

# Evaluation of the Toolkit to Support Evidence-Based Algebra Instruction in Middle and High School

## Part B: Collection of Information Employing Statistical Methods

June 6, 2023

---

# Evaluation of the Toolkit to Support Evidence-Based Algebra Instruction in Middle and High School

## Part B: Collection of Information Employing Statistical Methods

June 6, 2023

---

### **Submitted to:**

U.S. Department of Education  
Institute of Education Sciences  
550 12th Street, S.W.  
Washington, DC 20202  
Project Officer: Amy Johnson  
Contract Number: 91990022C0015

### **Submitted by:**

Mathematica  
P.O. Box 2393  
Princeton, NJ 08543-2393  
Telephone: (609) 799-3535  
Fax: (609) 799-0005  
Project Director: Phillip Herman  
Reference Number: 51410

## Contents

B. Collection of information employing statistical methods.....	1
Introduction.....	1
B1. Respondent universe and sampling methods.....	1
a. School districts.....	1
b. Schools.....	2
c. Instructional leaders.....	3
d. Teachers.....	3
e. Students.....	4
B2. Statistical methods for sample selection and degree of accuracy needed.....	5
a. Sample selection.....	5
b. Estimation procedures.....	6
c. Degree of accuracy needed.....	8
d. Unusual problems requiring specialized sampling procedures.....	9
B3. Methods to maximize response rates and deal with nonresponse.....	10
a. Maximizing response rates.....	10
b. Dealing with nonresponse.....	11
B4. Tests of procedures and methods to be undertaken.....	11
B5. Individuals consulted on statistical aspects of the design.....	11
References.....	12
APPENDIX A: STUDY RECRUITMENT LETTER	
APPENDIX B: STUDY RECRUITMENT FLYER	
APPENDIX C: STUDY RECRUITMENT SCREENER TOOL	
APPENDIX D: CONFIDENTIALITY PLEDGE	

## Exhibits

B.1	Respondent universe, sample, and expected response rate for study data sources.....	5
B.2	Type of analysis for each study research question.....	7
B.3	Minimum detectable effect (MDE) size for the toolkit efficacy study.....	8

## B. Collection of information employing statistical methods

### Introduction

The Institute of Education Sciences (IES) within the U.S. Department of Education (ED) requests clearance for activities to support the recruitment of school districts to participate in an efficacy study of a *Toolkit to Support Evidence-Based Algebra Instruction in Middle and High School* as part of the Regional Educational Laboratory Central contract (REL Central). 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.

This efficacy study will compare the outcomes of teachers and students in grades 8 and 9 Algebra I classrooms in which teachers have access to the toolkit resources and receive the toolkit’s professional development supports (the treatment group) with the outcomes of teachers and students in similar grades 8 and 9 Algebra I classrooms in which teachers receive their business-as-usual professional development supports (the control group).

### B1. Respondent universe and sampling methods

The efficacy study will employ a school-level randomized controlled design. Study respondents will include school districts, schools, instructional leaders, teachers, and students. The study will first recruit districts and schools to participate in the study from the Central region (Colorado, Kansas, Missouri, Nebraska, North Dakota, South Dakota, and Wyoming), collecting data from districts to determine their suitability for participation (the focus of this OMB package). To examine the implementation and efficacy of the toolkit, the study plans to collect additional data from the census of participating school districts, schools, instructional leaders, teachers, and students (the focus of a future OMB package). Because the ultimate sample could include any eligible respondents from the Central region, the respondent universe is defined as those that could have been recruited from the Central region. Below are descriptions of the respondent universe and the targeted sample size for each respondent type for both this OMB package and the future OMB package. Exhibit B.1 summarizes the respondent universe, sample, and expected response rate for study data sources.

