

Contract No.: ED-04-0112/0007
MPR Reference No.: 6398-050

MATHEMATICA
Policy Research, Inc.

**Responses to OMB
Questions for the
Evaluation Of Moving
High-Performing
Teachers To Low-
Performing Schools**

October 23, 2008

Submitted to:

Institute of Education Sciences
U.S. Department of Education
555 New Jersey Avenue, NW
Washington, DC 20208-5500

Submitted by:

Mathematica Policy Research, Inc.
600 Maryland Ave. S.W., Suite 550
Washington, DC 20024-2512
Telephone: (202) 484-9220
Facsimile: (202) 863-1763

Project Officer:

Elizabeth Warner

Project Director:

Steven Glazerman

OMB COMMENTS ON EVALUATION OF MOVING HIGH PERFORMING TEACHERS TO LOW-PERFORMING SCHOOLS

PILOT

QUESTION 1 - PILOT

What are the major pieces of information the study team expects to get out of the pilot test?

RESPONSE

The pilot study has several objectives, but the overall goal is to inform the full scale study and identify possible improvements to the evaluation design or the intervention itself. In particular, the pilot study aimed to test out the following procedures:

- District recruitment: develop and refine messages to maximize district participation and receive input on the program design from key decision-makers
- Value-added estimation: write computer programs for identifying high-performing teachers using district administrative records and evaluate internal procedures for gathering, safeguarding, and processing school district data for the purpose of value added analysis; also, examine the distribution of high-performing teachers by school type, to verify that there was an imbalance in the distribution of teacher quality by school performance
- Teacher recruitment: test messages, develop recruiting strategies, and refine planning assumptions regarding the conversion rate of transfer candidates to placements
- School recruitment: test messages and refine methods for identifying vacancies in a timely manner, explaining the intervention, and conducting random assignment
- Primary data collection instruments: pretest surveys using a realistic pool of teachers and principals

Perhaps the most critical input that the pilot study provides to the full scale study is the information on teachers' decisions to transfer when offered the incentive. In addition to simply observing and calculating the percentages of candidates (high-performing teachers who were invited to apply for a transfer) who applied, interviewed, and ultimately transferred, the researchers leading the pilot study seek to use the proposed Candidate Survey to examine the

types of teachers who respond to the transfer incentive and the factors that weigh in their decisionmaking process. For that reason, IES is seeking clearance to conduct a full Candidate Survey in the pilot instead of a pretest with less than 10 respondents.

QUESTION 2 - PILOT

Please explain how you will approach disclosure avoidance in developing releasable results on “...teachers, schools...” from your pilot study. Specifically, given 1 district and only about 20 applicants, how will you meet your commitment that the reports will not “associate responses with a specific district or individual?”

RESPONSE

While the study team will analyze the data gathered in the pilot to help identify changes needed to the survey instrument and also to inform the recruitment process, the study will not generate a pilot report. Pilot data will be used internally to inform the refinement of the study design.

QUESTION 3 - PILOT

Please confirm the apparent expectation (based on the burden table) that you will receive 100 percent response rate on the Candidate Survey in the pilot and explain why you anticipate this outcome.

RESPONSE

The study does not anticipate receiving a response rate of 100 percent for the pilot. However, since the study team did not have a contractual target response rate for the pilot and had not yet pretested the instrument, an estimate of the highest level of possible burden was presented in this table.

With more experience in the district and a more clearly defined data collection period, the expected pilot response rate is 80 percent. The burden table will be revised to reflect this, per OMB's direction.

TEACHERS

QUESTION 1 - TEACHERS

Please clarify that the only teachers eligible for the MTRP are those who have:

- a. taught the same subject for 3 or more years,
- b. within a single selected district,
- c. that has the granularity within its data systems to associate teachers with specific students and subjects for particular fractions of the year, and
- d. are deemed via the analysis to be high performing based on data solely from within that district.

RESPONSE

To be eligible, the teachers are required to have taught in one of the targeted grade/subject combinations within the district for each of the last three years. The three possible grade/subject combinations are: (1) middle school math; (2) middle school English/language arts; or (3) upper elementary math/reading. The program works with the district to select the combination(s) to target.

The district's data system requirements referenced above are correct. The high-performing designation is based on data solely from within that district, as stated in the question.

QUESTION 2 - TEACHERS

What does the study team know about the likely prevalence of teachers meeting all of these criteria?

