Three recent random assignment studies from the Institute of Education Sciences evaluated teacher professional development (PD) programs in different grades in reading and mathematics. The PD examined in the three studies emphasized building teachers’ content knowledge and their knowledge about content-specific pedagogy. The PD, which was based mostly on commercially available programs, combined summer institutes with periodic teacher meetings and coaching during the school year. Each program was substantially more intensive than the typical PD offered to teachers in study districts. All three studies examined the impact of the PD programs on teachers’ content knowledge and instructional practice, as well as their students’ achievement. The key findings from the three studies are as follows:

- **Intensive, content-focused PD improved teachers’ knowledge and some aspects of their practice.** The studies provide evidence that PD programs focused on improving teachers’ content knowledge and their knowledge about content-specific pedagogy can produce significant gains in teachers’ knowledge by the end of the year in which the PD program is implemented. The studies also provide evidence that a one-year PD program can improve some aspects of instructional practice.

- **Improving teachers’ knowledge or practice did not translate into improvements in student achievement.** None of the three studies showed a positive effect on student achievement at the end of the year that the PD was implemented, as measured by accountability tests or tests constructed specifically for the studies. The studies found that most of the measured aspects of teachers’ knowledge and practice were not associated with student achievement. The few that were had, at best, modest associations.

Federal and local governments continue to invest billions of dollars each year in professional development (PD) for teachers (U.S. Department of Education, 2014). Until recently, there has been little rigorous evidence to inform the design and delivery of these PD programs. Nevertheless, there has been growing consensus that deepening teachers’ content knowledge is an essential component of effective PD in both reading and mathematics (Ball, Thames, & Phelps, 2008; Conference Board of the Mathematical Sciences, 2012; Martin & Umland, 2008; Moats, 2009; Phelps, 2009; Wu, 2009).

Over the past decade, the Institute of Education Sciences (IES) conducted three large-scale random assignment studies of teacher PD in different grades in reading and math (Garet et al., 2008, 2010, 2016). These studies, which are the focus of this brief, reveal a common pattern of findings on the impact of intensive, content-focused PD on teaching and learning. The findings also highlight unresolved issues that future research might explore to advance our understanding and inform PD policy and practice.
Description of the PD Tested in the Three IES Studies

Exhibit 1 summarizes the key features of the PD programs tested in the three IES studies: the 2nd-Grade Reading study, the 7th-Grade Math study, and the 4th-Grade Math study.

<table>
<thead>
<tr>
<th>PD Features</th>
<th>2nd-Grade Reading Study&lt;sup&gt;a&lt;/sup&gt;</th>
<th>7th-Grade Math Study&lt;sup&gt;b&lt;/sup&gt;</th>
<th>4th-Grade Math Study&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Time span</strong></td>
<td>Summer and school year</td>
<td>Summer and school year</td>
<td>Summer and school year</td>
</tr>
<tr>
<td><strong>Total duration</strong></td>
<td>110 hours</td>
<td>68 hours</td>
<td>93 hours</td>
</tr>
<tr>
<td><strong>Components and format</strong></td>
<td>Summer institute (18 hours), school-year meetings (30 hours), and coaching (62 hours)</td>
<td>Summer institute (18 hours), school-year meetings (30 hours), and coaching (20 hours)</td>
<td>Summer institute (80 hours), school-year meetings (10 hours), and coaching with video-based feedback (3 hours)</td>
</tr>
<tr>
<td><strong>Primary emphasis</strong></td>
<td>Content knowledge</td>
<td>Pedagogical content knowledge</td>
<td>Content knowledge</td>
</tr>
<tr>
<td><strong>Secondary emphases</strong></td>
<td>Pedagogical content knowledge, connections to curriculum</td>
<td>Content knowledge, pedagogical knowledge</td>
<td>Pedagogical content knowledge</td>
</tr>
<tr>
<td><strong>Topic focus</strong></td>
<td>The five components of reading identified by the National Reading Panel (phonemic awareness, phonics, fluency, vocabulary, comprehension)</td>
<td>Rational numbers (fractions, decimals, ratio, proportion, percent)</td>
<td>K–8 mathematics (addition, subtraction, multiplication, division, fraction operations, rational numbers, linear relations, functions)</td>
</tr>
<tr>
<td><strong>Provider of summer institute and school-year meetings</strong></td>
<td>Language Essentials for Teachers of Reading and Spelling</td>
<td>America’s Choice (half of the districts) and Pearson Achievement Solutions (half of the districts)</td>
<td>Intel Math (summer institute); district-based staff trained by Mathematics Learning Community (school-year meetings)</td>
</tr>
<tr>
<td><strong>Provider of coaching</strong></td>
<td>Half-time district-based staff, trained to provide coaching in each school by the Consortium on Reading Excellence (CORE)</td>
<td>Facilitators who led the summer institute and school-year meetings</td>
<td>District-based staff who led school-year meetings, trained to provide video-based feedback by the research team</td>
</tr>
</tbody>
</table>

