Recruitment of Divisions and Schools Evaluation of the REL Appalachia Teaching Math to Young Children Toolkit

SUPPORTING STATEMENT
FOR PAPERWORK REDUCTION ACT SUBMISSION

PART B: Collection of Information Employing Statistical Methods

November 2023

Submitted to:
Institute of Education Sciences
U.S. Department of Education

Submitted by:
SRI International
333 Ravenswood Ave
Menlo Park, CA 94025
(650) 859-2000

Tracking and OMB Number: (XX) XXXX-XXXX
Revised: 11/14/2023

Overview

The U.S. Department of Education (ED), through its Institute of Education Sciences (IES), requests clearance for the recruitment materials under the Office of Management and Budget (OMB) clearance agreement (OMB Number (XX) XXXX-XXXX) for activities related to the Regional Educational Laboratory Appalachia (REL AP) program. A second OMB package, which will be submitted later this year, will request clearance for data collection instruments and the collection of district administrative data.

Mathematics knowledge acquired in early childhood provides a critical foundation for long-term student success in math as well as reading (Duncan et al., 2007; Watts et al., 2014), but the professional development (PD) and curricular support for preschool teachers often lack specific content and training on high-quality math instruction delivered by math content experts. To address this problem, REL Appalachia is developing a toolkit to help preschool teachers implement core teaching practices essential to promoting early math skills and knowledge in children. The toolkit is based on the IES *Teaching Math to Young Children* practice guide (Frye et al., 2013) and is being developed in collaboration with state and district partners in Virginia.

REL AP is requesting clearance to conduct an independent evaluation that will assess the efficacy of the school-based professional development resources included in the toolkit. The evaluation will also assess how teachers and facilitators implement the toolkit to provide context for the efficacy findings and guidance to improve the toolkit and its future use. The evaluation will take place in 50 schools across 10 divisions in Virginia and focus on early mathematics support for preschool teachers.

This package only requests clearance for data collection related to recruitment activities. A separate OMB package will request clearance for data collection procedures and activities related to addressing the study research questions (RQs). Additional details about the study goals and design are included below for context.

The impact and implementation RQs addressed in this study include the following:

1. Do teachers in intervention-assigned schools (that is, teachers who are offered the toolkit PD resources) report greater confidence in, and positive attitudes toward, using evidence-based practices in math compared to teachers in control-assigned schools?

2. Do teachers in intervention-assigned schools implement more math activities, spend more time on math through daily instruction, and include more instruction across settings and activities?

3. Do teachers in intervention-assigned schools demonstrate more frequent use of evidence-based math teaching practices than teachers in control-assigned schools?

4. Do preschool students in intervention-assigned schools score higher on measures of math achievement in the spring of preschool than students in control-assigned schools?

5. Did the professional development components of the toolkit implementation, classroom activities, and instruction occur as intended?

6. What are different ways that teachers engage with the toolkit PD resources? To what extent does teachers' use of the PD resources vary? What helps or hinders effective learning from the PD resources?

7. What challenges do teachers face in implementing the toolkit and how do teachers attempt to overcome those challenges? What additional supports are needed and what improvements do participants recommend for the toolkit?

B1. Respondent Universe and Sample Design

The evaluation will employ a school-level, cluster-randomized controlled design, and take place in 50 schools in approximately 10 divisions in the state of Virginia. The evaluation will examine the impact of the toolkit on student math achievement. The inference population is schools and preschool teachers in low-resourced communities. Specifically, the evaluation will examine whether the intervention impacts: Teachers' confidence in, and positive attitudes toward, using evidence-based practices in math; teachers' implementation of math activities and time spent on math through daily instruction and instruction across settings and activities; teachers' use of evidence-based math teaching practices; and students' math achievement. The evaluation will also address several questions about implementation including: Whether the professional development components of the toolkit implementation, classroom activities, and instruction occur as intended; how teachers engage with the toolkit PD resources; the extent to which teachers' use of the PD resources varies; what helps or hinders effective learning from the PD resources; what challenges teachers face in implementing the toolkit and how teachers attempt to overcome those challenges; and what additional supports are needed and what improvements do participants recommend for the toolkit.

