# EXAMINING PATTERNS AND PREDICTORS OF RESPONSE TO MATHEMATICS INTERVENTION

by

TAYLOR A. LESNER

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 2023

DISSERTATION APPROVAL PAGE

Student: Taylor A. Lesner

Title: Examining Patterns and Predictors of Response to Mathematics Intervention

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

| | |
|---|---|
| Ben Clarke, Ph.D. | Chairperson |
| Geovanna Rodriguez, Ph.D. | Core Member |
| Derek Kosty, Ph.D. | Core Member |
| Elizabeth Budd, Ph.D. | Institutional Representative |

and

| | |
|---|---|
| Krista Chronister, Ph.D. | Vice Provost for Graduate Studies |

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

Degree awarded June 2023

DISSERTATION ABSTRACT

Taylor A. Lesner

Doctor of Philosophy

Department of Special Education and Clinical Sciences

June 2023

Title: Examining Patterns and Predictors of Response to Mathematics Intervention

Concerns about low mathematics achievement have created a push to increase national mathematics proficiency levels through research and policy. Efforts have largely focused on early mathematics interventions for students with or at risk for mathematics learning difficulties implemented within multi-tiered response to intervention (RTI) frameworks. Within these models, important educational decisions are based upon instructional response, making meaningful categorization of student responsiveness to intervention paramount. Across research and practice, intervention outcomes are typically thought of as a binary, with students considered either responsive or non-responsive to intervention. However, defining and categorizing responsiveness in more complex ways may reveal important differences between subgroups of students who exhibit distinct patterns of responsiveness over time.

The present study explored patterns of response to an early mathematics intervention using data from the ROOTS Efficacy Project. Participants included kindergarten students at risk for mathematics difficulties who were randomly assigned to the ROOTS intervention condition (*n* = 880). Results of a latent profile analysis indicated that variability in response to a generally effective intervention was best captured by a more complex categorization framework encompassing four distinct response profiles: a moderate-risk, mildly responsive group; a moderate-risk, delayed response group, a high-risk, strongly responsive group; and a lower-risk, non-responsive group. Membership in each response profile group was predicted by pre-intervention performance on measures of both early mathematics and cognitive skills (visual-spatial, fluid reasoning, and working memory). On average, students with lower initial math skill and cognitive performance demonstrated stronger intervention response. Implications for future research and practice are discussed.

CURRICULUM VITAE

NAME OF AUTHOR: Taylor A. Lesner

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, Oregon

Vanderbilt University, Nashville, Tennessee

DEGREES AWARDED:

Doctor of Philosophy, School Psychology, 2023, University of Oregon

Master of Science, Special Education, 2021, University of Oregon

Bachelor of Arts, Psychology & Child Development, 2017, Vanderbilt University

AREAS OF SPECIAL INTEREST:

Integrating academic, behavioral, and social-emotional systems of support in schools

Prevention and intervention of mathematics difficulties

Improving outcomes for students with learning and attention difficulties

Supporting students who do not respond adequately to evidence-based interventions

PROFESSIONAL EXPERIENCE:

Clinical Extern, HEDCO Comprehensive Diagnostic Assessment Clinic, 2021-Present

School Psychology Advanced Practicum Student, Springfield Public Schools, 2020-2021

School Psychology Practicum Student, Springfield Public Schools, 2019-2020

Graduate Employee, University of Oregon Center on Teaching and Learning, 2019-Present

Research Assistant, Vanderbilt University, 2017-2018

GRANTS, AWARDS, AND HONORS:

Steven A. Raymund Doctoral Fellowship, University of Oregon, 2018

Graduated summa cum laude, Vanderbilt University, 2017

Jum C. Nunnally Award for Outstanding Honors Research, Vanderbilt University, 2017

Littlejohn Fellowship, Vanderbilt Undergraduate Summer Research Program, 2016

PUBLICATIONS:

**Peer-Reviewed**

Sutherland, M., **Lesner, T.**, Kosty, D., Lussier, C., Smolkowski, K., Turtura, J., & Clarke, B. (in press). Examining interactions across instructional tiers: Do features of Tier 1 predict student responsiveness to Tier 2 mathematics intervention? *Journal of Learning Disabilities.*

Clarke, B., Sutherland, M., Doabler, C.T., Lesner, T., Fainstein, D., Nolan, K., Landis, B., & Kosty, D. (2021). Developing and investigating the promise of early measurement screeners. *School Psychology Review.* https://doi.org/10.1080/2372966X.2021.1919493

Clarke, B., Nelson, N., Ketterlin-Geller, L., Kosty, D., Smolkowski, K., Lesner, T., Furjanic, D., & Fien, H. (2020). Investigating the promise of a tier two sixth grade fractions intervention. *Learning Disabilities Quarterly*. https://doi.org/10.1177/0731948720972411

Walden, T., Lesner, T., & Jones, R. (2020). Implicit and explicit attitudes toward stuttering among speech-language pathologists. *Journal of Communication Disorders, 83.* https://doi.org/10.1016/j.jcomdis.2019.105965

Walden, T. & Lesner, T. (2018). Examining implicit and explicit attitudes toward stuttering. *Journal of Fluency Disorders, 57,* 22-36. https://doi.org/10.1016/j.jfludis.2018.06.001

**In Preparation**

Clarke, B., Sutherland, M., & **Lesner, T.** (in production). Supporting student achievement in elementary mathematics: The role of the school psychologist. *Desk Reference in School Psychology* (L.A. Theodore, M. Bray, & B. Bracken, Eds.). Oxford University Press.

Clarke, B., Turtura, J., **Lesner, T.**, Cook, M., Smolkowski, K., Kosty, D., & Doabler, C. (submitted for review). A conceptual replication of a kindergarten math intervention within the context of a research-based core. *Exceptional Children.*

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

**Lesner, T.,** Clarke, B., Nelson, N., Kosty, D., Ketterlin-Geller, L., & Smolkowski, K. (under first review). Exploring the relationship between initial mathematics skill and a sixth-grade fractions intervention. *Remedial and Special Education.*

ACKNOWLEDGMENTS

I wish to express my deepest gratitude to my advisor, Ben Clarke, for his guidance throughout the dissertation process and my graduate school career. His advice, encouragement, and enthusiasm have been integral to my professional development, pushing me to grow as a scholar. I also wish to express special thanks to Derek Kosty for his time, patience, and methodological expertise, which were invaluable throughout this process. I would also like to express my gratitude to Geovanna Rodriguez and Elizabeth Budd for their service as members on my committee and for their support and feedback throughout the development and completion of this dissertation.

This investigation was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R324A120304 awarded to the Center on Teaching and Learning at the University of Oregon.

To my family, who taught me the power and value of education.

To my mother, for believing in my dreams and encouraging me to pursue them.

To my husband, for sharing in my ambitions, inspiring me to be my best self, and supporting me every step of the way.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

LIST OF TABLES

# CHAPTER I

# INTRODUCTION

In the past two decades, increasing mathematics proficiency has become a national priority. Students in United States schools consistently demonstrate poor mathematics performance compared to international peers, with the U.S. ranking among the lowest compared to other developed nations (Olson et al., 2008). Meanwhile, mathematics proficiency is critical to success in today's economy, with employment opportunities in the science, technology, engineering, and mathematics (STEM) fields projected to grow at over twice the rate of other sectors over the next decade, driven largely by expansion in mathematical sciences (e.g., mathematicians and statisticians; Zilberman & Ice, 2021). With growing recognition of the implications of mathematics learning across a range of short- and long-term outcomes, from school adjustment and academic achievement to employment opportunities and socioeconomic status (National Mathematics Advisory Panel, 2008; Morgan et al., 2009), recent years have brought several federal initiatives to improve mathematics instruction and achievement (e.g., Common Core State Standards Initiative, 2010; NMAP, 2008). Despite these efforts, National Assessment of Educational Progress (NAEP) data indicate that achievement levels are stagnating, and many students are not achieving proficiency with foundational mathematics content. In the most recent assessment, only 41% of fourth grade students scored at or above the proficient level while 19% of students performed at a level classified as below basic (NAEP, 2019). Students from low socioeconomic status backgrounds, English language learners, racial/ethnic minorities, and students with disabilities demonstrated poorer performance and were more likely to score in the below basic range, resulting in substantial achievement gaps.

Persistent low levels of math performance and achievement gaps are particularly troubling given findings that mathematics difficulties emerge early in schooling and lead to learning trajectories that are difficult to change (Duncan et al., 2007). At school entry, students demonstrate substantial variability in mathematics knowledge across a range of concepts and skills (Jordan et al., 2007), and this early knowledge strongly predicts later achievement. Converging evidence suggests that students with poor kindergarten mathematics performance demonstrate lower growth in mathematics skills and often experience continued learning difficulties, which become more difficult to remediate over time (Bodovski & Farkas, 2007; Duncan et al., 2007; Morgan et al., 2009). However, efforts to identify and provide intervention to students with early mathematics difficulties can significantly alter learning trajectories and improve long-term outcomes (Clarke et al., 2020; Morgan et al., 2009). Growth in mathematics knowledge during kindergarten and first grade has been shown to be a stronger predictor of long-term math achievement than skills at kindergarten entry (Watts et al., 2014), supporting the importance of providing effective early math intervention to students with low initial mathematics skill.

## Response to Intervention

Given the importance of developing early mathematics proficiency for long-term outcomes, recent years have seen increased interest in identifying effective strategies to support students with or at risk for mathematics difficulties (MD) and a growing research focus on developing and evaluating mathematics interventions. In the research literature, students are typically identified as at risk for MD based on performance at or below the $30^{th}$-$35^{th}$ percentile on one or more measures of mathematics knowledge and skill (e.g., L. S. Fuchs, Sterba, et al., 2016; Morgan et al., 2009). Thus, mathematics intervention studies typically include both students with math learning disabilities (defined as persistent, low math achievement that is discrepant from overall cognitive functioning; estimated at 5-10% of the school age population; Geary, 2011; Morgan et al., 2009) and those with low math performance defined exclusively by performance below a set achievement level. Much of the mathematics intervention literature has centered on early interventions targeting foundational, whole number concepts, thought to be a critical first step toward overall mathematics proficiency (Frye et al., 2013; Fuchs et al., 2021; NMAP, 2008). Typically, these programs are designed for implementation within multi-tiered systems of support (MTSS) or response to intervention (RTI) frameworks.

Although originally intended to codify response to research-based instruction as an alternative to traditional models for identifying students with specific learning disabilities, RTI frameworks have evolved into comprehensive, prevention-oriented models for delivery of instruction across multiple levels of support (Berkeley et al., 2020; Vaughn & Fuchs, 2003). RTI models are designed to support positive academic outcomes for all students and reduce achievement gaps by providing a continuum of evidence-based instruction and interventions, with intensity of supports matched to student need. They are frequently conceptualized as a pyramid consisting of three tiers of increasingly intensive supports (Gersten et al., 2009). Tier 1 encompasses core mathematics instruction provided to all students; Tier 2 consists of supplemental, small-group interventions provided to students at risk for MD; and Tier 3 consists of more intensive, individualized supports provided to students who do not make adequate progress toward instructional goals in Tier 2 (Gersten et al., 2009; National Center on Response to Intervention, 2010). A critical feature of RTI models is that instructional decisions are based on data: universal screening data are used to identify students at risk for mathematics difficulties who may benefit from more intensive supports, and progress monitoring data are used to determine whether students receiving intervention are making adequate progress toward instructional goals. Supports are then intensified when students are not making adequate progress and may be faded when students meet various criteria (Gersten et al., 2009; Mellard et al., 2010).

Within this context, several researchers have developed and evaluated supplemental (Tier 2) mathematics interventions for students in the early elementary school grades (e.g., Bryant et al., 2011; Clarke, et al., 2016a; Dyson et al., 2013; Fuchs et al., 2005; Gersten et al., 2015; Sood & Jitendra, 2011). Typically, these intervention programs target foundational, whole-number content and incorporate explicit and systematic instructional design in accordance with best practice guidelines for supporting students struggling in mathematics (e.g., Fuchs et al., 2021; Gersten et al., 2009). Results across these research efforts have been promising, documenting significant, positive impact on mathematics outcomes for students with MD.

Despite substantial progress toward developing and validating intervention programs, students at risk for MD demonstrate variability in response to validated mathematics intervention. A substantial minority of students at risk for learning difficulties do not respond adequately to generally effective academic intervention programs, posing a major challenge for practitioners and intervention researchers alike (L. S. Fuchs & Vaughn, 2012). Throughout the literature on mathematics and reading interventions, it is estimated that approximately 10-25% of students do not respond adequately to programs that are beneficial to many of their peers (D. Fuchs & Fuchs, 2019). Within an RTI framework, inadequate response to increasingly intensive, evidence-based instruction is viewed as an indicator of a learning disability; however the question of what constitutes inadequate response has been widely debated (Barth et al., 2008). As a growing number of school systems adopt RTI models of service delivery (Berkeley et al., 2020), important educational decisions are increasingly based upon instructional response. Multi-tiered models rely on evaluating instructional response to facilitate matching of students to tiers of instruction and specific supports within those tiers, decide when intervention should be intensified or faded, and determine eligibility for special education services. It is therefore critical to consider how instructional response is defined and the implications this may have for students.

**Defining Responsiveness to Intervention**

Fletcher and Miciak (2019) propose that some disorders are categorical in nature; individuals either have the disorder or do not. On the other hand, many disorders are dimensional in nature, existing along a continuum in which individuals may fall anywhere between the upper and lower extremes. They argue that intervention response is best understood as a continuous construct, with weaker response thought to reflect more severe academic difficulties (Fletcher & Miciak, 2019). As with all dimensional phenomena, no natural cut point distinguishes responders from non-responders. However, in the intervention literature as well as in practice, response to intervention is typically conceptualized as a binary outcome in which students are considered as either responding or not responding adequately to instruction (Fletcher & Miciak, 2019). This dichotomization of responsiveness aligns with the decisions educators must make within an RTI model, in which supports are maintained or faded for students who are responding to intervention or intensified for students who are not responding adequately (Barth et al., 2008).

Categorizing students according to a response/non-response binary facilitates the use of data to make decisions about movement between tiers. The literature in this area also suggests that adequate responders differ meaningfully from inadequate responders in domains related to, but distinct from, instructional response, including cognitive, behavioral, and academic characteristics (Al Otaiba & Fuchs, 2006; Fletcher & Miciak, 2019; Vellutino et al., 2006). However, the field has not yet arrived at a universally accepted method of operationalizing response to intervention, resulting in wide variability in criteria across both research and practice (Barth et al., 2008; Fletcher & Miciak, 2019). Three broad approaches to determining when student response is inadequate based on data have been proposed: 1) final status models, which compare postintervention performance to an established criterion, such as a benchmark score; 2) slope discrepancy models, which compare student growth over time to expected growth rates; and 3) dual discrepancy models, which consider both performance level and growth over time (D. Fuchs & Deshler, 2007). Within these broad approaches, infinite variations in specific measures and criteria are possible. Barth et al., (2008) examined agreement among different operationalizations of response within an RTI framework. The authors applied a range of different cut points, methods (i.e., criterion, slope discrepancy, and dual discrepancy), and measures to determine responsiveness within a data set and analyzed overlap in categorization when various combinations were applied. They found that decisions about how to operationalize response substantially impacted which students were identified as adequate and inadequate responders.

The widely used response/non-response binary has several key limitations that have implications for service-delivery within a RTI model. First, no matter where the cut point between adequate and inadequate response is set, scores that fall near the cut point are less reliable, making it difficult to distinguish between responders and non-responders whose scores fall in the borderline range. Furthermore, while distinctions between students whose scores fall near a cut point are necessary in practice, they are often arbitrary, since students who score just above or just below a cut point tend to demonstrate similar overall performance and learning needs (Barth et al., 2008; Fletcher & Miciak, 2019). Finally, in the intervention literature, response is often defined in terms of post-test scores that exceed those of a matched control group, with little attention to whether positive outcomes are sustained or fade over time (Peng et al., 2020). While relatively little research has examined long-term outcomes associated with mathematics intervention, the literature in this area suggests that even when significant impacts associated with intervention programs remain after a delay, gains often diminish over time (Bailey et al., 2020; Gersten, 2016). These fadeout effects are widespread and complex (Bailey et al., 2020), suggesting the importance of considering both short- and long-term patterns of responsiveness to intervention.

