Evaluation of Strategies to Address Unfinished Learning in Math (ReSolve Math Study)

Supporting Statement for Paperwork Reduction Act Submission
PART B: Collection of Information Employing Statistical Methods

May 2023

Contract 91990021D0001 (Task Order 91990021F0357)

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

Submitted by:
MDRC
200 Vesey Street, 23rd Floor
New York, NY 10281-2103
Tel: 212-532-0832
Fax: 212-684-0832

Table of Contents

Appendix Contents

Table of Exhibits

PART B. Collection of Information Employing Statistical Methods

Introduction

The Institute of Education Sciences (IES) within the U.S. Department of Education (ED) requests clearance to conduct new data collection activities for the *Evaluation of Strategies to Address Unfinished Learning in Math*. The evaluation (also referred to as the "ReSolve Math Study") will determine if intensive and consistent student use of digital math products during regular class time improves 4th and 5th grade students' math achievement. The study will also assess which of two possible instructional approaches that the digital products use to address unfinished learning is more effective at helping students catch up in math. This package requests clearance for study data collection activities, including: (a) teacher surveys; (b) school leader surveys; (c) extant school records and staff directory information from the study districts; and (d) parent/guardian consent forms for school records release. ED also plans to collect follow-up extant school records to assess long-term math outcomes for students in the study, but the clearance request for that collection will come as a separate package at a later date.

B.1. Respondent universe and sampling methods

The target population of schools for the study, from which a sample will be recruited, consists of all regular U.S. public schools that initially meet the following criteria:

- Are located in the continental U.S., to facilitate in-person implementation support, if needed;
- Serve grades 2 to 5, the grade levels in which the interventions will be implemented. (While the focus of the study will be on 4th and 5th graders, schools assigned to the "business as usual" group will have the opportunity to implement the products in the lower grades.);
- Demonstrate high academic need, having 50% or more of 4th graders scoring below proficient on the state math assessment and at least 30 students in 4th grade who scored below proficient on the state math test, to ensure that each school can provide a sufficiently large sample of students with unfinished learning to the study;
- Are not intensively using other math technology products—in grades 4 and 5—that are the same or similar to the study products.

Additionally, participating schools must be located in districts where at least 2 of the schools that meet the criteria above are also designated as Title I eligible to ensure that the study can achieve its purpose of informing ESSA Title I programs.

B.1.1. Recruitment and selection of districts and schools

From the initially identified target population described above, the study team expects to recruit 150 schools needed for the evaluation, split across two Cohorts. Based on study resources and past experience, the expectation is that the 150 schools will be distributed across approximately 15 school districts.

The study team will first reach out to districts that appear to have schools that meet the eligibility criteria above. After initial conversations appear to confirm eligibility and interest, the study team will then reach out to the district's potentially eligible schools to verify that they meet the study's eligibility criteria and establish whether they are willing and interested to participate in the study. Recruitment of districts/schools will be completed by the end of Summer 2023 for Cohort 1 and the end of Summer 2024 for Cohort 2.

While the study team will prioritize having a diverse study sample in terms of geography, locale, school size, racial composition, school performance, and business-as-usual conditions, among others, to help facilitate generalizability of the study findings, the primary goal of recruitment is to ensure sufficient sample size and statistical power for answering the study's research questions. As a result, the study sample will ultimately be a purposeful sample of districts/schools that volunteer to participate and are willing and able to meet the study requirements, including:

- Willingness to implement for two school years one of two digital math products selected for the study: Curriculum Associates LLC's i-Ready® Assessment and Personalized Instruction (i-Ready), and Renaissance Learning's Star Math and Freckle. Both products will have the ability to deliver math instruction to students using one of two instructional approaches or "modes": (1) instruction that supports building mastery in all unfinished learning (Broad Foundation Skill Building; BFSB) and (2) instruction that targets unfinished learning just-in-time on only the prerequisite concepts hypothesized to be necessary to succeed in grade-level content (Focused Just-In-Time Skill Building; FJIT). (See Supporting Statement A for more details on the digital math products and instructional approaches.);
- Willingness in both study years to work with the study team and the product vendors to participate in training and technical assistance for teachers and school administrators to support their use of the digital math products;
- Willingness to facilitate and implement the data collection needed to answer the evaluation's research questions;
- Willingness to allow school and student-level random assignment to study conditions.

B.1.2. Random assignment of schools and students

Once a district and its schools are recruited, the following procedures will be used to randomly assign eligible schools and students:

- **Stage 1: School-level random assignment:** Each recruited school district will be asked to select their preferred digital product (i-Ready Personalized Learning or Freckle). Immediately after recruitment (Spring-Summer 2023 for Cohort 1 and Spring-Summer 2024 for Cohort 2), the study team will randomly assign the study schools in the district to one of two research groups: a program group that will use the selected product in all 4th and 5th grade math classrooms for two school years (SY 2023-24 and 2024-25 for Cohort 1 and SY 2024-25 and 2025-26 for Cohort 2) <u>or</u> a "business-as-usual" (BAU) group that will continue to use the same instructional strategies they had been using in 4th and 5th grade. BAU schools will be given the option of using the product in their 2nd grade classrooms (non-study grade).