#### a. School districts

- *Respondent universe.* School districts in the Central region that have at least six middle and/or high schools are eligible to participate in the study. Based on data from the Common Core of Data (CCD) for the 2021–2022 school year, 100 school districts meet this criterion. (Note that the study will exclude statewide school networks—such as divisions of youth services—and one school district that is participating in usability testing as a part of the toolkit development process.)
- *Targeted sample size.* Up to 30 school districts will participate in recruitment data collection activities (this OMB package). Up to three school districts will ultimately participate in the study (data collection activities to be further described in a future OMB package).
- *Sampling method considerations.* Districts will be purposefully selected by the study team for participation in recruitment data collection (this OMB package) and final participation in the study.

For recruitment data collection activities, the study team will prioritize districts based on two criteria:

1. *District size.* The study team will group districts based on their size. Districts with at least 20 middle and/or high schools (14 districts) will be prioritized first, those with 12 to 19 middle and/or high schools (20 districts) second, and those with six to 11 middle and/or high schools (66 districts) third.
2. *Estimated likelihood of participation in the study.* The study team will draw on REL Central’s strong relationships and connections in the region to support the recruitment effort. When identifying districts to participate in recruitment data collection, the study team will consult with REL Central’s partners and members of the REL Central Governing Board to obtain districts that are best suited to participate in the study. Such districts are those with an interest in exploring professional development supports for algebra teachers and those that are more likely to be willing to participate in a study.

For ultimate participation in the study, the study team will prioritize districts based on three criteria:

1. *Number of schools eligible and willing to participate in the study.* Districts with at least six schools that are eligible and willing to participate in the study will be prioritized to help meet sample size targets.
2. *Proportion of students in schools eligible for free or reduced-price lunch subsidies.* To promote greater diversity in the sample, the study team will prioritize districts that have more schools in which at least half of the students are eligible for free or reduced-price lunch subsidies.
3. *Suitability for implementing the toolkit.* Given information collected through recruitment data collection activities, the study team will prioritize the districts that are best suited to implement the toolkit for final participation in the study. Because this is an efficacy study, ensuring that participating districts are well suited to implement the toolkit will assist the study team in assessing the toolkit’s impacts on student and teacher outcomes when the toolkit is implemented under favorable conditions. The study team will consider district-wide (rather than school-specific) Algebra I curricula and assessments, professional development structures aligned with toolkit implementation (such as regular professional development sessions and professional learning community meetings), and professional development supports that sufficiently contrast with the study’s supports.

**b. Schools**

- *Respondent universe.* Middle and high schools from eligible districts that have at least one Algebra I course are eligible to participate in the study. Based on CCD data for the 2021–2022 school year, the study team assumes that 1,310 schools from eligible school districts in the Central region may meet this criterion.
- *Targeted sample size.* Approximately 20 schools will participate in the study. Half of the participating schools (approximately 10) will be randomly assigned to the treatment group, and the other half (approximately 10) will be randomly assigned to the control group.
- *Sampling method considerations.* There are no school-level recruitment data collection activities (the focus of this OMB package). For final participation in the study, the study team will aim to include all middle and high schools within each recruited district, with three additional considerations. In particular, the study will take the following steps:
  - *Exclude schools unwilling to participate in the study.* School leaders will be asked to sign a school participation form, which commits the school and its Algebra I teachers to participate in the study. Schools in which principals refuse to participate will be excluded from participation.

- *Exclude schools that substantially differ from others in the district.* A middle or high school may be excluded from participation if it follows a different Algebra I curriculum, uses different student assessments, or has substantially different Algebra I professional development supports compared to the rest of the middle and high schools in the district.
- *Randomly select schools among particularly large districts.* Ten districts in the sampling universe have at least 25 middle or high schools. In these districts, the study team may need to randomly select a subset of willing and eligible schools to participate in the study to conserve study resources. The number of sampled schools will depend on how many schools are needed to meet the study’s sample targets as well as the other district-level sampling considerations (proportion of students in the schools who are eligible for free or reduced-priced lunch subsidies; suitability for implementing the toolkit).

