Measuring the Quantity and Quality of Explicit Instructional Interactions in an Empirically-Validated Tier 2 Kindergarten Mathematics Intervention

Authors

Christian T. Doabler, Ben Clarke, Derek Kosty, Hank Fien, Kieth Smolkowski, Meijia Liu, Scott K. Baker

Publication History

Article first published online: November 11, 2019; Issue published: February 1, 2021

Full Reference

Doabler, C. T., Clarke, B., Kosty, D., Fien, H., Smolkowski, K., Liu, M., & Baker, S. K. (2021). Measuring the quantity and quality of explicit instructional interactions in an empirically validated Tier 2 kindergarten mathematics intervention. *Learning Disability Quarterly*, *44*(1), 50–62. doi: 10.1177/0731948719884921

Funding Source

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education through Grant **R324A120304** and by the National Science Foundation through Grant **1503161**.

Abstract

Instructional interactions that occur between teachers and students around foundational mathematics topics are critical for supporting mathematical proficiency among students with mathematics learning disabilities (MLD). This study investigated whether the initial mathematics skill of 880 kindergarten students at risk for MLD predicted the quantity and quality of explicit instructional interactions (i.e., overt teacher modeling, student practice opportunities, and academic feedback) experienced during an empirically-validated, Tier 2 kindergarten mathematics intervention. It also examined whether the quantity and quality of such instructional interactions predicted gains in student mathematics achievement. Researchers conducted 740 direct observations of 255 intervention groups within a multi-year, randomized controlled trial. Results suggested that intervention groups with lower initial mathematics skill received higher rates of academic feedback and made more frequent errors. Additionally, more frequent and higher quality academic feedback and group practice opportunities predicted increased mathematics achievement. Implications for investigating the active ingredients of mathematics interventions are discussed.

Measuring the Quantity and Quality of Explicit Instructional Interactions in an Empirically-Validated Tier 2 Kindergarten Mathematics Intervention

A consistent finding of methodologically rigorous research is that explicit instruction leads to better student achievement for students who demonstrate academic risk (Agodini & Harris, 2010; Gersten et al., 2009; Vaughn & Swanson, 2015). In addition to helping establish the efficacy of explicit instruction, this body of empirical work has also begun to build consensus on the particulars of this instructional approach. It is evident from this literature base, for example, that explicit instruction is a systematic approach that utilizes empirically-validated instructional design and delivery principles to unambiguously teach fundamental concepts and skills that students would not otherwise acquire on their own (Deshler, 2015; Hughes, Morris, Therrien, & Benson, 2017; Simmons, 2015). This instructional approach also scaffolds instruction to promote a high success rate among at-risk students (Rosenshine, 2012).

Explicit instruction is known to facilitate purposefully-designed instructional interactions for students to learn foundational mathematics content (Hughes et al., 2017). Such interactions include overt teacher modeling, structured opportunities for students to practice with mathematical concepts, and teacher-provided academic feedback. Research conducted in core (Tier 1) mathematics settings suggests that early elementary students (i.e., kindergarten to second grade) benefit most from classrooms that offer more frequent explicit instructional interactions around critical mathematics concepts and skills (Author et al., 2015; Clements, Agodini, & Harris, 2013). Despite evidence that the explicit instructional interactions that occur between teachers and students in Tier 1 settings matter, little is known about how students with or at risk for mathematics learning disabilities (MLD) engage in mathematical content in Tier 2 mathematics interventions.

One gap in the Tier 2 mathematics literature concerns whether the quantity and quality of explicit instructional interactions predicts increased student mathematics achievement. Small-group mathematics instruction that provides frequent, high-quality instructional interactions may promote beneficial outcomes for students with MLD. Another gap in the knowledge base is whether the skill composition of small groups, based on each small group member's level of mathematical skill at the onset of Tier 2 mathematics interventions, influences the frequency and quality in which interventionists facilitate explicit instructional interactions. Group-level pretreatment mathematical performance may serve as a gauge for how and to what extent interventionists intensify instruction for students with or at risk for MLD. For example, Tier 2 interventionists may provide differentiated instruction, such as offering more overt teacher demonstrations, structured practice opportunities, and academic feedback, in small groups with lower preintervention skill levels. We sought to address these two blank spots in the empirical literature.

Tier 2 Kindergarten Mathematics Intervention Research

Researchers who have conducted recent randomized controlled trials (RCTs) involving Tier 2 kindergarten mathematics interventions have found that explicit instruction holds promise for increasing the mathematics achievement of kindergarten students at risk for MLD. The interventions targeted in this line of research are intended to supplement core mathematics instruction and designed to prioritize the development of early number sense (Berch, 2005). Such interventions, which are typically delivered in small group formats, are used to build and extend students' foundational numeracy knowledge and skills, such as using counting principles, making magnitude comparisons, solving number combinations, and composing and decomposing whole numbers.

In 2011, Dyson, Jordan, and Glutting used an RCT to test the impact of an 8-week intervention on the number sense knowledge of 121 kindergarten students at risk for MLD. The intervention offered 30-minute sessions of small-group instruction, 3 times per week. Medium to large effect sizes were observed, suggesting that intervention students outperformed their control peers on proximal and distal number sense measures. More recently, Jordan, Glutting, Dyson, Hassinger-Das and Irwin (2012) examined the efficacy of a Tier 2 kindergarten intervention focused on number concepts identified in the Common Core State Standards for Mathematics (CCSS-M, 2010). A total of 88 kindergarten students considered at risk for MLD participated. Results indicated a significant intervention effect on a set of distal outcome measures of whole number understanding, with effect sizes (Cohen's *d*) ranging from 0.30 to 2.64.