Notes. <sup>a</sup>The 2nd-Grade Reading study is Garet et al. (2008). The study evaluated two versions of the reading PD, one with a coaching component and one without. In this brief, we report results from the version with coaching; results were generally similar for the version without coaching.

<sup>b</sup>The 7th-Grade Math study is Garet et al. (2010). In this study, a second year of PD (46 hours) was offered in a subset of the study districts (see Garet et al., 2011). This brief focuses on results for the first year only; results for teacher knowledge and student achievement were generally similar for the second year. (Classroom instruction was not measured in the second year.)

<sup>c</sup>The 4th-Grade Math study is Garet et al. (2016).

The total amount of PD over a year varied across studies, from as few as 68 hours in the 7th-Grade Math study to as many as 110 hours in the 2nd-Grade Reading study. Still, all three PD programs were substantially more intensive than the PD teachers typically received. Compared with teachers who participated in business-as-usual PD, teachers who participated in the study PD reported having received, on average, 55 more hours of PD in the 7th-Grade Math study, 95 more hours in the 4th-Grade Math study, and 99 more hours in the 2nd-Grade Reading study.
Each PD program had three components: a summer institute, school-year meetings, and in-school coaching. In each study, the summer institute was intended to introduce new content, and the school-year meetings and coaching were intended to support teachers in putting their new knowledge into practice. In the 2nd-Grade Reading and 7th-Grade Math studies, the school-year seminars also introduced new reading and math content.

Although the PD tested in all three studies included similar structural components, the programs differed in the relative emphasis given to the components. In particular, the 4th-Grade Math study’s summer institute was appreciably longer than the summer institutes in the other two studies (80 hours versus 18 hours each), and its coaching was appreciably shorter (3 hours versus 62 hours in the 2nd-Grade Reading study and 62 hours in the 7th-Grade Math study). Furthermore, the 4th-Grade Math Study allocated 10 hours to school-year meetings, while the other two studies each allocated 30 hours.

In addition, although the PD tested in all three studies focused on teachers’ knowledge, the PD differed in the relative emphasis on teachers’ knowledge of the content they teach and their knowledge about teaching the content (pedagogical content knowledge). The 2nd-Grade Reading and 4th-Grade Math studies focused primarily on teachers’ knowledge of content. The 2nd-Grade Reading study emphasized knowledge of the five components of reading identified by the National Reading Panel (National Institute of Child Health and Human Development, 2000) as relevant to grades K–3 (for example, knowledge of the speech sounds and letter-sound correspondences), and the 4th-Grade Math study emphasized teachers’ knowledge of K–8 mathematics, not just the specific math content teachers taught. The 7th-Grade Math study also covered content, but focused more narrowly on particular content taught in grade 7 and gave greater emphasis than the other two programs to teachers’ knowledge of ways of teaching the content.

In all three studies, the summer institutes were led by external providers, listed in Exhibit 1, who had experience leading PD of similar focus and intensity in multiple school districts. The providers recruited, assigned, and trained the facilitators for the summer institutes. In most cases, the facilitators were experienced staff who had previously led PD offered by the providers.

The school-year meetings were led by the summer institute facilitators in the 2nd-Grade Reading and 7th-Grade Math studies and by district facilitators trained by an external provider in the 4th-Grade Math study. The approach taken to staffing the coaching component also differed across the three studies. In the 7th-Grade Math study, the coaching was provided by the institute facilitators. In the 2nd-Grade Reading study, district staff served as coaches and were trained by external organizations. In the 4th-Grade Math study, the coaches were also district staff, trained by the research team.