**c. Instructional leaders**

- *Respondent universe.* Each school that is randomly selected for the treatment group will identify one instructional leader to help facilitate the implementation of the toolkit. Thus, the study team assumes the universe of potential instructional leaders for the study equals the universe of potential schools: 1,310 instructional leaders.
- *Targeted sample size.* Approximately 10 instructional leaders will participate in the study (assuming one instructional leader for each treatment school). Control schools will not identify an instructional leader.
- *Sampling method considerations.* There are no instructional leader-level recruitment data collection activities (the focus of this OMB package). For final participation in the study, each school that is randomly selected for the treatment group will identify one instructional leader. The study team will assist treatment schools in identifying instructional leaders, using the following considerations:
  - Familiarity with and ability to facilitate professional development sessions and provide individualized support for treatment teachers. Instructional leaders will help teachers implement the toolkit and its recommended practices. Thus, instructional leaders should have the skills and experience to facilitate professional development sessions and individualized support.
  - Possibility of using the same instructional leader across treatment schools. Instructional leaders could either be district- or school-level staff. If the district structure allows, the same instructional leader could support implementation across multiple treatment schools. In this case, fewer than 10 instructional leaders will participate in the study.

**d. Teachers**

- *Respondent universe.* Teachers who teach at least one Algebra I course in eligible middle or high schools are eligible to participate in the study. Although the study team does not know the exact number of teachers who meet this criterion, it assumes three eligible teachers per eligible school, resulting in a total of 3,930 teachers who meet this criterion.
- *Targeted sample size.* Approximately 60 teachers will participate in the study (assuming three teachers per participating school). Because random assignment is at the school level, all teachers will be assigned to the treatment or control condition to which their school is assigned. This approach will result in half of the participating teachers (approximately 30) being randomly assigned to the treatment group and the other half (approximately 30) being randomly assigned to the control group.

- *Sampling method considerations.* There are no teacher-level recruitment data collection activities (the focus of this OMB package). For final participation in the study, the study team will aim to include all eligible teachers within each participating school, except that the study team will exclude teachers unwilling to participate in the study. Teachers will be asked to complete a consent form for their individual participation in the study, and teachers who do not consent to participate will be excluded from participation.

**e. Students**

- *Respondent universe.* Students enrolled in an Algebra I course taught by a participating Algebra I teacher are eligible to participate in the study. Although the study team does not know the exact number of students who meet this criterion, it estimates that 3,930 teachers are eligible to participate in the study, and each teacher will teach two courses with 22 students each. Based on these assumptions, the study team estimates that a total of 196,500 students could be eligible.
- *Targeted sample size.* Approximately 2,640 students will participate in the study (assuming 44 students per participating teacher and three teachers per participating school). Because random assignment is at the school level, all students within a school will be assigned to the treatment or control condition to which their school is assigned. Therefore, half of the participating students (approximately 1,320) will be randomly assigned to the treatment group, and the other half (approximately 1,320) will be randomly assigned to the control group.
- *Sampling method considerations.* There are no student-level recruitment data collection activities (the focus of this package). For final participation in the study, the study team will aim to include all students within each participating teacher's Algebra I courses, except that the study team will exclude students whose parents or guardians do not consent to their child participating in data collection activities. The study team will use passive parent/guardian consent for the student survey, when possible, as the student survey does not collect data on sensitive topics or student personally identifiable information. The study team will collect active consent from parents and guardians only in districts that require it. In districts requiring active consent, parents and guardians of students in participating teachers' Algebra I classrooms will be asked to complete a consent form for their child to participate in the study's data collection activities. Students whose parents or guardians do not provide consent will be excluded from the study.



