MATHEMATICA
Policy Research

Impact Study of Feedback for Teachers Based on Classroom Videos

Part B: Collection of Information Employing Statistical Methods

May 2, 2017

Submitted to:
U.S. Department of Education
National Center for Education Evaluation
Institute of Education Sciences
550 12th Street, S.W.
Washington, DC 20202
Project Officer: Elizabeth Warner
Contract Number: ED-IES-16-C-0021

Submitted by:
Mathematica Policy Research
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005
Project Director: Susanne James-Burdumy
Reference Number: 50330

This page has been left blank for double-sided copying.

CONTENTS

This page has been left blank for double-sided copying.

TABLES

This page has been left blank for double-sided copying.

PART B. SUPPORTING STATEMENT FOR PAPERWORK REDUCTION ACT SUBMISSION

This package requests clearance for data collection activities to support a rigorous evaluation of video-based observations and feedback for novice and early career teachers. This evaluation is being conducted by the Institute of Education Sciences (IES), National Center for Education Evaluation, U.S. Department of Education (ED). It is being implemented by Mathematica Policy Research, Inc. (Mathematica) and its partners: Clowder Consulting, LLC; Decision Information Resources, Inc. (DIR); Educopia; IRIS Connect; Pemberton Research; WestEd; and Teachstone.

The goal of this evaluation is to examine the impact of video-based observations and feedback on the classroom practices and student achievement of novice teachers (in their first year of teaching) and early career teachers (in their second through fourth years of teaching). This study provides an important test of whether intensive, individualized support for teachers improves their instructional practices and ultimately student achievement. By focusing on novice teachers, the study can inform both teacher induction policies and potentially teacher preparation programs. Examining the impact of this intervention on novice and early career teachers can also inform the effectiveness of providing individualized feedback as a model for teacher professional development programs.

We will implement two versions of the intervention – the full intervention and a less intensive version of the intervention. The full and less intensive versions of the interventions will differ in the number of feedback and coaching sessions that the teachers will receive from the coach. Teachers assigned to the less intensive version of the intervention will participate in 5 feedback cycles that include one-on-one sessions with a coach to review the teacher's performance and provide feedback based on videos of their teaching in their classroom. Teachers assigned to the full version of the intervention will participate in an additional 5 feedback cycles and one-on-one sessions with a coach to review the teacher's performance and provide feedback – for a total of 10 sessions over the course of the school year.

The evaluation will include implementation and impact analyses. The impact analysis will draw on data from teacher surveys, assessments of teachers' pedagogical knowledge and their attitudes towards teaching, video observations of their classroom practices[1], and district administrative records. The implementation analysis will use information on teachers' participation, the amount and type of feedback received, and teaching practices covered to document program implementation.[2] We will also use responses to the teacher survey to describe teachers' professional support and development experiences.

[1] We are not requesting OMB approval for the collection of this information because they will be collected by the study team and will not impose any burden on teachers or district staff.

[2] Ibid

Collection of information employing statistical methods

B1. Respondent universe and sampling methods

The evaluation will rely on a purposive sample of approximately 200 novice teachers and 300 early career fourth- and fifth-grade teachers from approximately 12 school districts in the United States. The study will not statistically sample districts or teachers, and thus we will not make statements that generalize beyond the districts and teachers in the study. By June 2017, we plan to recruit 200 novice teachers from 12 school districts from across the country who will be teaching fourth or fifth grade in the study districts during the 2017-2018 school year. We will group these 200 teachers into pairs based on similarity of their upcoming teaching assignments and route into teaching (traditional or alternative). We will then randomly assign teachers within each group to either the full intervention group or the control group.

The early career teacher sample will include 300 teachers who we will identify prior to the 2018 summer teacher training session. Within each district, we will group 4th and 5th grade early career teachers (entering their second, third, or fourth year of teaching) into clusters of 3 teachers each, based on the similarity of their upcoming teaching assignments for the 2018-2019 school year, their route into teaching, and their years of teaching experience. Within each group, we will randomly assign one teacher to the full intervention group, one teacher to the less intensive intervention group, and one teacher to the control group prior to the 2018 summer teacher training session.

