disabilities Research & Practice*, *22*(1), 36–46. https://doi.org/10.1111/j.1540-5826.2007.00229.x

Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850–867. https://doi.org/10.1037/a0014939

Karchmer-Klein, R., & Pytash, K. E. (2020). Effective Practices in Online Teacher Preparation for Literacy Educators. Advances in Higher Education and Professional Development (AHEPD) Book Series. In *IGI Global*. IGI Global. https://doi.org/10.4018/978-1-7998-0206-8

Kearney, C. A., & Childs, J. (2021). A multi-tiered systems of support blueprint for re-opening schools following COVID-19 shutdown. *Children and Youth Services Review*, *122*, 105919. https://doi.org/10.1016/j.childyouth.2020.105919

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding It Up: Helping Children Learn Mathematics*. National Academy Press.

Kim, Y., & Steiner, P. (2016). Quasi-Experimental Designs for Causal Inference. *Educational Psychologist*, *51*(3–4), 395–405. https://doi.org/10.1080/00461520.2016.1207177

Kisa, Z., & Correnti, R. (2015). Examining Implementation Fidelity in America's Choice Schools: A Longitudinal Analysis of Changes in Professional Development Associated With Changes in Teacher Practice. *Educational Evaluation and Policy Analysis*, *37*(4), 437–457. https://doi.org/10.3102/0162373714557519

Kleemans, T., Peeters, M., Segers, E., & Verhoeven, L. (2012). Child and home predictors of early numeracy skills in kindergarten. *Early Childhood Research Quarterly*, *27*(3), 471–477. https://doi.org/10.1016/j.ecresq.2011.12.004

Klein, A., Starkey, P., Clements, D., Sarama, J., & Iyer, R. (2008). Effects of a Pre-Kindergarten Mathematics Intervention: A Randomized Experiment. *Journal of Research on Educational Effectiveness*, *1*(3), 155–178. https://doi.org/10.1080/19345740802114533

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. Journal of Chiropractic Medicine, 15, 155-163.

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence. *Review of Educational Research*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Kraft, M. A., & Hill, H. C. (2020). Developing Ambitious Mathematics Instruction Through Web-Based Coaching: A Randomized Field Trial. *American Educational Research Journal*, *57*(6), 2378–2414. https://doi.org/10.3102/0002831220916840

Kurz, A., Reddy, L. A., & Glover, T. A. (2017). A Multidisciplinary Framework of Instructional Coaching. *Theory Into Practice*, *56*(1), 66–77. https://doi.org/10.1080/00405841.2016.1260404

Ledford, J. R., Hall, E., Conder, E., & Lane, J. D. (2016). Research for Young Children With Autism Spectrum Disorders: Evidence of Social and Ecological Validity. *Topics in Early Childhood Special Education*, *35*(4), 223–233. https://doi.org/10.1177/0271121415585956

Lee, S. C., Nugent, G., Kunz, G. M., Houston, J., & DeChenne-Peters, S. (2018). Case Study: Value-Added Benefit of Distance-Based Instructional Coaching on Science Teachers' Inquiry Instruction in Rural Schools. *Journal of Science Teacher Education*, *29*(3), 179–199. https://doi.org/10.1080/1046560X.2018.1432226

Lemons, C. J., King, S. A., Davidson, K. A., Berryessa, T. L., Gajjar, S. A., & Sacks, L. H. (2016). An Inadvertent Concurrent Replication: Same Roadmap, Different Journey. *Remedial and Special Education*, *37*(4), 213–222. https://doi.org/10.1177/0741932516631116

Lemons, C. J., & Toste, J. R. (2019). Professional Development and Coaching: Addressing the "Last Mile" Problem in Educational Research. *Assessment for Effective Intervention*, *44*(4), 300–304. https://doi.org/10.1177/1534508419862859

Libertus, M. E., Feigenson, L., & Halberda, J. (2011). Preschool acuity of the approximate number system correlates with school math ability. *Developmental Science*, *14*(6), 1292–1300. https://doi.org/10.1111/j.1467-7687.2011.01080.x