This random assignment will allow the study team to compare outcomes between BAU and program schools to determine the effectiveness of using the study's digital math products, which is a key research question for the study (See Supporting Statement A for more details on the study's research questions). Random assignment will be blocked by district, and the study team will use a 2:1 ratio when assigning schools to research groups – that is, one-third of the schools will be assigned to the BAU group, while the other two-thirds will be assigned to the program group, in large part to maximize precision for answering another key evaluation question, described below in Stage 2.

- **Stage 2: Student-level random assignment:** At the start of the first year of implementation (Year 1: SY 2023-24 for Cohort 1 and SY 2024-25 for Cohort 2), the study team and vendors will work with staff in the program schools to create and upload rosters of their 4th and 5th grade students to the product platforms, as is typical when schools are using a digital product. These rosters will be used to create student accounts and to randomly assign students to receive one of two instructional approaches or "modes" from the product. One group of students will receive the BFSB approach from the digital product and the other group will receive the FJIT instructional approach from the digital math product. Random assignment will be blocked by classroom, and students will be assigned to the two instructional approaches using a 1:1 ratio. In the second year of implementation (Year 2: SY 2024-25 for Cohort 1 and SY 2025-26 for Cohort 2), the study team and vendors will work with schools to upload updated 4th and 5th grade student rosters. Students who have already been randomly assigned (as 4th grade students in Year 1) will continue to receive the same instructional approach in 5th grade. Students who are new to the study (4th graders and students who transfer into the school in 5th grade) will be randomly assigned using the same approach as in Year 1. This random assignment of students within program schools will allow the study team to compare the outcomes of students using the BFSB mode and students using the FJIT mode of the same digital product, to determine if one instructional approach is more effective than the other, which is another key research question for the study (See Supporting Statement A for more details on the study's research questions).

B.2. Statistical methods for sample selection and degree of accuracy needed

B.2.1. Statistical methods for sample selection

As noted earlier, the selection of the study sample of districts and schools from the target population will aim to provide a generalizable sample but will not formally be a random (probability) sample. For all data collection activities for the evaluation itself, the study team will not use statistical sampling methods but will collect data from the census of respondents within the study sample. In other words, all participating districts, schools, 4th and 5th grade math teachers, and their students will be part of each relevant data collection. This approach will produce reliable estimates for the evaluation's research questions within the study sample, although these results may have limited generalizability to a larger population of schools and districts, depending on what the final recruited study sample of schools and districts looks like.

Exhibit B.1 provides an overview of the estimated respondent universe and sampling method to be used for each data collection activity for which the study team is requesting OMB clearance, as well as expected response rates. Data collection activities not requiring OMB clearance are also shown in Exhibit B.1 for context. Exhibit A.3 in Supporting Statement Part A provides additional information about the purposes and uses of the data.

Exhibit B.1. Respondent universe, sample, and expected response rate by data source

Data source	Timing	Respondent	Respondent universe (estimated per round)	Sampling approach	Expected response rate
Fall math teacher survey	Year 1 for all teachers (Fall 2023 for Cohort 1; Fall 2024 for Cohort 2)	Math teachers of 4th and 5th grade students in all participating schools	1,200 in Year 1 (Fall 2023 for Cohort 1; Fall 2024 for Cohort 2)	Census	85%
	Year 2 (Fall 2024 for Cohort 1; Fall 2025 for Cohort 2)	Math teachers of 4th and 5th grade students in all participating schools that are new to the school and study	240 in Year 2 (Fall 2024 for Cohort 1; Fall 2025 for Cohort 2)		
Spring math teacher survey	Year 1 (Spring 2024 for Cohort 1; Spring 2025 for Cohort 2); Year 2 (Spring 2025 for Cohort 1; Spring 2026 for Cohort 2)	Math teachers of 4th and 5th grade students in all participating schools	1,200	Census	85%
School leader survey	Year 1 (Spring 2024 for Cohort 1; Spring 2025 for Cohort 2); Year 2 (Spring 2025 for Cohort 1; Spring 2026 for Cohort 2)	School leader most familiar with math instruction in each participating school	150	Census	85%
District directory data about school staff	Year 1 (Fall 2023 and Spring 2024 for Cohort 1; Fall 2024 and Spring 2025 for Cohort 2); Year 2 (Fall 2024 and Spring 2025 for Cohort 1; Fall 2025 and Spring 2026 for Cohort 2)	One staff member at each district to provide staff directory data	15	Census	100%
District administrative records on students	Year 1 (Fall 2024 for Cohort 1 and Fall 2025 for Cohort 2); Year 2 (Fall 2025 for Cohort 1 and Fall 2026 for Cohort 2)	One staff member at each district to provide student records	15	Census	100%