Clarke, Doabler, Fien, Baker, and Smolkowski (2012) conducted a federally-funded, four-year, multisite efficacy trial aimed at testing the treatment effects of the ROOTS intervention. ROOTS is a Tier 2 kindergarten mathematics intervention that uses explicit instruction to promote early number sense among kindergarten students at risk for MLD. A primary aim of the ROOTS Efficacy Project (Clarke et al., 2012) was to experimentally manipulate group size to test whether this alterable variable intensified ROOTS instruction. Approximately 1,250 kindergarten students at risk for MLD were randomly assigned within 138 kindergarten classrooms to one of three conditions: (a) ROOTS with a 2:1 student-teacher ratio, (b) ROOTS with a 5:1 student-teacher ratio, or (c) a no-treatment control condition. Whereas control students received core (Tier 1) mathematics instruction only, treatment students received ROOTS in addition to Tier 1 instruction.

To build scientific credibility of ROOTS, our research team systematically replicated its impact across the larger efficacy project (Doabler et al., 2016). In all, four separate studies from two different geographical regions of the U.S. were conducted. Findings from two of the studies indicated that aggregated treatment students in the 2:1 ROOTS and 5:1 ROOTS groups significantly outperformed their control condition peers on a host of standardized mathematics achievement measures, with reported effect sizes (Hedges' *g*) ranging from .12 to .95 (Clarke et al., 2016; Doabler et al., 2016). In two other studies that examined the effect of group size on intervention impact, nonsignificant differences between the 2:1 ROOTS and 5:1 ROOTS groups were found, suggesting that the impact of ROOTS was essentially the same regardless of group size (Clarke et al., 2017; Doabler et al., 2019). While the collective results of this research program were encouraging, they in turn raised questions as to why ROOTS has led to increased student mathematics achievement in different instructional formats (2:1 and 5:1 group sizes) for a range of at-risk learners from different geographical regions of the U.S.

Unpacking the Black Box of Tier 2 Explicit Mathematics Interventions

It is encouraging that the findings of prior research on Tier 2 number sense interventions suggest explicit mathematics instruction is an effective means for increasing the mathematics achievement of kindergarten students at risk for MLD. And while this evidence is fundamental to our confidence in the effectiveness of explicitly-designed mathematics interventions, it is insufficient for informing the field as to "why" these types of mathematics interventions positively benefit at-risk learners. Ascertaining why Tier 2 mathematics interventions lead to desired student mathematics, such as those investigated by Clarke et al. (2012), Dyson et al. (2011) and Jordan et al. (2012), would align with the research agendas of major federal funding agencies, including the Institute of Education Sciences and the National Science Foundation (IES & NSF, 2013). However, unpacking the "black box" (McKinnon & Luecken, 2008) of these types of mathematics interventions requires investigating their active ingredients to determine their explanatory contributions to increased student mathematics outcomes.

For example, one plausible mediator of Tier 2 mathematics interventions is the explicit instructional interactions that they purposefully facilitate between teachers and students during instruction. These instructional interactions offer opportunities for at-risk kindergarten students to systematically and directly learn critical whole number concepts and skills. Comprised of overt teacher demonstrations (Gersten et al., 2009), group and individual student practice opportunities (Clements, Agodini, & Harris, 2013; NRC, 2001), and teacher-provided academic feedback (Hattie & Timperley, 2007), explicit instructional interactions often represent the underlying mechanisms (i.e., active ingredients) in which Tier 2 mathematics interventions are expected to achieve their desired effects on student mathematics achievement (Clarke et al., 2012). Given the plausible explanatory power of explicit instructional interactions, their investigations may help the field ascertain how and why Tier 2 mathematics interventions promote positive outcomes for students with MLD

Initial Mathematics Achievement and Explicit Instructional Interactions

Initial mathematics achievement in kindergarten is an important factor in understanding students' path for developing mathematical proficiency. Robust research findings suggest that students who begin their kindergarten year at risk for MLD are likely to demonstrate long-term difficulties in mathematics (Morgan, Farkas, & Wu, 2009). These findings are largely based on the readiness (or lack thereof) of many students to learn formal mathematics. Relative to their typically-achieving peers, students who begin their kindergarten year at risk for MLD often receive less support at home and in preschool to develop early number sense (Barnes et al., 2016). If initial mathematics achievement has important implications for students' immediate and long-term mathematics outcomes (Clarke et al., 2019; Powell, Cirino, & Malone, 2017; Morgan et al., 2009), then it is plausible that where groups of students at risk for MLD begin their kindergarten year in mathematics is also a significant predictor of the quantity and quality of explicit instructional interactions they engage in during Tier 2 mathematics interventions to learn foundational mathematical content. An aim of this study was to address this hypothesis.

Purpose of the Study

The purpose of this study was to investigate the active ingredients of the ROOTS intervention (Clarke et al., 2012) and explore why it has demonstrated strong efficacy across a recent program of research. Specifically, we sought to examine the extent to which students' group-level initial mathematics skill predicted the rate and quality in which they engaged in explicit instructional interactions during the ROOTS intervention. We also investigated the extent to which the quantity and quality of explicit instructional interactions predicted student mathematics achievement. Two research questions were addressed:

1. Does group-level initial mathematics achievement, as established by pretest performances on standardized measures of whole number understanding, predict the quantity and quality of explicit instructional interactions during ROOTS instruction?
2. Does the quantity and quality of explicit instructional interactions during ROOTS instruction predict gains in student mathematics achievement from pretest to posttest?

Method

Data analyzed in the current study were collected during the federally-funded ROOTS Efficacy Project (Clarke et al., 2012). Implementation of the ROOTS intervention occurred across three school years (2012-2015) at two different research sites: Oregon and Massachusetts. The ROOTS Efficacy Project employed a partially nested randomized controlled trial (Baldwin, Bauer, Stice, & Rohde, 2011), randomly assigning kindergarten students within classrooms to one of three conditions: 2:1 ROOTS group, 5:1 ROOTS group, and a no-treatment control condition. The current study focused specifically on student mathematics outcomes and direct observation data collected in the two ROOTS conditions (2:1 and 5:1 groups).