Data collected as part of the studies indicated that the PD was delivered as planned, with high fidelity, in all three studies. Teachers were offered stipends (consistent with district policy) to cover their time and were encouraged to participate by district and school staff. Perhaps as a result, teachers participated in most of the intended sessions. Across the three studies, treatment teachers, on average, attended at least 78 percent of the intended summer institutes and school-year meetings, and at least 93 percent of the planned coaching activities.
Description of the Design of the Three IES Studies

All three studies used an experimental design, with treatment (study PD) and control (business-as-usual) groups formed using random assignment. Each study’s design features, including sample and outcome measures, are summarized in Exhibit 2.

Exhibit 2. Design of the Three IES Studies

<table>
<thead>
<tr>
<th>Study Features</th>
<th>2nd-Grade Reading Study</th>
<th>7th-Grade Math Study</th>
<th>4th-Grade Math Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants:</td>
<td>2nd-Grade Reading Study</td>
<td>7th-Grade Math Study</td>
<td>4th-Grade Math Study</td>
</tr>
<tr>
<td>Districts</td>
<td>6</td>
<td>12</td>
<td>6</td>
</tr>
<tr>
<td>Schools</td>
<td>90</td>
<td>77</td>
<td>73</td>
</tr>
<tr>
<td>Teachersa</td>
<td>270</td>
<td>195</td>
<td>165</td>
</tr>
<tr>
<td>Students</td>
<td>5,530</td>
<td>11,479</td>
<td>3,677</td>
</tr>
<tr>
<td>Unit of random assignment</td>
<td>Schools (teachers in schools assigned to the treatment group were expected to participate in the study PD)</td>
<td>Schools (teachers in schools assigned to the treatment group were expected to participate in the study PD)</td>
<td>Teachers (teachers assigned to the treatment group were expected to participate in the study PD)</td>
</tr>
<tr>
<td>Measures of teacher knowledge</td>
<td>Knowledge of five components of reading, measured with a custom test producing (1) Total Score, (2) Word Subscore (phonemic awareness, phonics, fluency), and (3) Meaning Subscore (vocabulary, comprehension)c</td>
<td>Content knowledge and specialized knowledge of rational numbers, measured with a custom test producing (1) Total Score, (2) Content Knowledge Subscore, and (3) Specialized Knowledge Subscorec</td>
<td>Content knowledge of K–8 math, measured with a custom test, producing Total Scorec</td>
</tr>
<tr>
<td>Measures of instructional practice</td>
<td>Pedagogical strategies, measured once in fall and once in spring with a custom observation instrument, producing scores for frequency of (1) Explicit Instruction, (2) Independent Student Activity, and (3) Differentiated Instructiond</td>
<td>Pedagogical strategies, measured once in spring with a custom observation instrument, producing scores for frequency of (1) Elicit Student Thinking, (2) Use Representations, and (3) Math Reasoningd</td>
<td>Mathematical Quality of Instruction (MQI), measured once in fall and twice in spring with the MQI rubric, producing scores for (1) Richness of Mathematics, (2) Errors and Imprecision, and (3) Student Participation in Mathematicsd</td>
</tr>
<tr>
<td>Measures of student achievement</td>
<td>Second-grade reading achievement, measured with a district test (Terra Nova or Stanford) in five districts and a state accountability test in one district</td>
<td>Seventh-grade math achievement, measured with a custom test, yielding scores for (1) Total, (2) Fractions and Decimals, and (3) Ratio/Proportionc</td>
<td>Fourth-grade math achievement, measured with a custom test and a state accountability testc</td>
</tr>
</tbody>
</table>

Notes. aIn the 2nd-Grade Reading and 7th-Grade Math studies, which were based on school-level random assignment, all teachers present at the end of the school year were included in the analysis. In the 4th-Grade Math study, in which teachers were randomly assigned within each school, teachers were retained in the analysis if they completed all data collection activities and were in a school in which at least one other participating teacher in the opposite assignment condition also completed all data collection. There were 94 schools and 221 volunteer teachers in the 4th-Grade Math study at random assignment.

bThe timeframe shown is for the first year of the study only. The study continued for a second year in half of the districts.
The custom teacher knowledge assessments for the 2nd-Grade Reading and 7th-Grade Math studies were developed by the research team. For the 4th-Grade Math study, the custom teacher knowledge assessment was provided by the Northwest Evaluation Association (NWEA) based on specifications by the research team. The custom student tests used in the 7th-Grade Math and 4th-Grade Math studies were provided by NWEA based on specifications by the research team.