Data source	Timing	Respondent	Respondent universe (estimated per round)	Sampling approach	Expected response rate
Parent or guardian consent forms for student records release	Year 1 (Fall 2023 for Cohort 1 and Fall 2024 for Cohort 2)	All 4th and 5th grade students in all participating schools	22,500	Census	95%
	Year 2 (Fall 2024 for Cohort 1 and Fall 2025 for Cohort 2)	All 4th grade students enrolled in all participating schools; all 5th grade students who are new to the schools	13,613		
Product data	Year 1 (Summer 2024 for Cohort 1 and Summer 2025 for Cohort 2); Year 2 (Summer 2025 for Cohort 1 and Summer 2026 for Cohort 2)	One staff member at each product vendor to provide data about student and teacher usage of the digital math product	2	Census	100%
Product vendor interviews	Year 1 (Summer 2024 for Cohort 1 and Summer 2025 for Cohort 2); Year 2 (Summer 2025 for Cohort 1 and Summer 2026 for Cohort 2)	Up to three staff members at each product vendor who are working to support product implementation and knowledgeable about costs	6	Census	100%
Interviews with technical assistance provider	Year 1 (Summer 2024 for Cohort 1 and Summer 2025 for Cohort 2); Year 2 (Summer 2025 for Cohort 1 and Summer 2026 for Cohort 2)	Up to nine staff members at the technical assistance provider (Digital Promise) who are familiar with implementation across districts	9	Census	100%
Vendor progress reports & budget documents	Monthly	Up to three staff members at each vendor who are completing progress reports and budget materials	6	Census	100%

Notes: The shaded rows are the data sources for which clearance is not required. The size of the respondent universe in each implementation year represent totals across the two cohorts and are estimated based on the assumption of 15 school districts, 150 schools, four math teachers per grade level per school year (with 20% of teachers new to the study in Year 2, a total of 240 teachers distributed across Cohort 1 and Cohort 2), 75 students per grade level per school year (with 21% of 5th grade students new to the study in Year 2, a total of 2363 new 5th graders distributed across Cohort 1 and Cohort 2), and two digital product vendors. Assumptions about the number of students and teachers per school are based on the Common Core of Data for school year 2018-19 for the target population of schools for the study. Assumptions about teacher mobility rates are based on national rates for mid- and high-poverty schools (National Center for Education Statistics, 2022) and student mobility rates are based on school year and summer mobility rates reported in Potter *et al.* (2019, 2020). Assumptions about response rates are based on other recent ED evaluations in elementary schools (e.g., Impact Evaluation of Training in Multi-Tiered Systems of Support for Reading in Early Elementary School). The response rate for parent and guardian consent form refers to the share of students whose families return these forms, regardless of whether respondents agree or decline to the release of their child's records.

B.2.2. Estimation procedures

The data collection activities will allow the study team to characterize the students, teachers, and districts in the study to address the study's research questions listed in Exhibit B.2 and will contextualize the evaluation's findings. This information will play a critical role as the basis for the evaluation to assess the effect of using the digital math products and of the two instructional approaches for each product. For example, information from student records will be used to describe the characteristics of students and to identify baseline covariates for estimating impacts.

Exhibit B.2 summarizes the types of analytical methods that will be used to answer each research question, which includes:

- **Descriptive analyses (quantitative):** For continuous variables, the study team will use summary statistics including measures of central tendency and variation to report descriptive statistics on the study sample's characteristics and to describe implementation. For categorical variables, the study team will calculate the percentage of sample members in each category. When reporting these analyses by group or by study year, the study team will also report the magnitudes of differences and assess their statistical significance.
- **Regression analyses (quantitative):** The study team will use regression analyses in several ways. First, the study team will use regression analyses to estimate the effect of the digital math products, compared to "business-as-usual" (BAU), on classroom instruction and on student outcomes. Second, the study team will use regression analyses to estimate the differential effect of the two instructional approaches delivered by the products (BFSB *versus* FJIT) on student outcomes. Third, the study team will use regression analyses to understand for whom using the digital math product compared to BAU is most effective, and for whom each instructional approach is more effective than the other. To do so, the study team will examine how effects on student outcomes interact with students' characteristics, their baseline math achievement, and the characteristics of their teachers. In all regression analyses, the study team plans to control for covariates that represent the baseline characteristics of students, including students' demographic characteristics and their baseline scores on state math assessments.
- **Coding (qualitative):** The study team will employ qualitative analysis methods to assist in coding and analyzing of interview data and vendor progress reports, following established research procedures to ensure reliability in coding, including development of a codebook, meeting to discuss and resolve ambiguities and discrepancies in coding, and performing interrater reliability checks. Deductive coding will be employed to explore key aspects of the intended implementation of the digital products. Inductive coding of emergent themes will be used to capture information on facilitators and challenges of implementation that surface through progress reports and interviews.

