Abstract Body

Symposium Justification (Background/Context):

This symposium presents the design and results of the Middle School Mathematics Professional Development Impact Study, a large-scale, randomized experiment examining the impact of providing a professional development (PD) program in rational number topics to seventh-grade mathematics teachers. An interim report (Garet et al. 2010) described the findings after one year of PD. The final report, which is the primary focus of this symposium, documents the impact after providing a second year of PD in a subset of the original participating districts and includes supplemental analyses that use data from both years of the study.

The proposed symposium consists of three papers, each focusing on a different aspect of the study. The first paper describes the study design and measures; the second paper describes the PD intervention that was delivered to teachers; and the third paper discusses the impact results.

Given that the methodological and implementation complexities associated with this large-scale, multi-year randomized trial will be highly relevant to SREE participants, and that the final results will be available prior to the September conference, we believe that this project warrants a 2-hour symposium. We have secured a discussant and have planned the symposium to allow for in-depth presentations and discussions focused on the study’s design and measures, PD intervention and final results.

Paper 1: Middle School Mathematics Professional Development Impact Study: Design and Measures (James Taylor, Fran Stancavage and Fred Doolittle)

Purpose / Objective / Research Question / Focus of Study:
This paper describes the design and methodology of the study. The study randomly assigned 77 mid- and high-poverty schools from 12 districts to treatment and control conditions in the first year of implementation and to 39 schools and 6 districts in the second and final year of implementation. Outcome data on teachers and students were collected during both years of the study. The PD was delivered by two provider organizations, each of which worked with half of the participating districts in each year of the study. Seventh-grade teachers in the treatment schools had the opportunity to receive the PD program offered by the study and could also continue to participate in the PD activities that they would have received in the absence of the study. Seventh-grade teachers in the control schools received only the PD that they would have received in the absence of the study.

The interim report (Garet et al. 2010), which will be briefly described and referenced in the proposed symposium, addressed three central research questions:

1. What impact did the PD program provided in this study have on teacher knowledge of rational number topics? Teacher knowledge was measured in the fall and in the spring using a specially constructed teacher knowledge test. The test was designed to measure two constructs aligned with the purpose of the PD program: knowledge of rational numbers content typically taught in seventh grade (common knowledge of mathematics,
or CK) and additional knowledge that may be useful for teaching rational number topics (specialized knowledge of mathematics for teaching, or SK). Each form was equally divided between CK and SK and equally divided between: (1) fractions and decimals and (2) ratio, rate, proportion, and percent.

2. What impact did the PD program provided in this study have on teacher instructional practices? To measure instructional practice, one classroom observation was conducted for each teacher when they were teaching rational number topics. The observations produced three primary measures of instructional practice: teacher elicits student thinking, which encompassed such behaviors as asking other students whether they agree or disagree with a student’s response; teacher uses representations, which counted the number of times the teacher explained a visual representation of mathematics, and teacher focuses on mathematical reasoning, which counted the number of times the teacher asked questions such as “Why does this procedure work?” or “Why does my answer make sense?”

3. What impact did the PD program provided in this study have on student achievement in rational number topics? A customized, computer-adaptive student achievement test was constructed for the study by a major test publisher. The test was restricted to positive rational numbers content and drew on a customized item base that contained nearly 1,200 rational numbers items.

The final report, which will be the primary focus of this symposium, was designed to answer the following two questions:

1. What cumulative impact did providing two years of the specified PD program have on teacher knowledge of rational number topics?

2. What cumulative impact did providing two years of the specified PD program have on student achievement in rational number topics?

Setting:
The study was implemented in 77 mid- and high-poverty schools from 12 districts to in the first year of implementation. A sub-sample of 39 schools and 6 districts participated in the second and final year of the study (see Table 1 in Appendix B). All of the schools in the study used Connected Mathematics, Pearson-Prentice Hall or Glencoe Mathematics as their primary seventh grade mathematics curriculum.

Population / Participants / Subjects:
All eligible teachers teaching at least one regular seventh-grade mathematics class in spring 2009 were members of the second-year teacher impact sample, and a random sample of all seventh-grade students in their regular seventh-grade mathematics classes in spring 2009 were members of the second-year student impact sample (see Table 2 in Appendix B). This definition of the teacher and student samples implies that the study is a test of the impact of mandatory PD, as opposed to PD selected by individual teachers.