**Exhibit B.1. Respondent universe, sample, and expected response rate for study data sources**

Data source	Respondent	Respondent universe	Type of sample	Sample size	Expected response rate
<b>Recruitment data collection sources (this package)</b>					
Recruitment screener	District staff member	100	Prioritization based on size and likelihood of participation	30	100%
<b>Additional data collection sources (future package)</b>					
District administrative data	One staff member at each district to provide records	100	Census among participating districts	3	100%
Instructional leader implementation logs	Each treatment school's identified instructional leader	1,310	Census among participating instructional leaders in treatment schools	10	85%
Instructional leader interviews	Each treatment school's identified instructional leader	1,310	Census among participating instructional leaders in treatment schools	10	100%
Classroom rosters	All treatment and control teachers	3,930	Census among participating teachers	60	100%
Teacher survey	All treatment and control teachers	3,930	Census among participating teachers	60	85%
Professional development module attendance logs	Treatment teachers	3,930	Census among participating treatment teachers	30	100%
Teacher implementation logs	Treatment teachers	3,930	Census among participating treatment teachers	30	85%
Teacher focus group	Treatment teachers	3,930	Census among participating treatment teachers	30	85%
Student survey	All students taught by teachers participating in the study	196,500	Census among all students in Algebra I courses taught by participating teachers, excluding those without parent or guardian consent	2,640	85%

**B2. Statistical methods for sample selection and degree of accuracy needed**

**a. Sample selection**

The study team will select the study sample by prioritizing districts, schools, instructional leaders, teachers, and students based on the sampling methods considerations outlined in Section B1, *Respondent universe and sampling methods*. The study team will purposively select up to three school districts for

participation in the study and plans to select within participating districts all eligible schools, instructional leaders, teachers, and students for participation in the study. If a participating district has more eligible schools than the study can accommodate, the study team will randomly select schools for participation, using simple random sampling.

Because this is an efficacy study that seeks to assess the implementation and impacts of the toolkit when it is implemented under favorable conditions, an important consideration in selecting a sample for the study is ensuring that study participants exist in a setting in which implementation is more likely to be successful. For example, districts and schools where there is interest in additional Algebra I professional development resources, strong leadership, and no competing interventions present more favorable conditions for implementation than those in which these characteristics are not present. Thus, the sample will not necessarily be representative of districts and schools in the Central region.

#### **b. 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. Exhibit B.2 indicates the analysis method for each research question. To address the research questions, the study team will conduct two types of analysis:

- **Impact analysis.** The impact analysis will estimate the effects of the toolkit on teacher outcomes, including their knowledge and use of the recommended practices as well as their teaching self-efficacy. It will also estimate the effects of the toolkit on student outcomes, including student algebraic knowledge, student self-efficacy, mathematical mindset, and algebraic achievement. The study team will use regression analyses to compare student and teacher outcomes between the treatment and control groups. To estimate impacts, the study team will regress teacher and student outcomes on an indicator for whether their school was assigned to the treatment group, as well as baseline characteristics (see below for additional details). Additionally, the study team will use regression analysis to assess whether the toolkit is more effective for students and teachers with certain characteristics. To do so, the study team will run regression analyses with interactions between treatment status and individual characteristics.
- **Implementation analysis.** The implementation analysis will include several descriptive analyses that will aid in interpreting the impact findings. These analyses will include (1) assessing the fidelity of the toolkit's implementation, (2) describing teachers' and instructional leaders' perceptions of the toolkit and challenges with using it, and (3) assessing the contrast between the treatment and control groups in professional development supports and resources for algebra teachers. These analyses will rely on data from the recruitment screener, data on attendance in the training module, implementation log data from teachers and instructional leaders, teacher survey data, data from teacher focus groups and instructional leader interviews, and district administrative data. When presenting findings on an individual group, such as on the treatment teachers' implementation of the toolkit, the study team will calculate (1) the percentage of group respondents in each category for categorical variables and (2) the average values of continuous variables. To compare groups, such as comparing the receipt of support between the treatment and control teachers, the study team will report the magnitudes of

differences between groups and assess their statistical significance. The study team will also conduct a thematic analysis of the qualitative data.