The custom observation instruments used in the 2nd-Grade Reading and 7th-Grade Math studies were developed by the research team. The Mathematical Quality of Instruction (MQI) observation rubric used in the 4th-Grade Math study was developed by Heather Hill and colleagues (Mathematics Instrument Development Group, 2013).

For each study, the research team recruited a sample of districts, each with multiple schools meeting the study’s eligibility criteria (size and poverty status) and an interest in participating. Compared with the national population, the study districts tended to be larger and have a greater percentage of students eligible for free or reduced-price lunch. In the 2nd-Grade Reading study, the research team randomly assigned schools within each participating district to the treatment condition, which included a summer institute, school-year meetings, and coaching, or to the control condition, which included whatever PD was typical in the district. All regular classroom teachers who taught second-grade reading in participating schools were included in the study. The 7th-Grade Math study used a similar design, in which schools were randomly assigned to treatment or control conditions. In the 4th-Grade Math study, the team recruited schools that had at least two fourth-grade teachers interested in participating. Within each school, one or more of the volunteer teachers was randomly assigned to the treatment condition and one or more to the control.

In each study, the research team collected data on the fidelity with which the PD program was implemented, as well as on teachers’ attendance at each PD session. Teachers’ knowledge was assessed both before and after the PD. In addition, teachers’ instructional practice was measured by observing classrooms during the school year. (In the 4th-Grade Math study, classrooms were video-recorded.) The knowledge assessments and the classroom observation rubrics were designed to be aligned with the study PD, but also to tap aspects of knowledge and practice that all teachers could be expected to demonstrate.

Finally, in each study, student achievement test scores were obtained for students in the target grades. In the 2nd-Grade Reading study, the study relied on standardized reading tests already being administered in the study districts. In the 4th- and 7th-Grade Math Studies, student achievement was measured with assessments administered for the study that focused on topics covered in the teachers’ PD. The 4th-Grade Math study also used the state mathematics assessment scores as another measure of achievement.

**Finding 1: Intensive, content-focused PD improved teachers’ knowledge and aspects of their practice**

The PD programs tested in the three IES studies produced significant gains in teachers’ knowledge by the end of the year that the PD program was implemented.

Exhibit 3 shows the impact of the PD tested in each study on teachers’ knowledge in the spring of the year in which the PD was implemented. The graph displays the impacts in the form of an improvement index. The improvement index provides an intuitive indicator of the magnitude of an impact estimate. For example, in the 4th-Grade Math study, the impact on teacher knowledge scores was 21 percentile points. This means that the knowledge score for the typical control teacher would be expected to increase by 21 percentile points (from the 50th percentile to the 71st percentile) if the teacher received the study PD.
In two of the three studies, the PD had a statistically significant impact on teachers’ knowledge in the spring of the implementation year. In the 7th-Grade Math Study, the impact of the PD on teachers’ knowledge was not statistically significant, but it was in the hypothesized direction.

**Exhibit 3. Impact of PD on Teacher Knowledge in the Three IES Studies**

![Bar chart showing impact of PD on teacher knowledge](image)

**Notes.** The improvement index is the expected change in the percentile rank for an average control teacher if the control teacher had received the PD. For example, in the 2nd-Grade Reading study, the improvement index of 15 percentile points on Total Score implies an improvement from the 50th percentile to the 65th percentile. The improvement indices shown correspond to effect sizes of 0.38, 0.39, and 0.26 standard deviations for the 2nd-Grade Reading study; 0.19, 0.02, and 0.23 standard deviations for the 7th-Grade Math study; and 0.55 standard deviations for the 4th-Grade Math study. The Word Score focused on phonemic awareness, phonics, and fluency; the Meaning Score focused on vocabulary and comprehension; the Common Content Knowledge Score focused on concepts and operations in the area of rational numbers; and the Specialized Content Knowledge Score focused on specialized mathematical knowledge for teaching, such as identifying common student errors and selecting appropriate representations and explanations.