Participants

Districts and Schools. A total of 23 schools from four Oregon and two Massachusetts school districts participated in this study. Both Massachusetts districts were located in the metropolitan area of Boston. Three Oregon districts were located in rural and suburban areas of western Oregon, and one Oregon district was located in the metropolitan area of Portland. Enrollment for all districts ranged from 2,736 to 39,002. In each district, schools receiving Title I funding were targeted for participation. Within the 23 schools, 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. Within these same schools, 8%-25% of students received special education services, 1%-69% were ELs, and 17%-87% were eligible for free or reduced lunch.

Students. Participants were drawn from a total of 138 kindergarten classrooms. In each classroom, all students with parental consent were screened in the late fall of their kindergarten year. The screening process, which included 3,066 students, comprised two standardized measures of early mathematics proficiency: (a) Assessing Students Proficiency in Early Number Sense (ASPENS; Clarke, Gersten, Dimino, & Rolfhus, 2011) and (b) Number Sense Brief Screener (NSB; Jordan, Glutting, & Ramineni, 2008). Students were eligible for the ROOTS intervention if they received an NSB score of 20 or less and an ASPENS' Composite score in the strategic or intensive ranges. Prior research suggests that scores below these cutoffs are indicative of long-term MLD (Clarke et al., 2011; Jordan et al., 2008).

The screening process identified 1,251 students. For each eligible student, raw NSB and ASPENS scores were converted into norm-referenced standard scores and then combined to form an overall composite score. Composite scores within each classroom were then rank ordered, and the 10 ROOTS-eligible students with the lowest composite scores were randomly assigned to: (a) 2:1 ROOTS group (*n* = 258), (b) 5:1 ROOTS group (*n* = 622), or (c) a no-treatment control condition (*n* = 371). A total of 255 ROOTS groups were formed: 129 of the 2:1 groups and 126 of the 5:1 groups. Demographic data for the 880 ROOTS students indicated that 8% received special education services, 24% were English learners, and 51% were females. While the majority racial group of ROOTS students was White (64%), 24% were Hispanic, 6% were Black, 3% Asian, 1% were American Indian, and approximately 2% were Multiple Races.

ROOTS Intervention

Interventionists. District employees and interventionists hired specifically for the ROOTS project delivered the intervention. The majority of interventionists (93.5%) identified as female (93.5%) and White (76.1%), with 12.0% identifying as Hispanic. The remaining 11.9% identified as another race or ethnicity or declined to respond. Most interventionists (92.3%) had previous experience teaching small groups, and 60.5% held a bachelor's degree or higher. About half (56.5%) had taken a college level algebra course. On average, interventionists had 10.4 years of teaching experience (*SD* = 8.6) and 22.0% held a current teaching license.

ROOTS. ROOTS is a 50-lesson, Tier 2 kindergarten mathematics intervention that is centered on an explicit instructional framework and delivered in small-group instructional formats. Each 20-minute lesson focuses on critical whole number concepts and skills identified in the kindergarten CCSS-M (2010). Lessons 1-25 of the intervention initially focus on the relationship between numbers and quantities and then begin to incorporate instruction on early concepts of addition and subtraction (e.g., solving word problems). The second half of the intervention (Lessons 26-50) targets building students' understanding of place value with numbers 11-19. Each ROOTS lesson contains four mathematics activities with scripted guidelines for interventionists to (a) provide overt demonstrations and explanations of new mathematical content, (b) incorporate visual representations of mathematical ideas (e.g., base-ten blocks) into lesson activities, (c) facilitate frequent student practice opportunities, including student mathematics verbalizations, and (d) offer specific academic feedback to address student errors and affirm correct responses.

The opening activity in each ROOTS lesson consists of a brief warm-up (3-min), the "Nifty Fifty" activity, focused on number identification skills and use of efficient counting strategies with a (1–50) number chart. Each Nifty Fifty activity corresponds to the number of lessons completed in the intervention program. In Lesson 17, for example, interventionists use the Nifty Fifty activity to help children count and identify numbers up to 17. The warm-up activities also support students' knowledge of rational counting (i.e., one-to-one correspondence) and identifying whether one group of objects is greater than, less than, or equal to another group of objects. Next, interventionists deliver a 5-min activity that overtly introduces a new mathematical concept or skill that is central to the lesson's overall objective. For this activity, interventionists use concrete objects (e.g., counting blocks or number lines) to explicitly demonstrate and explain the targeted concept or skill. The third activity (7-min) involves either guided practice of content introduced in the second activity or a review of previously learned material. The final activity is a brief worksheet activity (5-min) that interventionists use to review the lesson's content. Worksheets contain a "note home" (in both English and Spanish) to provide students with additional practice opportunities outside of school.

In the current study, ROOTS was delivered in 20-minute sessions, 5 days per week for 10 weeks. ROOTS instruction began midway through the kindergarten school year to minimize the identification of typically-achieving students and allow at-risk students the opportunity to respond to initial core mathematics instruction. Because ROOTS is a supplemental mathematics intervention, treatment students continued to receive the Tier 1 core mathematics instruction delivered in their general education classrooms.

Implementation fidelity. Fidelity of ROOTS implementation was measured via direct observations by trained research staff, with each ROOTS group observed three times during the course of the intervention. On a 4-point scale (4 = all, 3 = most, 2 = some, 1 = none), observers rated the extent to which the interventionist (a) met the lesson's instructional objectives, (b) followed the lesson's teacher scripting, and (c) used the lesson's mathematics models. Observers also recorded whether the interventionist taught the number of activities prescribed in the lesson. Interventionists were found to meet instructional objectives (M = 3.49, SD = 0.69), follow scripting (M = 3.31, SD = 0.75), and use prescribed models (M = 3.61, SD = 0.64). The majority of prescribed activities were also taught (M = 4.14 out of 5 activities per lesson, SD = 0.77).