**Exhibit B.2. Type of analysis for each study research question**

Research question	Impact analysis	Implementation analysis
What is the impact of providing the toolkit to teachers on teachers' knowledge of the practices recommended in the practice guide and teaching self-efficacy?	X	
What is the impact of providing the toolkit to teachers on teachers' use of the practices recommended in the practice guide?	X	
What is the impact of providing the toolkit to teachers on students' understanding of how to solve algebraic problems, including identifying strategies for solving problems based on efficiencies and trade-offs, understanding the structure of a representation and using it to determine a solution strategy, and examining a solved problem to identify errors and alternative solution strategies?	X	
What is the impact of providing the toolkit to teachers on students' long-term outcomes, including self-efficacy, mathematical mindset, and algebra achievement as measured by student assessments and course passage rates for students in introductory algebra?	X	
Were all the toolkit components implemented as intended?		X
What were teachers' and instructional leaders' perceptions of how their capacity to implement the practice guide recommendations changed after using the toolkit? What aspects were perceived as most useful, and what improvements do they recommend for the toolkit? What challenges did they encounter, and how did they attempt to overcome those challenges?		X
How did the professional development supports and resources available to algebra teachers differ in treatment and control schools?		X

**Additional details on regression analyses.** The study team will use the following model to estimate the effects of the toolkit on student and teacher outcomes:

$$(1) \quad y_{ijb} = \sum_{k=1}^h \delta_k S_{ijb,k} T_{jk} + \sum_{k=1}^h \phi_k S_{ijb,k} + \beta X_{jb} + \gamma Z_{ijb} + e_{ijb}$$

where  $y_{ijb}$  is the outcome of interest for student or teacher  $i$  in school  $j$  in pair (block)  $b$ ;  $S_{ijb,k}$  is an indicator for whether an individual is in block  $k$  (in other words if  $b=k$ );  $T_{jk}$  is an indicator equal to 1 if school  $j$  was assigned to the group that implemented the toolkit, and 0 otherwise;  $X_{jb}$  is a vector of school-level covariates measured at the start of the school year;  $Z_{ijb}$  are individual-level covariates measured in the year prior to the study school year; and  $e_{ijb}$  is an individual-level error term.<sup>1</sup> The parameter  $\delta_k$  captures the average effect on the outcome for teachers or students of teachers assigned to the toolkit relative to the business-as-usual control group in block  $k$ . The study team will estimate an overall treatment effect,  $\delta$ , as an average across all blocks.

<sup>1</sup> The study team includes terms for all  $h$  blocks in the model and excludes the intercept term.

To estimate Equation 1, the study team will use weighted least squares and will calculate standard errors using design-based methods that account for clustering using the model residuals (Schochet et al. 2021). The study team will examine outcomes separately for student groups identified by race and ethnicity, eligibility for free or reduced-price lunch (which serves as a proxy for household income), and grade level. The study team will also estimate impacts separately for groups defined by the teacher’s level of experience. When estimating impacts, the null hypothesis significance testing framework will be used as the primary approach for conducting inference.

**c. Degree of accuracy needed**

The study sample will include 20 middle and/or high schools, corresponding to about 60 teachers and 2,640 students. To assess power levels with these study samples, the study team presents minimum detectable effect sizes (MDEs) on key study outcomes measured in standard deviation units (see Exhibit B.3).

**Exhibit B.3. Minimum detectable effect (MDE) size for the toolkit efficacy study**

Data sources and types of outcomes	MDEs for various outcomes (standard deviation units)
<b>Teacher survey</b> <ul style="list-style-type: none"> <li>• Teachers’ pedagogical content knowledge of the recommended practices</li> <li>• Teachers’ use of the recommended practices</li> <li>• Teachers’ self-efficacy</li> </ul>	0.68
<b>District administrative data</b> <ul style="list-style-type: none"> <li>• Students’ score on end-of-course algebra assessments</li> </ul>	0.25
<b>Student survey</b> <ul style="list-style-type: none"> <li>• Students’ self-efficacy and mathematical mindset</li> <li>• Students’ understanding of how to solve algebraic problems</li> <li>• Students’ perceptions of teacher practices</li> </ul>	0.39