*Impact is statistically significant at the 0.05 level, two-tailed test.

The PD programs tested in the three IES studies also produced significant improvements on some aspects of instructional practice that the PD targeted by the end of the year the PD was implemented.

Exhibit 4 shows the impact (in improvement index values) of the PD tested in each study on teachers’ instructional practice in the spring of the year in which the PD was implemented. The results show a positive, statistically significant impact on at least one of the three instructional practices in each study.

In the 2nd-Grade Reading study, the PD had a positive impact on the frequency with which teachers used “Explicit Instruction” (e.g., modeling the oral production of sounds and/or words in a decoding or spelling lesson). In the 7th-Grade Math study, the PD had a positive impact on the extent to which teachers demonstrated practices associated with “Eliciting Student Thinking” (e.g., probing a student for reasoning or justification of a solution and eliciting other students’ questions about the student’s response). In the 4th-Grade Math Study, the PD had a positive impact on the “Richness of Mathematics” observed in teachers’ math lessons (e.g., conceptual aspects of mathematics, such as the use and quality of mathematical explanations).

In summary, the PD programs examined had a significant positive impact on some but not all of the teacher knowledge and practice outcomes in each study.
Finding 2: Improving teachers’ knowledge or practice did not translate into improvements in student achievement

None of the three studies found a positive effect on student achievement by the end of the year in which the PD was implemented, as measured by accountability tests and tests constructed specifically for the studies (Exhibit 5). In the 4th-Grade Math study, there was a small statistically significant negative effect for the state test (one of the two measures of student achievement), although the result was not statistically significant in any of our sensitivity analyses. 7

Notes. The improvement index is the expected change in the percentile rank for an average control teacher if the control teacher had received the PD. For example, in the 2nd-Grade Reading study, the improvement index of 20 percentile points for “Explicit Instruction” implies an improvement from the 50th percentile to the 70th percentile. The improvement indices shown correspond to effect sizes of 0.53, 0.22, and -0.02 standard deviations for the 2nd-Grade Reading study; 0.36, 0.38, and 0.09 standard deviations for the 7th-Grade Math study; and 0.61, 0.22, and 0.19 standard deviations for the 4th-Grade Math study. Explicit Instruction refers to directly explaining the phonics patterns that are being practiced; Independent Student Activity refers to allowing students to apply what they have learned without direct teacher support; Differentiated Instruction refers to instructing individuals or small groups using tailored materials; Elicit Student Thinking refers to probing a student for reasoning or justification of a solution and eliciting other students’ questions about the student’s response; Use Representations refers to use and explanation of different representations; Math Reasoning refers to a teacher providing justification for a procedure or solution, or clarifying and extending what a student says; Richness of Mathematics refers to the conceptual aspects of math, including the use and quality of mathematical explanations; Errors and Imprecision refers to incorrect, unclear, and imprecise use of math; and Student Participation in Mathematics refers to student mathematical contributions, explanations, and reasoning.

*Impact is statistically significant at the 0.05 level, two-tailed test.
Notes. The improvement index is the expected change in the percentile rank for an average control teacher if the control teacher had received the PD. For example, in the 2nd-Grade Reading study, the improvement index of 1 percentile point on the District Test implies an improvement from the 50th percentile to the 51st percentile. The improvement indices shown correspond to effect sizes of 0.03 standard deviations for the 2nd-Grade Reading study; 0.04, 0.03, and 0.03 standard deviations for the 7th-Grade Math Study; and -0.05 and -0.06 standard deviations for the 4th-Grade Math study. *Impact is statistically significant at the 0.05 level, two-tailed test.

The PD model in all three studies assumed that improving teachers’ content knowledge and instructional practice would improve their students’ achievement. Because positive effects on student achievement were not realized, all three studies explored the premise behind content-focused PD. Specifically, they explored whether knowledge, practice, and student achievement were related to each other as hypothesized. The results in Exhibit 6 indicate that although a few measured aspects of teachers’ knowledge and instructional practice had statistically significant associations with student achievement, most had nonsignificant associations. 