Liou, S.-M., & Peng, C.-Y. J. (2006). The Use Of Hierarchical ANCOVA In Curriculum Studies. *Journal of Modern Applied Statistical Methods*, *5*(1), 230–247. https://doi.org/10.22237/jmasm/1146457080

Lukie, I. K., Skwarchuk, S.-L., LeFevre, J.-A., & Sowinski, C. (2014). The Role of Child Interests and Collaborative Parent–Child Interactions in Fostering Numeracy and Literacy Development in Canadian Homes. *Early Childhood Education Journal*, *42*(4), 251–259. https://doi.org/10.1007/s10643-013-0604-7

Lyons, I. M., Price, G. R., Vaessen, A., Blomert, L., & Ansari, D. (2014). Numerical predictors of arithmetic success in grades 1–6. *Developmental Science*, *17*(5), 714–726. https://doi.org/10.1111/desc.12152

Magill, N., Knight, R., McCrone, P., Ismail, K., & Landau, S. (2019). A scoping review of the problems and solutions associated with contamination in trials of complex interventions in mental health. *BMC Medical Research Methodology*, *19*(1), 4. https://doi.org/10.1186/s12874-018-0646-z

Major, C. S., Paul, J. M., & Reeve, R. A. (2017). TEMA and Dot Enumeration Profiles Predict Mental Addition Problem Solving Speed Longitudinally. *Frontiers in Psychology*, *8*. https://doi.org/10.3389/fpsyg.2017.02263

Makel, M. C., & Plucker, J. A. (2014). Facts Are More Important Than Novelty: Replication in the Education Sciences. *Educational Researcher*, *43*(6), 304–316. https://doi.org/10.3102/0013189X14545513

Malkus, N. (2020). Too Little, Too Late: A Hard Look at Spring 2020 Remote Learning. In *American Enterprise Institute*. American Enterprise Institute. https://eric.ed.gov/?id=ED610307

Maxwell, D. (2011). Influence of Instructional Coaching on Elementary Teachers' Instructional Practice: A Case Study. In *ProQuest LLC*. ProQuest LLC.

Maxwell, S. E., Lau, M. Y., & Howard, G. S. (2015). Is psychology suffering from a replication crisis? What does "failure to replicate" really mean? *American Psychologist*, *70*(6), 487–498. https://doi.org/10.1037/a0039400

Mazzocco, M. (2016). Mathematics Awareness Month: Why Should Pediatricians be Aware of Mathematics and Numeracy? *Journal of Developmental & Behavioral Pediatrics*, *37*(3), 251–253. https://doi.org/10.1097/DBP.0000000000000294

McCrary, M. (2011). Mapping the Road to Instructional Coach Effectiveness: Exploring the Relationship between Instructional Coaching Efficacy, Practices, and Outcomes. In *ProQuest LLC*. ProQuest LLC.

McIntosh, K., Horner, R. H., & Sugai, G. (2009). Sustainability of Systems-Level Evidence-Based Practices in Schools: Current Knowledge and Future Directions. In W. Sailor, G. Dunlap, G. Sugai, & R. Horner (Eds.), *Handbook of Positive Behavior Support* (pp. 327–352). Springer US. https://doi.org/10.1007/978-0-387-09632-2_14

McIntyre, L. L., Eckert, T. L., Fiese, B. H., DiGennaro Reed, F. D., & Wildenger, L. K. (2010). Family Concerns Surrounding Kindergarten Transition: A Comparison of Students in Special and General Education. *Early Childhood Education Journal*, *38*(4), 259–263. https://doi.org/10.1007/s10643-010-0416-y

McCrary, M. (2011). Mapping the Road to Instructional Coach Effectiveness: Exploring the Relationship between Instructional Coaching Efficacy, Practices, and Outcomes. In *ProQuest LLC*. ProQuest LLC.