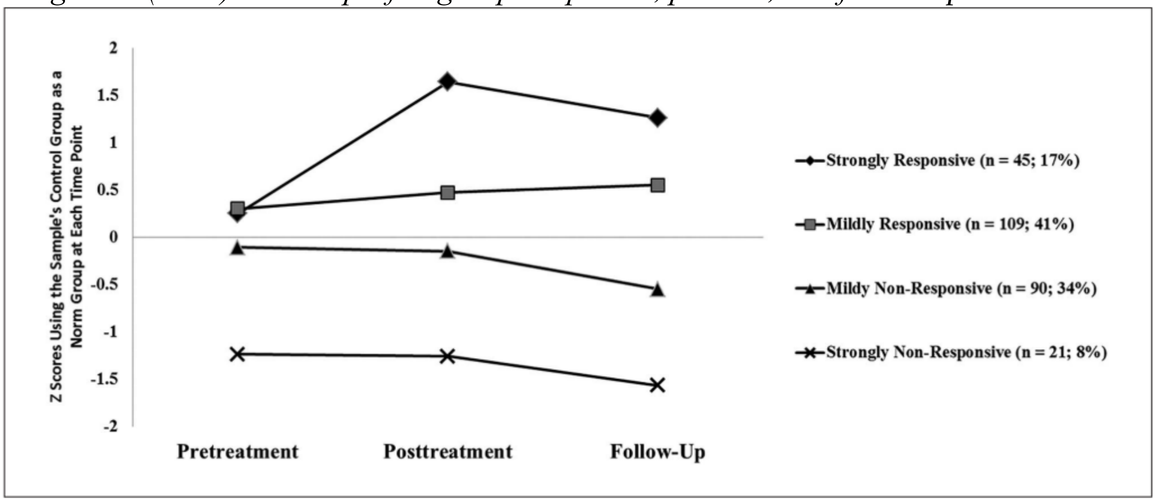
**Examining More Complex Categorizations of Intervention Response**

Peng and colleagues (2020) argued that the response/non-response dichotomy may be an oversimplification that obscures important differences between subgroups of students exhibiting different patterns of intervention response. As an example of the potential value of allowing for more than two response groups (i.e., responders and non-responders), the authors describe a more complex framework based on the early intervention literature that considers patterns of response to intervention across multiple timepoints, resulting in four possible response categories. First, students may significantly outperform the control group at immediate and delayed posttest. These students are sometimes considered false positives, since their initial low performance may be better explained by environmental factors (e.g., poor core instruction or classroom management, lack of exposure to critical content) than a learning disability. The majority of students in this group should experience continued growth if provided ongoing high-quality core instruction. Second, students may display significant gains at posttest that fade out, or attenuate, over time in the absence of further intervention. Students in this group may require ongoing support to make adequate progress toward educational goals. Third, students may experience outcomes comparable to peers who did not receive intervention across timepoints. Assuming the intervention was implemented with integrity, these students may be considered "true" non-responders to the program in question and require more intensive and/or individualized instruction to meet their learning needs. Finally, students may appear to be non-responders at posttest but display positive outcomes at delayed posttest. Students in this group may experience “sleeper” effects in which skills targeted in intervention continue to grow and generalize over time. The authors emphasize that the proposed typology is merely intended to illustrate one way that additional response categories may be used to understand variability in response to intervention and inform decision-making within an RTI framework.

To evaluate the utility of allowing for more than two categories of responsiveness, Peng and colleagues (2020) used latent profile analysis to determine the number of subgroups that best described variation in response to an evidence-based first grade reading intervention. The authors found that student response to intervention fit four distinct response types: Strongly responsive students (17% of the sample), who significantly outperformed the control group at both posttest and follow-up; mildly responsive students (41% of the sample), who significantly outperformed the control group across timepoints to a lesser degree; mildly non-responsive students (34% of the sample), who performed comparably to controls at posttest but ended up falling behind by follow-up; and strongly non-responsive students (8% of the sample), whose performance was significantly below that of the control group at both posttest and follow-up. Mean trajectories of each response category are shown in Figure 1 below. Overall, each group displayed relatively stable response (or non-response) across post-intervention timepoints, and results did not support either fade out or sleeper effects – however, the authors argue that this finding may be specific to the early literacy skills targeted by the intervention, which may be less susceptible to these temporal effects than learning across other skills or content areas. Critically, the results highlight the potential value of distinguishing between subgroups of non-responders who may require different next steps to achieve meaningful growth in academic skills. The authors argue that students in the mildly non-responsive group could be expected to respond positively to intensification of the program (e.g., higher dosage of the same intervention), whereas the strongly non-responsive group is likely to require specially designed instruction to make adequate progress.

**Figure 1**

*Peng et al. (2020)'s latent profile groups at pretest, posttest, and follow-up*




## Predictors of Intervention Response

Despite the substantial literature base on effective academic interventions, knowledge of how to effectively meet the needs of students who do not respond adequately to validated programs is lacking (Filderman et al., 2018; Gersten et al., 2009). To address this, there have been calls for intervention researchers to evaluate not only main effects, but also for which students and under what conditions programs are effective (D. Fuchs & Fuchs, 2019; Miller et al., 2014). Accordingly, researchers have begun to explore student-level cognitive, behavioral, and academic characteristics that predict differential treatment response, contributing to a more nuanced understanding of the impact of validated academic interventions. Extending this line of research to explore predictors of membership in various response categories may help link understanding of intervention impact with improved educational decision-making. For example, student-level predictors of intervention outcomes may be incorporated during screening to predict upfront which students are unlikely to respond adequately to a particular intervention (Miller et al., 2014). School teams may then be able to place some students directly into more intensive intervention programs or modify programs to better align with student needs (Al Otaiba et al., 2014; Lam & McMaster, 2014). Given the likelihood that effective methods of intensifying supports differ across subgroups of non-responders (Peng et al., 2020), student characteristics associated with response types may also inform next steps for students who do not respond adequately to intervention.

Since instructional response is fundamental to decision-making in an RTI framework, research in this vein is a critical step toward better understanding intervention effects, and ultimately supporting effective implementation of multi-tiered models of mathematics support in schools. The following sections review the literature linking student-level characteristics to mathematics development and intervention response. Prior research in this area suggests that domain-specific mathematics knowledge and skill, as well as domain-general cognitive abilities (e.g., working memory, fluid reasoning) are central to mathematics learning and may therefore predict differential response to mathematics intervention (Geary, 2004; Powell et al., 2017).

### ***Domain-Specific Predictors***

Domain-specific knowledge and skills support successful learning within a particular academic domain. Several longitudinal studies examining predictors of mathematics achievement have found that the impact of prior math performance on later math achievement increases throughout formal schooling (e.g., Geary et al., 2017; Lee & Bull, 2016). This suggests that, whereas other factors, including domain-general abilities, may be stronger predictors of achievement in the early grades, foundational mathematics knowledge is a critical prerequisite for later math success. Pre-intervention skills in an academic domain may also moderate intervention response. Two opposing hypotheses about the relationship between initial domain-specific skill and intervention response have been proposed in the literature. Fuchs, Sterba et al., (2016) outline the initial academic severity hypothesis, which poses that students with greater initial skill deficits (or lower pre-intervention performance) in a given domain are expected to benefit less from intervention due to more severe learning difficulties in that domain. On the other hand, Clarke et al., (2020) argued that students with greater initial skill deficits may benefit *more* from intervention than peers with higher pre-intervention performance when provided intensive intervention supports aligned with their learning needs. When initial skill does not moderate the effects of a validated intervention program, this is thought to support the robustness of the intervention. In the context of an effective intervention program, a lack of moderation indicates that treatment outcomes are comparably positive for all students receiving the intervention, regardless of their initial skill in the targeted domain (e.g., Fuchs, Sterba et al., 2016).

While research examining initial skill as a moderator of response to mathematics intervention has been somewhat limited compared to similar analyses in the reading literature (Lam & McMaster, 2014), divergent findings have emerged regarding the role of pre-intervention skill in predicting outcomes of mathematics intervention. Several studies have found comparable treatment effects for at-risk students across the range of initial skill level (L. S. Fuchs, Sterba et al., 2016; L.S. Fuchs et al., 2019), whereas others have documented differential benefits favoring students with lower (Clarke et al., 2019) or higher (Toll & Van Luit, 2013) initial skill. In a randomized controlled trial of a kindergarten mathematics intervention, Toll and Van Luit (2013) found that intervention was effective for students with early numeracy skills that fell between the 25th and 50th percentiles prior to intervention. However, the intervention had no significant impact for students with pre-intervention early numeracy skills below the 25th percentile. In contrast, Clarke et al., (2019) found that students with lower pre-intervention mathematics skill benefited more from a kindergarten mathematics intervention than their peers with higher initial skill. L. S. Fuchs et al., (2019), on the other hand, found that initial mathematics skill did not moderate response to a first-grade mathematics computation intervention, with at-risk students across the range of initial skill demonstrating comparable gains. Given these contrasting findings, further research is needed to better understand the relation between initial skill and differential response to mathematics intervention.

### ***Domain-General Predictors***

Domain-general abilities and skills support learning across many academic domains. Prior research indicates that several domain-general cognitive abilities play an important role in mathematics learning, including attention, processing speed, language skills, reasoning, and working memory (Geary, 2004; Powell et al., 2017). In a large longitudinal study of the influence of domain-general and domain-specific skills on mathematics achievement across grades 1-8, Geary and colleagues (2017) found that general intelligence, working memory, and reading achievement accounted for substantial variance in mathematics achievement. The effect of these domain-general abilities on later mathematics achievement was stable across grades. Additionally, difficulties with conceptual and procedural knowledge specific to mathematics, the hallmark symptoms of MD, are thought to result from underlying cognitive deficits, especially related to the central executive component of working memory (Geary, 2004). The literature suggests characteristic difficulties of MD, such as representing and manipulating mathematical information, remembering mathematical concepts and facts, and controlling attention to accurately use mathematical procedures, may be accounted for by these cognitive deficits (Geary, 2004).

The following sections summarize the literature on the role of three domain-general abilities that have been linked to mathematics learning—working memory, fluid reasoning, and visual-spatial skills—as well as research exploring these cognitive skills as moderators of response to mathematics intervention.

**Working Memory.** Working memory is the ability to maintain and manipulate mental representations of information to accomplish a task or goal. It is composed of two key systems, driven by attention control: the phonological loop, which represents linguistic information, and the visual-spatial sketchpad, which represents nonverbal information (Geary, 2011). Working memory is thought to support various mathematics skills, including computation and word problem solving, by facilitating retrieval of learned information from long-term memory and supporting the ability to sustain and manipulate representations of information critical to problem solving and learning tasks (Welsh et al., 2010). The role of working memory in mathematics learning has been studied more than any other domain-general predictor. In particular, the updating component of the working memory central executive, which allows individuals to process incoming information while holding previous information in mind, has emerged as a consistent predictor of math achievement (Geary, 2011; Geary et al., 2017).

In a longitudinal study examining the relative impact of various domain-general and domain-specific predictors, the central executive emerged as the strongest domain-general predictor of subsequent mathematics achievement, with a consistent impact on mathematics outcomes across grades 1-8 (Geary et al., 2017). Other longitudinal studies have also supported working memory as a strong, stable predictor of mathematics achievement over time (e.g., Geary, 2011; Lee & Bull, 2016). Van de Weijer-Bergsma et al., (2015) found that, while the overall relation between working memory and mathematics achievement remained stable over time, the impact of verbal working memory increased across grades 2-6, whereas the visual-spatial component of working memory became a weaker predictor of outcomes after grade 4. Welsh and colleagues (2010) found that working memory uniquely predicted end-of-kindergarten mathematics achievement, controlling for growth in domain-specific math skills during the kindergarten year, suggesting that this cognitive skill accounts for variability in math performance not accounted for by initial math skill.

Collectively, these findings suggest that working memory plays an important role in a range of mathematics learning processes, influencing outcomes related to word problem solving (Fuchs et al., 2010; Swanson, 2006; Zheng et al., 2011), calculation skills (L. S. Fuchs, Geary, et al., 2016), and broad mathematics achievement (Geary, 2011; Geary et al., 2017; Van de Weijer-Bergsma et al., 2015). Given these findings, it is perhaps unsurprising that working memory has emerged as a moderator of response to mathematics intervention in several recent studies. Powell et al., (2017) found that working memory moderated response to a supplemental, second grade mathematics intervention targeting calculation skills, such that students with stronger working memory skills demonstrated greater gains compared to those with lower working memory skills. Similarly, Fuchs et al., (2014) found that individual differences in working memory moderated response to a fraction intervention. Students with weaker working memory made greater gains when activities focused on development of fractions concepts were included in the intervention, whereas those with higher working memory benefitted more from the inclusion of activities targeting fluency.

**Fluid Reasoning.** Fluid reasoning, or fluid intelligence, encompasses the ability to identify patterns and relationships and apply logic to solve novel problems (Wechsler, 2012). Fluid reasoning skills are closely related to outcomes across a range of academic domains. In mathematics, fluid reasoning is thought to support key skills including development of schemas related to word problem solving and strategy selection and use (Floyd et al., 2003). Additionally, fluid reasoning and math problem solving are thought to share underlying cognitive processes. In particular, relational reasoning skills that support the ability to represent and consider multiple relationships between pieces of information at once may be critical to both math performance and general fluid reasoning skills (Green et al., 2017).

In several studies using large, nationally representative samples, fluid reasoning has emerged as a consistent predictor of mathematics performance in children and adolescents. Taub et al., (2008) used structural equation modeling techniques to examine relations between broad cognitive ability clusters and mathematics achievement using data from the Woodcock-Johnson III standardization sample. The authors found that fluid reasoning was a statistically significant, direct predictor of mathematics performance across age groups ranging from five to 19 years. Similarly, Floyd et al., (2003) found that fluid reasoning abilities moderately to strongly predicted performance on standardized, norm-referenced mathematics measures for individuals ages six to 19. Green et al., (2017) assessed the relative contributions of various cognitive ability clusters to future mathematics performance from age six to 21. They found that fluid reasoning was a stronger predictor of the following year's math achievement than visual-spatial skills, verbal reasoning, and mathematical reasoning, supporting the critical role of fluid reasoning skills in math development.

Fluid reasoning skills have also been linked to difficulties with mathematical reasoning (Proctor et al., 2005), suggesting that poor fluid reasoning skills may be a common cognitive weakness among children with MD who struggle with tasks requiring math reasoning (e.g., word problem solving). Fuchs and colleagues (2010) examined the relative impact of domain-general and domain-specific abilities on individual differences in early mathematics skill, including solving number combinations and word problems. The authors found that performance on measures of reasoning skills, including concept formation and nonverbal problem solving, uniquely predicted individual differences in development of word problem solving skills. These findings aligned with those of a previous study by the same author team in which reasoning skill was found to significantly predict word problem solving performance (L. S. Fuchs et al., 2006).

Overall, the literature suggests that fluid reasoning underlies skills critical to development of mathematics proficiency. However, research examining fluid reasoning as a moderator of response to mathematics intervention has yielded inconsistent results. Powell et al., (2017) found that fluid reasoning did not moderate response to two supplemental, second grade mathematics interventions targeting calculation skills and word problem solving. On the other hand, L. S. Fuchs, Malone, et al., (2016) found that individual differences in fluid reasoning skills significantly moderated response to a fraction intervention component consisting of schema-based word problem solving instruction. Students with stronger fluid reasoning skills benefitted more from instruction in fraction word problem solving compared to peers with less developed reasoning abilities, suggesting that students with low reasoning skills may require additional support to benefit from instruction in fraction word problem schemas.