In two of the studies (2nd-Grade Reading and 7th-Grade Math), teacher knowledge had a statistically significant association with student achievement (standardized regression coefficients of 0.07 in 2nd-Grade Reading and 0.05 in 7th-Grade Math). The 4th-Grade Math Study found no significant associations between teacher knowledge and student achievement (0.00 for the NWEA test and -0.02 for the state test).

In two of the studies (2nd-Grade Reading and 4th-Grade Math), one of the three instructional practices was statistically significantly associated with student achievement: “Differentiated Instruction” (0.07) and “Errors and Imprecision” (-0.21). Thus, most measured aspects of instructional practice did not show statistically significant associations with achievement. The two practices that were associated with achievement were not the practices on which the PD programs had a significant impact.
Exhibit 6. Statistically Significant Associations Between Teacher Outcomes and Student Achievement in the Three IES Studies

**Notes.** The associations shown are standardized regression coefficients, based on models controlling for student and teacher background characteristics. The specific models differed in the three studies. For more details, see Garet et al. (2008, 2010, 2016).

**Issues Needing Further Research**

Although the three IES studies differed in grade level, subject-matter focus, and the details of their designs, they produced a similar pattern of results. These results, put in the context of other recent research on PD, suggest two issues that could be more systematically explored in future studies of teacher PD and student achievement. These are not necessarily the only issues but are offered here as a starting point for further deliberation and discussion.

**Issue 1: Improving our understanding of the aspects of teacher knowledge and practice on which PD should focus**

The PD tested in the three IES studies was based on the hypothesis that improving teachers’ content knowledge and content-specific pedagogy would, in turn, boost student achievement. But the evidence suggests that teachers’ knowledge, as measured in the studies, is not strongly associated with student achievement (i.e., a standardized regression coefficient of less than 0.1). Likewise, all but one of the teachers’ instructional practices measured had an association with student achievement of less than 0.1. This suggests a potential need to
improve our understanding of the aspects of knowledge and practice that effective teachers should master. Doing so will better enable developers to design PD that focuses on improving those aspects of knowledge or practice that will most likely translate into improvements in student achievement.

Other studies have similarly found that teacher knowledge and practice in general do not appear to be strongly associated with achievement. For example, in the case of teacher knowledge, Carlisle, Kelcey, Rowan, and Phelps (2011) found an association of 0.08 between teacher knowledge and student reading comprehension in first grade, and associations of 0.02 and 0.01 in grades 2 and 3. In math, estimates of the association between teachers’ knowledge and student achievement from other studies range from 0.02 (Rockoff, Jacob, Kane, & Staiger, 2011) to 0.05 (Hill, Rowan, & Ball, 2005). In the case of instructional practice, Blazar (2015) found an association with achievement of -0.03 to -0.05 for “Errors and Imprecision” and 0.06 to 0.11 for “Richness of Mathematics.” (Note that the Blazar estimates contrast with those found in the 4th-Grade Math study, which observed an association of -0.21 for “Errors and Imprecision” and -0.04 to -0.05 for “Richness of Mathematics.”)

Although some of the estimates of association reported above were statistically significant, none appeared to be especially large in magnitude. This pattern also applies to estimates of association with achievement for more general instructional practices, not just subject-specific practices. For example, a recent study reported an association between teachers’ scores on the Classroom Assessment Scoring System (CLASS) and student achievement of 0.16 (Allen, Pianta, Gregory, Mikami, & Lun, 2011). But Kane and Staiger (2012) found a much weaker relationship between the CLASS and achievement in the Measures of Effective Teaching study (0.04 for math and 0.02 for English language arts [ELA]).

Taken together, teacher knowledge and practice, as measured in existing studies, do not appear to be strongly and consistently related to student achievement. This reinforces the need to improve our ability to identify and reliably measure the specific aspects of teacher knowledge and practice that PD should try to strengthen.