For recruitment purposes (the focus of this OMB package), the toolkit evaluation team will recruit school divisions of different sizes that have at least one preschool, prioritizing divisions with the highest percentages of students the Virginia Department of Education (VDOE) identifies as economically disadvantaged. Evaluating the toolkit's impact on teachers and students in low-resourced communities provides an opportunity to assess how this toolkit may improve early numeracy in this historically underserved population. VDOE identifies students as economically disadvantaged if they meet any of the following criteria: (1) eligible for free or reduced-price meals, (2) receive Temporary Assistance for Needy Families, (3) eligible for

Medicaid, or 4) identify as either migrant or experiencing homelessness, foster, or Head Start. The study team identified 12 divisions serving the highest percentages of students from low-resourced communities based on VDOE data from the 2022/23 school year. If obtaining enough schools from these divisions is not possible, we will recruit schools from the next most economically disadvantaged divisions until we reach our target of 50 schools.

The focus of the efficacy study will be preschool teachers in school-based settings under the auspices of a school division or Head Start program (that is, publicly funded preschool classrooms) and their students. We conducted power analyses estimating the minimum detectable effect size (MDES) for both student and teacher outcomes at the school level, given the nested structure of the data. Both analyses were performed with PowerUp! software (Dong & Maynard, 2013), using the MDES formula (Dong & Maynard, 2013, pp. 55–56). For the student math outcome measure, we used 4-Level Fixed Effects Blocked Cluster Random Assignment Designs (BCRA4_3f) — Intervention at Level 3 — to estimate the MDES with an assumption of an estimated school-level and teacher-level intraclass correlation coefficient (ICC) of 0.10 provided in the published estimates for kindergarten and student, teacher, and school and division R^2 of 0.5 (Bloom et al., 2007; Hedges & Hedberg, 2007). The study will include 50 schools (25 intervention and 25 control) from approximately 10 divisions, with an estimated two preschool teachers per school and 20 students per teacher after attrition. The power analysis assumes 10 percent attrition of students due to absences on assessment days and movement of students out of the participating divisions or out of state schools. Use of a type I error rate of .05 with a two-sided test of significance, with 80 percent power and with the intervention at the school level, yields an MDES of 0.24 for student math achievement outcomes.

Because of the lack of prior research on variance decomposition of teacher outcomes in preschool classrooms across schools, student variance estimates will be generalized to teacher outcomes. We additionally assume 5 percent teacher attrition during the study year based on our previous preschool evaluation work in Virginia. Using the same assumptions we applied to the student-level power analyses, the MDES is 0.43 for teacher outcomes. Although this MDES is high compared to some other education research studies, it is appropriate for this study. Previous PD evaluation studies conducted by the developers of the toolkit found significant impacts on preschool to grade 3 teachers' confidence (effect size [ES] ranges from 0.51 to 0.71) and instructional quality (ES ranges between 0.65 and 1.01) (Reid et al., 2020).

B2. Information Collection Procedures

a. Notification of the Sample and Recruitment

The study team will call or email preschool program and division administrators in the targeted divisions to inform them about the study. If division leaders are interested in participating, the evaluation team will ask them for help contacting schools and ideas for how the study might be a fit for their schools.