Outcome Measures

Our analyses focused on student performance data from three mathematics outcomes measures purported to assess students' whole number understanding. All ROOTS students were administered the measures at pretest and posttest by trained research staff. Inter-scorer reliability criteria were met for all assessments (i.e., > 95% agreement).

ROOTS Assessment of Early Numeracy Skills (RAENS; Doabler, Clarke, & Fien, 2012) is a researcher-developed, individually-administered measure. The RAENS consists of 32 items, assessing aspects of counting and cardinality, number operations, and the base-ten system.

In an untimed setting, students count and compare groups of objects, write, order, and compare numbers, label visual models (e.g. ten-frames), and write and solve single digit addition expressions and equations. RAENS' predictive validity ranges from .68 to .83 for the TEMA-3 and the NSB (Clarke et al., 2016; Doabler et al., 2016). Raw scores were used for analysis.

Assessing Student Proficiency in Early Number Sense (ASPENS: Clarke et al., 2011) is a set of three curriculum-based measures validated for screening and progress monitoring in kindergarten mathematics. Each 1-minute fluency-based measure assesses an important aspect of early numeracy proficiency, including number identification, magnitude comparison, and missing number. Test-retest reliabilities of kindergarten ASPENS measures are in the moderate to high range (.74 to .85). Predictive validity of fall scores on the kindergarten ASPENS measures with spring scores on the TerraNova 3 is reported as ranging from .45 to .52 (Clarke et al., 2011). Raw scores were used for analysis.

Number Sense Brief Screener (NSB; Jordan et al., 2008) is an individually administered measure with 33 items that assess counting knowledge and principles, number recognition, number comparisons, nonverbal calculation, story problems and number combinations. NSB has a coefficient alpha of .84 (Jordan et al., 2008). Raw scores were used for analysis.

Observations of ROOTS Instruction

Each ROOTS group was observed approximately three times (M = 2.9, SD = 0.8) over the course of the intervention, with approximately three weeks separating each observation occasion. A total of 740 observations were conducted, of which 139 (19%) included two observers who simultaneously evaluated inter-observer agreement. Observations were scheduled in advance and observers remained for the duration of ROOTS instruction, with an average observation lasting 20.8 minutes (*SD* = 3.8 min.). Trained observers, who were blind to our research hypotheses, conducted all observations using two observation measures.

Classroom Observations of Student-Teacher Interactions-Mathematics (COSTI-M). The COSTI-M (Doabler et al., 2015; Smolkowski & Gunn, 2012) is a low-inference observation instrument that has been empirically validated to document the frequency of teacher demonstrations, individual and group student practice opportunities, teacher-provided academic feedback, and student mistakes. As documented by the COSTI-M, teacher models represent a teacher's verbalizations of thought processes and physical demonstrations of mathematical content. For example, observers coded a teacher model if the teacher explicitly described the structural features of an "add to" word problem. Academic feedback was operationalized as a teacher's verbal reply or physical demonstration to affirm or correct a student response. For example, observers recorded an academic feedback code if the teacher restated an incorrect answer. Group practice opportunities were defined as a mathematics-related verbalization produced by two or more students in unison. Individual practice opportunities were coded whenever a single student had the opportunity to verbalize or physically demonstrate her mathematical thinking, such as when a teacher asked a specific student to answer a mathematical question (e.g., "Alejandro, use the place value bocks to show 19?"). Rates per minute for each targeted behavior were computed as the frequency of the behavior divided by the duration of the observation in minutes. Author et al. (2015) reported predictive validity of the COSTI-M with the TEMA-3 (p = .004, Pseudo-R^2 = .08) and the EN-CBM (p = .017, Pseudo-R^2 = .05).

Quality of Explicit Mathematics Instruction (QEMI). The QEMI (Doabler & Clarke, 2012) comprises seven items that target the quality of explicit instructional interactions, including group and individual practice opportunities, student participation, teacher modeling, academic feedback, efficiency of instructional delivery, and instructional scaffolding. Internal consistency of the measure was high, .93 (coefficient alpha). To rate the quality of each item, observers used a 4-point rating scale, with scores of 1–2 representing the lower quality range and 3–4 representing the upper quality range. Total QEMI scores were computed as the mean across all items. The mean across the three observations was used in subsequent analyses.

Inter-observer agreement and stability intraclass correlations coefficients (ICCs). To estimate inter-observer agreement in observation measures, we calculated ICCs to describe the proportion of variance in each observation measure occurring between versus within paired observation occasions. Inter-observer agreement ICCs for COSTI-M and QEMI scores ranged from .72 to .94, which based on guidelines proposed by Landis and Koch (1977) represented substantial to nearly perfect agreement. To estimate stability across time, we calculated ICCs to describe the proportion of variance in each observation measure occurring between versus within ROOTS groups. Stability ICCs were .14 for teacher demonstrations, .20 for individual practice, .26 for group practice, .20 for student mistakes, .39 for academic feedback, and .45 for the QEMI scale. Reliability of mean scores across the three observation occasions were fair and ranged from .33 (for teacher demonstrations) to .71 (for QEMI scores).

Statistical Analysis

Following the examination of descriptive statistics for the study variables, we performed a series of random coefficients analyses (RCAs; Snijders & Bosker, 2012) designed to address our research questions. The statistical models accounted for pretest and posttest measures of mathematics achievement nested within students and ROOTS groups. Specifically, we regressed mathematics achievement at pretest and posttest on time (coded 0 for pretest and 1 for posttest), a group-level quantity or quality of explicit instruction predictor variable (mean centered), and the cross-level Time × Predictor interaction. The effect of time represents the average change in outcome from pretest to posttest among groups given the average value of the predictor variable. The effect of the quantity or quality predictor variable addresses Research Question 1 and represents the association between group-level mathematics achievement at pretest and the specific measure of the quantity or quality of explicit instructional interactions. The Time × Predictor interaction addresses Research Question 2 and represents the difference in change in mathematics outcome from pretest to posttest due to a unit increase in the quantity or quality of explicit instructional interactions. To support interpretation of results, we reported r^2 equivalent (Rosnow & Rosenthal, 2003) for the fixed effects of the quantity or quality predictor variable (Research Question 1) and the Time × Predictor interaction (Research Question 2). Alpha was set to .05.