**Visual-Spatial Skills.** Visual-spatial skills underlie the ability to generate, recall, represent, and manipulate non-verbal, symbolic information. They include skills such as perception of spatial information and relationships and manipulation of spatial information, including mental rotation of figures (Cornu et al., 2017). Visual-spatial skills are more closely related to mathematics development than other domain-general abilities, and as such, some researchers have argued that skills in this domain may be better classified as domain-specific in the mathematics literature (e.g., Mix & Cheng, 2012). However, the non-numerical nature of visual-spatial skills lends support for their classification as domain-general abilities (Cornu et al., 2017). As with fluid reasoning, visual-spatial skills are thought to share underlying cognitive processes with a range of mathematics tasks. For example, visual-spatial skills may support the ability to create and manipulate mental representations of computation and word problems (Cornu et al., 2017). Across the literature, visual-spatial skills are consistently associated with mathematics learning and development beginning in early childhood (Cornu et al., 2017; Mix & Cheng, 2012), supporting the plausibility of visual-spatial skills as a moderator of response to mathematics intervention.

In summary, a handful of studies have supported an interaction between domain-general factors and treatment response, suggesting that students with stronger working memory and fluid reasoning skills may respond more positively to mathematics intervention than peers with lower performance in these cognitive domains (L. S. Fuchs et al., 2014; L. S. Fuchs, Malone, et al., 2016; Powell et al., 2017). However, there is a need for further research examining a range of general cognitive skills as moderators of intervention response, especially among young students with or at risk for MD. Critically, to further the field's understanding of the impact of validated mathematics intervention programs, domain-general skills with strong links to mathematics development, including working memory, fluid reasoning, and visual-spatial skills, should be investigated as potential predictors of response patterns.

### ***Predictors of Response Profiles***

In their analysis of typology of response to a validated early literacy intervention, Peng et al., (2020) examined a range of domain-specific and domain-general variables as potential predictors of response group membership. They found that domain-specific skills, including letter knowledge and passage comprehension, differentiated between students who responded adequately to intervention, based on membership in the strongly or mildly responsive group, versus students who did not respond adequately. Domain-specific skills also differentiated between students belonging in the strongly non-responsive versus the mildly non-responsive group, suggesting that pre-intervention literacy skills impact likely outcomes of early literacy intervention in the short- and long-term. On the other hand, domain-general skills, including working memory and non-verbal (fluid) reasoning, did not reliably differentiate between students belonging in the various response profiles. Given that the literature on early mathematics and literacy development suggest differential relationships between domain-general abilities, domain-specific skills, achievement, and response to intervention (Peng et al., 2020), it is likely that a different set of predictors differentiates between response groups associated with evidence-based early mathematics intervention.

## **Present Study**

RTI provides a promising framework for supporting students struggling in mathematics and increasing rates of math proficiency in U.S. schools. Intervention researchers have made significant progress toward developing and validating mathematics interventions designed for implementation within multi-tiered service delivery models, with a growing number of programs demonstrating positive outcomes for students at risk for MD. Meaningful categorization of responsiveness is central to effective instructional decision-making within an RTI framework. In both research and practice, the most common approach to categorizing treatment response creates a response/non-response binary. However, there are several important limitations of this approach, including a lack of consensus about how to define adequate response, low reliability near the cut point, and overemphasis on performance immediately after intervention completion at the expense of considering longer-term outcomes (Peng et al., 2020). Furthermore, recent findings by Peng and colleagues (2020) suggest that the traditional response/non-response binary may oversimplify more complex patterns of intervention response, whereas allowing for additional response categories may provide information about distinct subgroups of students who may require different instructional approaches following intervention completion. Given the implications of this work for implementation of multi-tiered service-delivery models in schools, further research exploring alternate methods of categorizing intervention response across academic domains and programs is needed.

Furthermore, a substantial minority of students with or at risk for MD do not respond adequately to generally effective mathematics intervention programs. Understanding variability in intervention response and how to effectively meet the needs of students who do not respond adequately is critical to successful implementation of RTI models in schools. Given recent calls in the field to understand how individual and contextual factors impact treatment response (D. Fuchs & Fuchs, 2019; Miller et al., 2014), there is a need for research examining student-level cognitive, academic, and behavioral characteristics that predict differential response to intervention. To date, few studies have examined domain-general moderators of response to early mathematics intervention (Powell et al., 2017; Shanley et al., 2021), and studies examining math-specific knowledge and skills as predictors of responsiveness have yielded inconsistent results.

The present study seeks to advance the knowledge base supporting RTI in early mathematics by exploring patterns of response to an evidence-based early mathematics intervention. ROOTS (Clarke et al., 2016a) is a Tier 2 kindergarten mathematics intervention program designed to build conceptual understanding of and procedural fluency with whole numbers. In previous studies, ROOTS has demonstrated effectiveness in improving outcomes for students with MD, with small to moderate effect sizes meeting What Works Clearinghouse criteria for educationally meaningful impacts (Clarke et al., 2016a; 2016b; 2017, 2019; 2020; Doabler et al., 2019). Across all cohorts of the efficacy project, students who participated in the ROOTS intervention demonstrated significantly greater gains from the fall to the spring of kindergarten compared to peers in the control condition on a range of proximal and distal measures of mathematics proficiency (Hedge's *g* effect sizes range from .18 - .81; Clarke et al., 2020). Differences between conditions did not persist at a first-grade follow-up.

While results have supported ROOTS as an effective program associated with improved math outcomes for students at risk for MD, as with any intervention program, ROOTS outcomes vary across students and contexts. In addition to examining main effects, the ROOTS project team has conducted various analyses to better understand treatment effects and response variability. To manipulate treatment intensity, students assigned to the treatment condition received ROOTS in small groups of either two or five at-risk students. In general, the data did not support differential outcomes by ROOTS group size (Clarke et al., 2020; Doabler et al., 2019). Across study cohorts, findings have consistently supported pre-intervention mathematics skill as a moderator of response to the ROOTS intervention, such that students with lower initial mathematics skill benefit more from ROOTS than their at-risk peers with higher initial skill (Clarke et al., 2019; 2020). Domain-general cognitive skills did not significantly moderate impact of ROOTS; however, overall patterns suggested that phonological working memory, fluid reasoning, and visual spatial skills may predict differential response to ROOTS with sufficient power to detect more subtle effects (Shanley et al., 2021).

The present study is a conceptual replication and extension of the approach utilized by Peng et al., (2020), expanding upon previous work through secondary analyses of data from the ROOTS Efficacy Project (Clarke et al., 2012). This study was intended to build upon the findings of Peng and colleagues by evaluating optimal response categorization with a larger sample, in a different academic domain, and within the context of a different evidence-based intervention in early mathematics. This study also included different outcome measures and potential predictors of intervention response theoretically linked to early mathematics outcomes. Alternate conceptualizations of intervention response have not previously been examined in the context of an early mathematics intervention. Given the limited prior research in this area, this study was intended to be exploratory in nature and addressed the following research questions:

**1. Is student response to a kindergarten mathematics intervention best categorized by a dichotomous, response/non-response framework, or are there more than two distinct response patterns, considering both immediate and long-term outcomes?**

Considering the findings of Peng et al., (2020) and the numerous limitations associated with the response/non-response binary, it was hypothesized that results would support a more complex response categorization framework consisting of several distinct response patterns. However, these patterns were expected to differ from the response profiles identified by Peng and colleagues in several ways. First, given the pervasiveness of fadeout effects in the math intervention literature (Bailey et al., 2020), it was expected that one or more response groups would demonstrate patterns characterized by post-intervention gains that diminished to some degree by first grade follow-up. Second, with respect to the growing body of research indicating that response to the ROOTS intervention is moderated by initial mathematics skill such that students who begin intervention with lower math skill demonstrate stronger response (Clarke et al., 2019; 2020), it was expected that one or more response profiles characterized by low pretest performance and strong growth from pretest to posttest would be identified.

**2. Do student-level domain-general and domain-specific skills differentiate between patterns of response to intervention? If so, which specific skills predict response patterns indicating positive treatment outcomes, and which predict inadequate response to intervention?**

Considering the results of prior research examining moderators of response to the ROOTS intervention, it was hypothesized that both domain-general cognitive skills and domain-specific mathematics skills would predict response profile membership. As discussed above, past findings consistently indicated that initial mathematics skill moderates ROOTS intervention response, so it was expected that pre-intervention mathematics skill would be a significant predictor of response profile membership, distinguishing between students demonstrating two or more distinct response trajectories. While a previous study found that cognitive skills did not significantly moderate response to the ROOTS intervention (Shanley et al., 2021), closer examination of data patterns suggests that working memory, fluid reasoning, and visual-spatial skills may predict differential response given adequate power to detect smaller effects. It was therefore hypothesized that the analysis methods used in the present study would reveal one or more cognitive variables as small but significant predictors of differential response patterns.

# CHAPTER II

# METHODS

This study analyzed a subset of data from the federally-funded ROOTS Efficacy Project (Clarke et al., 2012). A partially nested, randomized controlled trial was used to evaluate the impact of ROOTS on kindergarten mathematics outcomes across three school years (2012-2015) and four cohorts of students in Oregon and Massachusetts. The participants and intervention procedures for the full parent study were previously reported by Clarke et al., (2020). This study utilized data from all cohorts to address the first research question, and cohorts 2 and 3 to address the second research question.

## Participants

### *Districts and Schools*

Twenty-three elementary schools from school districts in Oregon ($n$ = 4) and Massachusetts ($n$ = 2) participated in the study. Massachusetts districts were located in the Boston metropolitan area; Oregon districts were located in rural/suburban western Oregon and in the Portland area. Enrollment across districts ranged from 2,736 to 39,002 students. Students' racial/ethnic backgrounds were reported by each school as follows: 0-12% of students were American Indian or Native Alaskan, 0-16% were Asian, 0-16% were Black, 0-83% were Hispanic, 0-2% were Native Hawaiian or Pacific Islander, 9-92% were White, and 0-15% were more than one race. Additionally, schools reported that 8-25% of students received special education services, 0-69% of students spoke a first language other than English, and 17-87% of students were eligible for free or reduced-price lunch.

### ***Classrooms and Teachers***

Across schools, 138 kindergarten classes were recruited for study participation. Over half of the classes (57%) were half-day kindergarten programs. Classes were taught by 75 certified kindergarten teachers, many of whom participated in the project across two consecutive years. Among participating teachers, 70 provided demographic information. All participating teachers identified as female, and the majority identified as White (88.6%). In terms of experience and qualifications, teachers had an average of 15.2. years of teaching experience ($SD$ = 9.1), and the majority had a master's degree in education (78.6%) and had taken a college or graduate level algebra course (58.6%).

### ***Students***

All students with parental consent ($n$ = 3,130) were screened for study eligibility during late fall of their kindergarten year. The screening process consisted of two standardized, individually administered measures of early mathematics proficiency, Assessing Student Proficiency in Early Number Sense (ASPENS; Clarke et al., 2011) and the Number Sense Brief (NSB; Jordan et al., 2008), described below. Students were considered at risk for MD, and therefore eligible for the ROOTS intervention, if their ASPENS composite score placed them in the *Strategic* (scores ranging from 17-59) or *Intensive* (scores less than 17) range, and their NSB score was 20 or lower. After screening was completed, the project's independent evaluator converted eligible students' ASPENS and NSB scores into standard scores and combined these into an overall composite score for each student. Within each classroom, the 10 eligible students with the lowest composite scores were selected for study participation and randomly assigned to one of two intervention conditions (ROOTS groups of two or five students) or a business-as-usual control condition. For the 33 classes that did not have at least 10 students who met eligibility criteria for the ROOTS intervention, the independent evaluator combined eligible students from multiple classes to facilitate random assignment. These procedures resulted in a sample of 1,251 at-risk students assigned to either ROOTS small-group intervention (*n* = 880) or no-treatment control (*n* = 371). For the purposes of this study, smaller and larger ROOTS groups (comprised of two and five students, respectively) were considered as one treatment condition, given that prior research has generally supported comparable outcomes regardless of group size (Clarke et al., 2020; Doabler et al., 2019). The mean age of participants across conditions was 5.3 years (*SD* = .4). District-provided demographic characteristics for the final sample are reported in Table 1.

**Table 1**

*Descriptive Statistics for Student Demographic Characteristics by Study Condition*

| | ROOTS<br>*n* (%) | Control<br>*n* (%) | Total<br>*n* (%) |
|---|---|---|---|
| Total sample *(n)* | 880 | 371 | 1,251 |
| Sex: | | | |
| Male | 425 (49) | 184 (50) | 609 (49) |
| Female | 441 (51) | 186 (50) | 627 (51) |
| Race: | | | |
| White | 500 (64) | 214 (64) | 714 (64) |
| Asian | 20 (3) | 12 (4) | 32 (3) |
| Black | 45 (6) | 15 (4) | 60 (5) |
| Hispanic | 185 (24) | 79 (24) | 264 (24) |
| Hawaiian/ Pacific Islander | 3 (0) | 0 (0) | 3 (0) |
| American Indian/ Alaskan Native | 11 (1) | 7 (2) | 18 (2) |
| Multiracial | 19 (2) | 7 (2) | 26 (2) |
| First Language: Spanish | 258 (34) | 117 (36) | 375 (35) |

*Note.* Sample *n*'s for each variable do not sum to the total sample *n* reported in the top row due to missing responses. Percentages represent the proportion of students for whom each variable was reported. Race and ethnicity categories were mutually exclusive.

### *ROOTS Interventionists*

ROOTS intervention groups were taught by employees of participating schools, as well as interventionists hired by the research team for the trial. The majority of ROOTS interventionists identified as female (93.5%) and White (76.1%). Interventionists had an average of 10.4 years of teaching experience (*SD* = 8.6). Over half had completed a bachelor's degree (60.5%), and 56.5% had taken a college or graduate level algebra course. Nearly all ROOTS interventionists reported previous experience delivering instruction in a small group setting (92.3%), and 22.0% held a teaching license or certification.

## ROOTS

ROOTS is a Tier 2 kindergarten mathematics intervention targeting development of whole number knowledge. The program includes 50 scripted, 20-minute lessons designed for delivery in a small-group format (2-5 students per 1 interventionist), five days per week, for approximately 10 weeks. ROOTS was implemented as a supplement to Tier 1 instruction, with sessions scheduled at times that did not conflict with core mathematics instruction. Intervention groups ran from late fall to spring to provide students the opportunity to respond to core mathematics instruction prior to intervention and therefore minimize identification of typically achieving students.

Based on expert recommendations to focus early math instruction on critical whole number concepts and skills (e.g., NMAP, 2008; Gersten et al., 2009), ROOTS content emphasizes standards from the Counting and Cardinality, Operations and Algebraic Thinking, and Number and Operations in Base Ten strands of the CCSS-M (2010). The ROOTS scope and sequence is designed so that skills are developed across lessons. Each lesson consists of 4-6 brief activities providing practice on multiple skills and concepts, including daily warm-up and guided math practice activities that provide frequent cumulative review. ROOTS incorporates explicit and systematic instructional design features aligned with best practice guidelines for supporting students with MD, including extensive teacher modeling, guided practice, immediate and specific academic feedback, frequent opportunities to respond, and use of mathematical language to illustrate reasoning processes (Coyne et al., 2011; Fuchs et al., 2021). Lesson scripting promotes the use of precise, consistent mathematical language and high-quality practice opportunities. Throughout the program, a wide range of mathematical models are incorporated, following the concrete, semi-concrete, abstract sequence to support deep conceptual understanding of mathematics content (Fuchs et al., 2021). In this sequence, students first represent numbers and concepts using concrete manipulatives (e.g., counters, base ten blocks), then visual representations (e.g., number lines, tally marks). Students use abstract, numerical representations alone only after having developed understanding of their meaning through use of alternate representations.