Note: MDEs are reported in effect-size units. Calculations assume that (1) desired power is 80 percent; (2) 20 schools will be assigned to a treatment or control group; (3) each school will have three Algebra I teachers, each of whom teaches two classes with each class having an average of 22 students; (4) the study team will obtain survey responses for 85 percent of teachers and students and outcome test score data for 95 percent of students; (5) the school intra-cluster correlation is 0.16 for student assessment outcomes, 0.10 for student survey outcomes, and 0.13 for teacher outcomes; and (6) covariates explain 80 percent of the between-school variance and 40 percent of the within-school variance of students’ assessment outcomes, 20 percent of the between-school variance and 20 percent of the within-school variance for student survey outcomes, and 30 percent of the within-school and between-school variance for teacher outcomes (Deke et al. 2010; Kautz et al. 2021; Schochet 2008).

The statistical power calculations suggest that the study sample will support meaningful statements about the efficacy of the toolkit in improving teacher outcomes and student algebra achievement. The MDEs are 0.25 for student assessments, 0.39 for student survey outcomes, and 0.68 for teacher survey outcomes. These effect sizes are meaningful and reasonable to detect. For example, Garrett et al. (2019) found an average impact of about 0.64 standard deviations on teachers’ instructional practices for a meta-analysis of interventions directed at classroom practice, similar to the study’s MDE for teacher survey outcomes.

The MDE for student assessment outcomes is within the range of the average student-level achievement impact found in studies of intensive coaching and professional development programs (for example, Kraft

et al. 2018). Across all studies considered in Kraft et al. (2018), the average effect size on student achievement was 0.18. However, when focusing on efficacy studies (as with this study), the effect size was 0.28 (above the MDE for this study). Similarly, the average of the effects reported for studies of programs in middle schools and high schools—the school levels relevant to this study—was 0.24, which is close to the MDE for this study. However, the efficacy study of the toolkit is not completely comparable to those summarized in Kraft et al. (2018). On the one hand, the toolkit is less intensive than those programs, suggesting the effect of the toolkit may be smaller. On the other hand, this study’s measure of achievement—end-of-course algebra assessments—is more directly tied to the content of the toolkit compared with the broader measures used in the studies in Kraft et al. (2018), suggesting the effect for this study might be higher. The student assessment MDE is also well below the student-level effect size of 0.49 found in a study of the impacts of a fractions resource kit for teachers (Lewis and Perry 2017). However, one limitation is that the previous study used a researcher-developed assessment closely aligned with the intervention, not a general fractions achievement test. Because the study team plans to use a general Algebra I assessment—rather than one that is aligned closely with the toolkit—a smaller effect than in this previous study might be expected.

In addition, past evidence suggests the MDE for student survey outcomes—such as self-efficacy and mathematical mindsets—is reasonable. Although no prior research has directly estimated the impact of a similar algebra toolkit on measures of these types of social and emotional skills, there is good reason to believe an impact this large is feasible. For example, a meta-analysis by Durlak et al. (2011) estimated that school-based interventions, on average, improved students’ social and emotional skills by 0.57 standard deviations, well above the study’s MDE of 0.39. As with the student assessment outcomes, some caution is warranted because past studies have examined interventions that focus more explicitly on these types of skills.

#### **d. Unusual problems requiring specialized sampling procedures**

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

**Use of periodic (less frequent than annual) data collection cycles to reduce burden.** To minimize burden, the study team is planning to collect data as infrequently as possible while fulfilling the study’s analytic requirements. For the recruitment activities—the focus of this OMB package—the study team will collect recruitment screener data from potential districts only once (starting in late 2023 and ending in early 2024).

As context, the data collection activities and their frequency are listed below and will be included in a future package for OMB clearance. The study team will conduct the following data collection activities only once:

- Collect classroom rosters from treatment and control classrooms (fall 2024).
- Administer the student survey to students in the treatment and control classrooms whose parents or guardians consented for them to participate in the survey (spring 2025).
- Conduct interviews with instructional leaders (spring 2025).
- Conduct focus groups with treatment teachers (spring 2025).
- Collect district administrative data from the 2024–2025 school year (summer 2025).