Upon division agreement, the team will contact all the division's school leaders in schools with at least one preschool classroom and invite them to learn more about the study and what will be required of school and division staff. Once school leaders agree to participation by their school

and teachers, the study team will schedule a series of webinars (at least one per division) with preschool teachers in participating schools to explain the study purpose, benefits, and time commitment. The team will then invite the teachers to join the study and data-collection activities by asking them to review and sign an online consent form, prior to random assignment. This nonbinding agreement will indicate that they understand the intervention and the study and will participate to the best of their ability, regardless of the condition to which they are assigned. Schools will be included in the random assignment pool if at least one preschool teacher in that school consents to participate in the study. Once the teachers have consented, the study team will follow division consent procedures for parents/guardians. The intention is to engage divisions that allow passive consent procedures for parents/guardians to opt out of their child's participation in the study if they choose. However, if divisions do not allow passive consent procedures, we will follow the division procedures for active consent. Based on prior studies that SRI has conducted in pre-K classrooms, we expect the SRI IRB to classify this study as exempt because it involves established educational settings and involves normal educational practices that are not likely to adversely impacts students' opportunities to learn required educational content. With this, we expect that passive consent will be sufficient for this study. Recruitment materials that will be used for this study are included in appendix A.

Data collection will occur in schools that consent to be part of the study. Each round of data collection will include an initial outreach and three followup emails for participants who have completed the consent form. Data-collection communication email texts will be submitted in a separate OMB package.

b. Statistical Methodology for Stratification and Sample Selection

The recruitment activities for which the study team requests clearance in this first OMB package will not be directly tabulated and published but rather will be used to facilitate sample selection for the efficacy study’s data collection activities.

For context, the study team has provided information below on the sample assignment the efficacy study will use.

Because the toolkit, centering on teacher knowledge and use of the practice guide recommendations for improving preschool students’ numeracy, is designed to be used as part of a school’s approach to improve math instruction by all preschool teachers within a school, the evaluation team proposes using the school as the unit of assignment. This level of assignment has multiple methodological benefits, including removing the within-school threat of diffusion of the toolkit use and crossovers.

Prior to randomization, the study team will create blocks of schools based on school size and/or percentage of student population identified as economically disadvantaged (if schools in a division vary on this variable), and the team will then randomly assign schools within these blocks to either participate in the intervention or serve as the business-as-usual control. If divisions with both Head Start and PreK classrooms sign up for the study, we will also create blocks and randomly assign schools within these blocks. Random assignment will occur approximately two to three weeks before the 2024/25 school year starts to respect teachers’ need

for time to prepare and plan for the upcoming year. Teachers will know their school has a 50 percent chance of being assigned to the intervention condition, and the study team will explain what will be required if teachers are assigned to the intervention in terms of meeting time, planning for instruction, and participation in the PD modules in addition to incorporating what they learn into their existing curricula, classroom instruction, and routines. Conducting random assignment before the school year (instead of after the school year begins) allows teachers to prepare and reserve time for the professional development activities soon after school begins. If we were to conduct random assignment earlier in the summer or spring, we would risk higher attrition of both teachers and students, especially given the timing of preschool enrollment, which typically is not firmly set until the first few weeks of the school year. At the time of random assignment, teachers and school principals will be notified of their study condition, and the toolkit (i.e., PD resources and implementation supports) will be made available to all participating intervention teachers within three weeks of the start of the school year. Control group teachers will be provided access to the toolkit after the study period has concluded.

c. Estimation Procedures

The recruitment activities for which the study team requests clearance in this first OMB package will not be directly tabulated and published but rather will be used to facilitate sample selection for the efficacy study's data collection activities.

For context, the study team has provided information below on the estimation procedures the efficacy study will use to estimate the impacts of the toolkit on student and teacher outcomes as well as to conduct the implementation analysis. Additional detail on these procedures will be provided as part of the second OMB package, focused on data collection.

To estimate the impact of the toolkit PD resources on teacher outcomes, the study team will use hierarchical linear modeling (HLM) to adjust standard errors associated with the clustering of observations within schools and divisions (Raudenbush & Bryk, 2002).