RESPONSE

In the pilot district the value added analysis was based on 518 eligible elementary teachers, among whom 52 (the top ten percent) among 92 schools were designated high-performing and 203 middle school math teachers, among whom 41 (the top twenty percent) among 32 schools were designated as high-performing. Teachers in non-tested grades and subjects or in non-selected grade-subject combination were not eligible. Because of some limitations in the data provided to us by the pilot district, we were not able to definitively classify all teachers by their reason for inclusion or exclusion or produce a definitive count of teachers who ever taught in the targeted grade-subject combinations during the three-year evaluation period. However, we know that the following teachers were excluded: those who retired during the three-year period, those who left the district for other reasons, those who took a leave of absence of one year or more, those who began teaching in the district or transferred into the targeted grade/subject combination after 2004-2005, and those who changed grades or subjects during this time period. We suspect that beginning teachers and retirees dominate this group of excluded teachers. To us, these seem like reasonable exclusions consistent with the purpose of the policy being tested.

QUESTION 3 - TEACHERS

Please also clarify that treatment teachers are limited to those who meet the criteria above AND also accept a MTRP slot within the same district as they had been working. If a treatment school were to hire a teacher outside of this program, would he/she be included or excluded from the analysis?

RESPONSE

The treatment teachers are indeed limited to those who meet the criteria above and also accept a designated program slot within the same district. It is possible that a match with a study candidate teacher is not made and a treatment school hires someone else instead. (The successful match rate is a key piece of knowledge that will emerge from the study). The individual who fills the position will be included in a survey of new hires (the survey that will be discussed in the forthcoming OMB clearance submission), but will be excluded, along with the school's control group counterpart, from the main impact analysis that focuses on the effectiveness of master teachers.¹ Each treatment school will have a control group counterpart because randomization will be done within matched pairs.

¹ The study has updated the language to drop references to "master" teachers when communicating with participating districts and schools, but the terminology is retained in this document for consistency.

QUESTION 4 - TEACHERS

Please also clarify who will be asked to complete the MTRP questionnaire. If only those candidates who have applied for the program, why does question B4 ask whether the person completed an application?

RESPONSE

If it was stated that the Candidate Survey population consists of only those teachers who apply, then that statement was an error. The Candidate Survey is designed to collect data from teachers who are “eligible” for the program – all high performers teaching the specified grade level/subject area. The Candidate Survey target population is all teachers invited to apply based upon the study determination that they are high-performing and still in the district when program implementation begins. This includes teachers who do not complete an application as well as those who do complete the application process.

QUESTION 5 – TEACHERS

NCES's Schools and Staffing Survey, Teacher Follow up estimates that more teachers move across districts than within districts in a given year. What are the implications of this mobility on the study?

Has the study team considered how to incorporate cross-district mobility into its study design?

RESPONSE

The Schools and Staffing Teacher Follow-up Survey is a nationally representative sample, whereas the IES study aims to focus on a set of districts that are much larger with many more intra-district mobility opportunities than the average school district in the U. S. Therefore, the MTRP was designed as an intra-district transfer incentive program. During the feasibility phase of the project, the study team carefully considered the possibility of including inter-district transfers in the program. However, it was judged too risky to start off with such a program/study design when there are many districts for which an intra-district transfer program could be beneficial. The principal threat to feasibility posed by the inter-district model would be the difficulty obtaining comparable data from sending districts that stood to lose their best teachers.

In terms of incorporating inter-district teacher mobility into the study design of an intra-district program, such mobility is inevitable and is part of the phenomenon under study. For example, does the presence of the program slow the transfer of high-performing teachers out of the system? Transfers of high-performing teachers into struggling schools in large urban districts is a more rare phenomenon, but the control group is designed to allow the research team to measure it along with the other ways in which principals in such schools fill their vacancies.

QUESTION 6 - TEACHERS

Even though teachers agree to remain in schools for 2 years in order to receive a bonus, some teachers may leave. How will the analysis address these leavers? Will the study collect additional data from leavers on their reasons for leaving?

RESPONSE

Teacher attrition is one of the key outcomes for the study. In that sense, teacher attrition is part of the study and not a hindrance to the design. The test score analysis will continue to include the departing teacher's classroom in estimating the effect of using the policy. In other words, the impact estimate will include the combined effect of having a high-performing teacher and having a replacement teacher, should the high-performing teacher leave. Data on the nature and timing of teacher attrition will be critical for interpreting such impact estimates. Data will be collected on the circumstances of teacher exits from the principal in a separate survey that is part of a second (forthcoming) submission to OMB for the full-scale study.

QUESTION 7 - TEACHERS

Receiving the intervention means that a school "may hire" a master teacher. How will the analysis address schools that do not hire teachers?