The following subsections describe the estimations procedures used to answer the research questions in more detail.

Exhibit B.2. Estimation methods for each study research question

Research question	Descriptive analyses	Regression analyses	Qualitative coding
1. Did regular use of the digital math products improve student achievement?		X	
2. Which instructional approach for the digital math products was more effective at improving student achievement?		X	
3. Which teachers and students benefit most from student use of the digital math products? Which instructional approaches worked best for underserved students, particularly students who were the furthest behind and from low-income families?		X	
4. Did students use the digital math products as intended?	X	X	X
5. Did use of the digital math products lead to greater individualization of the math instruction received by students? Did it improve student engagement?		X	
6. What was the cost per student of using the digital products? What was their cost-effectiveness?	X		

Estimation Procedures to Assess Effectiveness

To estimate the **pooled effect on student outcomes of implementing the digital math products in math classrooms, compared to business-as-usual math instruction**, the study team will estimate the following three-level model, with students (i) nested within classrooms I[1] nested within schools (j):

$$Y_{icj} = \theta_0 + \beta_0 TREAT_j + \sum_{k=1}^{14} \psi_k D_{k,j} + \sum_{1}^{Z} \vartheta_z X_{z,icj} + \phi M_{icj} + u_j + \tau_{cj} + v_{icj} \quad \textbf{(1)}$$

The variables in Model 1 are defined as follows:

Y_{icj}	=	The outcome of interest for student i
$TREAT_j$	=	An indicator for whether school j was randomly assigned to implement the digital math products (=1) or the BAU control group (=0)
$X_{z,icj}$	=	A set of student-level demographic characteristics for student i, where demographic characteristic z includes, among others, whether a student is eligible for free or reduced-price lunch, their race/ethnicity, their gender, whether the student has an individualized education plan (IEP), and whether the student is an English Language Learner (ELL)
M_{icj}	=	A student's (standardized) score on the state math assessment in the Spring before random assignment[2]
$D_{k,j}$	=	A set of indicators for school j's random assignment block (i.e., a particular school district) (=1 if school j is in district k; =0 otherwise).
u_j	=	Unexplained school-level variation in outcome Y (between-school residual)
τ_{cj}	=	Unexplained classroom-level variation in outcome Y (within-school and between-classroom residual)
v_{icj}	=	Unexplained student-level variation in outcome Y (within-classroom residual)

[1] Students' classrooms at the time of student-level random assignment (i.e., the classroom used for blocking the random assignment of students to approaches).

[2] The study team may also consider including classroom-level and school-level aggregate baseline measures of student achievement and demographic characteristics to improve precision.

The estimate of β_0 from this model is an estimate of the effect on student outcome *Y* of offering the digital math products to schools assigned to the program group, compared to business-as-usual math instruction.

To estimate the **relative effectiveness of the two instructional approaches,** the study team will fit the following model to the subset of students in the program schools:

$$Y_i = \theta_0 + \beta_1 FJIT_i + \sum_1^C \lambda_c B_{c,i} + \sum_1^Z \vartheta_z X_{z,i} + \phi M_i + v_i \quad \textbf{(2)}$$

The variables X_z and *M* and the error term v in Model 2 are defined as above, and the other variables are defined as follows:

$FJIT_i$	=	An indicator for whether student i was randomly assigned to receive the Focused Just-In-Time Skill Building approach from the digital product (=1) or the Broad Foundation Skill Building approach (=0)
$B_{c,i}$	=	A set of indicators for student i*'s* random assignment block (i.e., the student's classroom) (=1 if student *i* is in classroom *c*; =0 otherwise).

The estimate of β_1 from this model is an estimate of the effect on student outcome *Y* of offering the FJIT approach to students, compared to offering them the BFSB approach (the differential effect of the two instructional approaches).[3]

For both types of analyses, the study team will report statistical significance based on two-tailed t-tests. The standard errors used for hypothesis testing will account for the clustering of students within schools and classrooms by specifying a multi-level error structure (in Model 1) or by controlling for random assignment block indicators (in Model 2).

The study team will also estimate the effect of the digital math products and the relative effect of the instructional approaches on classroom-level outcomes (teacher practices and classroom-level student engagement), using similar equations adapted for the classroom level.

To estimate effects for subgroups of students or teachers defined by their baseline characteristics, the study team will adapt the above equations to include interactions between the research group indicators and subgroup indicators.

Estimation Procedures to Assess Implementation

The study team will describe the students participating in the study using measures of their characteristics and math achievement from the school year prior to random assignment, drawing

[3] Model 2 does not include unexplained variation between schools (u_j) or between classrooms (τ_{cj}) because school-level and classroom-level outcome variation will be explained by the random assignment block indicators $(B_{c,i})$.

on information from student records. The study team will calculate summary statistics for continuous measures and percentages for binary and categorical measures, by research group.