***Continuous Teacher Outcomes*.** The study team will examine the impact of the toolkit on post-intervention outcomes in the following teacher measures: math confidence and attitudes as measured by the Attitudes, Beliefs, and Confidence Survey (ABC, summary score for confidence items and summary score for attitudes items) (Reid & Melgar, 2018), quantity of math instruction as measured by the Instructional Log (total number of minutes of math lessons and activities will be summed across the logs to provide a daily and weekly average number of minutes across the two conditions), and quality of math instruction as measured by the TPOT (summary continuous score), controlling for baseline scores on each measure. HLM models estimating teacher-level impact will need to account for the nesting of teachers within schools. We plan to estimate multilevel models with two levels: teachers and schools, and we will also calculate and examine intraclass correlation coefficients to determine whether adding an additional level accounting for division-level nesting is necessary. Models will be estimated separately for each teacher outcome. The models will take the following form:

Level 1: Teachers

$$Y_{jk} = \beta_{0k} + \beta_{1k}Tch_{jk} + r_{jk}$$

where $_{0k}$ is the random adjusted mean outcome score of teachers in school k, Tch_{jk} is a vector of teacher-level baseline characteristics (education, years teaching, baseline score on the same outcome variable), and r_{jk} is a teacher-specific error term.

Level 2: Schools

$$_{0k} = \gamma_{00} + \gamma_{01}Schl_k + \gamma_{02}B_k + \gamma_{03}T_k + u_{0k}$$

where γ_{00} is the random adjusted mean outcome score across schools, $Schl_k$ is a vector of school-level baseline characteristics (urbanicity, percentage of students considered economically disadvantaged), B_k is a set of fixed effects for the randomization blocks, T_k is a binary intervention indicator, so that γ_{03} is the main effect of the intervention on outcome scores, and u_{0k} is a school-specific error term.

***Continuous Student Outcomes*.** To estimate the impact of the toolkit PD resources on students' Virginia Early Math Assessment System (EMAS) scores, controlling for baseline scores, we will use HLM to adjust standard errors associated with the clustering of observations within classrooms and schools (Raudenbush & Bryk, 2002). Additionally, student (age, gender, Individualized Education Program [IEP] or multilingual learner [MLL] status, fall EMAS score), teacher/classroom (education, experience), and school covariates will be included to reduce residual error and increase power.

Level 1: Students

$$Y_{ijk} = \pi_{0jk} + \pi_{1jk}St_{ijk} + e_{ijk}$$

where π_{0jk} is the random adjusted mean math outcome score of students with teacher j in school k, St_{ijk} is a vector of student-level baseline characteristics (age, gender, fall EMAS score), and e_{ijk} is a student-specific error term.

Level 2: Teachers

$$\pi_{0jk} = \beta_{00k} + \beta_{01k}Tch_{jk} + r_{0jk}$$

where β_{00k} is the random adjusted mean math outcome score of students in school k, Tch_{jk} is a vector of teacher-level baseline characteristics (education, years teaching preschool or kindergarten), and r_{0jk} is a teacher-specific error term.

Level 3: Schools

$$\beta_{00k} = \gamma_{000} + \gamma_{001}Schl_k + \gamma_{002}B_k + \gamma_{003}T_k + u_{00k}$$

where γ_{000} is the random adjusted mean math outcome score of students across schools, $Schl_k$ is a vector of school-level student baseline characteristics (school means), B_k is a set of fixed effects

for randomization blocks, T_k is a binary intervention indicator, so that γ_{003} is the main effect of the intervention on math scores, and u_{00k} is a school-specific error term.

***Strategies for Correcting for Multiple Hypothesis Testing*.** The proposed analysis does not require correction for multiple hypothesis testing, as it includes only one confirmatory analysis comparison within the math domain (treatment vs. control outcome for all students).