We performed analyses using SAS PROC MIXED version 9.4 (SAS Institute, 2016) and restricted maximum likelihood estimation. Maximum likelihood estimation uses all available data and produces potentially unbiased results even in the face of substantial missing data, provided the missing data were missing at random (Schafer & Graham, 2002). We considered this assumption tenable because missing data (10% across outcome measures) involved students who were absent on the day of assessment (e.g., due to illness) or transferred to a new school (e.g., family mobility). The statistical model also assumes independent and normally distributed observations. We addressed the first of these assumptions by modeling the multilevel nature of the data. The outcome measures in the present study also did not markedly deviate from normality; skewness and kurtosis fell within ± 1.2.

Results

Table 1 provides means, standard deviations, and sample sizes for each student outcome and observed measure of the quantity and quality of instructional interactions. Measures of the quantity and quality of explicit instructional interactions were correlated between $r = \pm.05$ to .39. Tables 2 to 4 summarize results of the RCAs designed to address our research questions.

Research Question 1 focused on the associations between group-level pretest mathematics achievement and measures of the quantity and quality of explicit instructional interactions. These associations were evaluated by the fixed effects of each quantity or quality predictor presented in the second row of data in Tables 2 to 4. Results demonstrated statistically significant associations between pretest mathematics performance and rates of student mistakes and academic feedback. Specifically, the second row of Table 2 shows that lower pretest RAENS scores were associated with higher rates of student mistakes ($p = .0020$, $r^2_{\text{equivelent}} = .037$) and academic feedback ($p = .0012$, $r^2_{\text{equivelent}} = .041$). Table 3 indicates that lower pretest ASPENS scores were associated with higher rates of academic feedback ($p = .0163$, $r^2_{\text{equivelent}} = .023$). Table 4 shows that lower pretest NSB scores were associated with higher rates of student mistakes ($p = .0370$, $r^2_{\text{equivelent}} = .017$) and academic feedback ($p = .0047$, $r^2_{\text{equivelent}} = .031$). No significant associations emerged between pretest mathematics performance and rates of teacher demonstrations (p's $\geq .838$), rates of individual practice opportunities (p's $\geq .661$), rates of group practice opportunities (p's $\geq .644$), or QEMI scores (p's $\geq .363$).

For Research Question 2, we examined whether the quantity or quality of explicit instructional interactions predicted gains in student mathematics achievement from pretest to posttest. We evaluated this question using the Time × Predictor interactions presented in the fourth row of data in Tables 2 to 4. Results indicated that gains in mathematics achievement were significantly associated with rates of individual and group practice opportunities, student mistakes, academic feedback, and QEMI scores. Specifically, the fourth row of Table 2 shows that greater gains in RAENS scores were associated with higher rates of group practice (p < .0001, $r^2_{\text{equivelent}}$ = .063), lower rates of student mistakes (p = .0080, $r^2_{\text{equivelent}}$ = .028), higher rates of academic feedback (p < .0001, $r^2_{\text{equivelent}}$ = .062), and higher QEMI scores (p = .0102, $r^2_{\text{equivelent}}$ = .026). Table 3 shows that greater gains in ASPENS scores were associated with lower rates of individual student practice (p = .0231, $r^2_{\text{equivelent}}$ = .020), higher rates of group practice (p = .0002, $r^2_{\text{equivelent}}$ = .053), lower rates of student mistakes (p < .0001, $r^2_{\text{equivelent}}$ = .129), higher rates of academic feedback (p = .0033, $r^2_{\text{equivelent}}$ = .034), and higher QEMI scores (p = .0073, $r^2_{\text{equivelent}}$ = .028). Table 4 shows that greater gains in NSB scores were associated with higher rates of group practice (p = .0041, $r^2_{\text{equivelent}}$ = .032) and higher rates of academic feedback (p = .0308, $r^2_{\text{equivelent}}$ = .018). No significant associations emerged between gains in mathematics achievement and rates of teacher demonstrations (p's ≥ .590).

Discussion

This study analyzed data collected during a recent efficacy project to explore the active ingredients of an empirically-validated, explicitly-designed Tier 2 kindergarten mathematics intervention. Specifically, we examined the extent to which group-level initial mathematics skill predicted the quantity and quality of explicit instructional interactions facilitated during ROOTS instruction. Additionally, we investigated the extent to which the quantity and quality of such interactions predicted mathematics outcomes for kindergarten students at risk for MLD who received the ROOTS intervention. Below, we summarize results for our two research questions.

Results Summary

Research Question 1. For our first research question, we used group-level initial mathematics achievement to examine its association with the quantity and quality of explicit instructional interactions facilitated during ROOTS instruction. Initial mathematics achievement was based on students' pretest performances on the three mathematics outcome measures focused on whole number understanding. Interestingly, results were mixed. Findings indicated that neither individual nor group practice rates were associated with any measure of mathematics performance at pretest. Similarly, rates of teacher models and the overall quality of explicit instructional interactions did not produce significant results. Rates of student mistakes and academic feedback, on the other hand, were significantly associated with group-level pretest scores, suggesting that ROOTS groups with lower level of initial mathematics achievement made more frequent errors and were recipients of more teacher-provided academic feedback.

In the current study, academic feedback was offered to students not only to rectify their errors but also affirm their correct responses. As such, the significant finding that ROOTS groups with lower mathematics achievement at pretest received higher rates of academic feedback than groups with stronger pretest performances gives an indication that interventionists, who were not privy to students' pretest data, may have attempted to use academic feedback as a way to better meet the needs of groups composed of at-risk students with more intensive learning needs. Interventionists, who ran groups with lower whole number understanding, may have recognized the need to deliver academic feedback above and beyond what was directly prescribed in the ROOTS lessons. While ROOTS interventionists were strongly encouraged to utilize the intervention's scripted guidelines, they were also given permission to provide additional practice opportunities and academic feedback when needed.