B2. Procedures for the collection of information

a. Statistical methods for sample selection

The study will include a purposive sample of approximately 12 districts that currently do not offer intensive teacher coaching. We will recruit teachers from those districts, focusing on districts with enough eligible novice teachers to meet our sample size targets of 200 novice teachers (approximately 17 per district) and 300 early career teachers (approximately 25 per district). Additionally, to provide a sufficient contrast to the study intervention, districts must not already be providing intensive feedback to novice and early career teachers. Finally, the evaluation will be limited to districts and teachers who are willing to participate. This will result in a purposive sample of eligible districts that are willing and eligible to participate, and teachers from within those districts that are willing and eligible to participate. Although we will not be able to generalize to all novice and early career fourth- and fifth-grade teachers, we will obtain valid estimates of the impact of the interventions for districts and teachers that meet our eligibility requirements and are willing to participate. Below we explain in more detail how we will select districts, teachers, and students for the study.

Selection of school districts. School districts must have at least 25 novice fourth- and fifth-grade teachers to be eligible to participate in this study. Using information from the Common Core of Data, we will begin with the 193 school districts that have at least 30 schools with fourth- or fifth-grade teachers, which we believe will roughly correspond to the set of districts with at least 25 novice fourth- and fifth-grade teachers. To help achieve geographic diversity, we will classify districts by region. We will then reach out to the largest districts in each region (as these are the districts most likely to have a sufficient number of eligible teachers), and ask further questions to verify (1) there are enough eligible teachers in their districts and (2) they don't currently offer similar intensive coaching and feedback to those teachers. We will recruit suitable districts until we reach our sample size target of 12 districts.

Selection of teachers. Within the participating districts we will invite eligible teachers to participate. We will include 200 novice fourth- and fifth-grade teachers for the 2017-2018 school year. After we meet our 2017-2018 sample targets we will cease enrolling new teachers. For the 2018-2019 school year, we will recruit 300 early career fourth- and fifth-grade teachers from the same districts. Teachers will be randomly assigned as described in B1 above.

Selection of students. We will include all students enrolled in the classes of teachers participating in the study. The study team will have access to administrative data on student characteristics and test scores through MOUs established with participating districts. Additionally, the study team will request parent consent for students to be included in video recordings of the study classrooms.

b. Estimation procedures

The evaluation will include four broad sets of analyses: (1) impact analyses, estimating the effects of the study interventions on student and teacher outcomes; (2) subgroup analyses, estimating the effects of the interventions on various subgroups of interest; (3) nonexperimental analyses, estimating the relative effectiveness of interventions based on key intervention and teacher characteristics; and (4) implementation analyses, to learn about implementation experiences and challenges.

Impact analyses. We will estimate the impact of the study interventions after both year 1 and year 2 of the evaluation, using a regression model to compare the outcomes of the teachers randomly assigned to the different intervention groups and the control group as well as the outcomes of their students.

To estimate impacts on student achievement, we will use the following model:

(1) $Y_{ij} = \alpha + \beta^1 T_j^1 + \beta^2 T_j^2 + X_{ij}^{'}\delta + \varepsilon_{ij}$,

where Y_{ij} is the outcome of interest for student *i* of teacher *j*; α is an intercept term; T_j^1 is an indicator for the full intervention group equal to one if the teacher is assigned to that treatment group and zero otherwise; T_j^2 is an indicator for the less intensive intervention group, equal to one if the teacher is assigned to that treatment group and zero otherwise; X_{ij}' is a vector of

baseline school, teacher, or student characteristics; δ is a coefficient vector; and ε_{ij} is a random error term. The baseline characteristics in X'_{ij} will include:

- student characteristics, such as test scores from the year before the intervention, gender, race/ethnicity, free- or reduced-price lunch eligibility, special education status, and English learner status
- teacher characteristics, such as demographic characteristics, age, experience and educational background
- school-level characteristics, such as school-level student achievement and demographic characteristics.

The coefficients β^1 and β^2 represent the average impacts of the full and less intensive feedback and coaching interventions, respectively. We will test the equivalence of β^1 and β^2 to compare impacts of the two versions of the intervention. We will estimate a version of the model that pools teachers from the novice and early career samples as well as separate models for the two samples. The model for the novice teacher sample will not include the T_j^2 indicator for the less intensive intervention group (or the associated β^2 coefficient) because no novice teachers will be assigned to that group.