Implementation Fidelity. This evaluation will examine implementation of the key school-based components of the toolkit, which are hypothesized as the primary mechanism for improving teacher knowledge, instruction, and student reading comprehension outcomes. The key components of the toolkit intervention — integrating both activities and materials — are planning, learning, and institutionalizing. Because these are three distinct phases of implementation, the evaluation team will look at implementation for each of these three components, constructing separate measures for planning, learning, and institutionalizing. To analyze the implementation fidelity research questions, the evaluation team will examine means, standard deviations, and frequencies (percentages) of items on the Implementation Checklist, track SAMI completion, and study responses to the ease, usability, and satisfaction questions on the Toolkit Satisfaction Survey. To the extent possible, we will examine how variations in teacher characteristics (education, experience) are associated with implementation fidelity.

Implementation Treatment Contrast. To understand the differences between the professional development experiences of teachers in the two groups, the research team will conduct descriptive analysis of use of toolkit-like activities. The study team will gather information on the non-intervention curricular material and PD resources available to and used by the control teachers by asking questions about these topics in conjunction with the administration of the online ABC Survey, collected post-test. Information about control teachers' curricular material and PD resources and experiences will be used to describe the contrast between the intervention and control conditions. To understand the differences in math teaching that students in the two groups experience, we will look at teacher outcomes on the Instructional Log and TPOT measures. Following completion of the efficacy study and publication of the toolkit PD resources online, teachers in control-assigned schools will be offered the materials and resources.

Implementation Challenges. To identify challenges for completing toolkit activities, the evaluation team will collect and analyze data from teachers, school leaders and facilitators, and district leaders. The survey will ask questions about potential challenges for toolkit implementation, and the answers will be summarized. The research team will conduct qualitative analyses of the open-ended responses on the Toolkit Satisfaction Survey to examine how teachers used the toolkit resources, how teachers approached challenges, and any suggested improvements teachers have for the toolkit. The evaluation team will analyze these responses qualitatively, coding each interview in Dedoose (Dedoose, 2021) and developing analytic summaries for the most commonly occurring challenges and supports.

d. Degree of Accuracy Needed

We conducted power analyses estimating the minimum detectable effect size (MDES) for both student and teacher outcomes at the school level, given the nested structure of the data. Both analyses were performed with PowerUp! Software (Dong & Maynard, 2013), using the MDES formula (Dong & Maynard, 2013, pp. 55–56). For the student math outcome measure, we used 4-Level Fixed Effects Blocked Cluster Random Assignment Designs (BCRA4_3f) — Intervention at Level 3 — to estimate the MDES, with an assumption of an estimated school-level and teacher-level ICC of 0.10 provided in the published estimates for kindergarten and student, teacher, and school and division R^2 of 0.5 (Bloom et al., 2007; Hedges & Hedberg, 2007). The study will include 50 schools (25 intervention and 25 control) from approximately 10 divisions, with an estimated two preschool teachers per school and 20 students per teacher after attrition. The power analysis assumes 10 percent attrition of students due to absences on assessment days and movement of students out of the study schools. Use of a type I error rate of .05 with a two-sided test of significance, with 80 percent power and with the intervention at the school level, yields an MDES of 0.24 for student math achievement outcomes.

Because of the lack of prior research on variance decomposition across schools of teacher outcomes in preschool classrooms, student variance estimates will be generalized to teacher outcomes. We additionally assume 5 percent teacher attrition during the study year based on our previous preschool evaluation work in Virginia. Using the same assumptions we applied to the student-level power analyses, the MDES is 0.43 for teacher outcomes. Although this MDES is high compared to some other education research studies, it is appropriate for this study. Previous PD evaluation studies conducted by the developers of the toolkit found significant impacts on preschool to grade 3 teachers' confidence (effect size [ES] ranges from 0.51 to 0.71) and instructional quality (ES ranges between 0.65 and 1.01) (Reid et al., 2020).

e. Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems requiring specialized sampling procedures.

f. Use of Periodic (Less than Annual) Data Collection to Reduce Burden

This project will collect data once for recruitment.