**Professional Development**

To support high-quality implementation, interventionists participated in ROOTS training provided by research team staff. Training consisted of two, 5-hour professional development workshops intended to familiarize interventionists with the structure, content, and instructional delivery features of the ROOTS intervention. The initial workshop addressed content covered in the first half of ROOTS, as well as strategies for effective instructional delivery and behavior management (e.g., academic feedback, signaling student responses, lesson pacing, group expectations). The second workshop, delivered partway through program implementation, targeted content covered in the second half of ROOTS. During workshops, project staff provided models of effective lesson delivery and time for interventionists to practice strategies and receive feedback. Outside of formal workshops, interventionists received additional coaching support. Each interventionist completed 2-4 coaching visits, during which ROOTS coaches observed lesson delivery and provided feedback on quality of instruction and fidelity of intervention implementation.

**Implementation Fidelity**

Trained project staff assessed fidelity of ROOTS implementation via direct observation. Each intervention group was observed on three occasions during ROOTS implementation. Observers rated the extent to which interventionists met instructional objectives, followed instructor scripting, and used provided mathematical models for the observed lesson using a 4-point scale (4 = *all*, 3 = *most*, 2 = *some*, 1 = *none*). Observers also recorded the number of planned activities that were completed within the session. Observation data indicate ROOTS was implemented with a high degree of fidelity. ROOTS interventionists generally met instructional objectives ($M$ = 3.49, $SD$ = 0.69), followed instructor scripting ($M$ = 3.31, $SD$ = 0.75), used provided mathematical models ($M$ = 3.61, $SD$ = 0.64), and completed most planned activities for the session ($M$ = 4.14 out of 5 activities, $SD$ = 0.77) during observations. Intraclass correlation coefficients (ICCs) indicated substantial agreement across observers, with ICCs of .82 (number of activities completed), .70 (instructional objectives met), .75 (scripting followed), and .70 (provided mathematical models used; Landis & Koch, 1977).

**Measures**

Trained project staff administered four measures of mathematics achievement intended to assess domain-specific knowledge at pretest and posttest. One additional measure of broad math achievement was administered at posttest and delayed follow-up. Additionally, four subtests from cognitive batteries, intended to measure domain-general cognitive skills, were administered at pretest. Interrater reliability of 95% or higher was met for all assessments.

### *Domain-Specific Measures*

**Assessing Student Proficiency in Early Number Sense (ASPENS).** ASPENS (Clarke et al., 2011) is a set of three brief, standardized, curriculum-based measures of early numeracy skills used for universal screening and progress monitoring in early mathematics. ASPENS measures include Number Identification, Magnitude Comparison, and Missing Number identification. Test-retest reliabilities of kindergarten ASPENS measures are in the moderate to high range (.74 to .85). Predictive validity of ASPENS scores in the fall of kindergarten with spring scores on the TerraNova 3 ranges from .45-.52.

**Number Sense Brief (NSB).** NSB (Jordan et al., 2008) is an individually administered, standardized measure of early mathematics proficiency. It consists of 33 items that assess knowledge and skills related to counting, number recognition and comparison, nonverbal calculation, number combinations, and word problems. NSB has a coefficient alpha of .84 at the beginning of first grade.

**ROOTS Assessment of Early Numeracy Skills (RAENS).** RAENS (Doabler et al., 2012) is a researcher-developed, proximal measure of skills taught in the ROOTS intervention. RAENS is individually administered and untimed, consisting of 32 items that assess knowledge of counting and cardinality, number operations, and the base-10 number system. Items require students to count and compare groups of objects; write, order, and compare numbers; label visual representations; and solve single-digit addition problems. RAENS' predictive validity with commonly used math achievement measures (e.g., the Test of Early Mathematics Ability—Third

Edition; the NSB) ranges from .68-.83. High interrater agreement (100%) and internal consistency (Cronbach's alpha = .91) have been reported for RAENS (Clarke et al., 2016b).

**Test of Early Mathematics Ability—Third Edition (TEMA-3).** TEMA-3 (Ginsburg & Baroody, 2003) is a norm-referenced, individually administered measure of early mathematics proficiency for children ages 3-8 years. It assesses whole number knowledge, including skills in counting and calculation. The TEMA-3 authors report test-retest and alternate-form reliabilities at or above .93. Concurrent validity with other mathematics achievement measures ranges from .54-.91.

**Stanford Early School Achievement Test—Tenth Edition (SESAT)**. SESAT (Harcourt Educational Measurement, 2002) is a standardized, norm-referenced test of mathematics ability. It is group administered and includes two subtests: Problem Solving and Procedures. Scores are typically reported at the overall and subtest level, but detailed reports of performance on specific skills are available. SESAT has adequate reliability ($r$ =.93) and validity ($r$ = .67). In the present study, the SESAT was administered at immediate and delayed posttest as a distal measure of intervention outcomes.

***Domain-General Measures***

**Wechsler Preschool and Primary Scale of Intelligence—Third Edition (WPPSI-III) Subtests: Matrix Reasoning and Block Design.** The WPPSI-III is a standardized, norm-referenced measure of cognitive skills for children ages 2 years, 6 months – 7 years, 7 months (Wechsler, 2012). The full WPPSI-III battery consists of 13 subtests; two subtests were administered in the present study. *Matrix Reasoning* was administered as a measure of fluid reasoning, conceptualized as the ability to apply logic and reasoning to solve novel problems (Wechsler, 2012). Matrix Reasoning is an untimed measure consisting of 26 items that require examinees to analyze incomplete matrixes and complete visual patterns by selecting from five response options. Test–retest reliability is reported at .86 and split-half reliability ranges from .88-.90 for 5 to 7-year-old examinees. Concurrent and predictive validity with other measures of fluid reasoning ranges from .48-.49. *Block Design* was administered as a measure of visual-spatial processing skills, including the ability to attend to and organize visual information and integrate it with motor functions (Wechsler, 2012). Block Design consists of 17 items that require examinees to use multi-colored blocks to replicate models within a set time limit. For initial items, the examinee uses blocks to replicate a two-dimensional image after the examiner first creates a three-dimensional model with another set of blocks; later items require examinees to use blocks to replicate two-dimensional designs without a three-dimensional model. Test-retest reliability is reported at .83 and split-half reliability ranges from .84-.86 for 5 to 7-year-old examinees. Concurrent and predictive validity with other measures of visual-spatial skills ranges from .51-.58.

**Comprehensive Test of Phonological Processing (CTOPP) Subtests: Memory for Digits and Nonword Repetition.** The CTOPP is a comprehensive, standardized assessment consisting of 13 subtests that measures skills in phonological awareness, phonological working memory, and rapid naming (Wagner et al., 1999). The Phonological Memory composite (consisting of the Memory for Digits and Nonword Repetition subtests), conceptualized as a measure of the phonological loop component of working memory, was administered for the present study. The *Memory for Digits* subtest consists of 21 items that require examinees to listen to a series of 2-8 numbers presented via audio recording and repeat the numbers in the same order. Items are administered until the examinee responds incorrectly on three consecutive items. Internal consistency for CTOPP Memory for Digits ranges from .78-.81, and test-retest reliability is .74 for 5 to 7-year-old examinees. Concurrent and predictive validity with other measures of phonological and reading skills ranges from .31-.49. The *Nonword Repetition* subtest consists of 18 items requiring examinees to listen to nonsense words presented via audio recording and repeat them orally. Nonsense words range from 3-15 sounds in length. Items are administered until the examinee responds incorrectly on three consecutive items. Internal consistency for CTOPP Nonword Repetition is .80, and test-retest reliability is .68 for 5 to 7-year-old examinees. Concurrent and predictive validity with other measures of phonological and reading skills ranges from .19-.41. Performance on nonsense word repetition tasks is correlated with recognition and recall of phonological information (Adlof & Patten, 2017) and other measures of phonological awareness skills (Clark et al., 2012).

**Digit Span Backward.** A Digit Span Backward task was administered as a measure of the central executive or updating component of working memory. The measure used in the present study was designed by the research team following standard administration procedures for digit span recall tasks. The task required examinees to listen to a series of numbers of varying lengths presented orally by the examiner, then repeat the numbers in the reverse order.

## Statistical Analysis

### *ROOTS Intervention Response Profiles*

To address the first research question, latent profile analysis (LPA) methods similar to those used by Peng et al., (2020) were used to determine the number of response types that best fit student mathematics scores across three time points: pre-intervention (i.e., fall of kindergarten), post-intervention (i.e., spring of kindergarten), and long-term follow up (i.e., spring of first grade). As a preliminary step, *z*-score transformations were performed for each mathematics performance measure using the mean and standard deviation of the control group for that measure at each time point, such that a score of 0 represents performance comparable to the control group average, a positive *z*-score represents stronger performance than students in the control group, and a negative *z*-score represents weaker performance than students in the control group. Transforming student scores in relation to the control group was intended to focus the analyses on gains of the intervention group in comparison to expected performance in the absence of intervention. A confirmatory factor analysis was then conducted to create latent mathematics scores for each student assigned to the ROOTS condition. Measures representing outcomes of interest were used to create latent mathematics performance variables at pretest (RAENS and TEMA-3) and posttest (RAENS, TEMA-3, and SESAT). The selected measures range from proximal (RAENS) to distal (TEMA-3 and SESAT). The confirmatory factor analysis was conducted using Mplus statistical software version 8.7 (Muthén & Muthén, 1998-2021). Fit indices for the measurement model were examined to ensure acceptable fit before using latent scores in subsequent analyses.

As described by Peng and colleagues, LPA is a modeling technique used to identify mutually exclusive classes based on some unobserved, categorical variable (in this case, intervention response class). Beginning with a two-profile model, increasingly more complex models are tested until model fit does not improve significantly by allowing for an additional response profile. To compare latent profile models, the Bayesian Information Criterion (BIC), Lo-Mendel-Rubin Likelihood Ratio Test (LMR-LRT), entropy, and posterior probability fit statistics were examined. The BIC (Schwarz, 1978) is an index of model fit that can be used to compare between two alternative models. Smaller BIC values indicate better model fit, and larger changes in BIC between competing models indicate greater differences in model fit. A sample-size adjusted BIC, intended to account for bias due to smaller sample sizes, was also reported for comparison. The LMR-LRT directly tests the fit of a model with *k* profiles compared to a model with *k* – 1 profiles; a significant value indicates that allowing for an additional profile results in a model that fits the data significantly better (Yungtai, Mendell, & Rubin, 2001). Entropy is an index that reflects how well profile membership in the model fits with the most likely profile membership for specific cases, with higher values indicating better model fit (Hart et al., 2016; Jung & Wickrama, 2008). Finally, average posterior probability for each profile was examined to assess the likelihood that individual students were placed in the profile with the highest probability of membership. LPAs were conducted using Mplus statistical software with adjusted standard errors (TYPE = COMPLEX) to account for the nested nature of the data. Full information maximum likelihood estimation procedures were used to reduce bias in results based on missing data. In the parent study data set, missing data largely involved students who were absent when assessments were conducted (e.g., due to illness or transferring schools) and are not believed to represent a meaningful violation of the missing-at-random assumption (Clarke et al., 2020).

***Predictors of Response Profile Membership***

To address the second research question, once the best fitting model had been selected, domain-general and domain-specific measures were examined as predictors of latent profile membership. Students were assigned to the profile with the highest probability of membership. Multinomial regression analyses were then used to examine the extent to which each domain-general and domain-specific measure predicted membership in the various response classes. One multinomial regression analysis was conducted to examine domain-specific skills as predictors of response class membership, including two measures of early mathematics proficiency – the NSB and ASPENS – that resemble typical school screening practices and were utilized as indicators of risk for mathematics difficulty in the parent study. A second multinomial regression analysis was run to examine domain-general skills as predictors of response class membership, using a subset of the sample for which these measures were collected. Regression analyses were conducted using SPSS version 28 statistical software (IBM Corp, 2021).

Assigning students to the profile with the highest probability of membership introduces additional measurement error that is not accounted for by the model because the probability that any given student belongs in a given response class is typically less than 1.0. In contrast, three-step modeling approaches introduce predictors of class membership while accounting for uncertainty associated with that membership (McIntosh, 2013). However, in regards to the parent study data set, the first approach confers the advantage of allowing for increased power to address the first research question by conducting LPA for the full sample across cohorts 1-4, while allowing for secondary exploration of domain-general predictors of response class membership that were only collected for a smaller subgroup of the parent study sample (cohorts 2 and 3).

# CHAPTER III

# RESULTS

## Descriptive Statistics

Inspection of histograms revealed that distributions of *z*-score transformed variables at each time point for students assigned to the ROOTS condition were approximately normal. Scores for the RAENS measure at posttest were somewhat negatively skewed, whereas scores for the ASPENS at pretest were somewhat positively skewed; however, skewness and kurtosis values all fell within the recommended range of –2 to 2. Several outlying data points exceeding three standard deviations below the mean were identified on the SESAT measure at follow-up; these data points were excluded from subsequent analyses to prevent disproportionate influence of a few cases on LPA results. Correlations among all variables are displayed in Table 2. Correlations between kindergarten math measures were significant and medium to large in magnitude (.36-.77), with weaker, small to medium correlations between kindergarten measures and the SESAT at first grade follow up (.28-.41.) Correlations between cognitive variables were generally weaker (.07-.37), and several were not significant. The majority of correlations between math measures and cognitive variables were significant, and small to medium in magnitude (.08-.45; Cohen, 1992).

## Latent Mathematics Performance Scores

Results of a confirmatory factor analysis were examined to ensure that the latent variable measurement model fit the data adequately. The RAENS and TEMA-3 were specified as measures of latent mathematics performance at pretest and the RAENS, TEMA-3, and SESAT as measures of latent mathematics performance at posttest. The root mean square error of approximation (RMSEA) was 0.12; values less than 0.1 are thought to indicate an acceptable fit.

**Table 2**

*Correlations Among Measures*

| Measures | 1. | 2. | 3. | 4. | 5. | 6. | 7. | 8. | 9. | 10. | 11. | 12. | 13. | 14. | 15. |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1. RAENS T1 | 1.00 | | | | | | | | | | | | | | |
| 2. RAENS T2 | .57** | 1.00 | | | | | | | | | | | | | |
| 3. ASPENS T1 | .68** | .47** | 1.00 | | | | | | | | | | | | |
| 4. ASPENS T2 | .54** | .69** | .52** | 1.00 | | | | | | | | | | | |
| 5. NSB T1 | .60** | .46** | .58** | .36** | 1.00 | | | | | | | | | | |
| 6. NSB T2 | .61** | .68** | .47** | .60** | .51** | 1.00 | | | | | | | | | |
| 7. TEMA T1 | .77** | .59** | .67** | .53** | .65** | .60** | 1.00 | | | | | | | | |
| 8. TEMA T2 | .67** | .74** | .60** | .71** | .54** | .70** | .73** | 1.00 | | | | | | | |
| 9. SESAT T2 | .63** | .59** | .54** | .54** | .57** | .65** | .64** | .65** | 1.00 | | | | | | |
| 10. SESAT T3 | .35** | .36** | .28** | .35** | .34** | .42** | .40** | .41** | .41** | 1.00 | | | | | |
| 11. Block Design | .19** | .17** | .18** | .18** | .15** | .14** | .15** | .12** | .20** | .08 | 1.00 | | | | |
| 12. Matrix Reasoning | .28** | .25** | .25** | .18** | .27** | .22** | .24** | .23** | .29** | .14* | .31** | 1.00 | | | |
| 13. Digit Memory | .40** | .34** | .35** | .21** | .38** | .38** | .45** | .41** | .40** | .25** | .11* | .20** | 1.00 | | |
| 14. Nonword Repetition | .20** | .16** | .19** | .15** | .20** | .20** | .20** | .27** | .19** | .12* | .08 | .12* | .34** | 1.00 | |
| 15. Digit Span | .43** | .28** | .38** | .27** | .41** | .41** | .44** | .44** | .39** | .18** | .07 | .10* | .37** | .33** | 1.00 |

*Note.* TI = Pretest (fall of kindergarten); T2 = posttest (spring of kindergarten); T3 = follow-up (spring of first grade). RAENS = ROOTS Assessment of Early Numeracy Skills; ASPENS = Assessing Student Proficiency in Early Number Sense; NSB = Number Sense Brief; TEMA = Test of Early Mathematics Ability—Third Edition; SESAT = Stanford Early School Achievement Test—Tenth Edition; Block Design = Wechsler Preschool and Primary Scale of Intelligence (WPPSI) Block Design subtest; Matrix Reasoning = WPPSI Matrix Reasoning subtest; Digit Memory = Comprehensive Test of Phonological Processing (CTOPP) Memory for Digits subtest; Nonword Repetition = CTOPP Nonword Repetition subtest; Digit Span = Digit Span Backward. $^*p < .05$. $^{**}p < .01$.