When estimating student achievement models, the outcome of interest will be a student's state standardized test score in reading or math. For comparability across states, we will convert state test scores to z-scores, subtracting off the mean and dividing by the standard deviation of scores for all students in that state and grade level. To estimate impacts on teacher-level outcomes, such as teachers' practices, pedagogical knowledge, and survey responses, we will estimate a similar model at the teacher level. When using continuous outcomes, such as student test scores, we will estimate Equation (1) using ordinary least squares. When using binary outcomes, such as yes/no answers to teacher survey questions, we will estimate Equation (1) using probit models. In all cases, we will give equal weight to each teacher (regardless of the number of students assigned to teachers) and cluster standard errors at the teacher level.

Subgroup analyses. To help districts and preparation programs develop plans for supporting teachers who might benefit most, we will estimate impacts for various teacher subgroups, including:

1. Teachers with different years of teaching experience (2, 3, or 4 years, for those in the early career teacher sample)
2. Teachers who received different amounts of support from their preparation programs
3. Teachers with higher versus lower baseline performance (as measured by the teachers' knowledge of teaching practices and classroom observation ratings)
4. Teachers prepared through different routes to certification (traditional versus alternative)

5. Teachers with higher versus lower levels of confidence in their ability to teach reading and math at baseline
6. Teachers with higher versus lower proportions of disadvantaged students
7. Teachers with higher versus lower quality coaches (as measured based on observations of coach feedback sessions)

Non-experimental analyses. We will conduct several types of non-experimental analyses to better understand the effectiveness of the studied interventions and their potential application for policy and practice.

1. **Correlational analyses to examine the relationship between teacher effectiveness and practices, attitudes, and beliefs.** We will conduct correlational analyses to explore which teaching practices, attitudes, and beliefs are most important for effective teaching. We will use baseline measures from the Praxis PLT, classroom observation rubric, Haberman Star Teacher Pre-Screener and the teacher survey to examine how baseline teacher knowledge of practices, skill in implementing those practices, and beliefs and attitudes relate to impacts on student achievement. We will examine how differences in teacher practices, knowledge, beliefs, and attitudes within matched groups of teachers relate to teacher effectiveness (as measured by differences in student achievement for teachers in these same matched groups), holding constant the teacher's treatment status. We will augment our main impact model as follows:

   (2) $Y_{ij} = \alpha + \beta^1 T_j^1 + \beta^2 T_j^2 + X_{ij}^{'}\delta + C_{ij}^{'}\gamma + \varepsilon_{ij}$

   where Y_{ij} is end-of-year student achievement; C_{ij} is a vector of teachers' baseline practices, attitudes, and beliefs; and all other terms are as previously defined. The coefficient vector γ will represent the correlations between these aspects of teachers' baseline characteristics and teacher effectiveness, over and above any effects of the interventions. These analyses will draw primarily from the novice teacher sample, for which we will have measures of baseline practices, knowledge, attitudes, and beliefs. Results may inform teacher preparation programs seeking to predict the future teaching effectiveness of their candidates and school districts seeking to predict teacher effectiveness at the time of hiring.

2. **Mediation analyses to examine the mechanisms through which the interventions affect student achievement.** If we find that the interventions affected student achievement, we will conduct mediation analyses to learn about the mechanisms through which this occurred. In these mediation analyses, which we will conduct on the novice and early career samples, we will determine if impacts on specific teaching practices (from classroom observation rubrics) or teachers' knowledge of practices (from the Praxis PLT) explain impacts on student outcomes. Results will be relevant to teacher preparation programs, coaching providers, and school districts interested in understanding the mechanisms through which individualized teacher coaching influences teacher effectiveness. Specifically, the analysis will include four stages.

i. Estimate the impact of the interventions on the mediators, using a model similar to the main impact model (Equation 1), with observation rubric scores or Praxis PLT scores as the outcome.

ii. Estimate the marginal effect of each mediator on student achievement, using a model similar to those used in other correlational analyses (Equations 2 and 3), where we include measures of teaching practices and pedagogical knowledge, as follows:

(3) $Y_{ij} = \alpha + \beta^1 T_j^1 + \beta^2 T_j^2 + X_{ij}^{'}\delta + M_{ij}^{'}\rho + \varepsilon_{ij}$

where Y_{ij} is end-of-year student achievement, $M_{ij}^{'}$ is a vector of mediators, and all other terms are as previously defined. The coefficient vector ρ will capture the mediators' marginal effects on student achievement. These coefficients will not necessarily reflect causal effects, as the mediators may be correlated with unmeasured teacher characteristics. They will, however, provide suggestive evidence of the relationship between mediators and student achievement.

iii. Estimate the implied contribution of each mediator to each intervention's total impacts on student achievement by multiplying the coefficient on each mediator (from stage ii) by the intervention's impact on that mediator (from stage i).

iv. Calculate the percentage of each intervention's total impact on student achievement that can be explained by impacts on each mediator. To do this, we will divide the implied contribution of the mediator (from stage iii) by the intervention's total impact on student achievement (from stage i). For example, if the total impact of an intervention on student achievement is 0.10 standard deviations and the implied impact of teachers' classroom management skills is 0.05 standard deviations, this would suggest roughly half of the intervention's impact is explained by its effect on teachers' classroom management skills.