The data collection activities and their frequency will be included in a future package for OMB clearance. Outcome data will need to be collected more frequently than annually because the evaluation is occurring within one school year, and some measures will need to be assessed at the beginning, middle, and end of that same school year. A longer period between data collection would make it difficult for the study team to meet the requirements for the efficacy study (by preventing baseline and followup data collection in the timeframe necessary for the evaluation). However, the study period will be one school year; there will not be recurring data collection in future school years.

B3. Methods for Maximizing the Response Rate

The study team does not anticipate problems contacting, gaining the cooperation of, and gathering information from division leaders during the recruitment activities. The study team will conduct outreach to division contacts via email and follow up with phone calls as necessary. The study team will conduct calls with division leaders during business hours at times that

coincide best with their schedules. The study team will also be flexible during recruitment, allowing divisions to provide the requested information either over the phone or by email.

For the data collection activities that will be included in the second OMB package, the study team is committed to obtaining complete data for this evaluation. A key to attaining complete administrative data is tracking the data components from each division with email and telephone contacts to the appropriate parties to resolve issues of missing or delayed data files. All administrative data files will be reviewed for consistency and completeness. If a data file has too many missing values or if an instrument in the implementation study has too few items completed to be counted as a response, the evaluation team will seek to obtain more complete responses by email or phone.

Based on our previous preschool evaluation work in Virginia, the evaluation team expects the response rate to be about 95 percent for teachers who consent to participate in the study. We will contact non-responding teachers up to four times to encourage participation. Three followup email reminders will be sent to individual teachers if responses are not obtained for online surveys. The evaluation team will consider other modes of follow-up, including reminder letters and phone calls if response rates are below expectation.

In addition, a number of steps will be taken to maximize response rates. For example, respondents will receive advance communications explaining the study, introducing REL AP, providing an assurance of confidentiality, and encouraging them to participate to help refine the toolkit. Respondents also will be given a contact number to reach the evaluation team with questions. Data collection (e.g., for the surveys and Instructional Logs) will be completed using web-based technology to minimize burden for participants.

Finally, respondents will receive an incentive for participating in the study. Teachers in the intervention group will either be directly paid $40 per hour of PD if completed outside of work hours (up to $800 total across the life of the study with an estimated 20 hours to participate in the professional learning activities), or the study team will provide funds to the school or division to pay for substitute days to allow intervention teachers to complete the PD modules. Intervention teachers will also be given $50 per data-collection wave for completing the implementation surveys. All intervention and control teachers will also be compensated for their time spent completing the surveys and teacher Instructional Logs ($50 per completed data-collection wave). The study team will also provide each division with $400 for its assistance with data exports from administrative data systems.

The evaluation team has multiple strategies to deal with missing data due to non-response. Prior to starting the analyses, the evaluation team will examine the extent of missing data overall and by treatment group. Starting from the What Works Clearinghouse (WWC, 2022) recommendations for dealing with missing data, researchers on the team will use appropriate analytic methods to account for missing data and will consider options such as complete case analyses with regression adjustment, maximum likelihood methods, or non-response weights. Implementation of the approach will follow requirements such as using one of the WWC acceptable approaches and assessing the analysis sample for low attrition before applying the acceptable missing data approach. The most recent statistical literature will also be considered to

examine additional methods. If such methods are necessary, results using data not adjusted for missingness will also be included in an appendix to the report.

B4. Test of Procedures

The recruitment activities for which the study team requests clearance in this first OMB package will not be directly tabulated and published but rather will be used to facilitate sample selection for the efficacy study's data collection activities. Additional detail about data collection activities and procedures will be included in a future package for OMB clearance.