By necessity, the study team will collect other data more frequently:

- The study team will administer the teacher survey to all treatment and control teachers twice: once in fall 2024 and once in spring 2025. This frequency will allow the team to collect both baseline and endline data. Without baseline data, the study team will not be able to accurately assess the impact of the toolkit on teacher practices.
- To ensure accurate reporting and allow for continuous implementation monitoring, the study team will collect professional development attendance logs after each of the four live professional development sessions, starting in fall 2024.
- To ensure accurate reporting and allow for continuous implementation monitoring, the study team will ask all instructional leaders and treatment teachers to complete an implementation log at the conclusion of each of the four professional development modules, starting in fall 2024. Less frequent collection of this implementation data could lead to errors in the logs because instructional leaders and teachers might not accurately recall their implementation experience, successes, and challenges.

### B3. Methods to maximize response rates and deal with nonresponse

#### a. Maximizing response rates

The study team does not anticipate problems contacting, gaining the cooperation of, and gathering information from district leaders during the recruitment activities. The study team will conduct outreach to district contacts via both email and phone. The study team will conduct calls with district leaders during business hours at times that coincide best with their schedules. The study team will also be flexible during recruitment, allowing districts 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 will employ several strategies to maximize the response rates. These include:

- **Secure school and teacher buy-in.** The study team will follow the districts' lead on how to best engage their schools and teachers in this study. The study will provide staff with an informative flyer about the study and the toolkit. The study team will conduct an orientation session (one for treatment teachers and their instructional leaders, one for control teachers) to explain the study and give them an opportunity to ask questions and fully understand what is involved in participating in the study. Staff will also receive the contact information of the study's data collection lead so they can ask questions at any point during the study.
- **Use various reminder strategies.** The study team will employ several strategies to encourage the completion of the teacher survey, teacher implementation logs, instructional leader implementation logs, and student survey, as well as the distribution and collection of parent consent for the student survey. These strategies could include sending periodic email reminders and text messages and leveraging the working relationship the study team has with the instructional leaders to remind teachers in person. In addition, the baseline teacher survey will be administered in person at the start of the study orientation session, a strategy that should help boost the response rate.
- **Provide incentives.** The study will provide incentives for following through with different data collection activities, including completing the teacher surveys and implementation logs, distributing and collecting parent consent, implementing the student survey, and participating in the teacher focus groups and instructional leader interviews. The second OMB package will provide further details on the incentives.

- **Flexibility in scheduling teacher focus groups and instructional leader interviews.** The study team will work with districts and their participating staff to schedule the teacher focus groups and instructional leader interviews at times that work best for their schedules. In addition, the focus groups and interviews will be conducted virtually, making it easier for staff to participate.

**b. Dealing with nonresponse**

The study team will conduct recruitment activities with districts on a rolling basis. Depending on the likelihood that previously screened districts will participate in the study, the study team will conduct outreach to additional districts, as needed, until the study achieves a study sample with approximately 20 middle and high schools.

For the data analysis activities that will be included in the second OMB package to address research questions 1 through 4, missing values of baseline covariates will be set to a single constant value, and the specification will include an indicator variable for missing values as additional covariates. This approach is appropriate when the covariates are not correlated with research groups, as is the case in evaluations with a random assignment design (Deke and Puma 2013; Puma et al. 2009). The study team will also report the attrition information needed for What Works Clearinghouse reviews of randomized controlled trials, including overall and differential attrition. The study team will report baseline differences in covariates as a way of checking whether attrition led to baseline imbalances between the treatment and control groups. The analytic models will control for these baseline differences.

For the scales based on the surveys, the study team will use an average scale score for the respondent if at least two-thirds of the items on a scale are not missing. Otherwise, the study team does not plan to impute outcome data. For the analyses to address research questions 5 through 7, the number of respondents contributing to findings will be reported (for example, the number of instructional leaders interviewed; the number of teachers who participated in focus groups) along with the total number of participants in the study (for example, the number of instructional leaders in the study; the number of teachers in treatment schools) to assist in interpreting findings given nonresponse.