3. **Cost-effectiveness analysis.** We will conduct cost-effectiveness analyses to determine each intervention's per-student cost of generating a given increase in student achievement. To do this, we will consider both tangible costs (for example, the cost of hiring coaches and providing feedback) and opportunity costs (such as the time teachers spend receiving feedback). This will allow us to compare the costs of the full and less intensive interventions to each other and to the costs of other similar interventions. We will conduct this analysis using novice and early career samples. Results will be relevant for teacher preparation programs and school districts that are considering implementing similar interventions; furthermore, if the interventions do improve teacher effectiveness, cost-effectiveness information will be critical for promoting their adoption.

Implementation analyses. Understanding the implementation experiences and challenges of districts, schools, and teachers participating in the intervention will provide important information for districts and teacher preparation programs considering similar interventions. The implementation analyses will support replication of the interventions in other districts and teacher preparation programs and provide necessary context for impact results.

We will conduct several implementation analyses. First, we will describe the interventions in terms of their intensity, teaching practices covered, use of videos and rubric ratings, and structure of the feedback sessions. Second, we will assess implementation fidelity, focusing on teacher participation, the amount and type of feedback received, and teaching practices covered. Third, we will measure the contrast in professional development and support experienced by teachers in the treatment and control groups. Finally, we will analyze teachers' perspectives on implementation challenges and on the quality of supports provided.

c. Degree of accuracy needed

We estimate that the targeted sample sizes for the study will achieve a minimum detectable effect (MDE) size of 0.09 standard deviations (SDs) on student achievement and 0.36 SDs on teacher classroom observation scores, as well as a minimum detectable impact of 19 percentage points on binary outcomes from the teacher survey. Using a 50 percent subsample – such as for subgroup analyses based on teacher preparation route or baseline performance – the study is designed to achieve MDEs of 0.13 SDs on student achievement, 0.51 SDs on teacher classroom observation scores, and 26 percentage points on teacher survey outcomes.

These target MDEs represent meaningful but realistic impacts, which balance policy relevance against the costs of data collection. Prior studies of coaching interventions have found effect sizes larger than these MDEs. For example, evaluations of the MyTeachingPartner coaching program found impacts of approximately 0.20 to 0.50 SDs on students' end-of-year test scores (Allen et al. 2011, 2015), and Kraft and Blazar (forthcoming) found impacts of about 0.60 SDs on teacher practice outcomes for a coaching intervention. Our proposed sample sizes will be sufficient to detect impacts of these magnitudes.

Table B.1 displays MDE sizes for the full sample of teachers as well as a 50 percent subsample. The full sample will include 100 teachers in each intervention arm for both the novice and early career teacher samples. Two key aspects of the study design maximize power to detect impacts. First, we randomize teachers, rather than schools, to the intervention groups. For a given number of teachers, teacher random assignment improves statistical power relative to school random assignment because outcomes will not be clustered within schools. Second, we will include only fourth- and fifth-grade teachers, because their students will have baseline test scores from third-grade assessments. We will use students' prior test scores as controls to increase statistical power.

The calculations in Table B.1 assume the following: (1) 80 percent power and a 5 percent significance level for a two-tailed test; (2) each teacher will have an average of 22 students; (3) we will obtain outcome test score data for 95 percent of students, (4) 85 percent of teachers will respond to the survey and have classroom observation ratings; (5) the teacher intracluster correlation is 0.16 for student outcomes; (6) covariates explain 80 percent of the between-school variance and 50 percent of the within-school variance of student test scores, 30 percent of the variance for classroom observation outcomes, and 20 percent of the variance for teacher survey outcomes; and (7) 64 percent of control group teachers will feel well prepared or very well prepared to handle a range of classroom management situations. Assumptions on the clustering of outcomes and the explanatory power of covariates for the student analyses are based on data from five large random assignment evaluations in K–12 education (Deke et al. 2010).