As indicated in Table 2 in Appendix B, on some key characteristics, the study sample schools were statistically different from the larger pool of eligible schools. The study sample schools were less likely to be in the South and more likely to be in the Northeast region and to be in
cities rather than in urban fringe communities, towns, or rural areas. On average, they had smaller enrollments than schools in the national sample (753 students vs. 920 students) and smaller teaching staffs (48.5 FTEs vs. 54.9 FTEs). The schools in the two-year districts also were less likely than schools in the national sample to be middle schools (67 percent vs. 95 percent) and more likely to serve a combination of elementary and middle school grades (33 percent vs. 3 percent).

**Intervention / Program / Practice:**
*Please see Paper 2.*

**Research Design:**
The study used an experimental design with random assignment of schools to treatment and control conditions within each participating district. Schools remained in the same treatment condition for both years of the study. The difference in outcomes between the treatment schools and the control schools can be interpreted as the effect of the study’s PD model relative to “business as usual” in each participating district.

Midway through the first implementation year, results from the NCEE study of PD in early reading became available (see Garet et al. 2008). The results showed that although the single year of PD tested in the study had a statistically significant impact on some dimensions of teacher knowledge and instructional practice at the end of the year in which the PD was implemented, the PD did not produce a statistically significant impact on student achievement and did not produce a statistically significant impact on teachers’ knowledge, teachers’ instructional practices, or student achievement in the year following the year of the PD. That is, the impact of the PD was not sustained.

Based on this information, NCEE elected to explore the effect of extending the implementation of the seventh-grade mathematics PD to two years. Because of resource constraints, the second year of PD was offered in only half of the originally participating districts.

**Data Collection and Analysis:**
Data were collected from teachers and students in the study schools in fall and spring of the 2007–2008 and 2008–2009 school years. The two main outcome measures used in the second year of the study were constructed as follows:

- **Teacher knowledge test.** Teacher knowledge was measured for all treatment and control teachers using a test constructed specifically for the study. The test consisted of multiple-choice and short-response items that were designed to measure knowledge of rational number topics. Three alternate forms of the test were administered so that individual teachers would receive different forms (i.e., different items) at each administration. In addition to a total score, the teacher knowledge test yielded two subscores for each participant, aligned with the two types of knowledge that were targeted by the PD: common knowledge of mathematics (CK) and specialized knowledge of mathematics for teaching (SK).

- **Student achievement test.** A customized, computer-adaptive rational number test was constructed for the study by the Northwest Evaluation Association (NWEA). The NWEA
Rational Number Test was restricted to positive rational number content and drew on a customized item bank of nearly 1,200 rational number items abstracted from the larger NWEA item bank of scaled, operational mathematics items. Three Item Response Theory (IRT)-based scores were computed for each participant: a total score, a fractions and decimals score, and a ratio and proportion score.

We also surveyed treatment and control teachers to gather data on their professional backgrounds and on the amount and type of PD in mathematics they participated in during the two-year study period. Study staff obtained information on the implementation of the PD by collecting attendance records, observing the institutes and seminars, and reviewing logs maintained by coaches that recorded the nature and extent of each coach interaction with each teacher.

The basic strategy for the impact analysis was to estimate the difference in outcomes between the treatment and control groups, adjusting for the blocking used in random assignment and for teacher- and student-level covariates. Because random assignment was conducted separately within each of the six school districts participating in the second year of the study, the study comprised six separate random assignment experiments. To obtain the impact estimates, we pooled the data for all six study districts in a single analysis, treating the districts as fixed effects. Separate program impact estimates were obtained for each district and then averaged across the six districts, weighting each district’s estimate in proportion to the number of treatment schools from the district in the study sample. Findings in this report therefore represent the impact on the performance of teachers and students in the average treatment school in the 6 two-year study districts. The results do not necessarily reflect what the treatment effect would be in the wider population of districts from which those in the study were selected.

The impact estimates provide an “intent to treat” analysis of the impact of the program; that is, the estimates reflect the program impact on all teachers and students in the targeted classrooms in the study schools, even though some of those teachers and students were not present for the full duration of the study and some of the teachers did not take full advantage of the opportunity to participate in the study-provided PD even though they were present.

A common way to represent statistical precision is as a minimum detectable effect size (MDES), which is the smallest true effect that an estimator has a good chance of detecting (Bloom 1995). The second year of the study was powered to detect an effect size of 0.59 for teacher knowledge and 0.20 for student achievement.

**Findings / Results:**
*Description of the main findings with specific details.*
*Please see Paper 3.*

**Conclusions:**
*Description of conclusions, recommendations, and limitations based on findings.*
*Please see Paper 3.*

---

1 Schools, classes, and students were treated as random effects.