McWayne, C. M., Fantuzzo, J. W., & McDermott, P. A. (2004). Preschool Competency in Context: An Investigation of the Unique Contribution of Child Competencies to Early Academic Success. *Developmental Psychology*, *40*(4), 633–645.

Mohr, D. C., Spring, B., Freedland, K. E., Beckner, V., Arean, P., Hollon, S. D., Ockene, J., & Kaplan, R. (2009). The selection and design of control conditions for randomized controlled trials of psychological interventions. *Psychotherapy and Psychosomatics*, *78*(5), 275–284. https://doi.org/10.1159/000228248

Montana Office of Public Instruction. *Growth and Enhancement of Montana Students (GEMS)*. Retrieved November 1, 2020, from https://gems.opi.mt.gov/StudentCharacteristics/Pages/Overview.aspx

Morgan, P. L., Farkas, G., Hillemeier, M. M., & Maczuga, S. (2016). Who Is At Risk for Persistent Mathematics Difficulties in the United States? *Journal of Learning Disabilities*, *49*(3), 305–319. https://doi.org/10.1177/0022219414553849

Morgan, P. L., Farkas, G., & Wu, Q. (2009). Five-Year Growth Trajectories of Kindergarten Children With Learning Difficulties in Mathematics. *Journal of Learning Disabilities*, *42*(4), 306–321. https://doi.org/10.1177/0022219408331037

Murray, D. M. (1998). *Design and Analysis of Group-randomized Trials*. Oxford University Press.

Napoli, A. R., & Purpura, D. J. (2018). The home literacy and numeracy environment in preschool: Cross-domain relations of parent–child practices and child outcomes. *Journal of Experimental Child Psychology*, *166*, 581–603. https://doi.org/10.1016/j.jecp.2017.10.002

National Assessment for Educational Progress. (2019). NAEP Report Card: Mathematics. Retrieved from https://www.nationsreportcard.gov/mathematics?grade=4

National Council of Teachers of Mathematics. (2006, October). *Curriculum Focal Points for Pre-K–Grade 8 Mathematics: A Quest for Coherence*. https://www.nctm.org/Publications/teaching-children-mathematics/2006/Vol13/Issue3/Curriculum-Focal-Points-for-Pre-K%E2%80%93Grade-8-Mathematics_-A-Quest-for-Coherence/

National Governors Association. (2020). *Common Core State Standards*. Common Core State Standards Initiative. http://www.corestandards.org/

Nelson, G., & McMaster, K. L. (2019a). Factors that May Influence Treatment Effects: Helping Practitioners Select Early Numeracy Interventions. *Learning Disabilities Research & Practice*, *34*(4), 194–206. https://doi.org/10.1111/ldrp.12208

Nelson, G., & McMaster, K. L. (2019b). The effects of early numeracy interventions for students in preschool and early elementary: A meta-analysis. *Journal of Educational Psychology*, *111*(6), 1001–1022. https://doi.org/10.1037/edu0000334

Nese, R. N. T., Meng, P., Breiner, S., Chaparro, E., & Algozzine, R. (2020). Using stakeholder feedback to improve online professional development opportunities. *Journal of Research on Technology in Education*, *52*(2), 148–162. https://doi.org/10.1080/15391523.2020.1726233

Nguyen, T., Watts, T. W., Duncan, G. J., Clements, D. H., Sarama, J. S., Wolfe, C., & Spitler, M. E. (2016). Which preschool mathematics competencies are most predictive of fifth grade achievement? *Early Childhood Research Quarterly*, *36*, 550–560. https://doi.org/10.1016/j.ecresq.2016.02.003

Niklas, F., Cohrssen, C., & Tayler, C. (2016). Parents supporting learning: A non-intensive intervention supporting literacy and numeracy in the home learning environment. *International Journal of Early Years Education*, *24*(2), 121–142. https://doi.org/10.1080/09669760.2016.1155147