Table B.1. Minimum detectable impacts with 100 teachers per treatment arm

Data source	Outcome	Minimum detectable impact	
		Full sample	50 percent sub-sample
District records	Students' reading and math test scores	0.09 SDs	0.13 SDs
Classroom observations	Teacher practices (e.g., scale measure of achievement of lesson aim)	0.36 SDs	0.51 SDs
Teacher survey	Percentage of teachers who felt prepared to handle a range of classroom management or discipline situations	19 percentage points	26 percentage points

SD = standard deviation.

d. Unusual problems requiring specialized sampling procedures

We do not anticipate any unusual problems that require specialized sampling procedures.

e. Use of periodic (less frequent than annual) data collection cycles to reduce burden

These data will be collected during the 2017-2018 and 2018-2019 school years and fall 2019 for 2018-2019 student test scores.

B3. Methods to maximize response rates and deal with nonresponse

There are multiple strategies to maximize response while minimizing burden on respondents. The following techniques will facilitate high response rates: establishing positive relationships with respondents and school and district staff; sending letters to teachers to alert them to an upcoming request to complete the survey; and establishing efficient and flexible scheduling. We will include a statement on confidentiality and data collection requirements (Education Sciences Reform Act of 2002, Title I, Part E, Section 183) in all letters and data collection instruments. We will include a statement indicating that participation is voluntary, yet emphasize the importance of each response for the study findings.

We anticipate full district participation for administrative records and their support for teacher participation. To further solidify administrators' cooperation, we will adhere to additional data collection requirements that districts may have such as preparing research applications and seeking institutional review board (IRB) approvals. Reducing districts' burden in the submission of study data will facilitate attaining a response rate of at least 85 percent on student records and educator administrative data. Federal rules permit ED and its designated agents to collect student demographic and existing achievement data from schools and districts without prior parental or student consent (Family Educational and Rights and Privacy Act (FERPA) (20 U.S.C. 1232g; 34 CFR Part 99)). To maximize the response rate and minimize burden on schools and parents, we will follow these federal rules.

Based on Mathematica's experience in conducting surveys with teachers, we expect at least an 85 percent response rate for the teacher survey and novice teacher assessments. Because teachers will receive full information on study commitments we anticipate high levels of cooperation. To ensure completion of surveys, we will take the following main steps. First, we

will send teachers an invitation letter both by mail and email with a link to the web-based survey. In previous studies in similar settings, we have found that some teachers do not check school email accounts frequently. Therefore, we will also give teachers the option of completing a hard-copy survey, which will be mailed to them at their schools. Over a 12-week data collection period, we will send teachers email and mail reminders (see Appendix B). We propose to offer $30 to teachers who complete the teacher survey. We will also coordinate in-person school visits with our field staff during the last four weeks of data collection to provide teachers with a hard-copy version of the teacher survey. This in-person connection has helped motivate teachers to participate in past surveys. By using these methods, we expect that 85 percent of sampled teachers will submit a teacher survey each round.

We have pretested the survey instrument for clarity, accuracy, length, flow, and wording. Based on the pretest, the instrument is estimated to take 30 minutes to complete. The web-based survey will not allow respondents to enter out-of-range or inconsistent responses, and data entry programs will also check for these. For surveys that are completed on paper, trained quality control staff will identify item nonresponse and reporting errors, by checking for complete and reasonable answers as soon as a hard copy questionnaire is received and follow up with respondents if problems are identified. Weekly reviews of web survey data will allow us to identify potential errors and follow-up with respondents prior to the end of data collection.

The baseline novice teacher assessments will be administered during the teacher training prior to the implementation of the evaluation. The follow-up assessment will be offered in a convenient location in the district. Follow-up will include administering assessments at teachers' schools as needed.

Finally, we will be courteous but persistent in our follow up with participants who do not respond quickly to our attempts to reach them.

B4. Tests of procedures or methods to be undertaken

As much as possible, the data collection instrument for the study draws on surveys and protocols that have been used successfully in previous studies. The pretest assessed the content and wording of individual questions, organization and format of the questionnaire, respondent burden time, and potential sources of response error. We piloted the teacher survey which is an adaptation and extension of existing surveys, and that has limited information on reliability and validity for the population in this study.