The study team will use descriptive analyses to describe the implementation of the digital products and assess fidelity to the model using progress reports from vendors, data from the products, and relevant items from the teacher and school leader survey. Examples of measures will include teacher participation in trainings; the amount of time that students and teachers spend engaging with the digital products; and students' progress through the product lessons. The study team will calculate and report summary statistics for continuous measures and percentages for binary and categorical measures, overall for the program schools, and by research group (BFSB and FJIT) where relevant.

The study team will also describe implementation facilitators and barriers using coded data from the vendor progress reports, interviews with the technical assistance providers, and relevant items from the teacher and school leader surveys. The study team will report on the prevalence of facilitators and barriers to implementing the digital math products as intended in classrooms.

Estimation Procedures to Assess Costs

The study team will use a resource cost model (RCM) to measure and analyze the costs of the two digital math products based on the ingredients approach (Levin & McEwan, 2001). The study team will use information from vendor budget documents and the vendor interviews to develop a list of "ingredients" for each digital math product and use the CostOut tool (Hollands *et al*., 2015) for pricing information. The team will then develop an RCM template to calculate the per-student costs of the products and analyze variation in costs across schools and districts. The study team will also produce cost-effectiveness estimates by dividing the average per-student cost of the products by the estimated effect of the products compared to BAU.

B.2.3. Degree of accuracy needed

The study is designed to reliably answer research questions about the effect of the digital math products compared to BAU as well as the effect of each instructional approach relative to each other. The study's sample sizes were chosen to yield sufficient statistical power for detecting effects that might plausibly be achieved and that are considered by experts and stakeholders to be meaningful and policy relevant.

To assess statistical power, the study team focused on the evaluation's ability to statistically detect effects on **math achievement** (the key outcome for the study) for the **focal sample of students** for the evaluation—defined as students who are in 4^{th} grade in Year 1 for each Cohort and who are below grade level in math. The study will focus on these students because they can be followed for two school years and because students who are below grade level are the main target population.[4] The evaluation's design will allow the study to detect policy-relevant and realistic effects of the instructional approaches on these students' math achievement (their performance on state math tests and on districts' formative math assessments), the key outcomes of the study, at the end of the second year of implementation.

With a target sample size of 150 schools, the study will be able to detect an **effect of the digital math products compared to BAU as small as 0.11 standard deviations** for the focal student sample, and 0.13 for a sample consisting of half of the focal students (such as in subgroup analyses), after two years of implementation. Past evidence suggests that an effect this large is realistic. A meta-analysis by Cheung and Slavin (2013) found that math educational technologies improved students' math achievement by 0.15 standard deviations, with larger effects for elementary school students (0.17), supplemental products (0.19), and well-implemented products (0.26), but smaller effects for studies based on randomized experiments (0.08). A meta-analysis by Williams *et al.* (2022) of math interventions evaluated using randomized experiments found an effect size of 0.11 standard deviations for technology-based interventions. Most of the reviewed studies in these two meta-analyses included shorter-duration interventions of less than two years, so the effects of the adaptive technologies in the proposed evaluation could be at least as large as those reported in the meta-analysis.

An estimated effect of 0.11 standard deviations represents 10% of typical two-year gains in math (or an additional 7 weeks of learning gains) (Bloom *et al.*, 2008). Such an effect seems meaningful, given that use of the digital products is embedded within regular math instruction, rather than requiring additional instructional time. It also represents an increase of five percentage points in the percent of students who are proficient in math, which would be enough to bring 4th grade math proficiency rates back to pre-pandemic levels (National Assessment of Education Progress, 2022).[5]

The study will also be able to detect a **differential effect of the two instructional approaches (BFSB compared to FJIT) as small as 0.06 standard deviations** for the focal student sample, and 0.09 for a sample consisting of half of the focal students, after two years of implementation. No prior research has rigorously compared the effect of these two instructional approaches, but a differential effect of 0.06 represents six percent of typical two-year gains in math (an additional 4 weeks of learning gains) (Bloom *et al.*, 2008). It also represents an increase of three percentage points in the percent of students who are proficient in math, which represents more than half of the amount by which 4th grade math proficiency rates would need to increase to get back to pre-pandemic levels (National Assessment of Education Progress, 2022). Such an

---

[4] Students will be defined as "below grade level" if they are not proficient on their state's math test in the Spring before student-level random assignment. The evaluation will also examine the effect of using the digital math products and the relative effect of the instructional approaches for students who are *not* below grade level, as an exploratory analysis.

[5] The percent of 4th grade students proficient on the National Assessment of Educational Progress (NAEP) math assessment decreased by 5 percentage points during the pandemic, from 41 percent in 2019 to 36 percent in 2022 (National Assessment of Education Progress, 2022).

effect seems meaningful, given that students are using the same underlying digital math product for the same amount of time under both instructional modes.