Collectively, our results add a complementary layer to the findings of recent moderation analyses of large-scale intervention studies (Fuchs & Fuchs, 2019). In a series of federally-funded investigations, different research teams examined the extent to which students differentially benefited from reading and mathematics interventions based on their pretreatment performances. Overall, findings were mixed. Coyne et al. (2019), for example, found that kindergarten students with higher initial vocabulary knowledge reaped greater benefit from a Tier 2 vocabulary intervention than their peers with lower preintervention vocabulary knowledge. Conversely, results from Clemens et al. (2019) suggested that students with lower oral reading fluency at pretest made stronger gains in reading comprehension than students with higher pretest reading fluency. Clemens et al. (2019), however, found no evidence that initial word identification efficiency or vocabulary knowledge moderated the treatment effects. Similarly, Fuchs, Fuchs, and Gilbert (2019) reported non-significant moderation effects for students' initial mathematics skill levels. In our own work (Clarke et al., 2019), we found moderation by initial skill such that students with lower mathematics skills at pretest showed greater response to the ROOTS intervention. Given that the results of these previous moderation studies varied (Clarke et al., 2019; Clemens et al., 2019; Coyne et al., 2019; Fuchs et al., 2019), findings from our first research question may provide a fuller picture for how, why, and to what extent such explicit, systematic interventions work across a continuum of at-risk students' pretreatment performances.

Research Question 2. Our second research question examined whether and to what extent the quantity and quality of explicit instructional interactions predicted gains in student mathematics achievement from pretest to posttest. In terms of quantity, one component of explicit instructional interactions that surfaced as a consistent predictor of student mathematics achievement was the frequency of group practice. Results suggested significant associations between rates of group responses and gains on all three mathematics outcome measures, with variance explained ranging from 3% to 6%. While group practice is an effective mechanism for fostering opportunities for all students to learn foundational mathematics concepts and skills, research suggests that group-level mathematics discourse can be difficult for teachers to manage, particularly in whole-class settings (Doabler et al., 2015; NRC, 2001). It may be, therefore, that small groups (≤ 5 students) are more conducive for facilitating these types of student practice opportunities. For example, when teaching in these instructional formats, teachers may be better positioned to initiate choral responses and gauge whether all students are responding and answering correctly.

Interestingly, higher rates of individual practice significantly predicted small negative gains or decreases on ASPENS and similar results were trending towards significance for RAENS. These findings run contrary to the results of previous observation research which suggest that how often students receive individual opportunities to practice with mathematical tasks and activities during Tier 1 mathematics instruction matters for increased student mathematics achievement (Clements et al., 2013; Doabler et al., 2015). Although the current negative effects of individual practice were small, they may suggest that individual practice does not have the same instructional value in small group (Tier 2) instruction compared to whole class (Tier 1) instruction. It is also plausible that individual practice in small groups contributes to a negative learning environment. For example, there may be stress or undue pressure involved for individual students, particularly students with MLD, with practicing mathematics skills in front of a small group of peers. This, in turn, could make individual practice in small groups less effective or even produce iatrogenic effects. Additional research is needed in this area.

Results also suggested that higher rates of academic feedback significantly predicted gains on all three mathematics outcome measures, with variance explained ranging from 2% to 6%. Research suggests that providing students with timely, informational feedback on their performance can decrease the likelihood of later misconceptions and address specific knowledge gaps (Hattie & Timperley, 2007; Vaughn & Swanson, 2015). In the current study, occurrences of academic feedback were documented when interventionists either addressed student errors or affirmed correct responses. It may be that when teachers immediately follow up practice opportunities with specific academic feedback in small-group interventions, they can improve students' opportunities to learn regardless for whether the prior response was correct.

Finally, our results suggested that higher quality explicit instructional interactions significantly predicted small gains or increases on all three outcome measures, with variance explained ranging from 2% to 3%. In recent years, a rapidly growing research base has documented a positive relation between the instructional quality of classrooms and improved student mathematics outcomes. Studies involving the Classroom Assessment Scoring System (CLASS; Pianta & Hamre, 2009), a validated a global-rating observation instrument, have reported much of these results (e.g., Howes et al., 2009; Mashburn et al., 2008). While our findings that higher quality explicit instructional interactions were associated with greater gains in mathematics align with the cogent work of Pianta and Hamre (2009), we contend the current research sheds light on a slightly different niche in the research area of instructional quality. Relative to the CLASS research (Pianta & Hamre, 2009), we took a more atomistic approach to documenting the quality of instruction, focusing specifically on principles of explicit instruction found beneficial for students who face MLD. While additional research on this approach is needed, the road to understanding the quality of instructional interactions that occur during Tier 2 mathematics instruction may run through investigations that are grounded in the preponderance of evidence supporting the use of explicit mathematics instruction (Agodini & Harris, 2010; Gersten et al., 2009; Hughes, et al., 2017).

Limitations and Directions for Future Research

We contend the findings of the current work yield important considerations for future research. However, several limitations should be considered when interpreting our results. One limitation is the outreach of our findings to other Tier 2 mathematics interventions. Data analyzed in the current study were generated from only one mathematics intervention at the kindergarten level. Given the recent spate of mathematics intervention research (National Center on Intensive Interventions, 2018), future research should investigate the active ingredients of other early elementary Tier 2 mathematics interventions. In particular, investigations of the quantity and quality of explicit instructional interactions may prove useful for understanding how and why these other interventions impact student mathematics achievement.