RESPONSE

The study will take two approaches to slots that fail to match with a master teacher. First, the analysis will drop the no-master teacher treatment schools and their control group counterparts (since randomization is done within matched pairs) to estimate the impact of master teachers in their new settings.

Second, the analysis will be repeated to include all schools that were part of random assignment, using an intent-to-treat model. In this model, the new hire will be part of the treatment group regardless of whether that teacher was hired through the program or in spite of it. To help readers interpret the intent-to-treat findings, the study team will carefully track and describe the background and pathway of the new hires in the treatment and control groups. This includes reporting the percentage of treatment new hires who were hired through the program (the teachers identified as high-performing based upon value added). This is expected to be a very high percentage, but it does not have to be 100 percent. If it is very low, then the analysis will have the Candidate Survey and the Principal Survey to help explain the low match rate. The study will also have the New Hire Survey data to explain who was hired instead (or moved from another grade within the school to fill the vacancy).

QUESTION 8 - TEACHERS

What if there are highly effective teachers already in low performing schools? Will they be eligible for bonuses?

RESPONSE

The study does expect that there will be some high-performing teachers already in low-performing schools. They will be eligible to receive a retention stipend, equal to half that of a transferring teacher.

QUESTION 9 - TEACHERS

Who will pay the bonuses-the study or the districts?

RESPONSE

In other NCEE studies the evaluation has paid for the intervention. Therefore, the study will pay the cost of the intervention (the pay incentive) in this evaluation as well.

STUDENT RECORDS

QUESTION 1 – STUDENT RECORDS

Please clarify whether the study team requires direct student IDs rather than just a linking ID. If needed, please justify. If not needed, we suggest incorporating this fact into the outreach to districts.

RESPONSE

The study will request unique student record identifiers that minimize disclosure risk while ensuring the integrity of the teacher-student match as well as the quality of the merge between test score data and other types of student data (enrollment and demographics). The study team always requests ID codes that cannot be linked back to the student by anyone outside the district and this information will be incorporated into the outreach with districts, per OMB's recommendation.

QUESTION 2 – STUDENT RECORDS

Why does NCEE believe that it will require 120 hours per district to pull student records? We have seen many other NCEE studies estimate burden at a small fraction of this amount of time.

RESPONSE

This conservative estimate was provided because there is a considerable degree of uncertainty about district's ability to compile historical data that links teachers to students and that links enrollment data to student records. However, the effort involved in creating or fixing links from students to teachers or teachers over time is unusual for this study because of the scale. It will be necessary to clean the data for each entire district for several years, including historical data which may have been linked using obsolete identifiers, in order to avoid disqualifying any potentially eligible teachers from being identified as high-performing. For many districts, the burden may be significantly lower because the study team would be able to absorb nearly all of the burden by working with raw data tables and doing the merging for the district, including sending staff on site to help with the programming components of the study. The study team expects that many district research and evaluation directors are eager to participate in this aspect in order to gain outside expertise on how to clean the data and construct the type of value added indicators that could be used for this proposed program as well as other district initiatives.

QUESTIONNAIRE

QUESTION 1 - QUESTIONNAIRE

Please justify the various sections of the questionnaire (including explaining why it collects some of the same information collected on the application). Please also provide the source of the listings within questions if from previously validated instruments. For items not previously validated, please indicate why you believe that pretesting on fewer than 9 teachers and principals will be adequate or preferable to requesting additional burden via this information collection.

RESPONSE

In the development of the Candidate Survey, an effort was made to minimize overlap with the application. However, the Candidate Survey will be collected from non-applicants as well as from applicants and certain key data will be needed from all respondents. If the applications are in and available prior to the fielding of the Candidate Survey, the survey team will create a separate questionnaire version for applicants and non-applicants, eliminating items that are identical between the final application and the questionnaire.

The Candidate Survey questionnaire includes the following sections/domains

Questionnaire Section/ Domain	Purpose
Teaching Experience Teachers' Education and Certifications	These items provide years of teaching experience used to confirm eligibility (and later compare treatment and control teachers), examine teacher mobility and satisfaction with teachers' previous year school (a possible factor in their decision to transfer), and education and other teacher quality indicators that may be associated with key outcome measures.
Experience with Program and Future Plans	Provide information about how teachers make decisions to transfer between schools and to identify factors associated with those who stay and leave
Compensation Demographic Characteristics	The study will look at previous compensation as well as compensation earned outside of teachers as a possible factor in decision-making and to help describe the sample of teachers participating in the study.
Family and Housing Commitments	These items ask about length and mode of the teacher's commute, home ownership, marital status and family status as items related to the degree of flexibility that may vary and influence rates of transfer.