Exhibit B.3. Minimum detectable effect size on math achievement outcomes for the primary student sample

Type of effect	Minimum detectable effect size (in standard deviation units) after two years of implementation	
	Full sample	50% subsample of students
Pooled effect of using the digital math products compared to BAU	0.11	0.13
Differential effect of the two instructional approaches (BFSB *vs.* FJIT)	0.06	0.09

Note: Calculations assume (1) 80 percent power and five percent significance level for a two-tailed test; (2) an average of 75 4th grade students per school in Year 1 of the study, 45 (62%) of whom are identified as not proficient on the state math test at baseline; (3) of these focal students, an average of 28 students per school (63%) who have math achievement data at the end of the second year of implementation, based the assumption that 80% of students will have parental consent for the release of school records and that 79% of students will still be enrolled in the district at the end of the second year or implementation; (4) a sample of 100 program school and 50 business as usual control schools (a 2:1 school-level random assignment ratio); and (5) a 1:1 student-level random assignment ratio for the randomization of students to the two instructional approaches (BFSB or FJIT). Assumptions about the number of focal students per school and proficiency levels are based on the Common Core of Data and EdFacts for school year 2018-19 for the target population of schools for the study. Assumptions about student mobility are from Potter *et al.* (2019, 2020), and assumptions about consent rates are from recent ED evaluations in elementary schools (e.g., Impact Evaluation of Training in Multi-Tiered Systems of Support for Reading in Early Elementary Schools). Assumptions about the clustering of outcomes and explanatory power of covariates come from estimates reported in Hedges and Hedberg (2013) for math achievement in grades 4-5: a between-school intracluster correlation coefficient (ICC) of 0.163, and a variance explained by the covariates (R^2) of 0.755 at the school level 0.535 at the student (within school) level.

B.2.4. Unusual problems requiring specialized sampling procedures

The study team does not anticipate unusual problems that will require specialized sampling procedures.

B.2.5. Use of periodic (less than annual) data collection to reduce burden

ED and the study team have carefully considered the frequency of each data collection activity under this request, and the study team is planning to collect the study's data as infrequently as possible while fulfilling the study's analytic requirements (see Exhibit B.1). In addition, for data collection activities that will be happening more than once, the forms and instruments have been designed to be very short and to gather only essential information to minimize respondent burden (see Supporting Statement Part A for time commitment estimates).

- The study will collect **parent/guardian consent forms** only in districts that require it for the release of student records, and we will only request permission once from each applicable respondent.

- The study team will collect **district administrative records** two times from each participating district, once after the end of each implementation year (Fall 2024 and Fall 2025 for Cohort 1 and Fall 2025 and Fall 2026 for Cohort 2). Less frequent collection of student records would mean not having up-to-date information needed to examine the characteristics of the study sample and interim findings at the end of the first implementation year. Moreover, some districts may not retain all of the required information if collection is deferred until the end of the study, which increases the risk of missing data.

- The study team will administer the math **teacher survey** to each teacher at most three times. The first time will be in the fall of a teacher's first year of implementation (Fall 2023 for Cohort 1; Fall 2024 for new Cohort 1 teachers and all Cohort 2 teachers; Fall 2025 for new Cohort 2 teachers[6]) to measure teacher characteristics and their self-efficacy and attitudes at baseline. Teachers will also be surveyed again at the end of each implementation year (Spring 2024 and 2025 for Cohort 1; Spring 2025 and 2026 for Cohort 2) to collect data on math instruction during the past school year. The fall (baseline) measure of teacher self-efficacy is critical to the analyses because it will allow the study team to examine whether the digital math products and instructional approaches are more effective for students taught by teachers with lower self-efficacy related to teaching math instruction. The Spring teacher surveys are critical for examining effects on math instruction and student engagement in the classroom, which are key intermediate outcomes. Less frequent collection of the Spring surveys would mean not being able to examine interim effects on these outcomes at the end of the first implementation year.

- The study team will administer the **school leader survey** twice to a school leader at each participating school, once at the end of each implementation year (Spring 2024 and 2025 for Cohort 1; Spring 2025 and 2026 for Cohort 2), to collect information about professional learning and schools supports for math instruction during the past school year. Having two years of data will make it possible to describe the contextual factors affecting math instruction across both implementation years.

- The study team will collect **staff directory data** four times in each participating district, once in the Fall and Spring of each implementation year, to define the target population for the teacher survey (Fall and Spring) and school leader survey (Spring). Less frequent data collection would mean possibly omitting staff in the target population from the survey administration, as staffing changes may occur intermittently throughout the school year.

[6] To minimize burden, only teachers new to the study in a given cohort will be surveyed in Year 2 (Fall 2024 for Cohort 1 and Fall 2025 for Cohort 2).

B.3. Methods to maximize response rates and address nonresponse

B.3.1. Methods to maximizing response rates

To maximize response rates, the study team will work closely with the study districts and schools and use strategies that have been successful in other large-scale evaluations (e.g., the Impact Evaluation of Training in Multi-Tiered Systems of Support for Behavior, and the Academic Language Impact Evaluation).

