Impact Evaluation of Math Professional Development

Responses to Questions from OMB

May 24, 2013

Study Design

1. Given that random assignment will take place within schools, is there any concern over possible spillover of treatment to control teachers which might dampen the treatment effect estimate? Is there statistical method to control for this possibility?

We believe the risk of spillover is low. Prior studies based on within-school random assignment of teachers in fact do not provide strong evidence of it (Allen, Pianta, Gregory, Mikami, & Lun, 2011; Dolan, Kellam, et al., 1993; Hitchcock, Dimino, Kurki, Wilkins & Gersten, 2011; Finkelstein, Hanson, Huang, Hirschman, & Huang, 2010). As part of designing this study we consulted with the study's Technical Working Group about spillover risks. Based on this discussion, we believe that the nature of the PD intervention likely makes the risk of spillover low. The PD includes 93 hours of structured, intensive instruction and discussion focused on complex mathematical content delivered by trained and qualified mathematicians, math educators, and expert math video coders. In addition, the video feedback component of the PD is tailored to each treatment teacher's needs. Given these core features of the PD, informal sporadic sharing of information alone, without intensive study, deliberate efforts, and formal support from trained instructors or facilitators is unlikely to have an appreciable impact on the math content knowledge and quality of instructional practice for control teachers. Potential spillover from sharing would likely result in at most gains in procedural knowledge or other types of more superficial changes. However, our teacher knowledge test, classroom practice measure, and study-administered student test were intentionally selected to be sensitive to changes in conceptual math knowledge and connections between math concepts. Thus, the gains from this type of sharing would not likely result in significantly higher scores on our study measures. Indeed, research indicates that even with targeted, intensive PD activities, changing outcomes for teachers and students is difficult (e.g., Garet et al., 2011).

Although the risk of contamination is believed to be low, we have a few strategies to minimize spillover risks. First, Access to the study PD will be tightly controlled. Attendance will be taken at all PD sessions to ensure that only treatment teachers participate. Developer-supplied, external instructors will teach Intel Math, so they will have no contact with control teachers. Facilitators for the Math Learning Community and Video Feedback Cycles are district based, however they are prohibited from working with control teachers during the study year. Treatment teachers will be asked to not share intervention materials (e.g., their Intel Math books) with control teachers during the study year. Second, the random assignment plan will minimize the number of treatment teachers in each school. For example, in schools with three teachers, we will select one treatment teacher and two control teachers. This unbalanced random assignment will minimally affect the study's statistical power, while minimizing the opportunity for control teachers to seek out the study PD from treatment teachers in their schools. Finally, we will focus on recruiting districts and schools that have limited existing structured opportunities for sharing, especially of math content, between teachers in the same grade (e.g., through common planning periods and professional learning communities).

In addition to minimizing spillover risks, we have a few strategies to try to detect the presence of potential spillover. First, we plan to measure both treatment and control teachers' content knowledge and classroom practice at the beginning of the 2013-14 school year, after the treatment teachers receive the core, 80-hour Intel Math PD but before any potential spillover could occur. Thus, this measure will provide a spillover-free impact estimate of the core piece of the study PD. Second, we also plan to measure both treatment and control teachers' content knowledge and classroom practice at the end of the 2013-14 school year. With two measures for both treatment and control teachers, we can examine the growth in each group's scores to better understand whether a null impact finding might have been due to spillover. For example, if treatment and control teachers both experience low growth during the year, this would suggest that the PD simply may be ineffective with no concerns about spillover. However, if treatment and control teachers both experience similarly high growth, then this could suggest that the PD may actually be effective, but we may not be able to detect treatment-control impacts due to potential spillover. Finally, the teacher survey asks both treatment and control teachers about their PD experiences during the 2013-14 school year, including the frequency, structure, and content of those experiences. Analysis of these data will provide evidence on whether the study achieved a sufficient PD service contrast between treatment and control teachers.

2. Is there any concern over bias associated with using the Mathematical Quality of Instruction (MQI) as both the tool that is part of the Treatment and as the tool to measure an actual outcome of interest (classroom practice)? That is, will the fact that Treatment teachers already have exposure to this assessment tool possibly bias the treatment effect on classroom practice upward (i.e. give them an unfair advantage since they know the tool so well)?