Items on teacher satisfaction, while based initially on the SASS items, have been modified to better reflect the interests of this study. These items have been reviewed by the survey team and their subcontractors and will be pretested in the pilot survey in order to provide an additional check on their validity and appropriateness to the context.

QUESTION 2 - QUESTIONNAIRE

To what degree has NCEE analyzed NCES's Schools and Staffing Survey (SASS) and SASS Teacher Follow up Survey (TFS) questionnaires as a source for these questionnaires? They cover many of the same topics. This approach would seem to allow both for previously validated instruments, as well as the ability to compare results to an external source. As a couple of examples of differences that may not be warranted, TFS uses a 1 year, rather than a 6 month, reference period in asking about job search, and TFS uses "full or part time" to calculate years of experience, while this instrument is silent on part time.

RESPONSE

The Schools and Staffing Survey (SASS) was a major source for the initial pool of items Mathematica compiled for the questionnaires and we looked primarily at the Teacher Follow-up Survey (TFS). Items from other NCES and IES surveys were also included in that pool (e.g., FRSS, Evaluation of Teacher Induction Programs, etc.). Over the course of the review process, some items were revised or dropped, while others were added to best serve the goals of this study.

In terms of the specific examples above, for years of teaching experience, we will modify the questionnaire including an instruction to report for "full or part time" in the years of teaching experience item. However, for the reference period, the research team is interested in teacher's more recent job searchers – those that roughly coincide with the recruitment period for the study – for the full-scale study that would be January – July 2009.

OTHER

QUESTION 1 - OTHER

Please clarify why the application asks about coaching and other extracurricular activities if they "will not affect the processing or consideration of your application in any way."

RESPONSE

While these activities do not affect the processing or consideration of the application to the study, these factors are of importance to the successful matching of teacher job candidates to low performing school that might provide similar coaching and other extracurricular activities.

QUESTION 2 - OTHER

A-6 states that the study will combine value-added data with "more readily observed proxies for teacher quality." Is one of the goals of the study to validate these proxies? What type of information will the study provide on these proxies (can you give us some examples)?

RESPONSE

IES, with input from the TWG, has since decided not to include such proxies. The experts on the TWG convincingly argued it was important to base eligibility on only objective measures and they thought that value added measures would be best.

QUESTION 3 - OTHER

Are principal interviews a reliable and sufficient source of information on the resource allocation effects? Would it be possible to collect school budgets to determine how resources are reallocated?

RESPONSE

The study proposes to conduct a principal survey (not interviews), which will be part of a future submission for clearance to OMB. The goal of the principal survey is to assess whether principals exhibit offsetting behaviors in response to the treatment, such as assigning more support to the other teachers in the building or assigning more difficult students to the master teacher. School budgets leave considerable discretion in the hands of principals, who are responsible for allocating the types of resources and setting the types of policies that we think are most likely to be adapted in response to the treatment. We believe that principal surveys represent the most cost-efficient method of gaining a handle on this potential issue. We will consider adding some questions to the New Hire Survey to gain a second perspective on how students and mentoring are distributed within the grade level.

QUESTION 4 - OTHER

Can we get more explicit information on the outcome variables and how they will be measured? Will direct effects on student achievement be measured as student gains? When looking at spillover effects, will the analysis look at spillover effects in specific affected grades or just in the whole school?

RESPONSE

The main outcome variable is student achievement, measured using test scores on the districts' usual NCLB assessments. Specifically, the research team will estimate a student achievement growth model that tracks test scores in the study's targeted schools and grades at the end of each of the two program years and uses a regression model to estimate the relationship between treatment status and test scores. The model will include pre-test scores (prior year test scores) and student background data as covariates in order to increase the precision of the estimates. By including the pre-test measure, one can interpret the impacts as effects on student achievement *growth* or gains during the year in question.

The analysis plan recognizes that master teachers may have an impact on their own students (direct effect) as well as on other students in the school (indirect effects) because of spillover. For example, the master teacher may help other teachers plan lessons or alternatively the presence of a highly paid master teacher may disrupt the school and hurt morale. Another type of spillover would result if the presence of a master teacher leads a principal to assign students or teacher supports to classrooms differently than he or she would have in the absence of the program.

To account for these different hypotheses, the analysis will be repeated three times. First, the regression will be estimated using only targeted classrooms in targeted grades (e.g. the master teachers in the treatment school and the corresponding new hires in the control schools) to estimate the direct effect under the assumption of zero spillover effects on the treatment

classrooms. Second, the regression will be estimated using the targeted grades in the treatment and control schools. This captures the total effect (direct and indirect) assuming no spillover into adjacent grades. Finally, the regression will be estimated using all tested grades in the treatment and control schools.