The general approach used by the study team will set high and clear expectations and processes for data collection. The team will execute a memorandum of understanding with each district that will specify their role and expectations related to data collection. The team will also communicate with respondents (districts, school leaders, teachers, parents) to establish procedures at the start of the study, emphasize the importance of following these procedures, and provide advance notifications and follow-up to remind respondents of the study's data collection expectations. The team will also increase completion rates by using forms and instruments that are intentionally short to reduce burden (see Supporting Statement Part A for burden estimates) and by pretesting the survey forms to ensure that they are concise and clear (see Section B.4).

Below are additional strategies for maximizing response rates for the specific data types in the clearance request:

- **Parent/guardian consent forms (95 percent expected response rate).**[7] Where required, the study team will work with schools to conduct a parent consent process for the release of student records. The following procedures will be used to maximize the rate at which consent forms are returned:
  - The study team will partner with school staff to conduct broad outreach, before distributing consent forms, to reduce potential confusion or misunderstandings about the study.
  - The consent materials will be brief – with one page of information and one page for signatures.
  - The materials will be written in clear, accessible language; they will be available in both English and Spanish, and the study team will offer additional translations if a participating district indicates that other languages are typically used when communicating with families.
  - The study team will identify liaisons in each school to distribute and gather consent forms as they are returned. These liaisons will distribute consent forms to students and families using the districts' preferred method of communicating with parents (e.g., back-to-school packets, "backpack mail," or at a back-to-school night). The liaisons will also

[7] This response rate assumes that the consent process will be active and refers to the share of students whose families return these forms, regardless of whether parents or guardians agree or decline to the release of their child's records.

conduct follow-up with parents who have not returned forms and will provide as-needed supports and encouragement to classroom teachers in facilitating form return.

- If district rules allow, the study will provide an incentive to classroom teachers for return of consent forms (irrespective of whether the parent or guardian agree or decline to allow the district to share data about their student). Classrooms that achieve a consent form return rate of 85% or higher will receive a gift valued at $10 per student in the class for use on behalf of all students in the class. The study team will work with each district to determine the most appealing form of incentive for classrooms, such as a gift card, and how best to distribute it during the study period.

- Where possible, the study will also send an electronic consent package/form to parents and guardians.

- Where possible, the study team will use a passive (opt-out) consent process instead of an active consent process, to reduce burden and improve response rates.

- **School leader survey and math teacher surveys (85 percent expected response rate):** To encourage survey completion, the web-based surveys will only gather essential information to minimize survey length. The study team will pretest the web-based surveys to maximize ease of completion and reduce respondent burden (see Section B.4). The following strategies will also be used to maximize survey response rates:
    - The study team will send respondents an invitation letter describing the study's purpose and the topics the survey will cover. The letter will also describe how to access the electronic survey. Where feasible, endorsement letters from the district encouraging the participation of school leaders and teachers will also be included in this advance communication. The invitation package will be emailed to respondents and a paper version will also be mailed to respondents' schools, to increase the likelihood that respondents will receive communications (in case spam filters block the email).
    - Throughout the data collection period, the study team will send email reminders to non-responders to encourage survey completion. The school liaison (who will be helping with parent/guardian consent forms) will also be asked to reach out to non-responders in person and/or by email to remind them to complete the surveys. Study staff will discuss all possible methods and determine the best course of action based on their experiences with the schools.
    - The study team is also proposing the use of incentives for the school leader and teacher surveys. If approved by OMB, the proposed incentive is $20 for school leaders, $15 for teachers in the Fall and $25 in the Spring, to be paid upon survey completion (see Supporting Statement A for more details on incentives).
- **Staff directory data and district administrative records (100 percent expected response rate):** The memorandum of understanding signed by each study district will include an agreement to provide records on the students in the evaluation and staff directory data. The study team will adhere to any district data requirements, such as preparing research applications. Using strategies the study team has carried out successfully on prior

evaluations for IES, the team will designate one liaison from the study team for each district who will communicate with the district staff person responsible for providing the data. Through phone calls and written explanations, the liaison will describe the data fields requested, data security procedures, and procedures for districts to submit the data. While the study's data request template will include file specifications, the study team will accept the data in any format in which they are provided (converting files to a consistent format for the analysis). The study team will also offer user-friendly instructions for encrypting files before uploading them to a secure file transfer system.

B.3.2. Methods to deal with issues of nonresponse

If response rates for the school leader and math teacher surveys fall below 85 percent, or if parent consent rates fall below 70 percent, the study team will examine response rates by research group (program and BAU schools), to explore whether large differences in response rates across groups could be biasing the estimation of effects on classroom and student outcomes. If these analyses point to the possibility of differential nonresponse bias, the study team will investigate whether differential response rates are being driven by particular school districts, and if so, conduct sensitivity analyses where these districts are excluded from the analysis.

To examine the representativeness (generalizability) of the study findings, the study team will also compare the characteristics of schools and students in the study to those in the target population. If these analyses indicate that the study sample is not representative of the target population, the study team will consider conducting sensitivity analyses that reweight the sample to make it more representative (Tipton & Olsen, 2022).