We conducted a pretest of the teacher survey with nine elementary school teachers (4th and 5th grade teachers) across nine districts. The purpose of the pretest was to identify problems that study respondents might have providing the requested information and to confirm the level of burden. We sent a full survey packet to pretest respondents and asked them to complete the survey. Respondents returned completed forms by mail. The study's instrument design team conducted a debriefing telephone interview with each respondent reviewing problems teachers may have encountered and following a protocol to probe on a number of items to be sure the survey questions were communicated clearly and collected accurate information. The results of the pretest were used to revise and improve the survey instrument. Respondent burden was estimated at 30 minutes to complete the survey.

The teacher survey was modeled on instruments used in previous studies, the Impact Evaluation of Teacher Preparation Models and the Impact Evaluation of the Teacher Incentive Fund. The school records data collection form was modeled on forms developed for the Impact Evaluation of the Teacher Incentive Fund. The school records data request will not be pretested for this study as it has been used effectively in previous studies for similar purposes. The teacher assessments will not be pretested in light of their established use as teacher assessments.

The parent consent forms will not be pretested as they were modeled on consent forms that were successfully used for the Impact Evaluation of Teacher Preparation Models and the Evaluation of the Impact of Teacher Induction Programs. They also meet requirements of the Institutional Review Board and individual district requirements.

We will provide a help desk for questions and our field staff will be available throughout the data collection period and trained to respond to frequently asked questions about the study and individual forms, so they can provide technical assistance and report any issues that come up in the field.

B5. Individuals consulted on statistical aspects of the design and on collecting and analyzing data

The following individuals were consulted on the statistical aspects of the study:

Table B.2. Individuals consulted on statistical design

Name	Title	Telephone Number
Hanley Chiang	Senior Researcher, Mathematica	617-674-8374
Melissa Clark	Senior Researcher, Mathematica	609-750-3193
Jill Constantine	Vice President of Human Services Research, Mathematica	609-716-4391
Mark Dynarski	Founder and President, Pemberton Research	609-443-1981
Brian Gill	Senior Fellow, Mathematica	617-301-8962
Susanne James-Burdumy	Senior Fellow and Area Leader, Mathematica	609-275-2248
Alison Wellington	Senior Researcher, Mathematica	202-484-4696

The following individuals will be responsible for data collection and analysis:

Table B.3. Individuals responsible for data collection and analysis

Name	Title	Telephone Number
Tim Bruursema	Survey Researcher, Mathematica	202-484-3097
Melissa Clark	Senior Researcher, Mathematica	609-750-3193
Kristin Hallgren	Senior Researcher, Mathematica	609-275-2397
Sheila Heaviside	Associate Director of Survey Research, Mathematica	202-484-3096
Mariesa Herrmann	Researcher, Mathematica	609-716-4544
Susanne James-Burdumy	Senior Fellow and Area Leader, Mathematica	609-275-2248
Jeffrey Max	Senior Researcher, Mathematica	202-484-4236
Catherine McClellan	Principal Scientist, Clowder Consulting	609-915-6676
Alison Wellington	Senior Researcher, Mathematica	202-484-4696
Emilyn Whitesell	Researcher, Mathematica	617-588-6691
Eric Zeidman	Associate Director of Survey Research, Mathematica	609-936-2784

REFERENCES

Allen, Joseph P., Christopher A. Hafen, Anne C. Gregory, Amori Y. Mikami, and Robert Pianta (2015). "Enhancing Secondary School Instruction and Student Achievement: Replication and Extension of the My Teaching Partner-Secondary Intervention." *Journal of Research on Educational Effectiveness,* vol. 8, no. 4, pp. 475–489.

Allen, Joseph P., Robert C. Pianta, Anne Gregory, Amori Yee Mikami, and Janetta Lun (2011). "An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement." *Science,* vol. 333, no. 6045, pp. 1034–1037.

Deke, John, Lisa Dragoset, and Ravaris Moore (2010). "Precision Gains from Publicly Available School Proficiency Measures Compared to Study-Collected Test Scores in Education Cluster- Randomized Trials." NCEE 2010-4003. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.

Kraft, Matt, and David Blazar (forthcoming). "Improving teachers' practices across grades and subjects: Experimental evidence on individualized teacher coaching." *Educational Policy.*

This page has been left blank for double-sided copying.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ■ ANN ARBOR, MI ■ CAMBRIDGE, MA ■ CHICAGO, IL ■ OAKLAND, CA ■ TUCSON, AZ ■ WASHINGTON, DC ■ WOODLAWN, MD

MATHEMATICA
Policy Research