QUESTION 5 - OTHER

In site selection, will the study team also look at stability of leadership and the term of the relevant union contract?

RESPONSE

Yes, both factors are part of the consideration (see also our response to Question 9 below). District leadership stability and the timing of union contract negotiation are taken into account during the district selection and recruitment process because the project requires buy-in from the teachers' union, the superintendent, and the key members of the district's senior leadership team. Most recruiting visits will include an in-person meeting with all of these parties represented. As a matter of routine, the study team will ask when the contract expires or is due to be re-negotiated and will ensure that the bargaining does not threaten the project (or vice versa). The team has already eliminated from consideration some districts that are in the process of replacing a superintendent, negotiating a difficult union contract, or are being run by an interim superintendent.

School building leadership is also important. IES recognizes that principals turn over somewhat regularly and has taken that into account by directing the study team to consider eliminating from the study sample any schools that are likely to lose a principal during the critical early phase of the program, when principals must interview and hire the master teacher candidates. Principal turnover later in the program can be disruptive to the school but does not threaten the feasibility of the program or the study.

QUESTION 6 - OTHER

Will the study team require districts to sign an MOU?

RESPONSE

Yes. Each district will be asked to sign an MOU.

QUESTION 7 - OTHER

On B-6, ED states "we have designed the study to be able to detect impacts of...between 15 and 20 percent of a standard deviation." Who is "the Government" in this sentence? If IES, please rewrite to the first person. How does this relate the impacts found in other studies looking at teacher related interventions and to studies of the distribution of teacher impacts on student learning? Will any effort be made to see if these impacts are significantly larger than those found when doing the original value added analysis to choose eligible teachers for the study?

In the quoted sentence, "the Government" was indeed referring to IES. The sentence will be revised as suggested.

Studies of the magnitude of teacher effects suggest that compared to having the average teacher, a student having a teacher who is in the top 16 percent of all teachers (a one standard deviation difference) could see an effect of as much as one-third of a standard deviation in student test scores (Nye et al. 2004). The range of teacher effect estimates includes values as low as 0.11 standard deviations. Accounting for the possibility that the teacher contrast is slightly less than a full standard deviation (e.g. a teacher at the 80th percentile would be 0.845 standard deviations above the mean), the selected threshold should put the study close to the middle of the range of expected impacts. Given the high cost of incentivizing teachers to move, impacts that are smaller than 15 percent of a standard deviation are not likely to be large enough for policymakers to justify the intervention.

The study team will compare the size of the estimated program effects – that is, the impact that master teachers have in their new schools – to the size of the difference in teacher effects estimated from the original value added analysis that was used to identify the teachers. This comparison will help readers of the study understand the degree to which the new settings might have been more challenging than the master teachers' previous settings.

QUESTION 8 - OTHER

How does this value added analysis compare to the work of established value-added methodologies (Goldhaber, Ladd, etc.)?

RESPONSE

Our methodology is consistent with that used by Robert Meyer of the Value Added Research Center at the University of Wisconsin. Dr. Meyer is a leader in the field of value added modeling and was the graduate school mentor of the study's principal investigator. Dr. Meyer and Tom Kane of the Harvard Graduate School of Education are on the study's TWG and have been advising the project on value added methodology.

QUESTION 9 - OTHER

How extensive is MPRs knowledge of all school districts' teacher union, reform, etc., climate? How can that knowledge be supplemented as needed to aid in district selection? How might these climate factors bias the results of the study?

RESPONSE

Mathematica has worked with several of the prospective districts in the study on past evaluations and is very familiar with the district leadership in those sites. In some cases the leadership has turned over and the research team has worked with the current leaders in their previous positions in another district. In addition, The New Teacher Project, a subcontractor to Mathematica, specializes in district-union relations and has extensive experience with research on and conducting direct negotiations over compensation, teacher transfer rules, and other human resource policies.

The study team acknowledges that districts with leadership turmoil, poor district-union relations, or hostility to reform will be under-represented in the study. The study makes no pretense to being nationally representative of all districts. Rather, the aim is to test a policy in districts that are disposed to adopt such a reform in the first place. The authors of reports will be very clear about this context when explaining the findings.

REFERENCES

Nye, Barbara, Spyros Konstantopoulos, and Larry Hedges. (2004) "How Large are Teacher Effects?" *Educational Evaluation and Policy Analysis*. 26:3, pp. 237-257.