A second limitation relates to the number of observations conducted with each ROOTS group. Due to resource constraints in the larger ROOTS Efficacy Project, the ROOTS groups were limited to only three observations. Additional observations for each group may have provided a richer perspective of explicit instructional interactions. For example, teacher demonstrations had a low stability ICC across time (.14). This ICC suggests that approximately eight observations per small group are required to obtain a reasonable estimate of teacher demonstrations (Shoukri, Asyali, & Walter, 2004). This type of measurement error could explain the non-significant association between teacher demonstrations and student outcomes. It may also suggest that future researchers of this topic should plan for conducting approximately 10 observations per group. However, real-time direct observations can be expensive, particularly in large-scale efficacy trials where vast distances can separate participating research sites.

Relatedly, the extent of evidence for the quantity and quality of explicit instructional interactions associating with student mathematics outcomes is a third limitation. While several statistically significant results were reported, the magnitude of our findings was somewhat small. Thus, one could question whether these effects are substantively important to the field. However, given the push by prominent funding agencies for researchers to explore potential mediators of educational interventions during efficacy trials (IES & NSF, 2013), we argue our findings shed initial light on the active ingredients of ROOTS and serve as a guidepost for other researchers interested in investigating the black box of Tier 2 mathematics interventions. While beyond the scope of the current study, formal mediation analyses might be better poised to understand the influence of explicit instructional interactions on the mathematics outcomes of at-risk kindergarten students who receive the ROOTS intervention.

Implications for Practice

Our findings, while preliminary, provide a number of practical considerations for teachers who implement Tier 2 mathematics interventions in the early elementary grades. Unique to the current study was a contemporary approach to measuring explicit instructional interactions. We explored not only whether the quantity and quality of explicit instructional interactions associated with student mathematics outcomes, but also how group-level initial mathematics skill predicted the frequency and quality in which they experienced such interactions. While the current study refrained from updating ROOTS interventionists on student pretest performances on the whole number assessments, it is plausible that such information in the hands of teachers would allow them to better provide differentiated instruction through facilitation of explicit instructional interactions (Coyne et al., 2013). For example, knowing that a group of kindergarten students struggle with rational counting and the principle of cardinality (Geary et al., 2018), a teacher might offer the group more guided opportunities to practice matching single counting objects with a specific number word, and telling how many.

Additionally, we contend that our findings on the quantity of explicit instructional interactions are reported in metrics that can be easily understood by teachers. Extant observational research, particularly studies conducted in recent years, is typically known for its singular focus on the quality of instructional interactions (e.g., Pianta & Hamre, 2009). Yet, data on the frequency of instructional interactions may provide teachers with information that can be more easily translated into classroom practice. Consider for example our finding on group practice opportunities. Results suggested that providing approximately two group response opportunities per minute had beneficial impact on students' whole number performances. Facilitating this rate of group response opportunities during Tier 2 instruction seems quite manageable for teachers, particularly when they accompany it with an effective response signal, such a verbal prompt or a finger snap.

Conclusion

Over the past decade, educational researchers have produced a number of empirically-validated mathematics interventions. Collectively, these efforts have significantly advanced the field's understanding of effective instructional practices for teaching students with or at risk for MLD. To further extend the current knowledge base, continued research is needed to understand the mechanisms through which these educational interventions operate and produce desired student outcomes, such as how often and at what quality students at risk for MLD receive opportunities to learn foundational mathematics content. Unpacking the "black box" of empirically-validated mathematics interventions could allow the field to better meet the instructional needs of students who face early and persistent difficulties in mathematics.

References

Agodini, R., & Harris, B. (2010). An experimental evaluation of four elementary school math curricula. *Journal of Research on Educational Effectiveness*, *3*, 199–253. doi:10.1080/19345741003770693

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

Barnes, M. A., Klein, A., Swank, P., Starkey, P., McCandliss, B., Flynn, K., . . . Roberts, G. (2016). Effects of tutorial interventions in mathematics and attention for low-performing preschool children. *Journal of Research on Educational Effectiveness*, *1*, 155–178. doi:10.1080/19345747.2016.1191575

Berch, D. B. (2005). Making sense of number sense: Implications for children with mathematical disabilities. *Journal of Learning Disabilities*, *38*, 333–339.

Clarke, B., Doabler, C. T., Fien, H., Baker, S. K., & Smolkowski, K. (2012). *A randomized control trial of a Tier 2 kinder- garten mathematics intervention* (Project ROOTS, CFDA No. 84.324A, 2012-2016, Funding No. R324A120304). Washington, DC: U.S. Department of Education, Institute of Education Sciences, Special Education Research.

Clarke, B., Doabler, C. T., Kosty, D., Kurtz Nelson, E., Smolkowski, K., Fien, H., & Baker, S. K. (2017). Testing the efficacy of a kindergarten mathematics intervention by small group size. *AERA Open, 3(2)*, 1-16. doi:10.1177/2332858417706899

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

Clarke, B., Doabler, C. T., Smolkowski, K., Turtura, J., Kurtz- Nelson, E., Fien, H., & Baker, S. (2019). Exploring the relationship between initial math skill and a kindergarten mathematics intervention. *Exceptional Children*, *85*, 129–146. doi:10.1177/0014402918799503

Clarke, B., Gersten, R. M., Dimino, J., & Rolfhus, E. (2011). *Assessing student proficiency of number sense (ASPENS)*. Longmont, CO: Cambium Learning Group, Sopris Learning.

Clemens, N. H., Oslund, E., Kwok, O., Fogarty, M., Simmons, D. and Davis, J. L. (2019). Skill moderators of the effets of a reading comprehension intervention. *Exceptional Children, 85*, 197-211. doi: 10.1177/0014402918787339

Clements, D. H., Agodini, R., & Harris, B. (2013). *Instructional practices and student math achievement: Correlations from a study of math curricula* (NCEE Evaluation Brief No. 2013- 4020). Washington, DC: National Center for Educational Evaluation and Regional Assistance, Institute of Education Sciences.