**Issue 2: Improving the impact of PD on teacher knowledge and practice**

Suppose we are unable to identify aspects of teacher knowledge and practice to target that are better associated with student achievement than those typically assessed (Issue 1). The results of the three IES studies then suggest we need to find ways to increase the impact of PD on the aspects of knowledge and practice that have been targeted, in order to have a chance of meaningfully improving student achievement. All but one of the teacher knowledge and practice associations with student achievement observed in the three IES studies was less than 0.1. Given the size of these associations, to obtain an impact on achievement of 0.1 standard deviations (the equivalent of moving a student’s test score from the 50th percentile to the 54th percentile), for example, would require an improvement in practice by about 1 standard deviation (the equivalent of moving a teacher’s practice from the 50th percentile to the 84th percentile). This required impact on practice of 34 percentile points is much larger than those observed in the three IES studies, where half of the impacts on knowledge and practice were less than 10 percentile points in those studies, and only three were 20 percentile points or higher. The largest impact was for “Richness of Mathematics” in the 4th-Grade Math study (+23 percentiles or an effect size of 0.61 standard deviations). Thus, although the PD tested in the three IES studies produced impacts on teachers’ knowledge and some aspects of practice, the impacts may have been too small to have a measurable effect on student achievement.

The three IES studies were not designed to generate definitive guidance on how to achieve larger impacts on teacher knowledge and practice. The following discussion is therefore meant only to generate hypotheses for
further exploration. For example, adding more PD time overall or more emphasis on coaching in general does not necessarily appear to be sufficient. The 7th-Grade Math study provided a second year of PD for half of the districts, and this did not yield impacts. Although the 2nd-Grade Reading study provided 62 hours of coaching and the 7th-Grade Reading study provided 20 hours, the PD in these studies did not produce larger impacts than the PD in the 4th-Grade Math Study, which provided only three hours of coaching. However, one possibility is for future studies to more systematically test the type of coaching used. In the 2nd-Grade Reading and 7th-Grade Math studies, the coaching involved a broad set of activities (e.g., coplanning, modeling instruction, and analysis of student work) with little systematic feedback on teachers’ actual instruction. In contrast, the coaching in the 4th-Grade Math Study focused on watching videos of the study teachers instructing their students and then using the videos to provide teachers with individualized feedback based on a structured observation rubric. Despite only three hours of coaching, the 4th-Grade Math Study obtained impacts on instructional practice that were at least as large as the other two studies (+23 percentiles on “Richness of Mathematics” versus +20 percentiles on “Explicit Instruction” for the 2nd-Grade Reading study and +14 percentiles on “Elicit Student Thinking” for the 7th-Grade Math study). Future studies might consider examining whether the type of coaching could have contributed to this outcome.

Other studies have provided limited evidence on the impact of PD programs that include more frequent video-based coaching using an observation rubric. For example, a recent experiment obtained a statistically significant positive impact on student achievement (+9 percentiles) for MyTeachingPartner, a video-based coaching system focused on improving secondary teachers’ general teaching practice using the CLASS observation rubric (Allen et al., 2011). Teachers in this study received video-based, individualized feedback approximately twice a month over a school year, plus a one-day workshop before and after the school year. This result suggests one possible avenue for further exploration, although it is important to reiterate that this is not the only possible avenue.

Taken together, the three IES studies suggest that existing models of PD could potentially be delivered in a way that supports teachers in improving their knowledge and practice to some degree. But the studies also indicate that the field does not yet fully understand how to ensure that teacher PD leads to measurable improvements in student learning.
12 DOES CONTENT-FOCUSED TEACHER PROFESSIONAL DEVELOPMENT WORK?

References


13 DOES CONTENT-FOCUSED TEACHER PROFESSIONAL DEVELOPMENT WORK?


American Educator, 3(3), 4–14.
Endnotes

1 The large investment in PD is one policy response to continuing concerns about students’ academic performance, as indicated by scores on domestic and international assessments. For example, 60 percent of fourth graders and 67 percent of eighth graders scored below the proficient level in mathematics on the 2015 National Assessment of Educational Progress. In reading, 64 percent of fourth graders and 66 percent of eighth graders scored below proficient (National Center for Education Statistics, 2015). The hypothesis is that PD will improve students’ academic performance by improving the quality of teaching.