However, all other fit indices were within acceptable ranges. The standardized root mean square residual (SRMR) was 0.02; values less than .05 indicate good fit. The comparative fit index (CFI) was 0.98 and the Tucker Lewis index (TLI) was 0.95; values greater than 0.90 are thought to indicate good fit. Since three of the four indices examined indicated acceptable fit for the measurement model, and examination of latent mathematics performance variables better accounts for measurement error associated with any one observed variable, estimates of latent mathematics performance at pretest and posttest were utilized in subsequent analyses.

**ROOTS Intervention Response Profiles**

Typology of patterns of response to the ROOTS intervention was explored in a series of LPAs with two, three, four, and five profiles included in the model. Table 3 displays fit statistics for each of the latent profile models evaluated. BIC and ABIC decreased with each subsequent model, such that the smallest BIC value was obtained for the five-profile model (BIC = 3,509.44). The greatest entropy value was also obtained for the five-profile model (entropy = .87). However, the Lo-Mendell-Rubin Likelihood Ratio Test was statistically significant for the three-profile model ($p < .001$), but not for the four-profile ($p = .116$) or five-profile ($p = .220$) models. This indicates that the three-profile model fit the data significantly better than the two-profile model, whereas goodness-of-fit did not increase significantly for the four-profile model compared to the three-profile model, or the five-profile model compared to the four-profile model. Considering all fit indicators, the four- and five-profile models both had better overall fit compared to the two- or three-profile models. Conceptually, the four- and five-profile models were similar; however, the five-profile model included a small class ($n < 50$) consisting of students with the highest pre-intervention mathematics skill, who benefitted minimally from the ROOTS intervention (see Clarke et al., 2020). Given these results, the four-profile solution was selected as the final model to balance the goals of parsimony, interpretability of the latent class response patterns, and adequate class sizes for subsequent analyses examining predictors of class membership. Average probabilities for class membership in latent profiles 1-4 were .89, .94, .90, and .92, respectively.

**Table 3**

*Latent Profile Analysis Model Fit Statistics for the Intervention Group*

| Number of Profiles | BIC | ABIC | Entropy | LMR LRT |
|---|---|---|---|---|
| 2 | 4,407.36 | 4375.61 | .77 | -- |
| 3 | 3,924.66 | 3880.20 | .86 | $p$ <.001 |
| **4** | **3,675.99** | **3618.82** | **.84** | ***p* = .12** |
| 5 | 3,509.44 | 3439.57 | .87 | $p$ = .22 |

*Note.* BIC = Bayesian information criterion; ABIC = sample-size adjusted Bayesian information criterion; LMR LRT = Lo-Mendell-Rubin likelihood ratio test. The four-profile model in bold was selected based on optimal model fit for subsequent analyses.

Results of the final model with four latent profiles are presented in Table 4, including sample sizes and mean math performance across time points for each profile. Four profiles of responsiveness to the ROOTS intervention were distinguished: a moderate-risk, mildly responsive group (Profile 1, $n$ = 259, 29.8% of the sample), a high-risk, strongly responsive group (Profile 2, $n$ = 121, 13.9% of the sample); a moderate-risk, delayed response group (Profile 3, $n$ = 321, 36.9% of the sample); and a lower-risk, non-responsive group (Profile 4, $n$ = 168, 19.3% of the sample). The mean trajectories for each response profile from pretest to follow-up are shown in Figure 2, and trajectories for individual students within each response profile group across time points are plotted in Figure 3.

Levene's test of homogeneity of variances was significant for pretest latent math scores, posttest latent math scores, and z-scores for the SESAT at follow-up ($p$'s < .001), so comparisons among the four response profiles at each time point were made using the Games-

**Table 4**

*Math Performance for Each Profile Group at Pretest, Posttest, and Follow-Up*

| | Profile 1: Moderate-Risk, Mildly Responsive | Profile 2: High-Risk, Strongly Responsive | Profile 3: Moderate-Risk, Delayed Response | Profile 4: Lower-Risk, Nonresponsive |
|---|---|---|---|---|
| *n* (%) | 259 (29.8%) | 121 (13.9%) | 321 (36.9%) | 168 (19.3%) |
| Mean Latent Math Performance (SE) | | | | |
| Pretest | -0.44 (0.12) | -1.24 (0.08) | 0.25 (0.11) | 1.14 (0.09) |
| Posttest | -0.34 (0.10) | -1.09 (0.09) | 0.24 (0.08) | 0.88 (0.06) |
| Follow-Up | -0.55 (0.11) | -0.80 (0.11) | 0.38 (0.16) | 1.04 (0.11) |

**Figure 2**

*Mean Latent Math Scores by Response Profile and Time Point*

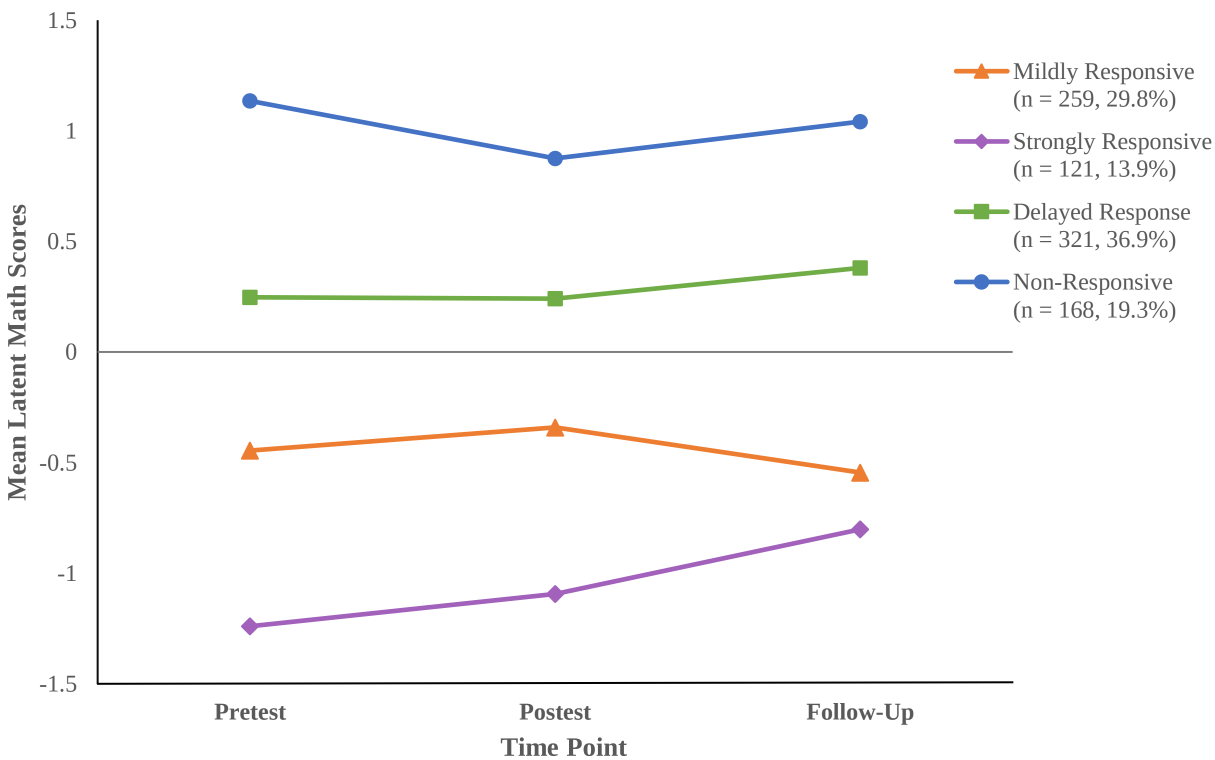
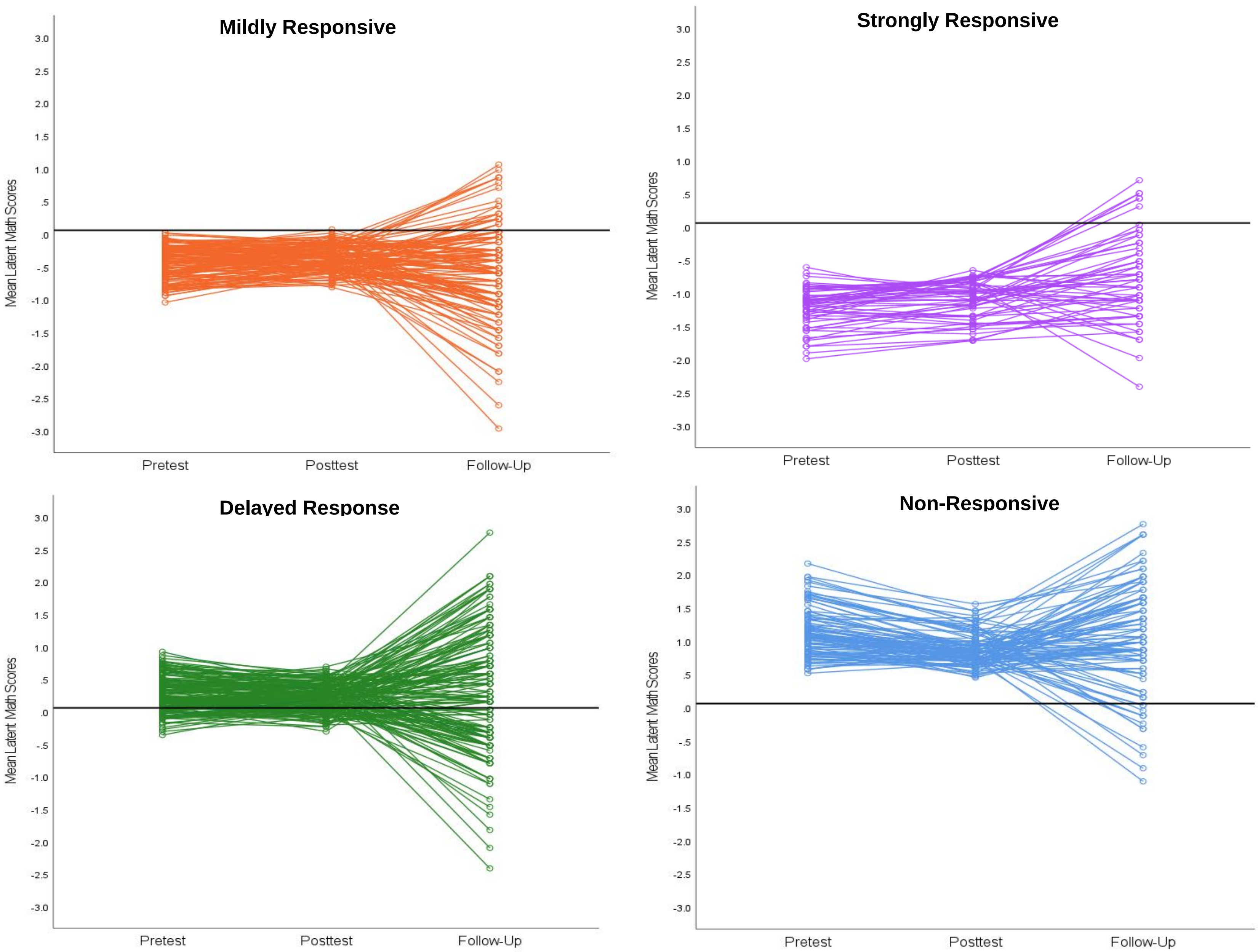



**Figure 3** (next page)

*Individual Math Performance Trajectories by Response Profile and Time Point*


Mildly Responsive
Strongly Responsive
Delayed Response
Non-Responsive
Mean Latent Math Scores
3.0
2.5
2.0
1.5
1.0
.5
.0
-.5
-1.0
-1.5
-2.0
-2.5
-3.0
Pretest
Posttest
Follow-Up

Howell test, which is robust to violations of the assumption of homogeneity of variance. Because these contrasts involved multiple statistical tests, the Benjamini-Hochberg correction procedure was applied to control the false discovery rate (Benjamini & Hochberg, 1995). *P* values were adjusted separately within the sets of analyses used to compare each pair of response profile groups across timepoints and predictors. Post-hoc comparisons revealed that students in the non-responsive group (Profile 4) demonstrated stronger math performance than students in all other profiles across timepoints, students in the delayed response group (Profile 3) displayed stronger math performance than students in the mildly responsive and strongly responsive groups (Profiles 1 and 2, respectively) across time points, and students in the mildly responsive group (Profile 1) displayed stronger math performance than those in the strongly responsive group (Profile 2) at pretest and posttest, with comparable performance to those in the strongly responsive group at follow-up.

To further contextualize the math performance levels of students in each response profile group over time, mean scores across measures of mathematics proficiency by response profile are presented in Table 5. On average, students in the mildly responsive group scored at the 9th percentile on the TEMA-3 at pretest and the 42nd percentile at posttest compared to peers in the TEMA normative sample. Students in the strongly responsive group scored at the 1st percentile at pretest and the 13th percentile at posttest; students in the delayed response group scored at the 25th percentile at pretest and the 63rd percentile at posttest; and students in the non-responsive group scored at the 55th percentile at pretest and the 84th percentile at posttest. It should be noted that these percentile ranks are based on the same age band (5 years, 6 months to 5 years, 8 months) on the TEMA-3 score tables at pretest and posttest due to the relatively short

**Table 5**

*Mean Pretest and Posttest Math Performance of Response Profile Groups Across Measures*

| Measure | Profile 1: Moderate-Risk, Mildly Responsive | | Profile 2: High-Risk, Strongly Responsive | | Profile 3: Moderate-Risk, Delayed Response | | Profile 4: Lower-Risk, Nonresponsive | |
|---|---|---|---|---|---|---|---|---|
| | Pretest | Posttest | Pretest | Posttest | Pretest | Posttest | Pretest | Posttest |
| NSB | 10.70 | 17.61 | 8.48 | 13.50 | 13.19 | 21.03 | 16.10 | 24.46 |
| | *(2.70)* | *(3.52)* | *(2.96)* | *(3.18)* | *(3.10)* | *(3.36)* | *(2.72)* | *(3.28)* |
| ASPENS | 13.57 | 75.09 | 5.01 | 41.22 | 26.44 | 95.11 | 40.68 | 117.40 |
| | *(11.98)* | *(27.51)* | *(7.73)* | *(26.81)* | *(15.12)* | *(26.99)* | *(12.95)* | *(24.23)* |
| TEMA | 13.53 | 22.80 | 6.93 | 14.92 | 18.94 | 29.30 | 26.37 | 35.69 |
| | *(3.39)* | *(3.62)* | *(3.54)* | *(3.95)* | *(3.50)* | *(3.74)* | *(4.55)* | *(3.69)* |
| RAENS | 8.25 | 22.38 | 4.63 | 14.02 | 12.81 | 26.14 | 18.92 | 28.78 |
| | *(2.97)* | *(3.97)* | *(2.47)* | *(5.39)* | *(3.49)* | *(2.96)* | *(4.08)* | *(2.05)* |
| SESAT | -- | 444.63 | -- | 419.39 | -- | 468.69 | -- | 501.79 |
| | | *(20.46)* | | *(26.93)* | | *(23.75)* | | *(25.59)* |

*Note*. RAENS = ROOTS Assessment of Early Numeracy Skills; ASPENS = Assessing Student Proficiency in Early Number Sense; NSB = Number Sense Brief; TEMA = Test of Early Mathematics Ability—Third Edition; SESAT = Stanford Early School Achievement Test—Tenth Edition.

intervention time frame and lack of consistently available testing dates in the parent study data set. While this age range roughly corresponds to the mean age of the sample, growth based on TEMA-3 percentiles may be somewhat artificially inflated as a result. Compared to peers in the normative sample for the SESAT, students in the mildly responsive group scored at the 13th percentile at posttest and the 7th percentile at follow-up; the strongly responsive group scored at the 6th percentile at posttest and the 5th percentile at follow-up, the delayed response group scored at the 26th percentile at posttest and the 19th percentile at follow-up, and the non-responsive group scored at the 47th percentile at posttest and the 31st percentile at follow-up.