Common Core State Standards Initiative. (2010). *Common core state standards for mathematics*. Retrieved from http://www.corestandards.org/assets/CCSSI_Math

Coyne, M. D., McCoach, D. B., Ware, S., Austin, C. R., Loftus-Rattan, S. M., & Baker, D. L. (2019). Racing against the vocabulary gap: Matthew effects in early vocabulary instruction and intervention. *Exceptional Children, 85*, 163-179. doi: 10.1177/0014402918789162

Coyne, M. D., Simmons, D. C., Hagan-Burke, S., Simmons, L. E., Kwok, O., Kim, M.,…Rawlinson, D. M. (2013). Adjusting beginning reading intervention based on student performance: An experimental evaluation. *Exceptional Children, 80,* 25-44.

Deshler, D. D. (2015). Moving in the right direction but at what speed, and how smoothly? *Remedial and Special Education*, *36*, 72–76. doi:10.1177/0741932514558093

Doabler, C. T., Baker, S. K., Kosty, D., Smolkowski, K., Clarke, B., Miller, S. J., & Fien, H. (2015). Examining the associa- tion between explicit mathematics instruction and student mathematics achievement. *Elementary School Journal*, *115,* 303–333.

Doabler, C. T., & Clarke, B. (2012). *Quality of explicit mathemat- ics instruction*. Unpublished observation instrument, Center on Teaching and Learning, University of Oregon, Eugene.

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

Doabler, C. T., Clarke, B., Kosty, D., Kurtz-Nelson, E., Fien, H., Smolkowski, K., & Baker, S. K. (2016). Testing the efficacy of a Tier 2 mathematics intervention: A concep- tual replication study. *Exceptional Children*, *83*, 92–110. doi:10.1177/0014402916660084

Doabler, C. T., Clarke, B., Kosty, D., Kurtz-Nelson, E., Fien, H., Smolkowski, K., & Baker, S. (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*, 168–180. doi:10.1177/00222194187893

Dyson, N. I., Jordan, N. C., & Glutting, J. (2011). A number sense intervention for low-income kindergartners at risk for mathematics difficulties. *Journal of Learning Disabilities*, *46*(2), 166–181. doi:10.1177/0022219411410233

Fuchs, D., & Fuchs, L. S. (2019). On the importance of moderator analysis in intervention research: An introduction to the special issue. *Exceptional Children, 85*, 126-128. doi: 10.1177/0014402918811924

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*, 147-162. doi: 10.1177/0014402918782628

Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., Nugent, L. (2018). Early conceptual understanding of cardinality predicts superior school-entry number-system knowledge. *Psychological Science, 29,* 191-205. doi: 10.1177/0956797617729817

Gersten, R., Chard, D. J., Jayanthi, M., Baker, S. K., Morphy, P., & Flojo, J. (2009). Mathematics instruction for students with learning disabilities: A meta-analysis of instructional components. *Review of Educational Research*, *79*, 1202–1242. doi:10.3102/0034654309334431

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–112. doi:10.3102/ 003465430298487

Howes, C., Burchinal, M., Pianta, R., Bryant, D., Early, D., Clifford, R., & Barbarin, O. (2008). Ready to learn? Children's pre-academic achievement in pre-kindergarten programs. *Early Childhood Research Quarterly*, *23*, 27–50.

Hughes, C. A., Morris, J. R., Therrien, W. J., & Benson, S. K. (2017). Explicit instruction: Historical and contemporary contexts. *Learning Disabilities Research & Practice, 32*, 140–148. doi:10.1111/ldrp.12142

Institute of Education Sciences and National Science Foundation (2013). *Common guidelines for education research and development*. Retrieved from http://ies.ed.gov/pdf/CommonGuidelines.pdf

Jordan, N. C., Glutting, J., Dyson, N., Hassinger-Das, B., & Irwin, C. (2012). Building kindergartners' number sense: A randomized controlled study. *Journal of Educational Psychology*, *104*(3), 647–660.

Jordan, N., 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). San Diego, CA: Academic Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

MacKinnon, D. P., & Luecken, L. J. (2008). How and for whom? Mediation and moderation in health psychology. *Health Psychology*, *27*(2s), 99–100.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., . . . Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, *79,* 732–749.

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*, 306–321. doi:10.1177/ 0022219408331037

National Center for Intensive Intervention (n.d.). Academic Intervention. Retrieved from http://www.intensiveintervention.org/

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*, 109–119.

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

Rosenshine, B. (Spring 2012). Principles of instruction: Research-based strategies that all teachers should know. *American Educator, 12-19, 39.*

Rosnow, R. L., & Rosenthal, R. (2003). Effect sizes for experimenting psychologists. *Canadian Journal of Experimental Psychology*, *57*(3), 221–237.

SAS Institute. (2016). *Base SAS® 9.4 procedures guide: Statistical procedures* (5th ed.). Cary, NC: SAS Institute, Inc. Retrieved from the SAS Product Documentation web site: http://support.sas.com/documentation/index.html

Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, *7*, 147–177.

Shoukri, M. M., Asyali, M. H., & Donner, A. (2004). Sample size requirements for the design of reliability study: Review and new results. *Statistical Methods in Medical Research, 13*, 251–271. doi:10.1191/0962280204sm365ra

Simmons, D. (2015). Instructional engineering principles to frame the future of reading intervention research and practice. *Remedial and Special Education*, *36*, 45–51.

Smolkowski, K., & Gunn, B. (2012). Reliability and validity of the Classroom Observations of Student-Teacher Interactions (COSTI) for kindergarten reading instruction. *Early Childhood Research Quarterly, 44*, 48–57.

Snijders, T., & Bosker, R. (2012). *Multilevel analysis: An introduction to basic and advanced multilevel modeling* (2nd ed.). Thousand Oaks, CA: Sage.

Vaughn, S., & Swanson, E. A. (2015). Special education research advances knowledge in education. *Exceptional Children*, *82*, 11–24.