2 Yoon, Duncan, Lee, Scarloss, and Shapely (2007) reviewed more than 1,300 studies of PD in mathematics, science, and English language arts, and found only nine that examined the impact of PD on student achievement and met the review’s criteria for rigorous research based on What Works Clearinghouse design standards (What Works Clearinghouse, 2014). Apart from the three IES studies summarized in this brief, we know of only two randomized trials testing PD in mathematics conducted after the Yoon review (Hammerman, Demers, & Higgins, 2015; Jacob, Hill, & Corey, 2015) and four in reading (Gersten, Dimino, Jayanthi, Newman-Gonchar, & Taylor, 2013; Matsumura, Garnie, Correnti, Bickel, & Junker, 2010; Sailors & Price, 2010).


4 For example, the average percentage of students eligible for free or reduced-price lunch was at least 66 percent in each study sample, compared with a national average of 48 percent for all U.S. public schools (Keaton, 2012).

5 In the 2nd-Grade Reading study, a third of the schools in each district were randomly assigned to a third condition, including the summer institute and school-year meetings but not the coaching (see Garet et al., 2008). This brief focuses only on the results from the full PD program that included the coaching component.

6 The improvement index can be calculated using the reported impact estimate in effect size units (d) and the standard normal cumulative distribution function (CDF). The percentile rank of the average control teacher is calculated by evaluating the CDF at 0 (which, by definition, is the standardized mean outcome for the control group). This yields a value of 0.50, meaning the average control teacher is at the 50th percentile, as we would expect. The percentile rank of an average treatment teacher is similarly calculated by evaluating the CDF at d (which is the impact estimate or, equivalently, the standardized mean outcome for the treatment group). This resulting percentile rank also can be interpreted as the percentile rank that an average control teacher would have achieved had that teacher received the treatment. For example, if d=0.2, the value of the CDF at 0.2 is 0.58. This implies that an average control teacher receiving the treatment would be at the 58th percentile, whereas an average control teacher not receiving the treatment would be at the 50th percentile. Thus, the improvement index value would be 8 percentile points. Note that the improvement index assumes that the outcome is normally distributed and that the variance of the outcome is similar between control teachers and treatment teachers. For more technical details, see What Works Clearinghouse (2014).

7 In addition to examining the impact of the PD on student achievement at the end of the implementation year, the 2nd-Grade Reading Study examined the impact of the PD a year later and found no significant impact. The 7th-Grade Math Study examined the impact on student achievement of a second year of the PD in half of the study districts and also found no significant impact.

8 For teacher knowledge, the exhibit shows only the total score, not the subscores.

9 The association between the Errors and Imprecision dimension of the Mathematical Quality of Instruction rubric and student achievement is negative, indicating that teachers with fewer errors and instances of imprecision in their math teaching had students with higher math achievement.
Although there are few rigorous studies of content-focused PD that report impacts on both content knowledge and student achievement, two such recent studies similarly found an impact on knowledge but not achievement, based on the results available to date (Hammerman et al., 2015; Jacob et al., 2015).

Kane and Staiger (2012) report the correlation of Classroom Assessment Scoring System observation scores and teachers’ value-added scores in the prior school year (0.18 for math and 0.08 for English language arts). We derived the correlation with student achievement from these estimates by multiplying them by 0.2, the approximate standard deviation of teacher value-added scores in student standard deviation units. This approximation was derived from Table 12 (page 43), which displays value-added estimates for teachers at the 25th and 75th percentiles.

If the impact of PD on practice is $a$, measured in standardized units, and the association between practice and achievement is $b$, also in standardized units, then the impact of PD on achievement operating through practice is $a*b$. If, as described in the text, $a=1$ and $b=0.1$, then the impact of PD on achievement would be $a*b=1*0.1=0.1$. This is a simplified example for illustrative purposes, and excludes interaction and indirect effects between teacher knowledge and practice.
For more information on the full studies cited, please visit:

http://ies.ed.gov/ncee/pubs/20084030

http://ies.ed.gov/ncee/pubs/20104009

and


This brief was prepared for NCEE by Michael S. Garet, Jessica Heppen, Kirk Walters, Toni Smith, and Rui Yang of American Institutes for Research under contract number ED-IES-12-C-0080; Project Officer: Thomas E. Wei.