Niklas, F., & Schneider, W. (2014). Casting the die before the die is cast: The importance of the home numeracy environment for preschool children. *European Journal of Psychology of Education*, *29*(3), 327–345. https://doi.org/10.1007/s10212-013-0201-6

Nozari, M. (2019). *Effects of early explicit strategic intervention on the mathematics performance of students at-risk for Mathematics Difficulties* [Thesis]. https://doi.org/10.26153/tsw/3275

Onken, L. S., Carroll, K. M., Shoham, V., Cuthbert, B. N., & Riddle, M. (2014). Reenvisioning Clinical Science: Unifying the Discipline to Improve the Public Health. *Clinical Psychological Science*, *2*(1), 22–34. https://doi.org/10.1177/2167702613497932

Powers, J. D., Bowen, N. K., & Bowen, G. L. (2010). Evidence-based programs in school settings: Barriers and recent advances. *Journal of Evidence-Based Social Work*, *7*(4), 313–331. https://doi.org/10.1080/15433710903256807

Pressley, T. (2021). Factors Contributing to Teacher Burnout During COVID-19. *Educational Researcher*, 0013189X211004138. https://doi.org/10.3102/0013189X211004138

Purtell, K. M., Valauri, A., Rhoad-Drogalis, A., Jiang, H., Justice, L. M., Lin, T.-J., & Logan, J. A. R. (2020). Understanding policies and practices that support successful transitions to kindergarten. *Early Childhood Research Quarterly*, *52*, 5–14. https://doi.org/10.1016/j.ecresq.2019.09.003

Quinn, D. M., Cooc, N., McIntyre, J., & Gomez, C. J. (2016). Seasonal Dynamics of Academic Achievement Inequality by Socioeconomic Status and Race/Ethnicity: Updating and Extending Past Research With New National Data. *Educational Researcher*, *45*(8), 443–453. https://doi.org/10.3102/0013189X16677965

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Reardon, S. F., & Portilla, X. A. (2016). Recent Trends in Income, Racial, and Ethnic School Readiness Gaps at Kindergarten Entry. *AERA Open*, *2*(3). https://doi.org/10.1177/2332858416657343

Reddy, L. A., Glover, T., Kurz, A., & Elliott, S. N. (2019). Assessing the Effectiveness and Interactions of Instructional Coaches: Initial Psychometric Evidence for the Instructional Coaching Assessments–Teacher Forms. *Assessment for Effective Intervention*, *44*(2), 104–119. https://doi.org/10.1177/1534508418771739

Reddy, L. A., Shernoff, E., & Lekwa, A. (2021). A randomized controlled trial of instructional coaching in high-poverty urban schools: Examining teacher practices and student outcomes. *Journal of School Psychology*, *86*, 151–168. https://doi.org/10.1016/j.jsp.2021.04.001

Reigosa-Crespo, V., & Castro Cañizares, D. (2015). *Dots and digits How do children process the numerical magnitude? Evidence from brain and behavior* (pp. 60–77).

Rimm-Kaufman, S. E., & Pianta, R. C. (2000). An Ecological Perspective on the Transition to Kindergarten: A Theoretical Framework to Guide Empirical Research. *Journal of Applied Developmental Psychology*, *21*(5), 491–511. https://doi.org/10.1016/S0193-3973(00)00051-4

Ritchie, S. J., & Bates, T. C. (2013). Enduring Links From Childhood Mathematics and Reading Achievement to Adult Socioeconomic Status. *Psychological Science*, *24*(7), 1301–1308. https://doi.org/10.1177/0956797612466268

Roberts, C., & Roberts, S. A. (2005). Design and analysis of clinical trials with clustering effects due to treatment. *Clinical Trials*, *2*(2), 152–162. https://doi.org/10.1191/1740774505cn076oa

Rodgers, W. J., Kennedy, M. J., VanUitert, V. J., & Myers, A. M. (2019). Delivering Performance Feedback to Teachers Using Technology-Based Observation and Coaching Tools. *Intervention in School and Clinic*, *55*(2), 103–112. https://doi.org/10.1177/1053451219837640