## Predictors of Response Profile Membership

A series of exploratory analyses was conducted to explore whether individual differences in domain-general and domain-specific skills predicted response profile membership. Odds ratios are reported for each predictor to characterize the relative likelihood of belonging to each response profile compared to the reference group. Odds ratios greater than 1 indicate that the likelihood of membership in the target response profile compared to the reference group increases as the value of the predictor variable increases, whereas odds ratios less than 1 indicate that the likelihood of membership in the target response profile decreases as the value of the predictor variable increases. As described above, the Benjamini-Hochberg procedure was used to control the false discovery rate within sets of analyses comparing each pair of response profile groups across timepoints and predictors. Adjusted *p* values were used to interpret significance of results.

### *Domain-Specific Predictors*

Table 6 summarizes results of a multinomial logistic regression evaluating whether domain-specific variables predicted the likelihood of belonging to each response class across the entire intervention sample (cohorts 1-4). Pretest performance on both measures of early mathematics proficiency – ASPENS and NSB – significantly predicted response group membership. Specifically, the mildly responsive group demonstrated stronger pre-intervention early numeracy skills than the strongly responsive, group (Profile 1 vs. Profile 2) and lower initial early numeracy skills than the delayed response and non-responsive groups (Profile 1 vs. Profiles 3 and 4, respectively). The strongly responsive group showed lower initial early numeracy skills on both pretest measures than the delayed response and non-responsive groups (Profile 2 vs. Profiles 3 and 4, respectively). Finally, the delayed response group demonstrated lower pre-intervention math skills than the non-responsive group across measures (Profile 3 vs. Profile 4).

***Domain-General Predictors***

Table 7 summarizes results of a multinomial logistic regression evaluating whether domain-general variables predicted the likelihood of belonging to each response class for a subset of the intervention sample (cohorts 2 and 3). WPPSI Matrix Reasoning was examined as a measure of fluid reasoning, WPPSI Block Design as a measure of visual-spatial processing, CTOPP Memory for Digits and Nonword Repetition as measures of the phonological loop of working memory, and Digit Span Backward as a measure of the central executive or updating component of working memory. Pretest performance on Matrix Reasoning, Block Design, Memory for Digits, and Digit Span Backward significantly predicted response group membership (the Nonword Repetition measure did not significantly predict membership for any response profile). Specifically, the mildly responsive group demonstrated lower pre-intervention fluid reasoning skills, phonological working memory (Memory for Digits), and updating working memory (Digit Span Backward) than the delayed response and non-responsive groups

(Profile 1 vs. Profiles 3 and 4, respectively). The strongly responsive group demonstrated lower visual-spatial processing skills than the delayed response group (Profile 2 vs. Profile 3) and lower fluid reasoning and phonological/updating working memory skills than the delayed response and non-responsive groups (Profile 2 vs. Profiles 3 and 4, respectively). Finally, the delayed response group demonstrated lower pre-intervention phonological and updating working memory skills than the non-responsive group (Profile 3 vs. Profile 4).

**Table 6**

*Domain-Specific Predictors of Response Profile Membership*

| Variable | Profile 1 vs. 2<br>OR (*p*)<br>[95% CI] | Profile 1 vs. 3<br>OR (*p*)<br>[95% CI] | Profile 1 vs. 4<br>OR (*p*)<br>[95% CI] | Profile 2 vs. 3<br>OR (*p*)<br>[95% CI] | Profile 2 vs. 4<br>OR (*p*)<br>[95% CI] | Profile 3 vs. 4<br>OR (*p*)<br>[95% CI] |
|---|---|---|---|---|---|---|
| NSB | 0.82 (<.001)**<br>[0.75, 0.90] | 1.22 (<.001)**<br>[1.14, 1.31] | 1.64 (<.001)**<br>[1.48, 1.81] | 1.48 (<.001)**<br>[1.33, 1.65] | 1.99 (<.001)**<br>[1.75, 2.27] | 1.35 (<.001)**<br>[1.24, 1.46] |
| ASPENS | 0.91 (<.001)**<br>[0.88, 0.94] | 1.06 (<.001)**<br>[1.04, 1.07] | 1.11 (<.001)**<br>[1.09, 1.14] | 1.17 (<.001)**<br>[1.12, 1.21] | 1.23 (<.001)**<br>[1.18, 1.28] | 1.05 (<.001)**<br>[1.04, 1.07] |

*Note.* OR = odds ratio; 95% CI = 95% confidence interval. Profile 1 = mildly responsive; Profile 2 = strongly responsive; Profile 3 = delayed response; Profile 4 = non-responsive. The first class of each contrast is the reference category. Original p-values are reported; tests that were significant after Benjamini-Hochberg correction are denoted with asterisks.

* Adjusted $p < .05$. ** Adjusted $p < .01$.

**Table 7**

*Domain-General Predictors of Response Profile Membership*

| Variable | Profile 1 vs. 2<br>OR (*p*)<br>[95% CI] | Profile 1 vs. 3<br>OR (*p*)<br>[95% CI] | Profile 1 vs. 4<br>OR (*p*)<br>[95% CI] | Profile 2 vs. 3<br>OR (*p*)<br>[95% CI] | Profile 2 vs. 4<br>OR (*p*)<br>[95% CI] | Profile 3 vs. 4<br>OR (*p*)<br>[95% CI] |
|---|---|---|---|---|---|---|
| Block Design | 0.84 (.06)<br>[0.71, 1.01] | 1.05 (.45)<br>[0.93, 1.18] | 1.04 (.66)<br>[0.89, 1.21] | 1.24 (.02)*<br>[1.04, 1.48] | 1.23 (.05)<br>[1.00, 1.50] | 0.99 (.86)<br>[0.87, 1.13] |
| Matrix Reasoning | 0.96 (.57)<br>[0.83, 1.11] | 1.21 (<.001)**<br>[1.08, 1.36] | 1.22 (<.01)**<br>[1.06, 1.42] | 1.26 (<.01)**<br>[1.09, 1.46] | 1.28 (<.01)**<br>[1.07, 1.52] | 1.01 (.88)<br>[0.89, 1.15] |
| Digit Memory | 0.87 (.06)<br>[0.74, 1.01] | 1.23 (<.001)**<br>[1.09, 1.38] | 1.44 (<.001)**<br>[1.24, 1.68] | 1.42 (<.001)**<br>[1.21, 1.66] | 1.67 (<.001)**<br>[1.38, 2.01] | 1.18 (.01)*<br>[1.04, 1.34] |
| Nonword Repetition | 0.98 (.75)<br>[0.85, 1.13] | 0.98 (.79)<br>[0.88, 1.10] | 0.97 (.71)<br>[0.84, 1.12] | 1.01 (.93)<br>[0.87, 1.17] | 1.00 (.96)<br>[0.84, 1.18] | 0.99 (.85)<br>[0.87, 1.12] |
| Digit Span | 0.64 (.15)<br>[0.35, 1.17] | 1.42 (.03)*<br>[1.04, 1.95] | 2.64 (<.001)**<br>[1.87, 3.72] | 2.21 (<.01)*<br>[1.23, 3.97] | 4.10 (<.001)**<br>[2.24, 7.51] | 1.86 (<.001)**<br>[1.43, 2.40] |

*Note*. OR = odds ratio; 95% CI = 95% confidence interval. Profile 1 = mildly responsive; Profile 2 = strongly responsive; Profile 3 = delayed response; Profile 4 = non-responsive. The first class of each contrast is the reference category. Original p-values are reported; tests that were significant after Benjamini-Hochberg correction are denoted with asterisks.

* Adjusted $p < .05$. ** Adjusted $p < .01$.

# CHAPTER IV

# DISCUSSION

The present study explored patterns of response to a kindergarten mathematics intervention through secondary analyses of data from the ROOTS Efficacy Project (Clarke et al., 2012), a large-scale randomized controlled trial. The primary objective was to determine whether variability in response to a generally effective intervention was better captured by the conventional response/non-response binary or a more complex framework allowing for additional response profiles. A secondary goal was to explore domain-general and domain-specific skills as predictors of intervention response profiles. Designed as a conceptual replication and extension of research conducted by Peng and colleagues (2020), this study expanded upon previous work by categorizing response profiles in the context of an evidence-based early mathematics intervention with a larger sample of students at risk for MD. Given that understanding and categorizing variability in intervention response is critical to effective instructional decision-making within multi-tiered service-delivery models in schools, results of this study contribute to the literature base supporting implementation of RTI in early mathematics.

## ROOTS Intervention Response Profiles

Similar to the findings of Peng et al., (2020), results of a series of LPAs identified four distinct profiles of response to the ROOTS intervention. Three groups with variable pre-intervention risk for MD demonstrated some degree of positive intervention response: a moderate-risk, mildly responsive group; a moderate-risk, delayed response group, and a high-risk, strongly responsive group. In line with previous findings indicating that students with lower initial mathematics skill (or greater risk for MD) benefit more from ROOTS than those with higher initial skill (Clarke et al., 2019; 2020), a non-responsive group, characterized by lower risk for MD during the fall of kindergarten, was also identified.

As expected, patterns characterizing performance for each of the response profiles over time differed from those described by Peng et al., (2020) in several ways. Their results identified two responsive groups, comprising 58% of students in the treatment condition, which performed similarly (and comparably to the control group) at pretest, but outperformed the control group at posttest and follow-up to varying degrees. In contrast, results of the present study identified three distinct responsive groups, comprising 81% of students in the treatment condition, all of which differed significantly from one another in terms of pre-intervention mathematics performance. The strongly responsive and mildly responsive groups were characterized by mathematics performance below the mean of the control group across timepoints ($z$ scores of -1.24 and -0.44 at pretest, -1.09 and -0.34 at posttest, and -0.80 and -0.55 at follow-up, respectively). However, the strongly responsive group demonstrated substantial growth in mathematics performance from pretest to posttest and posttest to follow up, with overall growth of 0.44 standard deviation units, reflecting a narrowing of the achievement gap between these high-risk students and their lower-risk peers. The mildly responsive group also demonstrated growth in mathematics performance from pretest to posttest corresponding to a 0.1 standard deviation unit increase, however intervention effects were not maintained at first grade follow-up, consistent with the fade-out effect detected in the parent study (Clarke et al., 2020). On average, students in the mildly responsive group fell further behind from spring of kindergarten to spring of first grade.

The delayed response group, on the other hand, demonstrated math performance slightly above the mean of the control group at pretest and posttest ($z$ scores of 0.25 and 0.24), with no evidence of growth beyond that expected in the absence of intervention. However, students in this group experienced greater-than-expected growth from posttest to first grade follow-up, consistent with the delayed or “sleeper” effect hypothesized by Peng et al. (2020). The literature suggests that delayed intervention effects may occur when targeted skills require more time to master and generalize to novel tasks and situations (e.g, Barnett, 2011). The distal nature of the SESAT measure used at follow-up, which requires students to transfer acquired math knowledge and skills to problems unlike those encountered in the ROOTS intervention, further supports this interpretation as a delayed response trajectory for this group. As discussed in the Introduction section, Peng and colleagues suggest that their finding of relative stability in response/non-response across post-intervention timepoints may be explained in part by the early literacy skills targeted by the intervention, which are relatively easy to acquire and used broadly enough outside of intervention to support maintenance. The authors argued that learning in other domains may be more susceptible to temporal effects, such as the fade-out and sleeper effects identified in the present study. These results lend support to that theory.

Peng et al., (2020) also identified two non-responsive profiles: a mildly non-responsive group, which performed comparably to controls at pretest and posttest, but fell behind by follow-up; and a strongly non-responsive group, which demonstrated performance significantly below the control group mean at all time points. In contrast, results of the present study identified a single non-responsive group comprised of students with higher initial math skill, and therefore lower risk for MD. On average, students in this response class demonstrated pre-intervention math skills that exceeded the control group mean by more than 1 standard deviation (*z* score of 1.14). This group continued to outperform the control group substantially across postintervention timepoints, with *z* scores of 0.88 and 1.04 at posttest and first grade follow-up, respectively, and demonstrated growth from pretest to posttest across all measures of mathematics performance (see Table 5). This group is considered non-responsive due to the decrease in mean latent mathematics performance relative to the control group from pretest to posttest, suggesting that their growth did not keep pace with the mean growth of students in the control group. This decrease also reflects a narrowing of achievement gaps between the four response groups from fall to spring resulting from accelerated growth of subgroups of students who were initially at higher risk for MD: the gap between the highest and lowest performing response profile groups at pretest is 2.38 standard deviation units, whereas the gap at posttest is 1.97 standard deviation units. One interpretation of this response trajectory is that students in the nonresponsive group experienced meaningful growth across timepoints, but did not benefit from the ROOTS intervention to the same degree as their peers who entered intervention with lower math skill or greater risk for MD. These findings add to converging evidence suggesting that the ROOTS program is best aligned with the needs of students with significant skill deficits in mathematics (e.g., Clarke et al., 2019, 2020). Across studies, results for at-risk students with higher initial math skill are consistent with a ceiling effect, suggesting that the needs of these students may be better met by a less intensive intervention, or a program targeting different math concepts and skills.

**Predictors of Response Profile Membership**

Results of regression analyses indicate that performance on measures of domain-specific early mathematics skills, as well as measures of domain-general cognitive skills, significantly predicted response profile membership. Results were consistent across both domain-specific measures examined, such that the strongly responsive group demonstrated lower initial early numeracy skills than all other groups, the mildly responsive group had lower pre-intervention math skills than the delayed response and non-responsive groups, and the delayed response group showed lower initial math skills than the non-responsive group. These results are perhaps unsurprising in light of the considerable variability in pre-intervention mathematics proficiency between response profile groups. In contrast, Peng et al., (2020) found that subgroups of students who were responsive to intervention generally had higher scores on measures of domain-specific early literacy skills prior to intervention.

However, the present findings are consistent with and add further nuance to previous research documenting initial mathematics skill as a moderator of response to the ROOTS intervention, such that students with lower initial skill benefit more from ROOTS. It is important to note that the domain-specific predictor measures examined functioned as screening measures in the present study, used to identify students at risk for MD. Therefore, performance on measures used to screen for mathematics risk predicted patterns of response to the ROOTS intervention across key proximal and distal outcome measures. On average, students with the highest initial risk demonstrated the strongest response to ROOTS, and those scoring closer to the cut point for ROOTS eligibility demonstrated the weakest response. For moderate-risk students falling between these extremes, outcomes were more promising for those with higher initial mathematics skill, who tended to show a delayed, positive response, than for those with lower initial skill, who on average demonstrated a weaker initial response that faded out over time. While a delayed intervention response may be interpreted as a sleeper effect as discussed above, the distinct trajectories of these moderate risk groups may also indicate an educational Matthew Effect (Walberg & Tsai, 1983), in which students who began and ended intervention with higher mathematics performance relative to the average control student were better able to build upon the skills and knowledge they gained during the ROOTS intervention to benefit from subsequent mathematics instruction, despite showing a smaller response within the intervention window.

Whereas Peng and colleagues found that domain-general skills did not predict response profile membership, in the present study, fluid reasoning, visual spatial skills, and the phonological loop and updating components of working memory distinguished between various profiles of intervention responsiveness. The strongly responsive group demonstrated lower visual-spatial processing skills than the delayed response group and lower fluid reasoning and working memory performance than the delayed response and non-responsive groups prior to intervention. The mildly responsive group showed lower fluid reasoning and working memory skills at pretest compared to the delayed response and non-responsive groups, and the delayed response group scored lower on measures of working memory than the non-responsive group at pretest. Measures of working memory skills, including phonological memory for digits and updating, were the most consistent predictors of response profile membership, distinguishing between all but the strongly and mildly responsive groups.