#### **B4. Tests of procedures and methods to be undertaken**

The study team will pre-test the recruitment screener tool with nine or fewer district leaders. The pre-tests will be conducted via telephone and will help the study team understand whether the recruitment screener tool requires terminology refinements for future conversations. The pre-tests will also confirm for the team whether the recruitment screener tool can be completed within the estimated 60-minute time frame.

The study team will pre-test the teacher and instructional leader implementation logs, the teacher survey, student survey, instructional leader interview protocol, and teacher focus group protocol with staff and students participating in the toolkit's usability testing in summer and fall 2023. Each protocol will be pre-tested with fewer than nine respondents. These instruments will be included in the second OMB package.

#### **B5. Individuals consulted on statistical aspects of the design**

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:

Name	Title	Telephone number
<b>Mathematica staff</b>		
Tim Kautz	Senior Researcher	(609) 297-4544
Hanley Chiang	Principal Researcher	(617) 674-8374
Philip Gleason	Senior Fellow	(202) 264-3443



## References

- Deke, J., and M. Puma. “Coping with Missing Data in Randomized Controlled Trials.” Evaluation Technical Assistance Brief, no. 3. Washington, DC: Administration for Children and Families, Office of Adolescent Health, 2013.
- Deke, J., M. Finucane, and D. Thal. “The BASIE (BAYesian Interpretation of Estimates) Framework for Interpreting Findings from Impact Evaluations: A Practical Guide for Education Researchers.” NCEE 2022-005. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2022.
- Durlak, J. A., R. P. Weissberg, A. B. Dymnicki, R. D. Taylor, and K. B. Schellinger. “The Impact of Enhancing Students’ Social and Emotional Learning: A Meta-Analysis of School-Based Universal Interventions.” *Child Development*, vol. 82, no. 1, 2011, pp. 405–432.
- Garrett, R., M. Citkowitz, and R. Williams. “How Responsive Is a Teacher’s Classroom Practice to Intervention? A Meta-Analysis of Randomized Field Studies.” *Review of Research in Education*, vol. 43, no. 1, 2019, pp. 106–137. <https://doi.org/10.3102/0091732X19830634>.
- Kautz, T., K. Feeney, H. Chiang, S. Lauffer, M. Bartlett, and C. Tilley. “Using a Survey of Social and Emotional Learning and School Climate to Inform Decision Making.” REL 2021-114. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Mid-Atlantic, 2021.
- Kraft, M., D. Blazar, and D. Hogan. “The Effect of Teacher Coaching on Instruction and Achievement: A Meta-Analysis of the Causal Evidence.” *Review of Educational Research*, vol. 88, no. 4, 2018, pp. 547–588. <https://doi.org/10.3102/0034654318759268>.
- Lewis, C., and R. Perry. “Lesson Study to Scale Up Research-Based Knowledge: A Randomized, Controlled Trial of Fractions Learning.” *Journal for Research in Mathematics Education*, vol. 48, no. 3, 2017, pp. 261–299.
- Puma, M. J., R. B. Olsen, S. H. Bell, and C. Price. “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, 2009.
- Schochet, P. Z. “Statistical Power for Random Assignment Evaluations of Education Programs.” *Journal of Educational and Behavioral Statistics*, vol. 33, no. 1, 2008, pp. 62–87.
- Schochet, Peter Z., Nicole E. Pashley, Luke W. Miratrix, and Tim Kautz. “Design-Based Ratio Estimators for Clustered, Blocked RCTs.” *Journal of the American Statistical Association*, vol. 117, 2021.

---

**Mathematica Inc.**

Princeton, NJ • Ann Arbor, MI • Cambridge, MA  
Chicago, IL • Oakland, CA • Seattle, WA  
Woodlawn, MD • Washington, DC



mathematica.org [website](#)

**EDI Global, a Mathematica Company**

Operating in Tanzania, Uganda, Kenya, Mozambique, and the United Kingdom

Mathematica, Progress Together, and the “spotlight M” logo are registered trademarks of Mathematica Inc.