Alignment between the intervention and outcome measures is desirable to a certain degree. In this study, the intervention aims to improve the general quality of mathematical instruction and so the outcome should measure the quality of mathematical instruction. Over-alignment becomes problematic when teachers are, either consciously or subconsciously, able to obtain a higher score on the instrument primarily because of familiarity with how the instrument is scored, rather than because of actual improvements in the general quality of their mathematical instruction. The study PD is intentionally aligned to some degree with the outcome measures, but we are not concerned about over-alignment (i.e., bias associated with using the MQI as part of the intervention and outcome), since treatment teachers will not know the ins and outs of the MQI after participating in just three video feedback sessions. The feedback is not a case of teaching to the test; that is, it is not focused on teaching teachers how to score highly on the MQI. Teachers will not receive their MQI scores, nor will they be instructed on how the MQI is scored. Instead, the feedback will focus on helping teachers improve their instruction in mathematics in three general areas that are most related to math content, which is also the focus of the other two PD components: (1) whether they provide rich mathematical explanations that are conceptual rather than simply procedural and include appropriate use of multiple representations; (2) whether they appropriately respond to student questions and promptly correct student misconceptions; and (3) whether the teacher's language is generally devoid of mathematical errors and imprecision. On the other hand, the outcome lessons will be scored using the full MQI, which includes a few additional areas such as student engagement. These areas of instruction captured by the MQI represent general, complex teaching behaviors and content knowledge that cannot be "staged"

even by those teachers who are familiar with the instrument. Moreover, the math content in the feedback lessons will be related but not identical to the math content in the outcome lessons, so it would be very difficult for teachers to boost their MQI scores on the outcome lessons simply by superficially replicating what they received feedback on.

Recruitment

3. Do you expect/have you factored in any effects on study recruitment and participation due to adverse effects from sequester-related budget cuts?

We do not anticipate adverse effects on the study of sequester-related budget cuts. We will focus on recruiting larger districts to ensure that even after accounting for budget cuts these districts have a sufficiently large enough pool of teachers to draw from to reach the target sample size of 200 teachers. In addition, by the time we move into the final stages of recruitment, we expect districts to be able to confirm that the volunteer teachers in our sample will not be adversely affected by the cuts next year. Many districts will not be able to offer substantial PD to teachers next year as a result of sequester-related budget cuts, so if anything we expect that the study's offer of high-quality PD to approximately 30 teachers per district will be appealing to many districts. If it turns out there are unusual circumstances related to budget cuts that lead to teacher attrition during the study year, our sample size of 200 teachers is large enough to absorb this attrition and still have sufficient statistical power to detect meaningful impacts.

Teacher Survey

4. Item 2d (and similarly Items 5d and 9d) is very complex and we are concerned that it may not yield valid responses. We think it might help to clarify it by adding an example of what is meant by the "underpinnings K-12 mathematics concepts."

As you suggested, we have added text providing a clarifying example of what "underpinning of K-8 mathematics concepts" means to item 2d, 5d, and 9d.

5. Item 15 is slightly confusing. For example, if a person were to have taken Calculus 1, Calculus 2, Calculus 3, where would they count those courses – it's not necessarily clear. Also, "differential equations" is listed in two places (d. and h.). Is that intentional?

The duplicate reference to differential equations was in error. We have removed the term from 15d (now 16d) and kept it in 15h (now 16h). To address the concern of possible confusion over how to mark courses with similar names, we have changed the heading of the table to "Content of Course" and retained the "Number of Courses Completed" heading. We think these headings will provide sufficient guidance to teachers in the scenario presented in the question above.

6. Did you consider asking whether a teacher was trained through traditional preparation program or an alternative route to certification explicitly?

We have added the Schools and Staffing Survey (SASS) version of this question to the survey; see item 14 in the revised teacher survey.

References

Allen, J. P., Pianta, R. C., Gregory, A., Mikami, A.Y., and Lun J. (2011). An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement. *Science,* 333, 1034-1037.

Dolan, L., Kellam, S. G., Brown, C. H., Werthamer-Larsson, L., Rebok, G. W., Mayer, L. S., Laudolff, J., Turkkan, J. S., Ford, C., & Wheeler, L. (1993). The short-term impact of two classroom-based preventive interventions on aggressive and shy behaviors and poor achievement. Journal of Applied Developmental Psychology, 14, 317-345.

Finkelstein, N., Hanson, T., Huang, C.-W., Hirschman, B., and Huang, M. (2010). *Effects of Problem Based Economics on high school economics instruction.* (NCEE 2010-4002). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Garet, M., Wayne, A., Stancavage, F., Taylor, J., Eaton, M., Walters, K., Song, M., Brown, S., Hurlburt, S., Zhu, P., Sepanik, S., and Doolittle, F. (2011). *Middle School Mathematics Professional Development Impact Study: Findings After the Second Year of Implementation* (NCEE 2011-4024). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

Hitchcock, J., Dimino, J., Kurki, A., Wilkins, C., and Gersten, R. (2011). *The Impact of Collaborative Strategic Reading on the Reading Comprehension of Grade 5 Students in Linguistically Diverse Schools* (NCEE 2011-4001). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.