Unexpectedly, domain-general cognitive skills predicted patterns of response to the ROOTS intervention in much the same way as domain-specific measures of risk for MD, such that students who scored lower on measures of fluid reasoning, visual spatial skills, and working memory demonstrated stronger response to the ROOTS intervention, on average. This suggests that, while subgroups of students who are or are not responsive to ROOTS are characterized by distinct cognitive skill profiles, pre-intervention cognitive performance measures do not predict response patterns as consistently as measures of early mathematics proficiency more commonly used to identify at-risk students in schools, nor do they provide additional information that further distinguishes between subgroups of students who respond adequately to ROOTS versus students who do not.

**Educational Implications**

In line with the conclusions drawn by Peng and colleagues (2020), a key takeaway of the present findings is that looking beyond the response/non-response binary may be instrumental in identifying appropriate next steps for subgroups of students at risk for MD. In this case, distinguishing between the responsive groups may be most informative, given the substantial variability in mean trajectories for various subgroups of students who demonstrated significant growth over the course of the ROOTS intervention. For example, the mean trajectory of the strongly responsive group is promising, representing meaningful growth in math skills and a substantial narrowing of achievement gaps with lower-risk peers. This pattern suggests that the ROOTS intervention aligned well with the educational needs of the strongly responsive group. However, these students' mathematics performance across postintervention timepoints remains low compared to that of other students at risk for MD. These students may respond positively to subsequent mathematics intervention, but may ultimately require more intensive, longer-duration supports to further accelerate their growth if they are to catch up to their lower-risk peers.

The mean trajectory of students in the delayed response group is also promising. Despite a relatively weak initial response compared to the control group, students in this group experienced meaningful gains in mathematics skills. At posttest, their mean performance on the TEMA-3 fell within the average range, with low-average performance on the SESAT. The delayed response group also demonstrated accelerated growth over the following school year. Collectively, performance indictors across postintervention time points suggest that students in the delayed response group experienced a reduction in overall risk for MD and were likely better able to access and benefit from core mathematics instruction over the course of the study, potentially mitigating the need for ongoing supplemental mathematics supports. On the other hand, the relatively weak initial response and fade-out effect demonstrated by the mildly responsive group may suggest a mismatch between student needs and the focus or intensity of the ROOTS intervention, a need for ongoing mathematics support to maintain gains and facilitate further growth, or both.

Identifying appropriate next steps for the lower-risk, non-responsive group is more nuanced than the standard response/non-response binary might suggest, as well. Whereas typically, a lack of response indicates the need to intensify intervention to better align with student needs, the existing body of research on the ROOTS intervention suggests this response pattern is more consistent with a ceiling effect. The trajectory and overall performance of the non-responsive group suggest a meaningful reduction in risk for MD, with mean performance on broad measures of mathematics skill across postintervention timepoints falling in the average range. The non-responsive group may not require further supports to benefit from core mathematics instruction and make progress toward grade level standards. This finding also has broader implications for service delivery at a systems level. Matching intensity of supports provided with intensity of student need is essential to optimize resource allocation and maximize overall impact within multi-tiered educational frameworks. Findings of the present study contribute to a growing body of evidence suggesting that a less intensive intervention may be sufficient to meet the needs of at-risk students with higher initial mathematics skill, whereas providing an intensive, whole-number focused intervention such as ROOTS to these students may not be the best use of schools' limited resources (e.g., Clarke et al., 2019; 2020; 2022; Sutherland et al., in press).

Finally, while exploratory, findings of the present study suggest that mathematics performance measures like those typically used to screen for risk for MD in schools can consistently predict how at-risk students are likely to respond to ROOTS, distinguishing between students who will display meaningfully different response trajectories over time. Knowledge of how students are likely to respond to a given intervention prior to implementation may ultimately support more effective instructional decision making, allowing educators to match students to interventions they are more likely respond to, or group at-risk students who are likely to respond similarly together and adapt intervention delivery in ways that are tailored to their unique needs. This study contributes to a growing literature base supporting the importance of examining how student-level factors predict patterns of response to generally effective academic interventions, and utilizing that knowledge to inform how finite school resources are allocated across the range of students at risk for academic difficulties (Lam & McMaster, 2014; Miller et al., 2014). While not the primary purpose of the study, the present findings further suggest that initial mathematics skill is the strongest and most consistent predictor of responsiveness to early mathematics intervention, suggesting that school screening data may be used not just to determine risk for MD, but also to inform instructional decisions around program placement, pacing, or grouping.

In contrast, the present findings do not provide strong support for incorporating measures of domain-general cognitive skills in school screening procedures. While cognitive measures did predict membership in the various response profiles, they did so less consistently than and largely followed the same pattern as domain-specific measures, such that students who scored lower on measures of mathematics performance *or* cognitive skills prior to intervention tended to demonstrate stronger response to the ROOTS intervention. It is therefore likely that administering cognitive measures as part of the screening or curriculum placement process would not yield enough additional information to merit the increased time and resources that would be required to do so. However, it is important to note that this conclusion is specific to the ROOTS intervention, which was designed to support the mathematics learning of a wide range of at-risk students. The literature has consistently documented strong associations between domain-general cognitive skills and mathematics learning (Geary, 2004; Powell et al., 2017); as a result, there have been calls in the field to consider both expanded screening batteries to identify students with cognitive weaknesses and modifications to existing programs to boost effectiveness by addressing these weaknesses (e.g., Fuchs et al., 2016; Powell et al., 2017). Yet, similar to the findings of Shanley et al., (2021), results of the present study suggest that the ROOTS intervention is accessible to and meets the needs of students at risk for MD with lower fluid reasoning, visual-spatial, and working memory skills. Therefore, cognitive screening measures may be of little use in identifying which students are likely to benefit from ROOTS, and modifications targeting cognitive skills may have little impact on its effectiveness. A promising implication is that the instructional design of mathematics intervention programs has the potential to mitigate the impact of cognitive skills on mathematics learning when curriculum designers attend to domain-general skills theoretically linked to targeted mathematics outcomes at the program development stage.

## Limitations and Future Research Directions

Findings of the present study should be considered in the context of several limitations. First, it is important to reiterate the exploratory nature of this study. Limited prior research has examined more complex categorizations of intervention response beyond the response/non-response binary. Additionally, the response profiles identified in the present study differed from those found by Peng et al. (2020), limiting the conclusions that can be made regarding what a more complex categorization of intervention response might look like. Further research is needed to determine whether examining response profile patterns and predictors is a useful step toward addressing calls in the field to explore variability in response to generally effective intervention programs and understand how to support students who do not respond adequately to those programs (D. Fuchs & Fuchs, 2019; Miller et al., 2014). As researchers work to build a more nuanced understanding of the impact of validated interventions, thinking about and defining response in more complex ways may help link that knowledge to the classifications and decision rules schools use to match supports to student needs. Toward this end, intervention researchers are encouraged to expand upon the findings of the present study and those of Peng et al., (2020) by exploring intervention response profiles within extant data sets from previous studies of evidence-based academic interventions. Given the potential for this work to inform decision making within multi-tiered models of service delivery in schools, it is critical to establish a broad evidence base regarding meaningful categorization of intervention response across content areas and grade levels.

Relatedly, findings of the present study are inherently limited to the specific evidence-based early mathematics intervention that was implemented and may not generalize to other content areas or math intervention programs. As discussed above, the ROOTS program is more intensive than a typical Tier 2 mathematics intervention and was designed to support students with lower mathematics skill through its instructional architecture. A growing body of evidence suggests ROOTS is best aligned with the needs of students with lower initial mathematics skill, and the response profile patterns identified in the present study reflect this. It is likely that replicating this approach with response data from other validated mathematics interventions would yield different results. For example, patterns of response to less intensive early mathematics interventions may be more similar to those identified by Peng and colleagues (2020) in early literacy, such that students with stronger initial mathematics skill are more likely to show some degree of responsiveness to intervention whereas those with lower initial skill are more likely to show a non-responsive trajectory. Findings to date suggest that the role of initial skill in predicting responsiveness may vary considerably across validated math intervention programs (e.g., Clarke et al., 2019; L.S. Fuchs et al., 2019; Toll & Van Luit, 2013). Furthermore, while in the present study ROOTS students with lower cognitive skills tended to demonstrate stronger intervention response, this may not be the case for other math intervention programs. The literature has established a strong theoretical and empirical link between cognitive skills and math performance, and cognitive variables have been found to moderate response to math intervention in a handful of studies ( L. S. Fuchs, Malone, et al., 2016; Powell et al., 2017). It is likely that differences in the features of validated math intervention programs (e.g., instructional design elements, breadth versus depth of content coverage) contribute to differences in program "fit" for students with different cognitive skill profiles. Further research is needed to allow for clear conclusions regarding patterns and predictors of response across mathematics intervention programs spanning different concepts and skills, grade levels, and instructional design features.

Third, while the present study examined math trajectories over time with the inclusion of a follow-up data point around one year after students had completed the ROOTS intervention, it was not possible to examine latent mathematics performance at first grade follow up or explore students' mathematics trajectories between post-intervention timepoints. A single mathematics indicator—specifically, the most distal measure relative to the early numeracy skills targeted by the ROOTS intervention—was available for a single follow-up time point. Given the temporal response patterns identified in the present study (i.e., delayed and fadeout effects), as well as the literature indicating that treatment effects often diminish over time even in the context of validated mathematics interventions (Bailey et al., 2020), more in-depth examination of both short- and long-term patterns of responsiveness to early mathematics intervention is an important direction for future research. Studies that include multiple and varied measures of mathematics skills across a larger number of post-intervention timepoints may provide further insight into the trajectories of students who display a delayed response or fail to maintain gains over time, further clarifying how best to support these students to ensure continued growth and prevent fadeout.

Fourth, during the course of this study, a number of challenges arose related to the LPA approach itself. While this methodological approach allows for exploration of interesting and important educational research questions, challenges related to selecting the model with the optimal number of latent classes are common in the literature (e.g., Bray & Dziak, 2018). In this case, selecting the best fitting model was not straightforward, as different indicators of goodness of fit favored different models. While this provided flexibility to select a model that fit the data adequately while maintaining parsimony, interpretability, and adequate class sizes for subsequent analyses, it also introduced an element of subjectivity. An effort was made to provide a clear rationale for selecting the four-profile model, but valid arguments can be made in favor of alternate models. As described by Bray and Dziak, the use of LPA and related methods in applied research is both an art and a science.

Finally, the fact that cognitive measures were only available for a smaller subsample of students who participated in cohorts 2 and 3 necessitated the introduction of additional measurement error. To explore whether these measures predicted response profile membership, students were assigned to the profile with the highest likelihood of membership. It was not possible to model the uncertainty associated with response profile membership using a three-step modeling approach without sacrificing power to address the primary research question. Additionally, examining cognitive predictors for a smaller subsample made it challenging to directly compare the utility of domain-general versus domain-specific predictors of intervention response patterns. Since domain-specific initial mathematics skill measures were examined as predictors for the entire treatment group, separate regression analyses were run for each group of predictors, and there was greater power to detect effects for domain-specific predictors compared to those for domain-general predictors. These limitations point to the need for continued exploration of the complex interactions between initial skill, cognitive variables, and response to mathematics intervention.

**Conclusion**

With a national focus on increasing mathematics proficiency through implementation of evidence-based practices, meaningful categorization of responsiveness to intervention is essential to decision making within multi-tiered RTI frameworks. The present study explored patterns of response to the ROOTS kindergarten mathematics intervention, as well as predictors of those patterns, in a replication and extension of research conducted by Peng and colleagues (2020). Results indicated that variability in response to a generally effective intervention was best captured by a more complex categorization encompassing four distinct response profiles: a moderate-risk, mildly responsive group; a moderate-risk, delayed response group, a high-risk, strongly responsive group; and a lower-risk, non-responsive group. Membership in each response profile group was predicted by pre-intervention performance on measures of both early mathematics and general cognitive skills, including visual-spatial skills, fluid reasoning, and working memory. Specifically, students with lower initial math skill and cognitive performance demonstrated stronger intervention response, suggesting that the ROOTS intervention effectively mitigated the impact of cognitive and learning challenges among at-risk students.

While the response profiles identified in the present study differed from those described by Peng et al., findings of both studies suggest the potential utility of looking beyond the widely used response/non-response binary and thinking about intervention response in more complex, nuanced ways. In particular, distinguishing between subgroups of students with distinct response trajectories may inform decisions about next steps to ensure positive learning outcomes for individual students, such as when to discontinue, maintain, intensify, or modify intervention supports. Future research may clarify how more complex response categorization frameworks may inform instruction decisions within an RTI framework.

## REFERENCES CITED


Adlof, S. M., & Patten, H. (2017). Nonword repetition and vocabulary knowledge as predictors of children's phonological and semantic word learning. *Journal of Speech, Language, and Hearing Research*, *60*(3), 682–693. https://doi.org/10.1044/2016_JSLHR-L-15-0441

Al Otaiba, S., Connor, C. M., Folsom, J. S., Wanzek, J., Greulich, L., Schatschneider, C., & Wagner, R. K. (2014). To wait in tier 1 or intervene immediately: A randomized experiment examining first rrade response to intervention (RTI) in reading. *Exceptional Children*, *81*(1), 11–27. https://doi.org/10.1177/0014402914532234.To

Al Otaiba, S., & Fuchs, D. (2006). Who are the young children for whom best practices in reading are ineffective? An experimental and longitudinal study. *Journal of Learning Disabilities*, *39*(5), 414–431. https://doi.org/10.1177/00222194060390050401

Bailey, D. H., Duncan, G. J., Cunha, F., Foorman, B. R., & Yeager, D. S. (2020). Persistence and fade-out of educational intervention effects: Mechanisms and potential solutions. *Psychological Science in the Public Interest*, *21*(2), 55–97. https://doi.org/10.1177/1529100620915848

Barnett, W. S. (2011). *Effectiveness of Early Educational Intervention*. *333*, 975–979. doi: 10.1126/science.1204534

Barth, A. E., Stuebing, K. K., Anthony, J. L., Denton, C. A., Mathes, P. G., Fletcher, J. M., & Francis, D. J. (2008). Agreement among response to intervention criteria for identifying responder status. *Learning and Individual Differences*, *18*(3), 296–307. https://doi.org/10.1016/j.lindif.2008.04.004

Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society*, *57*(1), 289–300.

Berkeley, S., Scanlon, D., Bailey, T. R., Sutton, J. C., & Sacco, D. M. (2020). A snapshot of RTI implementation a decade later: new picture, same story. *Journal of Learning Disabilities*, *53*(5), 332–342. https://doi.org/10.1177/0022219420915867

Bodovski, K., & Farkas, G. (2007). Mathematics growth in early elementary school: The roles of beginning knowledge, student engagement, and instruction. *Elementary School Journal*, *108*(2), 115–130. https://doi.org/10.1086/525550

Bray, B. C. & Dziak, J. J. (2018). Commentary on latent class, latent profile, and latent transition analysis for characterizing individual differences in learning. *Learning and Individual Differences, 66*, 105-110. https://doi.org/10.1016/j.lindif.2018.06.001

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

Clark, N. B., McRoberts, Gerald, W., Van Dyke, J. A., Shankweiler, D. P., & Braze, D. (2012). Immediate memory for pseudowords and phonological awareness are associated in adults and pre-reading children. *Clinical Linguistics and Phonetics*, *26*(7), 577–596. https://doi.org/10.3109/02699206.2012.673045.Immediate

Clarke, B., Doabler, C. T., Fien, H., Baker, S. K., & Smolkowski, K. (2012). *A randomized control trial o fa Tier 2 kindergarten mathematics intervention* (Project ROOTS; CFDA No. 84.324A, 2012–2016, Funding No. R324A120304, awarded $3,338,552). US Department of Education, Institute of Education Sciences, Special Education Research.