Romano, E., Babchishin, L., Pagani, L. S., & Kohen, D. (2010). School readiness and later achievement: Replication and extension using a nationwide Canadian survey. *Developmental Psychology*, *46*(5), 995–1007. https://doi.org/10.1037/a0018880

Rosenquist, B. A., Henrick, E. C., & Smith, T. M. (2015). Research–Practice Partnerships to Support the Development of High Quality Mathematics Instruction for All Students. *Journal of Education for Students Placed at Risk*, *20*(1–2), 42–57. https://doi.org/10.1080/10824669.2014.988335

Rubin, D. B. (1986). Comment. *Journal of the American Statistical Association*, *81*(396), 961–962. https://doi.org/10.1080/01621459.1986.10478355

Sanetti, L., & Luh, H.-J. (2019). Fidelity of Implementation in the Field of Learning Disabilities. *Learning Disability Quarterly*, *42*(4), 204–216. https://doi.org/10.1177/0731948719851514

Sanetti, L. M., & Collier-Meek, M. A. (2019). Increasing implementation science literacy to address the research-to-practice gap in school psychology. *Journal of School Psychology*, *76*, 33–47. https://doi.org/10.1016/j.jsp.2019.07.008

SAS Institute Inc (2010). SAS/STAT Software, Version 9.3. SAS Institute Inc., Cary, NC. URL http://www.sas.com/.

Schmidt, S. (2009). Shall we Really do it Again? The Powerful Concept of Replication is Neglected in the Social Sciences. *Review of General Psychology*, *13*(2), 90–100. https://doi.org/10.1037/a0015108

Shanley, L., Cary, M. S., Clarke, B., & Jungjohann, K. (2013). Teaching Early Knowledge of Whole Number Concepts through Technology: Findings from a Feasibility Study of an iPad Delivered Kindergarten Mathematics Intervention. In *Society for Research on Educational Effectiveness*. Society for Research on Educational Effectiveness. https://eric.ed.gov/?id=ED564109

Shanley, L., Clarke, B., Doabler, C. T., Kurtz-Nelson, E., & Fien, H. (2017). Early Number Skills Gains and Mathematics Achievement: Intervening to Establish Successful Early Mathematics Trajectories. *The Journal of Special Education*, *51*(3), 177–188. https://doi.org/10.1177/0022466917720455

Simons, D. J. (2014). The Value of Direct Replication. *Perspectives on Psychological Science*, *9*(1), 76–80. https://doi.org/10.1177/1745691613514755

Skwarchuk, S.-L. (2009). How Do Parents Support Preschoolers' Numeracy Learning Experiences at Home? *Early Childhood Education Journal*, *37*(3), 189–197. https://doi.org/10.1007/s10643-009-0340-1

Slavin, R. E., & Storey, N. (2020). *The US Educational Response to the COVID-19 Pandemic* (SSRN Scholarly Paper ID 3652585). Social Science Research Network. https://papers.ssrn.com/abstract=3652585

Sloboda, Z., Dusenbury, L., & Petras, H. (2014). Implementation Science and the Effective Delivery of Evidence-Based Prevention. In Z. Sloboda & H. Petras (Eds.), *Defining Prevention Science* (pp. 293–314). Springer US. https://doi.org/10.1007/978-1-4899-7424-2_13

Smolkowski, K., Crawford, L., Seeley, J. R., & Rochelle, J. (2019). Introduction to Implementation Science for Research on Learning Disabilities. *Learning Disability Quarterly*, *42*(4), 192–203. https://doi.org/10.1177/0731948719851512

Snodgrass, M. R., Chung, M. Y., Meadan, H., & Halle, J. W. (2018). Social validity in single-case research: A systematic literature review of prevalence and application. *Research in Developmental Disabilities*, *74*, 160–173. https://doi.org/10.1016/j.ridd.2018.01.007

Sood, S., & Jitendra, A. K. (2013). An Exploratory Study of a Number Sense Program to Develop Kindergarten Students' Number Proficiency. *Journal of Learning Disabilities*, *46*(4), 328–346. https://doi.org/10.1177/0022219411422380

Sood, S., & Mackey, M. (2015). *Examining the Effects of Number Sense Instruction on Mathematics Competence of Kindergarten Students* (Vol. 2). Arc Journals.