The student outcome measure used to analyze the impact of the toolkit will be the Virginia Kindergarten Readiness Program (VKRP) Early Math Assessment System (EMAS), which has demonstrated validity and reliability and will not require pretesting. Teachers will complete the Attitudes, Beliefs, and Confidence (ABC) Survey to capture their attitudes and beliefs. The ABC Survey has demonstrated validity and reliability and will not require pretesting. Teacher Instructional Logs, used to capture data on teachers' instructional practices relating to mathematics, will be pilot tested in 2023 with fewer than nine respondents. If the school or preschool program participates in Virginia's Quality Rating Improvement System, we will obtain teachers' Classroom Assessment Scoring System (CLASS)® scores to provide a measure of overall teaching quality practices in addition to the math-specific practices. The instruments will be submitted in a separate package for OMB approval.

B5. Names and telephone numbers of individuals consulted on statistical aspects of the design and the names of the contractors who will actually collect or analyze the information for the agency

The study's analytical plans were reviewed by the Regional Educational Laboratory Peer Review contract with IES, which supports quality assurance of REL applied research studies. In addition, the following people were consulted on the statistical aspects of the study:

Todd Grindal (SRI) served as the subject matter expert. Mary Klute (SRI) and Julie Harris (SRI) were consulted on the design and statistical aspects of this study. The following staff are responsible for collecting and analyzing the study data:

Name	Project role	Organization	Phone number
Erika Gaylor	Evaluation Lead	SRI	650 859 2110
Sarah Nixon Gerard	Evaluation Co-Lead	SRI	703 247 8545
Mary Klute	REL AP Research Lead	SRI	650 859 2380
Julie Harris	REL AP Deputy Research Lead	SRI	703 247 8619
Marta Mielicki	Project Coordinator	SRI	703 247 8430

References

Bloom H. S., Richburg-Hayes, L., & Black A. R. (2007). Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educational Evaluation and Policy Analysis, 29(1),* 30–59. doi:10.3102/0162373707299550

Dedoose Version 9.0.17, cloud application for managing, analyzing, and presenting qualitative and mixed method research data (2021). Los Angeles, CA: SocioCultural Research Consultants, LLC www.dedoose.com.

Dong, N., & Maynard, R. (2013). PowerUp! A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. *Journal of Research on Educational Effectiveness*, *6*(1), 24–67. doi:10.1080/19345747.2012.673143

Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., & Klebanov, P. (2007). School readiness and later achievement. *Developmental Psychology*, *43*, 1428−1446. https://doi.org/10.1037/0012-1649.43.6.1428

Frye, D., Baroody, A. J., Burchinal, M., Carver, S. M., Jordan, N. C., & McDowell, J. (2013). *Teaching math to young children practice guide* (NCEE 2014-4005). U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. https://ies.ed.gov/ncee/wwc/practiceguide/18

Hedges, L. V., & Hedberg, E. C. (2007). Intraclass correlations and covariate outcome correlations for planning two-and three-level cluster-randomized experiments in education. *Evaluation Review*, *37*(6), 445–489.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data\ analysis methods* (vol. 1). Sage Publications.

Reid, E. E., McCray, J. S., Gaylor, E., & Dominguez, X. (2020). “Innovations in early mathematics professional development: Benefits to teachers.” Poster presentation at the Society for Research in Educational Effectiveness (SREE) Spring 2020 conference, online program, March 11–14, 2020, Crystal City, VA.

Reid, E. E., & Melgar, C. (2018, March). “A center-based approach to changing teacher math attitudes in Head Start centers.” Poster presented at the annual Erikson Research Symposium. Chicago: Erikson Institute, Early Math Collaborative

Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. (2014). What’s past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher, 43(7),* 352–360.

What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook*, version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance (NCEE). https://ies.ed.gov/ncee/wwc/Handbooks

Appendix A – Recruitment Materials

- Appendix A1 – School Division Recruitment Flyer and Follow-up Phone Call Talking Points
- Appendix A2 – School Leader Recruitment Email, Follow-up Phone Call Talking Points, and Agenda for School Staff Informational Webinar
- Appendix A3 – Teacher Recruitment Email and Consent Form