For the analysis of effects on classroom and student outcomes, outcomes data will not be imputed. However, the study team will impute missing data on the baseline covariates in the regression models using the indicator variable approach.[8] Studies have shown that in a randomized experiment, this approach produces impact estimates that are as unbiased and as precise as other imputation methods (Puma *et al.*, 2009).

B.4. Tests of procedures

The study team conducted a pilot test of the fall math teacher survey and spring math teacher survey with 8 elementary school teachers and the school leader survey with 5 current or former elementary school leaders. The study team reviewed the completed instruments and conducted virtual debriefing interviews with each respondent to inquire on the time it took them to complete each instrument, to identify any issues they may have encountered with providing the requested information and to probe on specific items and wording that may have been

[8] This approach consists of (1) imputing a value of "zero" for the missing values in each of the covariates, (2) creating a dichotomous indicator of missingness for each covariate, and (3) including these indicators alongside the imputed covariates in the statistical model.

unclear. Respondents reports on time to complete the instrument confirmed the study team's burden estimates. The study team used feedback from the pilot test interviews to revise and improve the wording of specific survey questions to ensure that respondents can understand and complete the survey accurately and as intended.

The study team will not pretest the requests for district administrative records and staff directory information – or the parent/guardian consent form – because these forms are closely modeled on forms that have been effectively used for other studies. The study team has confirmed that the reading level for the parent/guardian consent form is accessible (9th grade reading level).

B.5. Individuals consulted on statistical aspects of the design

The following study team members provided primary consultation for ED about the design of the study and data collection plan and/or will lead the data collection and analysis.

Name	Role(s)	Title	Telephone Number
John Pane	Co-principal investigator, consultation on statistical design, lead for data collection/analysis	Senior Scientist, Rand Corporation	412-683-2300 ext 4619
Patty Troppe	Lead for data collection	Vice President for Education Studies, Westat	301-294-3924
Marie-Andree Somers	Consultation on statistical design, lead for data collection/analysis	Senior Research Associate, MDRC	212-340-8825
Lauren Decker-Woodrow	Lead for cost analysis	Research Associate, Westat	210-558-4148

References

Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness*, *1*(4), 289-328.

Cheung, A. C., & Slavin, R. E. (2013). The effectiveness of educational technology applications for enhancing mathematics achievement in K-12 classrooms: A meta-analysis. *Educational research review*, *9*, 88-113.

Hedges, L. V., & Hedberg, E. C. (2013). Intraclass correlations and covariate outcome correlations for planning two-and three-level cluster-randomized experiments in education. *Evaluation Review*, *37*(6), 445-489.

Hollands, F. M., Hanisch-Cerda, B., Levin, H. M., Belfield, C. R., Menon, A., Shand, R., Pan, Y., Bakir, I., & Cheng, H. (2015). CostOut®. Teachers College, Columbia University. Available online at https://www.cbcse.org/costout.

Levin, H., & McEwan, P. (2001). *Cost-effectiveness analysis: Methods and applications*. Sage.

National Center for Education Statistics. (2022). Teacher Turnover: Stayers, Movers, and Leavers. *Condition of Education*. U.S. Department of Education, Institute of Education Sciences. Retrieved 12/14/22, from https://nces.ed.gov/programs/coe/indicator/slc/

National Center for Education Statistics. (2022). National Achievement-Level Results. *NAEP Report Card.* U.S. Department of Education, Institute of Education Sciences. Retrieved 1/17/22, from https://www.nationsreportcard.gov/mathematics/nation/achievement/?grade=4

Potter, D., Alvear, S., Bao, K., & Min, J. (2019). *Changing Schools, Part 2: Student Mobility During the School Year in Texas and the Houston Area*. Houston, TX. Houston Education Research Consortium, Kinder Institute for Urban Research, Rice University. Retrieved 12/19/22, from https://rice.app.box.com/s/num2aubml7err9i0fs18sacrhg5dlhiy

Puma, M. J., Olsen, R. B., Bell, S. H., & Price, C. (2009). *What to do when data are missing in group randomized controlled trials* (NCEE 2009-0049). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Tipton, E., & Olsen, R. B. (2022). *Enhancing the Generalizability of Impact Studies in Education* (NCEE 2022-003). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. Retrieved from http://ies.ed.gov/ncee.

Williams, R., Citkowicz, M., Miller, D. I., Lindsay, J., & Walters, K. (2022). Heterogeneity in Mathematics Intervention Effects: Evidence from a Meta-Analysis of 191 Randomized Experiments. *Journal of Research on Educational Effectiveness*, 1-5.

U.S. Department of Education. (2018). *Table 6. Average Class Size in Public Schools, by Class Type and State: 2017–18.* Available at https://nces.ed.gov/surveys/ntps/tables/ntps1718_fltable06_t1s.asp