Spear, C. F., Strickland-Cohen, M. K., Romer, N., & Albin, R. W. (2013). An Examination of Social Validity Within Single-Case Research With Students With Emotional and Behavioral Disorders. *Remedial and Special Education*, *34*(6), 357–370. https://doi.org/10.1177/0741932513490809

Strand Cary, M. G., Clarke, B., Doabler, C. T., Smolkowski, K., Fien, H., & Baker, S. K. (2017). A Practitioner Implementation of a Tier 2 First-Grade Mathematics Intervention. *Learning Disability Quarterly*, *40*(4), 211–224. https://doi.org/10.1177/0731948717714715

Susperreguy, M. I., Douglas, H., Xu, C., Molina-Rojas, N., & LeFevre, J.-A. (2020). Expanding the Home Numeracy Model to Chilean children: Relations among parental expectations, attitudes, activities, and children's mathematical outcomes. *Early Childhood Research Quarterly*, *50*, 16–28. https://doi.org/10.1016/j.ecresq.2018.06.010

Thompson, R. J., Napoli, A. R., & Purpura, D. J. (2017). Age-related differences in the relation between the home numeracy environment and numeracy skills. *Infant and Child Development*, *26*(5), e2019. https://doi.org/10.1002/icd.2019

Throndsen, J. E., Shumway, J. F., & Moyer-Packenham, P. S. (2020). The relationship between mathematical literacy at kindergarten entry and public preschool attendance, type, and quality. *Early Childhood Education Journal*, *48*(4), 473–483. https://doi.org/10.1007/s10643-019-01014-7

Toll, S. W. M., & Van Luit, J. E. H. (2012). Early Numeracy Intervention for Low-Performing Kindergartners. *Journal of Early Intervention*, *34*(4), 243–264. https://doi.org/10.1177/1053815113477205

Vernon-Feagans, L., Kainz, K., Hedrick, A., Ginsberg, M., & Amendum, S. (2013). Live webcam coaching to help early elementary classroom teachers provide effective literacy instruction for struggling readers: The Targeted Reading Intervention. *Journal of Educational Psychology*, *105*(4), 1175–1187. https://doi.org/10.1037/a0032143

Visscher, A. J. (2020). On the value of data-based decision making in education: The evidence from six intervention studies. *Studies in Educational Evaluation*, *69*. https://doi.org/10.1016/j.stueduc.2020.100899

Wade, M., Prime, H., & Browne, D. T. (2020). Why we need longitudinal mental health research with children and youth during (and after) the COVID-19 pandemic. *Psychiatry Research*, *290*. https://doi.org/10.1016/j.psychres.2020.113143

Walsh, M., Matsumura, L. C., Zook-Howell, D., Correnti, R., & DiPrima Bickel, D. (2020). Video-based literacy coaching to develop teachers' professional vision for dialogic classroom text discussions. *Teaching and Teacher Education*, *89*. https://doi.org/10.1016/j.tate.2019.103001

Weiss, M. J. (2010). The Implications of Teacher Selection and the Teacher Effect in Individually Randomized Group Treatment Trials. *Journal of Research on Educational Effectiveness*, *3*(4), 381–405. https://doi.org/10.1080/19345747.2010.504289

Wolf, M. M. (1978). Social Validity: The Case For Subjective Measurement or How Applied Behavior Analysis is Finding its Heart. *Journal of Applied Behavior Analysis*, *11*(2), 203–214. https://doi.org/10.1901/jaba.1978.11-203