Clarke, B., Doabler, C., Smolkowski, K., Kurtz Nelson, E., Fien, H., Baker, S. K., & Kosty, D. (2016a). Testing the immediate and long-term efficacy of a tier 2 kindergarten mathematics intervention. *Journal of Research on Educational Effectiveness*, *9*(4), 607–634. https://doi.org/10.1080/19345747.2015.1116034

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), 1–16. https://doi.org/10.1177/2332858417706899

Clarke, B., Doabler, C. T., Smolkowski, K., Baker, S. K., Fien, H., & Strand Cary, M. (2016b). 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., 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. *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 of number sense (ASPENS)*. Cambium Learning Group, Sopris Learning.

Clarke, B., Nelson, N. J., Ketterlin Geller, L., Kosty, D., Smolkowski, K., Lesner, T., Furjanic, D., & Fien, H. (2020). Investigating the Promise of a Tier 2 Sixth-Grade Fractions Intervention. *Learning Disability Quarterly*. https://doi.org/10.1177/0731948720972411

Clarke, B., Turtura, J., Lesner, T., Cook, M., Smolkowski, K., Kosty, D., & Doabler, C. T. (2022). A conceptual replication of a kindergarten math intervention within the context of a research-based core. *Exceptional Children*, Advance online publication. doi: 10.1177/00144029221088938

Cohen, J. (1992). *A Power Primer*. *112*(1), 155–159.

Common Core State Standards Initiative. (2010). Common Core State Standards for Mathematics. http://www.corestandards.org/the-standards/mathematics

Cornu, V., Hornung, C., Schiltz, C., & Martin, R. (2017). How do different aspects of spatial skills relate to early arithmetic and number line estimation? *Journal of Numerical Cognition*, *3*(2), 309–343. https://doi.org/10.5964/jnc.v3i2.36

Coyne, M. D., Kame'enui, E. J., & Carnine, D. W. (2011). *Effective teaching strategies that accommodate diverse learners* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Doabler, C. T., Clarke, B., & Fien, H. (2012). *ROOTS assessment of early numeracy skills (RAENS)* [Unpublished measurement instrument]. Center on Teaching and Learning, University of Oregon.

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

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School Readiness and Later Achievement. *Developmental Psychology*, *43*(6), 1428–1446. https://doi.org/10.1037/0012-1649.43.6.1428

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

Filderman, M. J., Toste, J. R., Didion, L. A., Peng, P., & Clemens, N. H. (2018). Data-Based Decision Making in Reading Interventions: A Synthesis and Meta-Analysis of the Effects for Struggling Readers. *Journal of Special Education*, *52*(3), 174–187. https://doi.org/10.1177/0022466918790001

Fletcher, J. M., & Miciak, J. (2019). *The Identification of Specific Learning Disabilities: A Summary of Research on Best Practices*. 1–32. www.texasldcenter.org

Floyd, R. G., Evans, J. J., & McGrew, K. S. (2003). Relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and mathematics achievement across the school-age years. *Psychology in the Schools*, *40*(2), 155–171. https://doi.org/10.1002/pits.10083

Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). *Teaching Math to Young Children*. 165. http://whatworks.ed.gov

Fuchs, D., & Deshler, D. D. (2007). What we need to know about responsiveness to intervention (and shouldn't be afraid to ask). *Learning Disabilities Research & Practice*, *22*(2), 129–136. https://doi.org/10.1111/j.1540-5826.2007.00237.x

Fuchs, D., & Fuchs, L. S. (2019). On the Importance of Moderator Analysis in Intervention Research: An Introduction to the Special Issue. *Exceptional Children*, *85*(2), 126–128. https://doi.org/10.1177/0014402918811924

Fuchs, L. S., Compton, D. L., Fuchs, D., Paulsen, K., Bryant, J. D., & Hamlett, C. L. (2005). The prevention, identification, and cognitive determinants of math difficulty. *Journal of Educational Psychology*, *97*(3), 493–513. https://doi.org/10.1037/0022-0663.97.3.493

Fuchs, L. S., Fuchs, D., Compton, D. L., Powell, S. R., Seethaler, P. M., Capizzi, A. M., Schatschneider, C., & Fletcher, J. M. (2006). The cognitive correlates of third-grade skill in arithmetic, algorithmic computation, and arithmetic word problems. *Journal of Educational Psychology*, *98*(1), 29–43. https://doi.org/10.1037/0022-0663.98.1.29

Fuchs, L. S., Fuchs, D., & Gilbert, J. K. (2019). Does the Severity of Students' Pre-Intervention Math Deficits Affect Responsiveness to Generally Effective First-Grade Intervention? *Exceptional Children*, *85*(2), 147–162. https://doi.org/10.1177/0014402918782628

Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child Development*, *81*(5), 1520–1533. https://doi.org/10.1111/j.1467-8624.2010.01489.x.The

Fuchs, L. S., Geary, D. C., Fuchs, D., Compton, D. L., & Hamlett, C. L. (2016). Pathways to third-grade calculation versus word-reading competence: Are they more alike or different? *Child Development*, *87*(2), 558–567. https://doi.org/10.1111/cdev.12474.Pathways

Fuchs, L. S., Malone, A. S., Schumacher, R. F., Namkung, J., Hamlett, C. L., Jordan, N. C., Siegler, R. S., Gersten, R., & Changas, P. (2016). Supported self-explaining during fraction intervention. *Journal of Educational Psychology*, *108*(4), 493–508. https://doi.org/10.1037/edu0000073

Fuchs, L. S., Newman-Gonchar, R., Schumacher, R., Dougherty, B., Bucka, N., Karp, K. S., Woodward, J., Clarke, B., Jordan, N. C., Gersten, R., Jayanthi, M., Keating, B., & Morgan, S. (2021). *Assisting Students Struggling with Mathematics: Intervention in the Elementary Grades*.

Fuchs, L. S., Schumacher, R. F., Sterba, S. K., Long, J., Namkung, J., Malone, A., Hamlett, C. L., Jordan, N. C., Gersten, R., Siegler, R. S., & Changas, P. (2014). Does working memory moderate the effects of fraction intervention? An aptitude-treatment interaction. *Journal of Educational Psychology*, *106*(2), 499–514. https://doi.org/10.1037/a0034341

Fuchs, L. S., Sterba, S. K., Fuchs, D., & Malone, A. S. (2016). Does evidence-based fractions intervention address the needs of very low-performing students? *Journal of Research on Educational Effectiveness*, *9*(4), 662–677. https://doi.org/10.1080/19345747.2015.1123336

Fuchs, L. S., & Vaughn, S. (2012). Responsiveness-to-intervention: A decade later. *Journal of Learning Disabilities*, *45*(3), 195–203. https://doi.org/10.1177/0022219412442150

Geary, D. C. (2004). Mathematics and learning disabilities. *Journal of Learning Disabilities*, *37*(1), 4–15.

Geary, D. C. (2011). Cognitive predictors of achievement growth in mathematics: A five year longitudinal study. *Developmental Psychology*, *47*(6), 1539–1552. https://doi.org/10.1037/a0025510.Cognitive

Geary, D. C., Nicholas, A., Li, Y., & Sun, J. (2017). Developmental change in the influence of domain-general abilities and domain-specific knowledge on mathematics achievement: An eight-year longitudinal study. *Journal of Educational Psychology*, *109*(5), 680–693.

Gersten, R. (2016). What We Are Learning About Mathematics Interventions and Conducting Research on Mathematics Interventions. *Journal of Research on Educational Effectiveness*, *9*(4), 684–688. https://doi.org/10.1080/19345747.2016.1212631

Gersten, R., Beckmann, S., Clarke, B., Foegen, A., Marsh, L., Star, J. R., & Witzel, B. (2009). Assisting Students Struggling with Mathematics:Response to Intervention (RtI) for elementary and middle schools. *What Works Clearinghouse*, 1–98. https://doi.org/10.1016/j.jhazmat.2011.04.026

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

Ginsburg, H.P., & Baroody, A.J. (2003). *Test of Early Mathematics Ability—Third Edition (TEMA-3)*. ProEd.

Green, C. T., Bunge, S. A., Briones Chiongbian, V., Barrow, M., & Ferrer, E. (2017). Fluid reasoning predicts future mathematical performance among children and adolescents. *Journal of Experimental Child Psychology*, *157*, 125–143. https://doi.org/10.1016/j.jecp.2016.12.005

Harcourt Educational Measurement. (2002). *Stanford Achievement Test (SAT-10)*.

Hart, S. A., Logan, J. A. R., Thompson, L., Kovas, Y., McLoughlin, G., & Petrill, S. A. (2016). A latent profile analysis of math achievement, numerosity, and math anxiety in twins. *Journal of Educational Psychology*, *108*(2), 181–193. https://doi.org/10.1037/edu0000045

IBM Corp. Released 2021. IBM SPSS Statistics for Windows, Version 28.0. Armonk, NY: IBM Corp.

Jordan, N. C., Glutting, J., & Ramineni, C. (2008). A number sense assessment tool for identifying children at risk for mathematical difficulties. In A. Dowker (Ed.), *Mathematical difficulties: Psychology and intervention* (pp. 45–57). Academic Press.

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

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, *2*(1), 302–317. https://doi.org/10.1111/j.1751-9004.2007.00054.x

Lam, E. A., & McMaster, K. L. (2014). Predictors of responsiveness to early literacy intervention: A 10-year update. *Learning Disability Quarterly*, *37*(3), 134–147. https://doi.org/10.1177/0731948714529772

Lee, K., & Bull, R. (2016). Developmental changes in working memory, updating, and math achievement. *Journal of Educational Psychology*, *108*(6), 869–882.

McIntosh, C. N. (2013). Pitfalls in subgroup analysis based on growth mixture models: a commentary on van Leeuwen et al. (2012). *Quality of Life Research*, *22*(9), 2625-2629

Mellard, D., McKnight, M., & Jordan, J. (2010). RTI tier structures and instructional intensity. *Learning Disabilities Research & Practice*, *25*(4), 217–225. https://doi.org/10.1111/j.1540-5826.2010.00319.x

Miller, B., Vaughn, S., & Freund, L. (2014). Learning disabilities research studies: Findings from NICHD funded projects. *Journal of Research on Educational Effectiveness*, *7*(3), 225–231. https://doi.org/10.1080/19345747.2014.927251

Mix, K. S., & Cheng, Y. L. (2012). The relation between space and math: Developmental and educational implications. *Advances in Child Development and Behavior*, *42*, 197–243). https://doi.org/10.1016/ B978-0-12-394388-0.00006-X

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.

Muthén, L.K. and Muthén, B.O. (1998-2021). *Mplus User's Guide*. Eighth Edition. Los Angeles, CA: Muthén & Muthén.

National Assessment of Educational Progress. (2019). Results from the 2019 Mathematics and Reading Assessments (NCES No. 2020012). National Center for Education Statistics, U.S. Department of Education, Institute of Education Sciences. https://www.nationsreportcard.gov/mathematics?grade=4

National Center on Response to Intervention. (2010). Essential components of RTI: A closer look at response to intervention. Washington, DC: U.S. Office of Special Education Programs, National Center on Response to Intervention. http://www .rti4success.org/

National Mathematics Advisory Panel. (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel.* Washington, DC: U.S. Department of Education. DOI: 10.3102/0013189X08329195

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (2008). *TIMSS 2007 technical report.* Chestnut Hill, MA: International Association for the Evaluation of Educational Achievement.

Peng, P., Fuchs, D., Fuchs, L. S., Cho, E., Elleman, A. M., Kearns, D. M., Patton, S., & Compton, D. L. (2020). Is "response/no response" too simple a notion for RTI frameworks? Exploring multiple response types with latent profile analysis. *Journal of Learning Disabilities*, *53*(6), 454–468. https://doi.org/10.1177/0022219420931818

Powell, S. R., Cirino, P. T., & Malone, A. S. (2017). Child-level predictors of responsiveness to evidence-based mathematics intervention. *Exceptional Children*, *83*(4), 359–377. https://doi.org/10.1177/0014402917690728

Proctor, B. E., Floyd, R. G., & Shaver, R. B. (2005). Cattell-Horn-Carroll broad cognitive ability profiles of low math achievers. *Psychology in the Schools*, *42*(1), 1–12. https://doi.org/10.1002/pits.20030

Schwarz, G., (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Shanley, L., Clarke, B., Smolkowski, K., Doabler, C. T., Kurtz-Nelson, E. C., & Fien, H. (2021). Examining the role of domain-general skills in mathematics learning and intervention response in kindergarten. *Learning Disabilities Research & Practice*, *36*(4), 330–352. https://doi.org/10.1111/ldrp.12267

Sutherland, M., Lesner, T., Kosty, D., Lussier, C., Turtura, J., Doabler, C. T., & Clarke, B. (in press). Examining interactions across instructional tiers: Do features of Tier 1 predict student responsiveness to Tier 2 mathematics intervention? *Journal of Learning Disabilities*, Advance online publication.

Sood, S., & Jitendra, A. K. (2011). 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

Swanson, H. L. (2006). Cross-sectional and incremental changes in working memory and mathematical problem solving. *Journal of Educational Psychology*, *98*(2), 265–281. https://doi.org/10.1037/0022-0663.98.2.265

Taub, G. E., Floyd, R. G., Keith, T. Z., & McGrew, K. S. (2008). Effects of general and broad cognitive abilities on mathematics achievement. *School Psychology Quarterly*, *23*(2), 187–198. https://doi.org/10.1037/1045-3830.23.2.187

Toll, S. W. M., & Van Luit, J. E. H. (2013). Accelerating the early numeracy development of kindergartners with limited working memory skills through remedial education. *Research in Developmental Disabilities*, *34*(2), 745–755. https://doi.org/10.1016/j.ridd.2012.09.003

Van de Weijer-Bergsma, E., Kroesbergen, E. H., & Van Luit, J. E. H. (2015). Verbal and visual-spatial working memory and mathematical ability in different domains throughout primary school. *Memory and Cognition*, *43*, 367–378. https://doi.org/10.3758/s13421-014-0480-4

Vaughn, S., & Fuchs, L. S. (2003). Redefining Learning Disabilities as Inadequate Response to Instruction: The Promise and Potential Problems. *Learning Disabilities Research and Practice*, *18*(3), 137–146. https://doi.org/10.1111/1540-5826.00070

Vellutino, F. R., Scanlon, D. M., Small, S., & Fanuele, D. P. (2006). Response to intervention as a vehicle for distinguishing between children with and without reading disabilities: Evidence for the role of kindergarten and first-grade interventions. *Journal of Learning Disabilities*, *39*(2), 157–169. https://doi.org/10.1177/00222194060390020401

Wagner, R. K., Torgesen, J. K., & Rashotte, C. A. (1999). *Comprehensive Test of Phonological Processing (CTOPP)*. Pro-Ed.

Walberg, H. J., & Tsai, S. (1983). “Matthew” Effects in Education. *American Educational Research Journal*, *20*(3), 359–373. https://www.jstor.org/stable/1162605

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What’s Past Is Prologue: Relations Between Early Mathematics Knowledge and High School Achievement. *Educational Researcher*, *43*(7), 352–360. https://doi.org/10.3102/0013189X14553660

Wechsler D. (2012). *Wechsler Preschool and Primary Scale of Intelligence*, 4th edition. Psychological Corporation

Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, *102*(1), 43–53. https://doi.org/10.1037/a0016738

Yungtai, L. O., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*(3), 767-778.

Zheng, X., Swanson, H. L., & Marcoulides, G. A. (2011). Working memory components as predictors of children's mathematical word problem solving. *Journal of Experimental Child Psychology*, *110*, 481–498. https://doi.org/10.1016/j.jecp.2011.06.001

Zilberman, A., & Ice, L. (2021). Why computer occupations are behind strong STEM employment growth in the 2019-29 decade. *Beyond the Numbers: Employment & Unemployment*, *10*(1). Washington, DC: U.S. Bureau of Labor Statistics. https://www.bls.gov/opub/btn/volume-10/why-computer-occupations-are-behind-strong-stem-employment-growth.htm